

Lenger, Alexander; Wolf, Stephan; Goldschmidt, Nils

Article — Published Version

Choosing inequality: how economic security fosters competitive regimes

The Journal of Economic Inequality

Provided in Cooperation with:

Springer Nature

Suggested Citation: Lenger, Alexander; Wolf, Stephan; Goldschmidt, Nils (2020) : Choosing inequality: how economic security fosters competitive regimes, The Journal of Economic Inequality, ISSN 1573-8701, Springer US, New York, NY, Vol. 19, Iss. 2, pp. 315-346, <https://doi.org/10.1007/s10888-020-09472-5>

This Version is available at:

<https://hdl.handle.net/10419/288285>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Choosing inequality: how economic security fosters competitive regimes

Alexander Lenger, et al. *[full author details at the end of the article]*

Received: 8 June 2018 / Accepted: 12 October 2020 / Published online: 22 December 2020
© The Author(s) 2020

Abstract

In a novel experimental design, we study how social immobility affects the choice among distributional schemes in an experimental democracy. We design a two-period experiment in which subjects first choose a distributional scheme by majority voting (“social contract”). Then subjects engage in a competitive real-effort task to earn points. Based on production success, participants are ranked from best to worst. In combination with the initially chosen scheme, these ranks determine the final payout of the first round, leading to a pattern of societal stratification. Participants are informed individually about points and rank, before the same sequence of voting, production and payoff determination is repeated in a second round. To test the effect of social immobility on choosing distributional regimes the experiment is conducted with and without a social immobility factor, i.e. a different weighting of the two rounds. In our standard scenario, payoffs are simply added. In our “social immobility setting”, we alter the game as follows: the actual income in round 2 is calculated by adding 0.2 times the raw payoff from the second production game and 0.8 times the income from round 1. With the higher importance of round 1 success, we simulate the fact that economic movement upwards and downwards in societies (“social mobility”) is a de facto rigid constraint: high and low incomes tend to reproduce themselves. Our main findings are that in the *Equal Weight Treatment*, most groups opt for complete equality in both rounds, while in the unequal weight setting the initial choice of equality is followed by a shift to the most competitive regime. In both treatments, we observe that those performing well in round 1 tend to vote for unequal schemes in round 2, while low-performers develop an even stronger “taste for equality”. This supports a central Rawlsian idea: behind an (experimental) “veil of uncertainty”, the lack of idiosyncratic information is strong enough to let people decide as if driven by social preferences. The different group decisions in round 2 suggest that for this to happen, stakes need to be sufficiently high. To our surprise, other factors like gender, social background or real-life income have hardly any impact on unveiled decision making. We conclude that in our experimental democracy, competition based income allocation (a “market economy”) finds support only if people are sufficiently well off. Hence, increasing inequality perpetuated by social immobility is likely to undermine the general support for market-based systems.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10888-020-09472-5>) contains supplementary material, which is available to authorized users.

Keywords Decision making · Experiment · Inequality · Information · Justice · Social choice · Social contract · Social immobility · Stratification

Jel classification C91 · C92 · D31 · D63 · D71 · D81 · D90

1 Introduction

Income and wealth inequality have increased significantly since the 1980s (OECD 2008, 2011, 2015; Wolff 2002). There is a substantial amount of academic and political literature, showing that we face a period of new polarization: A significant increase in income and capital revenues at the top of the social hierarchy coincides with a pauperization and increasingly precarious conditions of a broader level of the population (Giesecke et al. 2015; Milanovic 2005; OECD 2015; Piketty 2014; Piketty and Saez 2003). The result is success concentration in markets and a so-called “Winner-Takes-All”-society (Frank and Cook 1996; Hacker and Pierson 2010). In summary, within academia and the public there is hardly another issue scholars agree on more commonly (and despite disciplinary backgrounds): social inequality is on the rise (again) (Lenger and Schumacher 2015; Piketty 2014; Stiglitz 2012).

In this paper, we present a novel experimental design that captures social inequality by incorporating the effect of social immobility: precisely speaking, we study how income immobility between two rounds in a dynamic real-effort economy affects the democratic choice of distributional regimes. Our work is based on established findings in behavioral economics, especially on inequality aversion in the lab (Blinder and Choi 1990; Camerer 2003; Güth et al. 1982; Kahneman et al. 1986), or inequality aversion (Bolton and Ockenfels 2000; Fehr and Schmidt 1999; Fischbacher and Gächter 2010; Konow 2000). However, none of these papers have accounted for the effect of social immobility on distributive justice under laboratory conditions – a gap we fill by our own experiment. Our design builds on the traditional Frohlich and Oppenheimer (Frohlich and Oppenheimer 1990; Frohlich and Oppenheimer 1992) laboratory setting, which we amplify as follows: In a two-period experiment, subjects first have to agree on a distributional scheme (“social contract”). Thereafter, we let subjects participate in a competitive real-effort task to earn points. Subsequently, these points are conveyed into ranks determining, depending on the scheme chosen, each participant’s first round endowment. Each participant then is informed about his or her success in the production phase. Then, the steps of voting, producing and payoff determination described above are repeated in a second round.

In our “standard setting”, the *Equal Weight Treatment (EWT)*, total income is determined by adding up the payoff from both rounds. In our immobility treatment, the *Unequal Weight Treatment (UWT)*, we change the game by giving different weights to both rounds: the income individuals achieve from production in round 2 is multiplied by a factor of 0.2. To this, we add 0.8 times the income of the first round. Effectively, total income is hence determined largely by individual success in round 1 simulating some form of social immobility in an experimental setting. Movements upwards or downwards on the “social ladder” are seriously inhibited in this treatment. Like in other such dynamic laboratory settings (Gächter et al. 2017), inequality arises endogenously in our experiment, allowing us to address new issues of the interplay and effects on egalitarian choice and social preferences within a stratified experimental democracy.

Results from 23 groups with 213 students suggest the following patterns: In our experimental democracy *with* social mobility (*EWT*), people opt for an egalitarian regime in both

rounds. In the immobility treatment (*UWT*), people only opt in the first round for an egalitarian regime and switch to maximum inequality in the second round. Our analysis shows that in both treatments the success (or failure) in the first round production significantly affects the voting decision in round 2: high-performers develop a tendency towards unequal schemes that favor the more productive ones, while low-performers develop an even stronger inclination towards equality that would ensure a higher income for low-level achievers. From a Rawlsian perspective (1971), we can interpret these results as a combination of veiled decision-making and changing stake sizes: not knowing how well they would perform, people tend towards complete equality, which ensures a sufficient income from round 1 for everyone. Indeed, choosing the egalitarian scheme coincided with the Rawlsian command to “maximize the worse-off position” (Rawls 1971). Lifting the veil in round 2 leads to the above-described self-selection for or against equality in line with pure self-interest. This on the other hand implies: the “Rawlsian maximin choice” in round 1 was the result of the veil, with stakes being large enough to make people concerned about low-income levels. In the *Equal Weight Treatment*, the concern in round 2 was still large enough to counterbalance the shift towards inequality by some high-performers. In the *Unequal Weight Treatment*, in contrast, the importance of round 2 was low enough so that enough voters were ready to implement the most risky option, opting for maximum competition.

In other words, contrary to reality, giving people the ability to vote directly for allocative patterns in partially veiled settings, they do only opt for competitive regimes and unequal distributions once a considerable safe income is guaranteed. Once the stakes are small and there is ‘not much to lose’ anymore, people switch and start favoring unequal distributions. Relating this back to social immobility: the actual inverse development of rising income inequality might erode public support for market based income allocation over time. Hence, the inequalities produced by a market system may in the end lead to a political abolition of the same market economy. However, as soon as a reasonable income is ensured a priori, a strong preference for unequal distributional schemes arise, thereby fostering and stabilizing competitive regimes. In fact, we observe a disintegration of long-term distributional preferences and short-term voting behavior in our experimental democracy. Consequently, inequality and its reproduction by social immobility must be of major concern for political economists.

The paper is organized as follows. Section 2 briefly reviews the literature on economic inequality and its treatment in experimental economics. In Section 3, we present the experimental design. Section 4 contains our hypothesis and main results. In Section 5, we discuss the findings and Section 6 conclude the paper. An appendix contains experimental instructions and supplementary tables and figures.

2 Literature review

2.1 Inequality in experiments

The issue of economic inequality and distributional preferences has gained new interest in economic research recently. Most prominently, a huge branch of literature on distributional preferences has been published identifying social preferences for redistribution and egalitarian regimes (for an overview, see Binmore and Shaked 2010; Bolton and Ockenfels 2000; Charness and Rabin 2002; Ellingsen et al. 2012; Fehr and Fischbacher 2002; Fehr and Schmidt 1999). Additionally, it is well documented that individuals from different countries

(Alesina and Fuchs-Schündeln 2007; Alesina and Glaeser 2004; Osberg and Smeeding 2006; Shayo 2009); different cultural backgrounds (Henrich 2000; Henrich et al. 2005; Oosterbeek et al. 2004), and different sexes (Saad and Gill 2001; Solnick 2001) differ significantly in their support for redistribution of wealth and income. Mostly, the negative effects of inequality have been addressed. For example, research shows that in terms of pay inequality – when made public and people know their ranks – lower paid workers report less job satisfaction while higher paid workers stay unaffected (Card et al. 2012; see also Cohn et al. 2014). However, also negative secondary effects of inequality have been discussed, i.e. that unequal societies are more violent, have lower quality of public goods and have a generally less satisfying standard of living (Wilkinson and Pickett 2009).

In summary, an important, established and fast growing strain of literature can be identified dealing with the issue of inequality and fairness often using experiments to test its relevance (for an overview see e.g. (Cardenas et al. 2011; Fischbacher and Gächter 2010; Gächter et al. 2013; Gerber et al. 2013; Güth and Kocher 2014; Güth et al. 2003; Kittel et al. 2015; Konow 2003; Traub et al. 2009). Altogether, behavioral economics demonstrates that on the micro level in an experimental setting, people gravitate towards a largely equal distribution, making a strong argument for egalitarian policies. At the same time, however, contradicting findings in surveys show that in large group settings people prefer unequal societies rejecting an equal distribution (Norton 2014; Norton and Ariely 2011; Norton et al. 2014). “The data suggests that when it comes to real-world distributions of wealth, people have a preference for a certain amount of inequality.” (Starmans et al. 2017) Asking what the right level of inequality is, scholars find evidence that people indeed endorse some level of basic inequality (Kiatpongsan and Norton 2014; Norton 2014; Norton and Ariely 2011). Norton summarizes the current findings in a laconic comment: “People strongly believe that the current levels of inequality are unfair, but they rarely want perfect equality. [...] People exhibit a desire for *inequality* – not too equal, but not too unequal.” (2014: 152) Starmans et al. (2017) argue that people are not concerned about inequality at all but simply care about economic unfairness. Consequently, what scholars observe is nothing else than an “equality bias in the lab” (Starmans et al. 2017).

2.2 Equity vs. efficiency

A major insight of economics is that there exists a fundamental conflict between economic efficiency and social justice. Since Arthur Okun (1975), this conflict has commonly been known as “the big tradeoff between efficiency and equity”. The argument is simple: On the one hand, markets generate efficient outcomes on the aggregated level. On the other hand, markets produce unequal results on the individual level, which are frequently perceived as being unfair.

In light of this normative issue, three types of inequality on the individual level must be distinguished: First, unequal market incomes can be perfectly compatible with distributional justice. As such, people tend to approve unequal outcomes if differences can be attributed to *effort* providing a legitimate base for discrimination. High effort leading to high incomes is a generally accepted inequality; pure luck driven inequalities on the other hand are not widely acknowledged (see e.g. Cappelen et al. 2013; Mollerstrom et al. 2015).¹ Second, the existing inequality can be ascribed to personal factors given by birth like e.g. *talents* and IQ (Lynn and Vanhanen 2006; Zagorsky 2007). Great talent may lead to high performance in markets. Just

¹ Lottery games might be an exception, but even there, people invest some money, and technically, everyone has the chance to participate or not. However, this case is of little importance for the overall argumentation.

like effort, this type of inequality is traced back to individual factors and mostly accepted among society (even if it is accepted less than inequality as a consequence of effort). Third, inequalities resulting from *social factors* like heritage or discrimination. Inequalities being attributable to such social factors usually are not considered just and are consequently not accepted. For example, being born into a high-status family and having access to certain educational resources and other opportunities is not open to everyone. If everyone does not enter the competitive field with similar starting conditions, individuals with otherwise similar talents, e.g. the equally intelligent or industrious, may realize different incomes (see Davis and Moore 1945), which then may reinforce and even widen the initial unequal starting conditions. Market based income determination then loses its approval by those negatively affected from inequality reproduction. From this type of inequality, either the claim for adjusting starting conditions *ex ante* (e.g. Sen 1999; Sen 2010) or *ex post* through income equalization (e.g. Rawls 1971) arises.

2.3 Stratification and the reproduction of social inequality

Modern societies are stratified societies characterized by social and economic inequalities (Lenger and Schumacher 2015). In fact, inequalities tend to reproduce themselves (Boudon 1974; Butler and Watt 2007; Johnson 1996; Marger 2005; Neckerman 2004; Shavit and Blossfeld 1993; Solon 2015). The observed reproduction of high and low incomes necessarily undermines market efficiency, since high incomes are not allocated according to current, but past success. In other words, reward is detached from performance. Similarly, keeping low-income earners from entering higher positions, e.g. via social exclusive mechanisms like a specific habitus of upper class members, employment bans or lack of educational qualifications, keeps potential high-performers away from productive occupations. Apart from being inefficient, this conflicts with most justice criteria, such as merit based criteria, equality of opportunities, or responsibility-sensitive theories which demand that no one should have a benefit or disadvantages due to factors beyond her or his own control, e.g. being born into a specific social stratum. Especially based on the last ideal, the above-cited studies often develop some hidden criticism concerning the “unfairness” of income inequality (Blackburn and Prandy 1997): In democratic and meritocratic societies, everyone should be able to participate in society at least to some extent (Nussbaum 2006; Sen 1999). If this is not possible, political measures are demanded, reestablishing “procedural fairness”. Hence, it is not surprising that empirical survey data shows widespread dissatisfaction with existing market-based income allocation whenever people have the impression that hard work does not pay off. The perspective of being part of “the working poor” (Shipler 2005) or of being permanently excluded from the labor market undermines the incentives to try hard and perform well (Alesina and Rodrik 1994). If the widespread feeling is that hard work – including individual investment in education – does not pay off in the market, then political participation and support for the market economy is going to drop (Solt 2008). This is especially reinforced if, at the same time, high-level income groups seem to be able to maintain their living standard and social status without productive efforts, since they have the necessary political and economic measures at their disposal (e.g. heritage, lobbying, tax evasion, and job polarization). If the belief in markets giving to the industrious and punishing the lazy erodes, the very fundamentals of widespread support for markets is at stake.²

² The “tolerance premium”—buying support from systematic market losers by welfare measures (Buchanan 1975; Wyss 2011)—might simply rise to unsustainable levels. The necessary redistribution might achieve such high levels of taxation and direct payments that there is not much space left for actual market activity.

2.4 Choosing justice: normative justification and experimental evidence

Many modern theories of justice are built on the principle of impartial decision-making (Buchanan 1975; Buchanan and Brennan 1985; Kolm 1996; Rawls 1971). While in ancient or medieval societies, individuals shared (at least group-specific) commonly accepted norms, which often rooted in religious dogma, the rise of the modern society, brought the challenge of ideological pluralism. In reply, modern philosophers like Adam Smith (2011), Jean-Jacques Rousseau (2003), Immanuel Kant (2008) presented philosophical concepts all aiming at one thing: starting from whatever individual perspective, people shall be – detached from their idiosyncratic bias – enabled to take a universal perspective, the only position from which morally relevant judgments can be made. The respective approaches suggested are known e.g. as the impartial spectator, the *volonté general*, and the categorical imperative.

John Rawls (1971) took up the idea of a hypothetical “veil of ignorance” behind which individuals should make impartial decisions on the basic structure of society. Rawls assumed the deciders behind the veil not to possess any social preferences, but to be purely selfish and, important for our later discussion, risk neutral. Briefly speaking, we are dealing with a *homo oeconomicus* character here. The crucial difference is, however, that Rawls detaches any person behind the veil completely of idiosyncratic knowledge about his or her personality, i.e. gender, race, age, talents, tastes, etc. In consequence, each person behind the veil would be identical to a generic impartial decider. Rawls concludes that this decision maker would demand (1) the largest degree of equal legal freedom for all, and then, that (2a) “primary goods” (property, income, etc.) should be distributed to the advantage of the worst off, while (2b) offices and position should be open to competition under “equality of opportunity”. These principles would find unanimous support since *any* person would support them once he or she is put behind the veil. In a similar vein, Buchanan (1975) suggests how a “more natural” veil of uncertainty would create similar, yet not identical, results in a hypothetical “constitutional stage”: When requested to introduce rules jointly for society, individuals would only agree on rather general, non-discriminatory rules, even though they are aware of individual characteristics. For Buchanan, such knowledge is of limited value only since an individual might now want certain privileges, but is uncertain if later in life, she would be disadvantaged by the same discriminating rule. The only set of rules people can *unanimously* agree on are very general, non-discriminatory ones.

However, the dispute between Rawls (1971), Harsanyi (1975) and the subsequent discussion of Binmore (1994, 1998, 2005) illustrates that there is too much left unclear to call the theoretical case settled. Consequently, Buchanan and Mathieu (1986) suggest integrating empirical evidence to resolve such a priori unclear arguments. In this line Frohlich and Oppenheimer (1990, 1992) set up an economic experiment to approximate what people would choose in an experimental democracy with joint self-government. Frohlich/Oppenheimer conducted an experiment addressing the distribution issue in a veil-approximating setting. Participants, before knowing their individual position, had to choose among different distributional schemes. They let groups of five persons unanimously decide on distributional principles. In total, 85 student groups from the U.S. and Poland participated in their test.

The experiment was structured as follows: The participants had to choose from four different income schemes, each containing five income classes ranging from low to high. Before face-to-face discussion took place, students were introduced to the main distributional

rules as discussed by Rawls (1971). Frohlich and Oppenheimer informed all participants about the four possibilities, but also allowed for the development of any other solution by the participants themselves. The four principles were

1. Maximizing average income.
2. Maximizing the average, subject to a minimum income (floor).
3. Maximizing the average, subject to a maximum distance between highest and lowest income (range).
4. The Maximin solution (corresponding to Rawls's Difference Principle).

In case group discussion and voting would not lead to consensus and people would declare unanimity impossible, a random selection of one principle would take place. In case unanimity was found, each person would be allotted one of the above income groups by chance, and a fraction of the hypothetical yearly pay was received.³ The veil aspect of the experiment, as Frohlich and Oppenheimer argue, results from the *ex ante* lack of knowledge in which position one later will end. Hence, individuals are forced to take *all* different positions into consideration when negotiating a solution.

The results from all 85 sessions were clear (Frohlich and Oppenheimer 1992: 60): a majority of 78% opted for a maximum average constrained by an income floor in the experiment. Maximizing the minimum income, in contrast, was almost never chosen (1%); neither did maximizing average with range (9%) or unconditional average maximization (12%) do well. Overall, participants were strongly concerned about guaranteeing a decent income for everyone, but still did not opt for pure egalitarianism for both efficiency and justice reasons. In summary, Frohlich and Oppenheimer provide empirical evidence that most participants reject both high-income differentials and pure equalization. This result was independent of cultural differences.

However, participants received income without any effort. Hence, the argument stands that the entire experiment is more a gamble than a decision on how to distribute market incomes. Therefore, Frohlich and Oppenheimer (1990, 1992) augmented the initial design and let students participate in a production game (counting spelling mistakes). The more points achieved in the production game, the higher the final income. Like in the initial experimental setting, the participants had to decide on potential redistribution of earned incomes after production via a tax and a transfer system. Despite the fact that redistribution was no longer affected by luck, but money had to be earned, students still decided for distributional schemes realizing a minimum income with otherwise unconstrained income maximization (Frohlich and Oppenheimer 1990).

Rawls uses the veil argument to derive social preferences from purely selfish behavior (Müller 2002). Indeed, people may possess genuinely social preferences behind an experimental veil, i.e. decisions for or against certain distributional norms are not necessarily driven by risk aversion and self-interest only but rather must be seen as a characteristic of human beings. Inequality aversion can be part of individual preferences, leading people to choose more egalitarian outcomes, even if people are generally risk neutral (cf. Kroll and Davidovitz 2003). Of course, just by putting people into a lab situation does not deprive them, unlike in Rawls' hypothetical original position, of their individual

³ That way, the hypothetical income level was coupled with actual payoffs, creating the usual incentives in economic experiments. Unlike in psychological experiments, people are not paid a flat fee for participation, but most often can affect the height of individual or group payoffs (sometimes both) by their own behavior.

preferences. Distributional norms, for example, or their knowledge of how well they do – or are most likely to do – in real society still exist in experimental settings. An engineering student is still aware of his mathematical talents and that he most probably will find a high-paying job outside the laboratory world, while the situation for an art student is expected to be different. Hence, socio-economic factors can be expected to have an impact on decision making behind the veil.

3 The experiment: choosing justice in experimental democracies with income immobility

Once the veil is used to limit the influence of individual characteristics in an experimental democracy, one may ask how income immobility affects distributional choice. In Section 2, we have shown that income immobility undermines the adjustment of markets relative to changes in individual effort. Lagged adjustment benefits high-income groups and harms low-level earners, but no single member of society is to be blamed for this. Hence, income immobility can be considered a factor beyond individual control. If people possess social preferences (Fehr and Fischbacher 2002; Fehr and Gintis 2007; Fehr and Schmidt 1999), responsibility-sensitivity would suggest that people want to neutralize this effect, independent of the positive or negative effects (producing a form of inequality aversion). However, behind an experimental veil, where people do not know how they will be affected, risk aversion might work exactly the same way: since income immobility inflates difference, accepting more inequality is rather risky since the negative effect from low incomes is even stronger. Hence, in a veiled situation, income immobility should be conducive to egalitarian choice. Without a veil, social preferences should let people want to reduce inequality, especially if individuals cannot be made responsible for differences. If people are simply risk averse and otherwise mostly selfish, lifting the veil should produce a separation: better-off individuals should try to increase inequality (benefiting from higher incomes at the upper end), while the worse-off are expected to opt for redistribution from the rich to the poor, closing the income gap.

Thus, choosing justice and egalitarian regimes in experimental settings might have two very different explanations. Unfortunately, the Frohlich-Oppenheimer design is not able to distinguish between risk aversion and social preferences, since egalitarian *choices* may be, behind the veil, *motivated* by both concerns. To overcome these shortcomings, we set up a dynamic design of one veiled and one unveiled decision round to be able to isolate these two factors. Our own experimental design builds up on both versions published by Frohlich and Oppenheimer. We used a similar income matrix (cf. Table 1), but ran a production game for the different positions. The overall results are not as dynamic and open as in a system in which income is randomly distributed and *ex post* changed by taxes and income subsidies in accordance with an *ex ante* chosen principle. The advantage, however, lies in a clear picture how the chosen redistribution affects the income levels of different income classes.

To isolate the effect inequality has on distributional choices in experimental democracies we created one income level per participant, and offered again four different distributional schemes. Based on Frohlich and Oppenheimer's initial schemes and increasing the spread but abstracting from efficiency aspects, we constructed the income matrix shown in Table 1.⁴

⁴ The slight differences in averages were the result of sticking to round numbers, and at the same time having certain differences between income classes. Those differences were often noticed by participants, but played only a minor role in discussions about which scheme to choose. Although in retrospect, we would have preferred avoiding these small differences, they were not crucial for our experiment.

Table 1 Income schemes available to deciding groups

Position	Scheme A	Scheme B	Scheme C	Scheme D
1	300.000	240.000	250.000	120.000
2	240.000	230.000	190.000	120.000
3	190.000	140.000	130.000	120.000
4	140.000	130.000	120.000	120.000
5	120.000	120.000	100.000	120.000
6	60.000	110.000	90.000	120.000
7	50.000	50.000	90.000	120.000
8	30.000	40.000	70.000	120.000
9	10.000	30.000	50.000	120.000
Average	126.666	121.111	121.000	120.000

Source: Authors' compilation

Like Frohlich and Oppenheimer (1990), we used a production game to determine income positions within a given scheme. Simulating free choice of occupation, participants either had to answer simple multiplication problems (as many as possible within 90 s), or find words in a crossword puzzle (as many as possible within 5 min). The more correct calculations respectively words found, the higher one's income class.⁵

The first round of a session comprised the following steps:

1. We presented the participants the four different distributional schemes as in Table 1.
2. The group could discuss the schemes for up to 30 min.
3. Once the group as a whole decided to end discussion, a secret vote took place. This either directly determined a scheme, or a run-off between the two most popular alternatives followed.
4. After the distributional scheme had been chosen collectively, individuals chose which production game to play, i.e. between a crossword puzzle and mathematical problems. Puzzles and problems were identical for all players.
5. After the production phase, the points were counted (and normalized between crossword puzzle and problem set) and the corresponding ranking determined. In case of identical point numbers, we tossed a coin.
6. The rank number determined the income position, i.e. the best performing students yielded position 1, the second best one position 2, etc. Together with the initially determined income scheme, the position determined the hypothetical income according to Table 1. The information on both rank and hence income was communicated privately to each participant.

In our *Equal Weight Treatment (EWT)*, this round was simply repeated. The overall income consequently was the sum of the two rounds, which was divided by 10,000 to determine the actual payoff:

$$\text{Payoff Equal Weight Treatment} : Y_{\text{real,EWT}} = (Y_1 + Y_2)/10,000 \text{ [in EUR]}$$

⁵ Each word or solution yielded one point, and we added four points to the crossword results of each student who had opted for the language task. Earlier tests had shown that this way, an average participant would yield close to identical point numbers.

To investigate the potential effect of income immobility, we recruited new participants to run a variant of this experiment—which we call *Unequal Weight Treatment (UWT)*—with the following extension: Unlike in *EWT*, the second round income counts only with 20% weight in treatment *UWT*. The remaining 80% of the income from round two is inherited from the first one. So parts of the income were transferred from round 1 to round 2 “as-if inherited”, thereby simulating the existence of income immobility (as observable for Germany and other market societies; cf. Atkinson and Bourguignon 2000; OECD 2008, 2011, 2015). Since no common measures for social mobility are available, we opted for a 0.2 social mobility factor in approximation to the Gini Coefficient of currently 0.28, intergenerational income elasticity of demand of 0.32, and the poverty rate of 0.1 (Bundesministerium für Finanzen 2017; OECD 2015).

A feature shared by both treatments is that in round 1, participants decide behind an “experimental veil”, which is lifted in round 2: Before choosing the scheme for the first time, no one knows how well he or she will perform in the production game. People might speculate on their talents and try to steer the outcome somewhat by choosing one’s preferred occupation, but as such, individuals are ignorant about their future success. Naturally, the imperfection of the veil remains in any experimental setting, but at least a very fundamental part of idiosyncratic knowledge is not available during decision-making. Compared to a pure discourse ethic environment, where people debate while having full knowledge about their capacities and situation, our setting is far closer to “ideal” veiled decision making.

In the second round, people already received information about their overall performance, since they privately know how well they performed in round one. Hence, the veil is largely lifted. Still, some uncertainty remains concerning the *future* results in round two, but it seems very unlikely for someone ranking two out of nine, for example, to fall down to position eight or nine.

Despite the lower impact of income from round 2, the second round comprises the same steps as the first one: discussion and voting, production, and information about individual success. Note that participants were free to stay with the production game type of round 1 or change it.

Total payoff in treatment *UWT* is again determined by dividing the – now weighted – sum of first and second round income by 10,000:

$$\text{Payoff } \textit{Unequal Weight Treatment} : Y_{\text{real},UWT} = (1 \cdot Y_1 + 0.8 \cdot Y_1 + 0.2 \cdot Y_2)/10,000$$

The entire sequence of one experiment session as well as the questionnaire and the quiz used for knowledge test is summarized graphically in Appendix 1. To ensure that instructions were understood by all, each participant had a few minutes to answer the quiz. Before starting a session, we went through the room and checked answers. In case a participant had answered wrong, we explained this issue again and let the participant repeat this point in her own words, making sure that at the end, everyone was fully informed about the procedure.

4 Results

4.1 Descriptive statistics and treatment randomization

4.1.1 Participants’ characteristics and treatment randomization

The experiment was conducted between February 2012 and June 2014 in Frankfurt, Freiburg and Siegen. Overall, 213 participants took part in our experiment: 108 were subjects in the *Equal Weight Treatment*, 105 in the *Unequal Weight Treatment*, forming 12 groups for the

former and 11 groups for the latter treatment. Group size was usually 9 persons, with the exception of one early 16 person group in the immobility treatment plus one where only 8 instead of 9 participants showed up. Allocation to the treatments was as follows: We scheduled several sessions (usually around 6) and distributed flyers in the in front of cafeterias. In Freiburg, we recruited at two cafeterias, one being located at the natural science/medicine campus, the other one on the social science and humanities campus. If it was, for organizational reasons, not possible to hand out recruitment flyers at both cafeterias, we alternated location on a daily base over several days. The other university had one central cafeteria for students of all disciplines. Students were only informed about possible time slots (we varied days and early, middle-of-day, and late hours) and registered, often expressing their flexibility with respect to several dates. Of course, each subject was only invited to one session. We then filled up sessions on a “first come, first serve” base and reserved sessions on an alternating base for either the *EWT* or *UWT*. As to be expected from such randomized process, we had a few sessions dominated by either “hard” or “soft” sciences, or gender-wise unbalanced, but on average, sessions turned out to be balanced. Since we sometimes could not fill all session slots (and were restricted in total number of session due to room and time constraints), we had to run three attempts until we had reached over 100 participants per treatment.

At the beginning of the experiment, we asked participants to provide the following background information, which we summarize under the categories *personal attributes*, *social background*, and *income*:

- *personal attributes*: sex, subject of studies, age, number of semesters completed at university
- *social background*: secondary and tertiary degree as well as the job of mother and father
- *income*: total disposable income per month and its sources: parents, own job, state funding, scholarship, or private loan

For social background variables, we use the coding developed by Lenger et al. (2013). It assigns zero to degrees or jobs with the lowest social status and 9 to the highest ones, as a proxy for a participant’s social status.

Table 2 provides and overview, including the above data on treatment composition, the parameter values for *EWT* und *UWT* and the aggregate values for both treatments. In summary, the average participant is 23 years old, has been studying for about 6 semesters, his or her parents have an academic background (the father’s degree is slightly higher than that of the mother), receives substantial support from his or her parents, and is likely to have her own job, too, such that the available monthly amount of money sums up to almost 700€. In terms of gender, our sample is almost balanced 1:1; about 40% are students of “hard sciences” (natural sciences, engineering, medicine, etc.), about 25% study law or economics, and the remaining 33% are enrolled in a social sciences or humanities programs.

The first step of our analysis is to test if subjects in both treatments are statistically identical. Respective test details are found in Appendix 2. In sum, the tests show that for none of the subject related variables, there is any significant difference, assuming a 5% significance level. On a 10%-level, a significant difference occurs for *age*: immobility treatment students are on average 0.7 years older than those in the *Equal Weight Treatment*. Overall, we conclude that both groups are statistically identical, hence, we can safely assume that different results for treatments *EWT* and *UWT* are due to design variation, i.e. caused by the “80:20 immobility factor”.

Table 2 Summary of treatment and participants' data

Category	Variable	Values/Categories	<i>EWT</i>	<i>UWT</i>	Both treatments
Treatment data	Number of participants		108	105	213
	Number of groups		12	11	23
Personal attributes	Gender	Proportion females	49.1%	50.5%	49.8%
	Subject	Proportion "Natural sciences/engineering/medicine, etc."	35.5%	48.6%	41.4%
		Proportion "Economics/law"	26.9%	22.9%	25.2%
		Proportion "Social sciences/humanities"	37.6%	28.6%	33.7%
	Age	Mean	22.88	23.63	23.18
		(stand. dev.)	(2.63)	(3.27)	(2.92)
	Semesters at university	Mean	5.65	6.31	5.92
		(stand. dev.)	(3.56)	(3.55)	(3.57)
Social background	Secondary degree, mother	Mean	6.48	6.58	6.51
		(stand. dev.)	(3.15)	(3.01)	(3.09)
	Tertiary degree, mother	Mean	4.06	4.43	4.20
		(stand. dev.)	(2.60)	(2.72)	(2.64)
	Job, mother	Average	4.98	5.06	5.01
		(stand. dev.)	(2.36)	(1.63)	(2.13)
	Secondary degree, father	Mean	6.82	6.72	6.80
		(stand. dev.)	(3.06)	(3.00)	(3.03)
	Tertiary degree, father	Mean	4.97	5.47	5.12
		(stand. dev.)	(2.52)	(2.35)	(2.47)
	Job, father	Mean	5.72	5.56	5.69
		(stand. dev.)	(2.14)	(2.36)	(2.17)
Income	Source of income				
	parents	Proportion "yes"	84.2%	91.5%	87.2%
	own job	Proportion "yes"	63.6%	54.9%	60.1%
	State funding scholarship	Proportion "yes"	28.7%	31.0%	29.6%
	Private loan	proportion "yes"	9.3%	11.3%	10.1%
	Available monthly amount	proportion "yes"	3.7%	5.6%	4.4%
	Available monthly amount	mean (stand. dev.)	675.94 (249.17)	699.92 (165.65)	685.45 (219.66)

4.1.2 Distributional choices in *equal* and *unequal weight* treatment

In this section, we describe the decisions made by the participants in both treatments, both on the level of individual and group level decision-making. Tables 3 and 4 summarize group data on vote distribution and the elected scheme, points generated in the productions games, and they additionally provide the ratio between lowest and highest actual payoff generated in each group, i.e. showing the real level of inequality reached in a group.

Note: The vote distribution aggregated over all groups are given both in absolute and relative terms. For example, 4:2:28:74 in the last row means that a total of 4 votes went to Scheme A, 2 to B, 28 to C, and 74 to D. This implies that Scheme A gained the support from 3.7% of all voters, Scheme B yielded a 1.9% support rate, etc.

Figure 1 illustrates how many participants (white bars) and in the end groups (grey bars) opted in total for Scheme A, B, C, or D, separated according to treatment (*EWT* and *UWT*) and round (1 and 2). In summary, we find that the decision among distribution schemes is only partly affected by the lifting of the veil but strongly affected by the treatment design. While in a world with equal weight of each round the preference is for egalitarian regimes, in an

experimental democracy with unequal weighting of rounds a shift from egalitarian to unequal schemes can be observed.

Following this, Figs. 2 and 3 cross-tabulate the decision made by all groups in round 1 and round 2 (cf. column 4 in Table 3 and column 8 in Table 4). The resulting matrix illustrates – once for *EWT*, once for *UWT* – if and how groups switch among distributional schemes from R1 to R2. In the *Equal Weight Treatment*, a maintenance of equality and only a partial movement towards moderate unequal schemes is observed. In the *Unequal Weight Treatment*, however, a movement from egalitarian regimes towards the most unequal schemes can be observed and the maintenance of egalitarian systems builds the exemption.

4.2 Choosing justice behind the veil: results from round 1

In line with Rawls and Frohlich/Oppenheimer, a strong tendency towards equality in the first round is expected in comparison to the second round. Since people decide behind the experimental veil, they stay in a situation of relative uncertainty about future outcomes in the first round, i.e. their (relative) individual success in the production games. Consequently, risk adverse individuals as well as people with social preferences and inequality aversions would choose equality. However, because the veil is revealed after the first round, people have enough individual knowledge about their relative position on which they will (at least to some extent) base their second decision for a distributional scheme. In contrast to Frohlich/Oppenheimer without an immobility factor, people are expected to choose more egalitarian schemes since their consequences are more substantial. Therefore, we can formulate:

Hypothesis 1: In both treatments, participants vote more often for complete equality (Scheme D) in round 1 than in round 2.⁶

We test this hypothesis with a χ^2 -test, in which we compare the voting decisions in round 1 against those in round 2. Table 5 summarizes individual and group choices in rounds 1 and 2 for *EWT* and *UWT*. Here, we compare “Individual votes” (upper part of Table 5) between round 1 (regular font) and round 2 (in italics) once for *EWT*, and separately for *UWT*. Figure 4 presents the p values of the respective χ^2 -test (first and second row).

With p values below 5%, we see that voting results differ significantly among rounds, both in treatment *EWT* and *UWT*. We therefore conclude: participants choose the egalitarian scheme significantly more often in R1 than in R2. This result is perfectly in line with the hypothesis that in round 1, participants decide behind an experimental Rawlsian veil: in round 1, the concern for a safe minimum makes people tend towards complete equality, while in round 2, knowledge about one’s position lets some people opt for less equality. We will find additional evidence for this interpretation in Section 4.5.

⁶ Of course, we statistically test the corresponding null hypothesis that there is no significant difference. We formulate the hypothesis as presented such that it is in accordance with our theory. We will follow this convention throughout the paper.

Table 3 Group specific results for EWT

Group ID	Round 1				Round 2				Both rounds	
	Votes for schemes, A:B:C:D	Run-off votes	Scheme chosen	Production point avg.	Votes for schemes, A:B:C:D	Run-off votes	Scheme chosen	Production point avg.	Maximum payoff: minimum payoff (€)	
C1	0:0:0:9		D	15.3	1:0:0:8		D	16.6	24.00;24.00	
C2	1:0:0:8		D	15.2	3:0:0:6		D	21.1	24.00;24.00	
C3	0:0:9:0		C	19.8	1:5:3:0		B	21.2	43.00;10.00	
C4	0:1:4:4	C:D=5:4	C	21.8	0:1:4:4	C:D=5:4	C	23.0	50.00;10.00	
C5	1:0:0:8		D	23.2	0:2:6:1		C	23.8	37.00;17.00	
C6	1:0:1:7		D	20.4	0:7:1:1		B	25.3	36.00;15.00	
C7	0:0:1:8		D	24.6	0:3:0:6		D	17.0	24.00;24.00	
C8	0:0:1:8		D	14.7	0:0:0:9		D	14.7	24.00;24.00	
C9	0:1:2:6		D	22.4	1:3:0:5		D	22.3	24.00;24.00	
C10	0:0:1:8		D	15.8	0:0:3:6		D	16.2	24.00;24.00	
C11	0:0:3:6		D	15.9	0:1:4:4	C:D=4:5	D	17.8	24.00;24.00	
C12	1:0:6:2		C	17.1	2:4:1:2 ^a	B:D=7:2	B	18.7	49.00;9.00	
All groups	4:2:28:74=3.7:1.9:25.9:68.5(%)		0:0:3:9=0:0:25:0:75.0(%)	18.9	8:26:22:52=7.4:24.0:20.3:48.1(%)		0:3:2:7=0:25:0:16.7:58.3(%)	19.8	31.92;19.08	

The vote distribution aggregated over all groups are given both in absolute and relative terms. For example, 4:2:28:74 in the last row means that a total of 4 votes went to scheme A, 2 to B, 28 to C, and 74 to D. This implies that scheme A gained the support from 3.7% of all voters, scheme B yielded a 1.9% support rate, etc

^a In this case, we let participants vote again between options A and D to determine which one would enter the run-off against B. With a 5 to 4 majority, students opted for D to compete against B

Table 4 Group specific results for UWT

Group ID	Round 1				Round 2				Both rounds	
	Votes for schemes, A:B:C:D	Scheme chosen	Production point avg.	Votes for schemes, A:B:C:D	Scheme chosen	Production point avg.	Run-off votes	Scheme chosen	Production point avg.	Maximum payoff: minimum payoff (€)
	I1	1:1:3:11	D	19.1	7:4:2:3	A	14.8	A:B = 9:7	A	14.8
I2	0:0:0:9	D	20.0	6:2:1:0	A	25.1		A	25.1	27.60:21.80
I3	1:2:0:6	D	17.6	3:1:2:3	A	24.1	A:D = 5:4	A	24.1	27.60:21.80
I4	0:0:8:1	C	20.0	0:1:2:6	D	19.0		D	19.0	47.40:11.40
I5	0:1:2:6	D	13.4	6:1:2:0	A	17.2		A	17.2	27.60:21.80
I6	0:2:2:5	D	18.0	4:2:3:0	A	18.3	A:C = 5:4	A	18.3	27.60:21.80
I7	0:0:0:8	D	13.5	3:5:0:0	B	15.1		B	15.1	26.40:22.20
I8	0:0:1:8	D	16.4	0:4:5:0	C	18.2		C	18.2	26.60:22.60
I9	0:3:6:0	C	20.8	6:0:0:3	A	22.4		A	22.4	27.60:21.80
I10	0:0:9:0	C	19.4	1:2:6:0	C	22.2		C	22.2	50.00:10.80
I11	0:0:4:5	D	17.0	5:1:2:1	A	20.0		A	20.0	27.60:21.80
All groups	2:9:35:59= 1:9:8.6:33.3:56.2(%)	0:0:3:8= 0:0:27.3:72.7(%)	17.9	43:22:23:17= 41.0:21.0:21.9:16.2(%)	7:1:2:1= 63.6:9.1:18.2:9.1 (%)	19.4			19.4	31.2:4:19.96

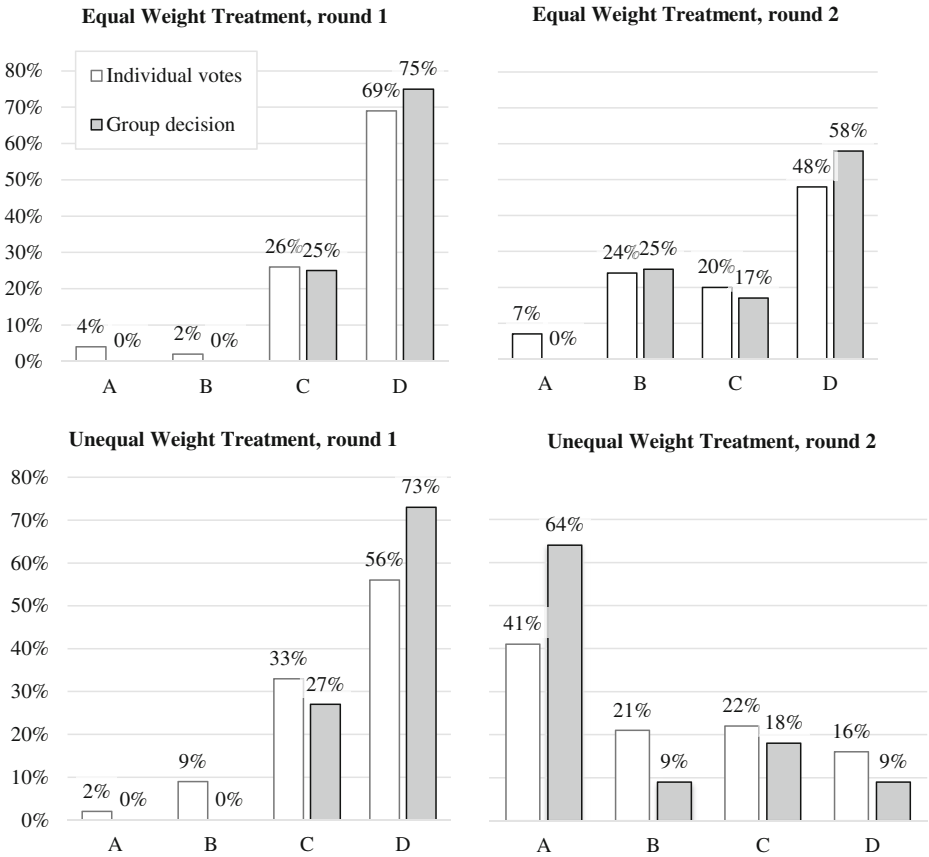


Fig. 1 Relative distribution of votes (white bars) going to Schemes A to D, disaggregated according to treatment (*Equal Weight* or *Unequal Weight*) and to Round (1 or 2). Grey bars indicate the percentage of groups having chosen Scheme A to D, respectively. Source: author’s compilation

R1 \ R2	A	B	C	D
A				
B				
C			0.17	0.08
D			0.08	0.08
				0.58

below diagonal: switch from equality to inequality
above diagonal: switch from inequality to equality

Fig. 2 Cross-tabulation of round 1 and round 2 choices in the *Equal Weight Treatment*. The figure shows that 58% of all groups (7 out of 12) chose Scheme D in both rounds. Two groups (17% of all groups) switched from Scheme C to B. One group each (8%) moved from D to B and from D to C. One group respectively opted twice for C. All ‘switching cells’ lie to the lower left of the diagonal, hence the scheme chosen in round 2 was always more unequal than the one selected in the first round

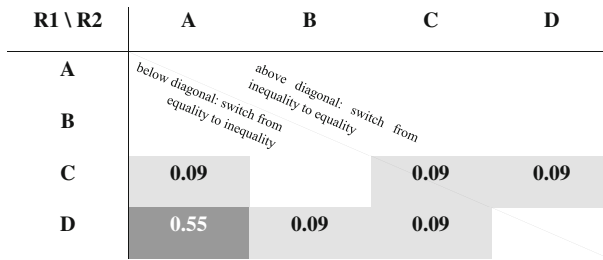


Fig. 3 Cross-tabulation of round 1 and round 2 choices in the *Unequal Weight Treatment*. The figure shows that 55% of all groups (6 out of 11) chose Scheme D in the first round and switched to Scheme A in the second round. One group each (9%) moved from D to B and from D to C. Another Group switched from C to A. Two other groups stayed with their initially chosen scheme

4.3 Dumping equality in a stratified experimental democracy: comparing results from round 1 and round 2

Looking at individual choice data of round 1 and 2 in Table 5, we see that while in round 1 participants decide similarly in both treatments (compare R1 values (regular font) between *EWT* and *UWT* rows), the round 2 voting results (compare R2 values (in italics) between *EWT* and *UWT*) seem to diverge strongly. We therefore formulate

Hypothesis 2: In round 1, individual voting results are similar in both treatments.

Hypothesis 3: In round 2, individual voting results differ significantly between both treatments: in the *Equal Weight Treatment*, equality dominates, while in the *Unequal Weight Treatment*, voters shift towards (more) inequality.

Respective χ^2 -test show that on the individual voting decision level, we must reject Hypothesis 2, but not Hypothesis 3 (at 5% significance level, see Round 1 and Round 2 columns in Fig. 4). In other words: In both rounds, *EWT* and *UWT* subjects vote differently.

Given that the choice of the distributional scheme is determined by the aggregate group decision, we test if we find the expected pattern on the level of the group choices. Correspondingly, we reformulate

Hypothesis 4: In round 1, group choices are similar in both treatments.

Hypothesis 5: In round 2, group choices differ significantly between both treatments: in the *Equal Weight Treatment*, equality dominates, while in the *Unequal Weight Treatment*, voters shift towards (more) inequality. Respective χ^2 -test for aggregate group choices are conducted

Table 5 Individual votes and groups decision in *EWT* and *UWT*, rounds 1 and 2

		A = 1	B = 2	C = 3	D = 4	Total
Individual votes						
EWT	R1 R2	4 8	2 26	27 22	75 52	108 108
UWT	R1 R2	2 41	9 23	35 25	59 16	105 105
Group choices						
EWT	R1 R2	0 0	0 3	3 2	9 7	12 12
UWT	R1 R2	0 7	0 1	3 2	8 1	11 11

		Round	
		1	2
Treatment	EWT	Individual votes in EWT.R1	Individual votes in EWT.R2
	\neq $p = 0.000$	\neq $p = 0.046$	\neq $p = 0.000$
	UWT	Individual votes in UWT.R1	Individual votes in UWT.R2
		\neq $p = 0.000$	

Fig. 4 Respective p values indicate that individual voting decisions are significantly different when we compare rounds within one treatment, but also when we compare single rounds among treatments

similar to the above tests on individual votes. The lower part of Table 5 presents groups summarizes groups choices in the same format as individual votes were given. Figure 5 presents respective p values of tests comparing group choice between rounds (treatment fixed) and treatments (round fixed) similar to Fig. 4 results on individual choices.

On the group level, we indeed cannot reject either hypothesis 4 or 5: the decision pattern in round 1 is similar in both treatments, but in round 2, equality still dominates in *Equal Weight Treatment*, while in the *Unequal Weight Treatment*, we observe a significant shift towards equality.

4.4 Distributional schemes and productive efficiency

Directly linked to the question of distributive justice and the effect of unequal societies is the issue of efficiency, which is affected by the changing incentives from different distributional norms. In our case, we would expect – in line with standard economic theory – that the egalitarian Scheme D should generate virtually no incentive to be productive, while Scheme A is expected to induce the largest degree of productive effort, with C and B ranging in between.

		Round	
		1	2
Treatment	EWT	Group choice in EWT.R1	Group choice in EWT.R2
	$=$ $p = 0.178$	$=$ $p = 0.901$	\neq $p = 0.002$
	UWT	Group choice in UWT.R1	Group choice in UWT.R2
		\neq $p = 0.003$	

Fig. 5 Respective p -values indicate that group choices in EWT and UWT are not significantly different in round 1 (Scheme D dominates in both), while round 2 results are different (EWT: D prevails, UWT: A now dominates; 1% significance level). Similarly, we find that round 1 and round 2 choices in EWT are not significantly different (Scheme D dominates in both), while in UWT, the shift from mostly D in round 1 to mostly A in round 2 is significant on a 1% level

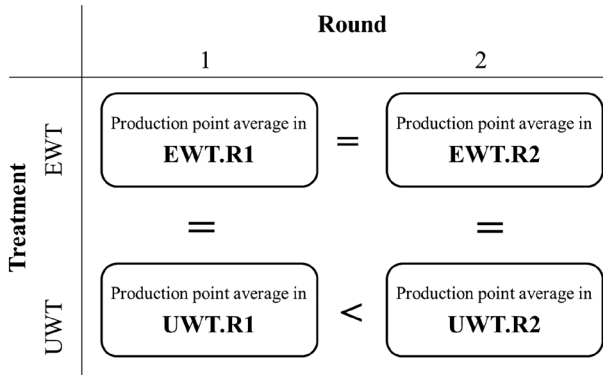


Fig. 6 Comparison of production point averages among treatments and rounds. The statistics in Appendix 4 and Table 4 results imply that in the immobility treatment, production increases from R1 to R2 (at 10% significance level). All other averages do not differ significantly (Note that no significant difference between EWT.R1 and EWT.R1, EWT.R1 and EWT.R2, and also EWT.R2 and EWT.R2 does not imply that EWT.R1 and EWT.R2 cannot also be significantly different. Slight increases from EWT.R1 to EWT.R1, then to EWT.R2, and finally to EWT.R2 might all be insignificant, while the direct difference between EWT.R1 and EWT.R2—the sum of the three smaller differences—might be large enough to be significant.)

Remember that in *EWT*, participants chose the strictly egalitarian option in both rounds, while in the immobility treatment the egalitarian choice in R1 was followed by maximum inequality in R2. Hence, we may expect that in R1 participants on average yield a similar amount of production points. In round 2, though, the more competitive scheme in *UWT* should induce more effort, and average production points therefore increase. In the *Equal Weight Treatment*, no such effect is expected.

Nevertheless, these scheme-based incentive effects are not the only possible factors driving productivity. Given the low impact of round 2 in the *Unequal Weight Treatment*, participants may also decide to exert less effort. Especially below-average performers might rationally decide not to invest too much effort given their low prospects in round 2, the disappointing feedback from round 1, and the considerable income already guaranteed. Also for high-performers, the lower impact of round 2 may lead them to invest less effort. Additionally, they may still receive an acceptable second round payoff given their own high productivity, which makes lower effort a less hazardous strategy.

Apart from incentive effects which increase or decrease effort, we can also expect a learning effect from round 1 to round 2 in both treatments, especially since the majority of individuals did not switch the production game type (*EWT*: 98% kept their “occupation”, *UWT*: 86%). This learning effect might increase productivity, i.e. the average number of production points achieved in R2, in both treatments. Given the existence of ambiguous effects on productivity – especially in the immobility treatment – on top of the usual random noise, we formulate:

Hypothesis 6: In both treatments, the average level of production points does not increase from R1 to R2.

Hypothesis 7: The production level in R1 is, statistically speaking, the same in both treatments. The same holds for R2 production levels: they do not differ among treatments.

Given that socioeconomic background data and other factors may have a systematic impact on production points, we apply regression analysis to identify potential difference in productivity.

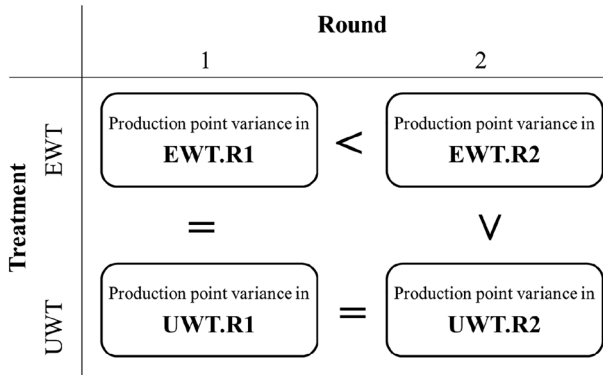


Fig. 7 Summary of comparing production point variance levels among treatments and rounds. The symbols are to be interpreted as in Figs. 4 and 5, with the “V” standing for a vertical “>” sign. The figure shows that in *EWT*, we observe an increase of production point variance from R1 to R2. In *UWT*, no such increase is measured. In round 1, there is no significant difference in variance among both treatments, whereas in round 2, the control treatment variance exceeds that of the immobility setting

One additional problem is that data was collected over a longer period (more than two years). Therefore, we include experiment session based fixed effects and cluster standard errors accordingly. The results are as well robust to other specifications.

Regression details are found in Appendix 4. Figure 6 summarizes the comparison of production point averages among rounds and treatments.

We find that there is no production effect in the *Equal Weight Treatment*. However, a production effect in the *Unequal Weight Treatment* can be observed. The production point average in round 2 is significantly higher than in round 1. Based on our test results, we reject hypothesis 6: unlike in the *EWT*, the productivity increase in the *Unequal Weight Treatment* was significant. Hypothesis 7 cannot be rejected: R1 and R2 productivity is the same in both treatments. Our main explanation for the productivity increase in the immobility treatment is the positive incentives created by Scheme A. Since average production points did not increase significantly in the *Equal Weight Treatment*, we can rule out a substantial learning effect. Due to the similarity of both treatments, any such effect should have occurred in both treatments to similar extents. This leaves the incentive effect as the main driver. The fact that productivity increased from round 1 to round 2 in the immobility treatment indicates that the motivational forces of Scheme A overcompensated the potential disincentives (Fig. 7).

In order to see whether the increase in productivity in the immobility treatment is accompanied by a drop of production effort by some participants—most likely, below-average performers—we test if the variance of production point level in the *Unequal Weight Treatment* increases from round 1 to round 2. We again compare immobility treatment values of R1 and R2 with respective *Equal Weight Treatment* results in order to capture potential effects between rounds that occur although the distributional scheme remains egalitarian. Our hypotheses within and between treatments, respectively, are:

Hypothesis 8: In the *Unequal Weight Treatment*, the variance of production point increases from R1 to R2; in the *Equal Weight Treatment*, the variance stays the same.

Hypothesis 9: In round 1, both treatments have the same variance; in round 2, the production point variance of the *Unequal Weight Treatment* exceeds that of the *Equal Weight Treatment*.

The results of the respective variance ratio tests (see Appendix 5) are summarized in Fig. 5.

As Fig. 5 shows, we need to reject both Hypothesis 8 and Hypothesis 9. The results are indeed opposite to what Hypothesis 8 suggests: it is in the *Equal Weight Treatment* where we observe an increase of the production point variance. In the *Unequal Weight Treatment*, there is no such change. Starting from similar variance levels in round 1 of both treatments, we correspondingly see that the production point variance of the *Equal Weight Treatment* exceeds that of the *Unequal Weight Treatment* in the second round.⁷

4.5 Support from winners and losers for unequal distribution

When we relate the regression results from Section 4.1 with the issue of social immobility, the most interesting effect identified is how information about round 1 production success shaped voting behavior in round 2. In short, we find that in both treatments unequal distributions obtain approval from winners *and* losers from round 1 share. More precisely, in the *Unequal Weight Treatment* we find eight supporters for Scheme A in positions 7 to 9, while in the *Equal Weight Treatment*, there is only one (even none for positions 8 and 9). Also at the upper end, support for inequality is stronger in the *Unequal Weight Treatment* than in the *Equal Weight Treatment*: looking at positions 1 to 3, we find 18 scheme-A-voters in the immobility treatment, but only 4 in the *Equal Weight Treatment*. Surprisingly, we see that in the *Equal Weight Treatment*, ranks 1 to 3 more often support Scheme A than Scheme D. Ranks 7 to 9, though, show the expected tendency towards Scheme D. In the immobility treatment, ranks 1 to 3 vote as expected with a clear preference for A. Here, it is ranks 7 to 9 producing an unexpected pattern: support for Scheme A is slightly stronger than for Scheme D (see additionally Appendix 6) (Table 6).

4.6 Veiled convergence, unveiled divergence, and the socio-economic (non-)impact

Finally, the real world effects of personal characteristics were tested. Table 7 shows the results of ordered probit-model regressions for both treatments. Since most of the six social background parameters are strongly correlated,⁸ we create an index that sums up all six social background variables (education and job of both father and mother). In case of missing values, the index has been normalized to match the final 0 to 54 points scale. Income related variables showed an intermediate degree of correlation. Hence, we excluded only those income

⁷ In both treatments, the number of points in round 2 and the *rank number* in round 1 are negatively related. Although the magnitude of the “post1” coefficient in the control treatment regression is larger than in the immobility setting regression (1.58 vs. 1.18), no significant correlation is observed. Moreover, no significant effect can be observed for the reaction to information about round 1 success (see Appendix).

⁸ This is no surprise given that partners are usually chosen from the same social stratum, explaining the correlation between social background variables of fathers and mothers. The fact that higher ranked jobs require a higher level of tertiary education, which again requires a higher secondary degree, explains why job and education variables of a single individual are strongly related. Other authors, especially in educational science, use similar composite indices as well due to the high level of correlation among single elements like mother’s and father’s education, occupation, etc. (cf. Dupriez et al. 2012; Ehmke and Siegle 2005; for an earlier, theoretical contribution, see Osborn and Morris 1979).

Table 6 Round 2 voting decisions according to position (production rank) in round 1, comparison between both treatments

Vote in Round 2 EWR UWT	Position in Round 1									Total
	1	2	3	4	5	6	7	8	9	
A	2 6	1 8	1 4	1 7	0 4	2 4	1 2	0 3	0 3	8 41
B	3 0	6 1	4 5	5 2	2 4	0 4	3 1	2 5	1 1	26 23
C	4 3	3 2	3 3	0 1	1 2	3 1	3 5	1 3	4 5	22 25
D	3 2	2 1	4 0	6 2	9 1	7 3	5 4	9 1	7 2	52 16
Total	12 11	12 12	12 12	12 12	12 11	12 12	12 12	12 12	12 11	108 105

Grey: Number of votes for scheme A in Round 2 from individuals who scored better or worse than median (position 5). In frame: Number of votes for scheme D from individuals who scored better or worse than median in the first production round

parameters that proved to be highly insignificant to avoid over-specification problems. We kept those variables that are indispensable both due to theoretical reasons and for hypothesis testing. Cronbach’s alpha of the final index is 0.879, which shows that the specification is reasonable. Except for “Job, mother”, dropping single items produces lower alpha values. We still kept “Job, mother” since, firstly, the difference was negligible (0.8798 instead of 0.8793), and secondly, because theory suggest including this parameter.

The model finally employed is:

$$voteR2_i^* = \beta_0 + \beta_1 \cdot posR1_i + \beta_2 \cdot gender_i + \beta_3 \cdot sem_i + \beta_4 \cdot nat_i + \beta_5 \cdot econlaw_i + \beta_6 \cdot socbackgr_i + \beta_7 \cdot scholarship_i + \beta_8 \cdot loan_i + \beta_9 \cdot income_i$$

$$\text{with } y_i = 1 \quad \text{if } \mu_0 < y_i^* \leq \mu_1$$

$$= 2 \quad \text{if } \mu_1 < y_i^* \leq \mu_2$$

$$= 3 \quad \text{if } \mu_2 < y_i^* \leq \mu_3$$

$$= 4 \quad \text{if } \mu_3 < y_i^* \leq \mu_4$$

So the model calculates the cut-off values μ_j between individuals voting for A, B, C, and D in round 2, respectively, with explanatory variables being the position achieved in round 1, gender, the number of semesters enrolled at university, whether a subject studies natural sciences or law/economics, the social background index value, if the subjects receives a scholarship or holds a loan and, finally, how much money she has at disposal per month. Alternative models containing more or fewer variables tended to produce similar results, which suggests that our regression results are sufficiently robust, especially concerning the importance of “posR1”, Table 7 summarizes the regression results for both EWT and UWT (for complete statistics, see Appendix 7).

Since ordered probit models do not allow a direct interpretation of the coefficient magnitudes, we calculated for both treatments the marginal effects for the average individual. This is also necessary to identify effects occurring only in subsets of our sample.

Table 8 summarizes the parameters that significantly influence the voting decision in round 2, once participants were grouped into A-voters, B-voters, etc. (by design of the ordered probit model). The complete statistics are again found in Appendix 7.

The regression analysis and marginal effects analysis indicate that our central hypothesis should not be rejected: lifting the veil causes people to sort, in accordance with their self-interest, into equality and inequality supporters; the correlation between position in round 1

Table 7 Summary of ordered probit regression results, *EWT* and *UWT*, for average individual

Variable	Equal weight treatment			Unequal weight treatment		
	Coefficient	Std. Error	p-value	Coefficient	Std. Error	p-value
Position in R1	0.145 ***	0.034	0.000	0.127 ***	0.060	0.034
Gender	-0.498 ***	0.171	0.004	-0.154	0.325	0.636
Semester at university	-0.098 ***	0.261	0.000	0.033	0.048	0.493
Student of economics or law	0.055	0.219	0.802	0.071	0.367	0.847
Student of social sciences or humanities	0.658 ***	0.202	0.001	-0.130	0.338	0.700
Social background	0.006	0.007	0.359	0.017	0.012	0.161
Scholarship receiver	0.787 **	0.307	0.010	-5.317	345.872	0.998
Loan borrower	1.232 **	0.529	0.020	-5.640	244.518	0.982
Available monthly income	-0.000	0.000	0.585	0.001 **	0.001	0.099

and voting decision in round 2 is positive and significant. This self-selection is – with one exception discussed below – most salient for those voting for A or D, but less pronounced for B- and C-voters. In the *Equal Weight Treatment*, doing one rank better increases the likelihood that a subject votes for A instead of B (a “one step” move towards maximum inequality) by 1.1 percentage points (pp). Similarly, the likelihood to choose B instead of C increases by 3.3 pp.; doing one position worse, though, makes it 5.2 pp. more probable to vote for complete equality (D) than for the more unequal option C. The effects observed in the *Unequal Weight Treatment* are similar: each position improved in production round 1 makes it 7.5 pp. more likely to vote for A instead of B; one position worse increases the likelihood to choose the more egalitarian Scheme C over the less egalitarian option B. Surprisingly, performing one rank worse *decreases* the likelihood to choose D instead of C by 4.0 pp. A possible explanation for individual and group voting behavior—including this specific anomaly—will be provided in Section 5.

Table 8 Significant marginal effects in control and immobility treatment

Vote in Round 2	Coefficient: Value			
	Equal weight treatment		Unequal weight treatment	
A	Position in R1:	-0.012**	Position in R1:	-0.075***
	Semester at univ.:	+0.007*	Loan:	+0.474*
	Scholarship receiver:	-0.033*		
	Loan:	-0.040**		
B	Position in R1:	-0.038***	-	
	Gender:	+0.119**		
	Semester at univ.:	+0.021**		
	Scholarship receiver:	-0.014		
	Loan:	-0.226***		
C	-		Position in R1:	+0.034**
			Loan:	-0.195**
			Scholarship:	-0.173**
D	Position in R1:	+0.059***	Position in R1:	-0.040**
	Gender:	-0.183**	Loan:	-0.129**
	Semester at univ.:	-0.033**	Scholarship:	-0.124**
	Loan:	+0.417***		

***, **, * = significant at 1% level, 5%, and 10% level, respectively

A gender effect can only be observed for the *Equal Weight Treatment*, where male participants show a significantly stronger inclination towards inequality. Looking at the marginal effects, we see that being male increases the probability to vote for A by 4.7 pp. and to choose B even by 13.3 pp., for each rank improved in R1. Voting for complete inequality (D) is 20.8 pp. less likely for male participants doing one position better compared to females. In the *Unequal Weight Treatment*, though, we find no significant gender effect. There, the sample size possibly was too small and the gender effect was hence obscured by the very salient “position 1” impact.

Concerning age and the number of semesters completed at university, we find that in the *Equal Weight Treatment*, the number of semesters at university and voting decision are negatively correlated, i.e. the longer someone has been at university, the more likely he or she prefers inequality. For each additional semester enrolled, we find that the likelihood to vote for A increases by 0.6 pp., that for B by 1.7 pp., while electing D is 2.6 pp. less probable.

With respect to sources of income, both “loan” and “scholarship” have significant effects for single voter categories in the immobility treatment, although no effect is measured in the general model. Individuals who (partially) finance their studies via a loan are 47.4 pp. (!) more likely to vote for the most unequal Scheme A, while voting for C or D becomes 19.5 pp. or 12.9 pp. less probable, respectively. Similarly, scholarship receivers are less prone to vote for C (17.3 pp) or D (12.4 pp). Hence, we conclude: the sources of income (a) own job, (b) parents, and (c) state funding have no effect on distributional choice; for (d) scholarship, we find a negative correlation, and in case of (e) loan, there is a negative effect.

5 Discussion

The experiment we ran allows us to investigate how social immobility affects distributional choices in an experimental democracy. Consequently, it must be mentioned that our findings are limited to this very specific experimental setting and must not be extrapolated to “real life”. Furthermore, it must be stressed that our experimental findings remain valid only for the tested distributional schemes. In this section, we discuss the potential mechanisms underlying our main results. Section 5.1 discusses results on choosing justice behind a veil. In Section 5.2, we discuss the effect of dumping equality in a stratified experimental democracy. In Section 5.3, we provide explanations for the relationship between voting behavior and productivity. In Section 5.4, we present a possible explanation for the stability and collapse of equality in the second round of our social contract democracy. Finally, in Section 5.5, we address the socio non-impact of participants’ personal characteristics.

5.1 Choosing justice

The main finding is that in our setting, participants behind an experimental veil of uncertainty (uncertainty about their individual future success, cf. Rawls 1971) vote for equality, which in our case coincides with the Rawlsian difference principle. This finding matches previous findings from behavioral economics on social contracting behind such a veil (Cappelen et al. 2013; Charness and Rabin 2002; Frohlich and Oppenheimer 1990; Frohlich and Oppenheimer 1992).

Two possible explanations for such a behavior arise. Either people opt for egalitarian regimes in such experimental settings due to some form of other-regarding preferences. In

our case, distributional preferences in the form of inequality aversion (Blinder and Choi 1990; Bolton and Ockenfels 2000; Camerer 2003; Fehr and Schmidt 1999; Fischbacher and Gächter 2010; Güth et al. 1982; Kahneman et al. 1986; Konow 2000) may provide a good explanation for the observed choice behavior. Alternatively, people opt for egalitarian regimes due to risk (and especially loss) aversion (Arrow 1965; Holt and Laury 2002; Kahneman and Tversky 1979; Rabin 2000).

5.2 From *egalitarian* to non-egalitarian choice

Lifting the veil, i.e. repeating choice and production in a second round, produced the interesting effect that people self-selected into supporting equality and inequality according to self-interest. In both treatments, support for equality decreased while unequal regimes became more popular. This shows that other-regarding preferences like inequality aversion cannot be the driver of round 1 choices. If that had been the case, the second round should have produced almost the same results as before. In the *UWT*, this shift towards inequality was so strong that the majority of groups even switched to maximum inequality. The individual change in voting behavior as well as the fact that only in the *UWT* treatment did a shift in group-choice occur can both be explained by loss aversion (Kahneman and Tversky 1979). In accordance with Rawls's explanation why a veil forces social preferences on selfish individuals (cf. Rawls 1971), we argue that people in an experimental democracy are over-proportionally concerned with *not* ending up in a very low income group under gross inequality, and put much less weight on the potential gains from ending up in a high income group. In our setting, equality is the safe option for loss averse individuals. Additionally, it might be possible that we observe a group-specific behavior. Within a group of unknown individuals, egalitarian choices are obviously socially desirable and agreeable.

However, once the veil is lifted, individuals use private information about their first round success to either support inequality (low-performers) or equality (high-performers). In the *EWT*, we observe movements of votes from D to A, but it is not strong enough to trigger a large-scale change in group choices, though majorities erode. In the *UWT*, though, the low impact of round 2 seems to make the shift towards maximum inequality in this second round worth taking a risk: even if a participant ends up on the lower end under inequality, this "fall" is cushioned by the safe income from round 1. This safety net seems to be even large enough to incentivize the *UWT* low performers from round 1 to opt for inequality in round 2 – an actually striking phenomenon. In summary, we argue that the initial choice of equality in both rounds is driven by risk aversion, and that stake size explains why loss averse individuals realize the shift towards inequality only in the immobility treatment, but not in the *Equal Weight Treatment*.

Both results stand in contrast to Frohlich and Oppenheimer (Frohlich and Oppenheimer 1992: 474) who argue that people's distributional choice in the lab is driven by concern "for the poor and weak, a desire to recognize entitlements, and sensitivity to the need for incentives to maintain productivity." Bluntly speaking, not knowing their talents and prospects, veiled participants, in their own best interest, choose the safe option: equality. However, as soon as the stakes are reduced in the second round and people have gained a substantial coverage from the first round people drop equality and choose inequality.

5.3 Inequality increases productivity

As pointed out we found no production effect in the *Equal Weight Treatment*. Interestingly, a production effect in the *Unequal Weight Treatment* can be observed. The production point

average in round 2 is significantly higher than in round 1. We conclude that in fact unequal distribution regimes foster productivity efforts, as most economists would argue. This incentive even works with very small stakes. Surprisingly, though, there is no *drop in productivity* in round 2 of the *UWT*. Exerting effort in round 1 can be rationalized, since truthful performance – assuming that there is no large number of strategic underperformers – produces relevant private information about one’s productive position in the group. In round 2, though, there is no such benefit, and we would actually expect that people rationally reduce their performance effort – in the fully rational case, down to zero effort. A part of the remaining incentive may be explained by the fact that some groups switched toward more unequal – and hence pro-competitive – schemes in round 2 (see the cells below the diagonal in Figure 2). Yet, most groups in the *EWT* kept equality in round 2, so a substantial part of effort must have been based on intrinsic motivation. This explanation is compatible with the early findings of Deutsch (1985) who argues that in most student experiments, participants seemed to derive utility from seriously fulfilling the tasks themselves, rather than remaining “rationally” passive during the course of the experiment. In short, there is a “game” component in such experiments. But as our own results and the many economic experiments show, participants still respond (strongly) to the offered monetary incentives. In our case, the findings demonstrate that the positive effects of markets generate positive outcomes only if a substantive minimum floor income is created in the first round and the risk of losing out is reduced significantly. Overall our experimental findings can be interpreted as a strong argument for having created an “experimental social market economy” in the sense that people opt for a minimum wage and an artificial “welfare state” before entering into an experimental “competitive market economy”. Consequently, our findings might be taken to argue in favor of a substantial coverage to profit from the efficiency of a market.

5.4 Deselection from winners and losers

In round 2, individuals know their round 1 production rank, so they are able to take a decision more tailored to their self-interest. Ranking better than median position 5 makes it more likely to actually achieve an income above the 12 € average. If expected gains are sufficiently higher than expected losses, even loss averse participants will prefer inequality. The better the rank, the larger the expected gains and the lower the expected losses, and hence the stronger the tendency to opt for (maximum) inequality out of rational self-interest. For the same reason, subjects ranking below the median position should find it rational to vote for equality: the closer they come to position 9, the lower the expected gains and the higher – since they are more probable – the expected losses. This tendency towards equality is driven by loss aversion: even if a risk neutral calculation would yield a positive net result from shifting towards inequality, sufficient loss aversion would let the individual still opt for equality to realize the safe income. If loss aversion is strong enough, even individuals ranking better than the median producer might find it better to vote for equality. The data above suggest exactly this interpretation. On the aggregate level, this self-selection into equality and inequality supporters after lifting the veil necessarily leads to voter polarization in round 2, which explains the drop of majority levels.

The only effect not explained is that those doing worse in round 1 of the immobility treatment are *less* likely to later vote for equality. This “outlier” might appear more plausible, though, once we have provided an explanation why in round 2, *Equal Weight Treatment*-participants largely maintained their egalitarian choice while we observed the shift presented above to the other end of the distributional spectrum in the *Unequal Weight Treatment*. Given that treatment populations

are statistically (almost) identical, the differences in round 2 voting behavior must be caused by the difference in treatment design. For *Equal Weight Treatment*-participants, both rounds are equally important concerning their effect on final payoff. According to our above argument, voting for an unequal scheme is only a rational option for those who have a good chance of benefitting. Low ranking individuals would rationally prefer equality anyways, and those close to the median position would – assuming a considerable amount of loss aversion – find equality more attractive. In sum, the rather few individuals benefitting from inequality are outnumbered by a majority that prefers a safe round 2 income over a risky gamble. Note that in Scheme A, ending up in position 9 means losing 11 Euros compared to the same position in Scheme D while the respective difference for position 1 is 18 Euros (see Table 1).

In the *Unequal Weight Treatment*, though, all losses and gains are deflated by a factor of 0.2, while round 1 income is effectively multiplied by 1.8. Although even for these far lower stakes the same logic as above should apply, we find that *Unequal Weight Treatment*-participants react differently. In some groups, the discussions before the first round of voting yielded that group members would be willing to opt for a “package solution”: a considerable “basic income” in round 1 in combination with free competition in round 2. Since all groups did not arrive at this consensus, and that even in those cases where this “social contract” was agreed on there was no effective enforcement mechanism, we can also expect – and see – the above voter polarization in the *Unequal Weight Treatment*. Yet, polarization was not as strong as in the *Equal Weight Treatment*, and there were enough middle and low position participants to support a shift towards inequality. The fact that both rounds are weighed unequally made it possible to agree on this compromise between equality and competition, as implemented by most groups in the *Unequal Weight Treatment*. This is in line with the hypothesis that people value a rightsized dose of equality: neither too little, nor too much. Additionally, the fact that everyone already had gained a substantial safe income made the “gamble” in round 2 less risky. Those who had not done too well in round 1 might have been ready for some competition to improve their situation in round 2.

It is different with the fact that losers from the first round also deselect equality in the second round. Two possible explanations arise why losers choose inequality in the second round. First, the second round is regarded as largely insignificant, leading participants to gamble; second, choosing an unequal scheme would be the only way to win back at least something. These patterns, however, could be found with winners and losers of the first round. Obviously, some behavioral attitude exists leading to an *ex post* revision of equality, leading people to choose a situation they would not choose in the beginning. Such a problem only becomes visible if the rounds – like in reality – are unequally weighted. Otherwise, the public discourse established after the second round would miss the legitimation (and a majority vote in the election) if some “losers” did not also choose inequality instead of equality.

5.5 The socio-economic non-impact

Interestingly, we found that social background has no influence on voting patterns in round 2. Neither the parents’ socio-economic parameters nor a student’s subject-specific background showed robust correlation patterns. Students of the “hard” sciences, of law or economics showed no tendency towards more equality. In summary, the outcome of the experiment is independent from socio-economic factors. One explanation for this finding could be that people are indeed motivated largely by risk-aversion. This explanation is, first of all, compatible with the observed selection into equality and inequality choosers in round 2, depending on low and high performance in round 1.

Factors showing an impact on voting in round 2 are if a participant holds a scholarship or took a loan to finance her or his studies. Concerning loans, it is possible that having borrowed a loan is a signal for meritocratic preferences, since a loan implies people being willing to engage in personal sacrifice, i.e. working hard to improve their social position via education. Yet, we would not put too much emphasis on these effects given the limited number of participants. Before finally drawing a conclusion, we would first broaden our database – especially since these findings were not robust throughout all our econometric model specifications, while “first round success” was always highly significant.

Concerning age and semester, we find that in the *Equal Weight Treatment*, the number of semesters at university and the voting decision are negatively correlated, i.e. the longer someone has been at university, the more likely he or she prefers inequality in the experimental setting. The reason could be that students are nowadays more exposed to competitive forces than, say, two or three decades ago, and over time also adapt to these norms. This would also explain why including *age* instead of *sem* does not produce an effect: it is the time being exposed to the competitive university environment that creates this effect, not growing older per se. Finally, the “semester effect” might also be explained by the fact that more advanced students are more likely to have gained experience with experiments during their studies whether as students or participants, which would represent some form of learning effect (cf. Wolf and Lenger 2014).

6 Conclusion

Overall, our data is in line with the hypothesis that under social immobility and strong uncertainty in the first round, individuals tend towards income equalization in an experimental democracy. As soon as economic security is provided, participants choose distributions that are more unequal. These findings are in line with the findings of Norton et al. that despite the desire for greater equality people are still in favor of some level of inequality in reality (Kiatpongsan and Norton 2014; Norton 2014; Norton et al. 2014).

Regarding the relevance of this paper for public policies and justice theories, a major problem of such experiments is the restricted external validity. In fact, the findings of our experiment neither can be fully extrapolated to reality nor are suited to measure society-wide preferences. However, we have argued elsewhere that given the normative implications of justice theories and behavioral ethics, it is highly relevant that the underlying assumptions are sufficiently realistic and robust (Wolf and Lenger 2014: 99). In the words of Buchanan and Mathieu (Buchanan and Mathieu 1986): [I]t would be a mistake to assume that progress toward a convergence of belief what is the most reasonable philosophical theory of justice can be achieved simply by refinements in philosophical thinking. On the contrary, it is becoming increasingly clear that philosophical disputes about justice cannot be resolved without significant contributions from the social sciences. “Such input from social sciences could be provided by empirical economic methods, above all economic experiments applied to moral questions. In experimental economics, hypotheses are systematically deduced from theory, which are then tested by an adequately designed experiment. The respective empirical results then allow for improving the theory, for example by increasing the validity of positive assumptions, which affect normative theory or political decisions. Unlike other empirical methods like descriptive survey statistics experiments enable the researcher to construct settings very close to the initial theoretical question, which seems especially interesting for political economy. Consequently, even though the experimental findings must be considered carefully acknowledging the short range, we can interpret and transfer our result this way: If people can freely choose the level of competition and hence how insecure and

unequal incomes will be, the risky market game is only accepted once a considerable safe income is guaranteed to all. Only then, a majority supports some mild income variation. For income immobility, our experimental findings might mean that the perpetuation and even aggravation of income inequality is likely to undermine the democratic support of the market system. Redistributive measures closing the widening income gap therefore should also find support from those who generally prefer markets over politically steered allocation: if the income gap reaches large levels, people might simply opt against further market allocation – be it via democratic channels, or by less peaceful means. This suspicion against pure market systems, however, is neither the result of distributional preferences or inequality aversion, but rather stems from individual self-interest.

Following our experimental results, this article argues for the systematic integration of social inequality and social immobility into experimental justice research. From an efficiency perspective, social immobility drives a wedge between marginal product and respective payment, distorting scarcity signals. Justice considerations highlight that people usually have a problem when income levels are too much detached from factors within individual control, of which income immobility is an important one. In the long-run, income immobility has the potential to undermine public support for market economies if markets are both inefficient and tend to aggravate inequalities at the expense of the already less well-off. Therefore, it seems surprising how little attention has been directed to a systematic investigation of how social immobility may adversely affect the legitimation of market economies.

Testing the effects of an unequal distribution on competition in the laboratory is an important task. Our empirical data shows that at the end of the day the majority – in contrast to their hypothetical decision-making in an original position – would choose competitive regimes on the level of action. Thus, the status quo of unequal distribution is being reproduced through the voluntary acceptance of the individuals involved. Still, it is not at all clear how such a self-selection of competition and social inequality as well as its consequences for preference formation work, but its potential impact is worth examining. Consequently, further experiments must be conducted modifying the degree of inequality to test for the effect of existing social inequality on justice perception and individual behavior.

Funding Open Access funding enabled and organized by Projekt DEAL.

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

References

- Alesina, A., Fuchs-Schündeln, N.: Good-bye Lenin (or not?): the effect of communism on people's preferences. *Am. Econ. Rev.* **97**(4), 1507–1528 (2007)
- Alesina, A., Glaeser, E.L.: *Fighting Poverty in the US and Europe*. Oxford University Press, Oxford, A World of Difference (2004)
- Alesina, A., Rodrik, D.: Distributive politics and economic growth. *Q. J. Econ.* **109**(2), 465–490 (1994)
- Arrow, K.J.: *Aspects of the Theory of Risk Bearing*. Academic Bookstores, Helsinki (1965)

- Atkinson, A. B., Bourguignon, F.: *Handbook of Income Distribution*. Elsevier, Amsterdam (2000)
- Binmore, K.: *Game Theory and the Social Contract*. Volume 1: *Playing Fair*. MIT Press, Cambridge (1994)
- Binmore, K.: *Game Theory and the Social Contract*. Volume 2: *Just Playing*. MIT Press, Cambridge (1998)
- Binmore, K.: *Natural Justice*. Oxford University Press, New York (2005)
- Binmore, K., Shaked, A.: Experimental economics: where next? *J. Econ. Behav. Organ.* **73**(1), 87–100 (2010)
- Blackburn, R.M., Prandy, K.: The reproduction of social inequality. *Sociology*. **31**(3), 491–509 (1997)
- Blinder, A.S., Choi, D.H.: A shred of evidence on theories of wage stickiness. *Q. J. Econ.* **105**(4), 1003–1016 (1990)
- Bolton, G., Ockenfels, A.: ERC: a theory of equity, reciprocity and competition. *Am. Econ. Rev.* **90**(1), 166–193 (2000)
- Boudon, R.: *Education, Opportunity, and Social Inequality*. *Changing Prospects in Western Society*. Wiley, New York (1974)
- Buchanan, J.M.: *The Limits of Liberty*. *Between Anarchy and Leviathan*. University of Chicago Press, Chicago (1975)
- Buchanan, J.M., Brennan, J.: *The Reason of Rules: Constitutional Political Economy*. Cambridge University Press, Cambridge (1985)
- Buchanan, A., Mathieu, D.: Philosophy and justice. In: Cohan, R.L. (ed.) *Justice*. *Views from the Social Sciences*, pp. 11–45. Plenum Press, New York (1986)
- Bundesministerium für Finanzen: *Einkommensungleichheit und soziale Mobilität*. *Gutachten des Wissenschaftlichen Beirats beim Bundesministerium der Finanzen*. Berlin (2017)
- Butler, T., Watt, P.: *Understanding Social Inequality*. Sage Publications, London (2007)
- Camerer, C.F.: *Behavioral Game Theory*. *Experiments in Strategic Interaction*. Princeton University Press, Princeton (2003)
- Cappelen, A.W., Konow, J., Sørensen, E.Ø., Tungodden, B.: Just luck: an experimental study of risk-taking and fairness. *Am. Econ. Rev.* **103**(4), 1398–1413 (2013)
- Card, D., Mas, A., Moretti, E., Saez, E.: Inequality at work: the effect of peer salaries on job satisfaction. *Am. Econ. Rev.* **102**(6), 2981–3003 (2012)
- Cardenas, J.C., Rodriguez, L.A., Johnson, N.: Collective action for watershed management: field experiments in Colombia and Kenya. *Environ. Dev. Econ.* **16**(3), 275–303 (2011)
- Charness, G., Rabin, M.: Understanding social preferences with simple tests. *Q. J. Econ.* **117**(3), 817–869 (2002)
- Cohn, A., Fehr, E., Götte, L.: Fair wages and effort provision: combining evidence from a choice experiment and a field experiment. *Manag. Sci.* **61**(8), 1777–1794 (2014)
- Davis, K., Moore, W.E.: Some principles of stratification. *Am. Sociol. Rev.* **10**(2), 242–249 (1945)
- Deutsch, M.: *Distributive justice: a social-psychological perspective*. Yale University Press, New Haven (1985)
- Dupriez, V., Monseur, C., van Campenhout, M., Lafontaine, D.: Social inequalities of post-secondary educational aspirations: influence of social background, school composition and institutional context. *Eur. Educ. Res. J.* **11**(4), 504–519 (2012). <https://doi.org/10.2304/eej.2012.11.4.504>
- Ehmke, T., Siegle, T.: ISEL, ISCED, HOMEPOS, ESCS—Indicators of social background for quantifying social disparity. *Z. Erziehungswiss.* **8**(4), 521–539 (2005). <https://doi.org/10.1007/s11618-005-0157-7>
- Ellingsen, T., Johannesson, M., Mollerstrom, J., Munkhammar, S.: Social framing effects: preferences or beliefs? *Games Econ. Behav.* **76**(1), 117–130 (2012)
- Fehr, E., Fischbacher, U.: Why social preferences matter. The impact of non-selfish motives on competition, cooperation and incentives. *Econ. J.* **112**(478), C1–C33 (2002)
- Fehr, E., Gintis, H.: Human motivation and social cooperation: experimental and analytical foundations. *Annu. Rev. Sociol.* **33**, 43–64 (2007)
- Fehr, E., Schmidt, K.M.: A theory of fairness, competition and cooperation. *Q. J. Econ.* **114**(3), 817–868 (1999)
- Fischbacher, U., Gächter, S.: Social preferences, beliefs, and the dynamics of free riding in public good experiments. *Am. Econ. Rev.* **100**(1), 541–556 (2010)
- Frank, R.H., Cook, P.J.: *The Winner-Take-All Society*. *Why the Few at the Top Get So Much More Than the Rest of Us*. Penguin Books, New York (1996)
- Frohlich, N., Oppenheimer, J.A.: Choosing justice in experimental democracies with production. *Am. Polit. Sci. Rev.* **84**(2), 461–477 (1990)
- Frohlich, N., Oppenheimer, J.A.: *Choosing Justice*. *An Experimental Approach to Ethical Theory*. University of California Press, Berkeley (1992)
- Gächter, S., Nosenzo, D., Sefton, M.: Peer effects in pro-social behavior: social norms or social preferences? *J. Eur. Econ. Assoc.* **11**(3), 548–573 (2013)
- Gächter, S., Mengel, F., Tsakas, E., Vostroknutov, A.: Growth and inequality in public good provision. *J. Public Econ.* **150**(1), 1–13 (2017)
- Gerber, A., Neitzel, J., Wichardt, P.C.: Minimum participation rules for the provision of public goods. *Eur. Econ. Rev.* **64**, 209–222 (2013)
- Giesecke, J., Heisig, J.P., Solga, H.: Getting more unequal: rising labor market inequalities among low-skilled men in West Germany. *Res. Soc. Stratif. Mobil.* **39**, 1–17 (2015)

- Güth, W., Kocher, M.G.: More than thirty years of ultimatum bargaining experiments: motives, variations, and a survey of the recent literature. *J. Econ. Behav. Organ.* **108**, 396–409 (2014)
- Güth, W., Schmittberger, R., Schwarze, B.: An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* **3**(4), 367–388 (1982)
- Güth, W., Kliemt, H., Ockenfels, A.: Fairness versus efficiency. An experimental study of (Mutual) gift giving. *J. Econ. Behav. Organ.* **50**(4), 465–475 (2003)
- Hacker, J.S., Pierson, P.: *Winner-Take-All Politics. How Washington Made the Rich Richer – And Turned Its Back on the Middle Class.* Simon & Schuster, New York (2010)
- Harsanyi, J.: Can the Maximin principle serve as a basis for morality? A Critique of John Rawls's Theory. *Am. Polit. Sci. Rev.* **69**(2), 594–606 (1975)
- Henrich, J.: Does culture matter in economic behavior? Ultimatum Game Bargaining among the Machiguenga of the Peruvian Amazon. *Am. Econ. Rev.* **90**(4), 973–979 (2000)
- Henrich, J., et al.: 'Economic Man' in cross-cultural perspective: behavioral experiments in 15 small-scale societies. *Behav. Brain Sci.* **28**(6), 1–61 (2005)
- Holt, C.A., Laury, S.K.: Risk aversion and incentive effects. *Am. Econ. Rev.* **92**(5), 1644–1655 (2002)
- Johnson, D.T.: *Poverty, Inequality and Social Welfare in Australia.* Physica-Verlag, Heidelberg (1996)
- Kahneman, D., Tversky, A.: Prospect theory. An analysis of decisions under risk. *Econometrica.* **47**(2), 263–291 (1979)
- Kahneman, D., Knetsch, J.L., Thaler, R.: Fairness as a constraint on profit seeking: entitlements in the market. *Am. Econ. Rev.* **76**(4), 728–741 (1986)
- Kant, I.: *Grundlegung zur Metaphysik der Sitten.* Reclam, Stuttgart (2008 [1785])
- Kiatpongsan, S., Norton, M.I.: How much (more) should CEOs make? A universal desire for more equal pay. *Perspect. Psychol. Sci.* **9**(6), 587–593 (2014)
- Kittel, B., Paetzel, F., Traub, S.: Competition, income distribution and the middle class: an experimental study. *J. Appl. Math.* **2015**, 1–15 (2015)
- Kolm, S.-C.: *Modern Theories of Justice.* MIT Press, Cambridge (1996)
- Konow, J.: Accountability and cognitive dissonance in allocation decisions. *Am. Econ. Rev.* **90**(4), 1072–1091 (2000)
- Konow, J.: Which is the fairest one of all? A positive analysis of justice theories. *J. Econ. Lit.* **41**(4), 1188–1239 (2003)
- Kroll, Y., Davidovitz, L.: Inequality aversion versus risk aversion. *Economica.* **70**(277), 19–29 (2003)
- Lenger, A., Schneickert, C., Schumacher, F.: Pierre Bourdieu's Konzeption des Habitus. In: Alexander Lenger, Christian Schneickert und Florian Schumacher (Hg.): *Pierre Bourdieu's Konzeption des Habitus. Grundlagen, Zugänge, Forschungsperspektiven.* Springer VS, S. 13–41 (2013)
- Lenger, A., Schumacher, F.: *Understanding the Dynamics of Global Inequality. Social Exclusion, Power Shift, and Structural Changes.* Springer, Heidelberg (2015)
- Lynn, R., Vanhanen, T.: *IQ and Global Inequality.* Washington Summit Publishers, Augusta (2006)
- Marger, M.: *Social Inequality. Patterns and Processes.* McGraw-Hill, Boston (2005)
- Milanovic, B.: *Worlds apart. Measuring International and Global Inequality.* Princeton University Press, Princeton (2005)
- Mollerstrom, J., Reme, B.-A., Sørensen, E.Ø.: Luck, choice and responsibility—an experimental study of fairness views. *J. Public Econ.* **131**, 33–40 (2015)
- Müller, C.: The methodology of Contractarianism in economics. *Public Choice.* **113**(3/4), 465–483 (2002)
- Neckerman, K.M. (ed.): *Social Inequality.* Sage Publications, New York (2004)
- Norton, M.I.: Inequality: who gets what and why it matters. *Policy Insights Behav. Brain Sci.* **1**(1), 151–154 (2014)
- Norton, M.I., Ariely, D.: Building a better America – one wealth quintile at a time. *Perspect. Psychol. Sci.* **6**(1), 9–12 (2011)
- Norton, M.I., Neal, D.T., Govan, C.L., Ariely, D., Holland, E.: The not-so-common-wealth of Australia: evidence for a cross-cultural desire for a more equal distribution of wealth. *Anal. Soc. Issues Public Policy.* **14**(1), 339–351 (2014)
- Nussbaum, M.C.: *Frontiers of Justice. Disability, Nationality, Species Membership.* The Belknap Press, Cambridge (2006)
- OECD: *Growing Unequal? Income Distribution and Poverty in OECD Countries.* OECD Publishing, Paris (2008)
- OECD: *Divided we Stand. Why Inequality Keeps Rising.* OECD Publishing, Paris (2011)
- OECD: *In It Together: Why Less Inequality Benefits All.* OECD Publishing, Paris (2015)
- Okun, A.: *Equality and Efficiency: the Big Tradeoff.* Brookings, Washington D.C (1975)
- Oosterbeek, H., Sloof, R., Van den Kuilen, G.: Cultural differences in ultimatum game experiments: evidence from a meta-analysis. *Exp. Econ.* **7**(2), 171–188 (2004)
- Osberg, L., Smeeding, T.: 'Fair'-inequality? attitudes towards pay differentials: the United States in comparative perspective. *Am. Sociol. Rev.* **71**(3), 450–473 (2006)
- Osborn, A.F., Morris, T.C.: The rationale for a composite index of social class and its evaluation. *Br. J. Sociol.* **30**(1), S. 39 (1979). <https://doi.org/10.2307/589500>
- Piketty, T.: *Capital in the Twenty-First Century.* Harvard University Press, Harvard (2014)

- Piketty, T., Saez, E.: Income inequality in the United States, 1913–1998. *Q. J. Econ.* **118**(1), 1–39 (2003)
- Rabin, M.: Risk aversion and expected-utility theory: a calibration theorem. *Econometrica.* **68**(5), 1281–1292 (2000)
- Rawls, J.: *A Theory of Justice*. Harvard University Press, Cambridge (1971)
- Rousseau, J.-J.: *Vom Gesellschaftsvertrag*. Reclam, Stuttgart (2003 [1762])
- Saad, G., Gill, T.: Sex differences in the ultimatum game: an evolutionary psychological perspective. *J. Bioecon.* **3**(2), 171–193 (2001)
- Sen, A.K.: *Development as Freedom*. Oxford University Press, New Delhi Oxford (1999)
- Sen, A.K.: *The Idea of Justice*. Allen Lane, London (2010)
- Shavit, Y., Blossfeld, H.-P.: *Persistent Inequality. Changing Educational Attainment in Thirteen Countries*. Westview Press, Boulder (1993)
- Shayo, M.: A model of social identity with an application to political economy: nation, class, and redistribution. *Am. Polit. Sci. Rev.* **103**(2), 147–174 (2009)
- Shipler, D. K. *The Working Poor. Invisible in America*. Vintage Books, New York (2005)
- Smith, A.: *The theory of moral sentiments*. Empire Books (2011 [1759])
- Solnick, S.J.: Gender differences in the ultimatum game. *Econ. Inq.* **39**(2), 189–200 (2001)
- Solon, G.: What Do We Know So Far about Multigenerational Mobility? National Bureau of Economic Research Working Paper 21053. Available at <http://www.nber.org/papers/w21053>. Last accessed 2 Jun 2017 (2015)
- Solt, F.: Economic inequality and democratic political engagement. *Am. J. Polit. Sci.* **52**(1), 48–60 (2008)
- Starmans, C., Sheskin, M., Bloom, P.: Why people prefer unequal societies. *Nat. Hum. Behav.* **1**(0082), (2017)
- Stiglitz, J.E.: *The Price of Inequality. How Today's Divided Society Endangers our Future*. W. W. Norton & Co, New York (2012)
- Traub, S., Seidl, C., Schmidt, U.: An experimental study on individual choice, social welfare, and social preferences. *Eur. Econ. Rev.* **53**(4), 385–400 (2009)
- Wilkinson, R.G., Pickett, K.: *The Spirit Level. Why More Equal Societies Almost Always Do Better*. Allen Lane, London (2009)
- Wolf, S., Lenger, A.: Utilitarianism, the difference principle, or else? An experimental analysis of the impact of social immobility on the democratic election of distributive rules. In: Lütge, C., Rusch, H., Uhl, M. (eds.) *Experimental Ethics. Toward an Empirical Moral Philosophy*, pp. 94–111. Palgrave Macmillan, Houndmills (2014)
- Wolff, E.N.: *Top Heavy. Increasing Inequality of Wealth in America and What Can Be Done About It*. New Press, New York (2002)
- Wyss, R.: *The Tolerance Premium as a Constitutional Element of the Protective and Welfare State*. Constitutional Economics Network Working Papers 01–2011 (2011)
- Zagorsky, J.L.: Do you have to be smart to be rich? The impact of IQ on wealth, income and financial distress. *Intelligence.* **35**(5), 489–501 (2007)

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

Affiliations

Alexander Lenger¹ · Stephan Wolf² · Nils Goldschmidt²

✉ Stephan Wolf
stephan.wolf@uni-siegen.de

Alexander Lenger
alexander.lenger@kh-freiburg.de

Nils Goldschmidt
goldschmidt@wiwi.uni-siegen.de

¹ Catholic University of Applied Sciences Freiburg, Freiburg, Germany

² Contextual Economics and Economic Education, University of Siegen, Siegen, Germany