Motivation

The legitimacy, or lack thereof, of taxing high earnings is a topic of perennial import to both students and practitioners of politics. Citizens tend to evaluate high earnings and earnings inequalities based on often implicit moral rules of thumb about both work effort and the societal practices and procedures that translate work effort, via payrates and tax regimes, into earnings outcomes (Bokemper, DeScioli, and Kline 2019; Lind 2006). “Equal pay for equal work” is a popular moral criterion in this regard—and a cornerstone of international human rights law.1 It prohibits discriminatory wage policies, but justifies wage inequalities based on performance differences. Whether this criterion is embodied in societal practices is a question of significant salience in larger debates about societal inequalities and meritocracy (Almas et al. 2010; Almas, Cappelen, and Tungodden 2017; Autor 2014; Cappelen et al. 2013; Markovits 2019).

“Merit versus luck” is another key moral criterion for the normative acceptance of wage inequalities. A long line of “luck egalitarians” in political philosophy (e.g., Dworkin 2002; Rescher 1995) and political economists (e.g., Arrow, Bowles, and Durlauf 2000; Frank 2016; Roemer 1998) has addressed its equity implications. They argue that the task is to distinguish to which degree outcome inequalities derive from luck (e.g., birth, random encounters) or merit (e.g., risk taking, hard work), then to correct the former through state intervention (as individuals are not responsible) and accept the latter as just.

Survey evidence indicates that subjective beliefs about responsibility, merit, and luck are important determinants of popular support for state redistribution (e.g., Alesina and Glaeser 2006; Esarey, Salmon, and Barrilleaux 2012b; Fong 2001; van Oorschot 2000, 2006). People who believe that income inequality or poverty result from bad luck, rather than lack of effort, are more willing to redistribute. This pattern is consistent with cumulative experimental evidence from voting on redistribution based on the Meltzer and Richard

1University of Oldenburg, Germany
2University of Southern Denmark, Denmark

Corresponding Author:
Markus Tepe, Department of Social Sciences, University of Oldenburg, Ammerländer Heerstraße 114-118, Oldenburg 26111, Germany.
Email: markus.tepe@uni-oldenburg.de
(1981, henceforth MR) model (Agranov and Palfrey 2015; Durante and Puttermann 2014; Esarey, Salmon, and Barrilleaux 2012a; Krawczyk 2010; Lefgren, Sims, and Stoddard 2016). However, these results conflate political reasoning about two very distinct mechanisms in the production of pre-tax income inequalities. Analytically, there are two stages in this process. First, individuals make efforts to deliver a performance (individual “contributions” to society). Then, societal norms, institutions, and practices, such as the prevailing cultures of esteem, remuneration, power, and discrimination (societal “reward rules”), translate those efforts, often unequally, into actual (pre-tax) earnings.

This article studies the latter stage only. After all, certain forms of persistent pay inequality (say, between genders/ethnicities within the same profession or, across professions, between hedge fund managers and surgeons) may evoke strong negative popular reactions, not because individual contributions to society are perceived to be different (“equal work”), but rather because societal reward rules are perceived to remunerate similar efforts differently (“unequal pay”). To separate out the latter element, we design a laboratory experiment in which we manipulate the payrates that translate subjects’ performance on a real-effort counting task into pre-tax income. This allows us to create an unusually clear causal belief in the minds of our subjects about whether “luck or merit” in being rewarded for effort, not effort as such, was the true reason behind earnings inequality.

There are five main steps in our procedure. First, subjects are matched in groups of three and must earn tokens in a counting task. Second, the number of tokens is then multiplied by the payrate to determine subjects’ pre-tax income: these payrates proxy the different “reward rules.” Here we distinguish three treatment conditions: payrates are (1) the same among all three group members (“equal pay”), (2) unequal (low, medium, high) and randomly distributed among the members (“luck rule”), and (3) unequal and distributed according to their performance on a learning and knowledge test (a quiz on the White House) following a week-long voluntary opportunity to prepare for the test (“merit rule”). Third, subjects are then asked to vote on a linear tax rate to redistribute the resulting earnings (flat tax system). Fourth, we explore how “merit information signals” may matter. Here we investigate whether knowing precisely how the other two group members received their pre-tax income affects tax voting (“full information”) affects tax voting, compared to having no such merit information. Finally, we then test how self-interest and self-reported political ideology affect support for redistribution under these different rules of reward.

State of the Art: Merit Reasoning and Inequality Acceptance

Beliefs about luck and merit reasoning are central to both political ideology (Gromet, Hartson, and Sherman 2015) and political behavior (Alesina and La Ferrara 2005). Outcomes that are believed to stem from one’s own performance rather than luck tend to create a sense of just entitlement and to reduce the willingness to redistribute or pay taxes. This is true even when such beliefs are false, as the sense of entitlement over one’s income often exhibits a self-serving bias (Frank 2016). Attitudinal research finds strong evidence that support for income redistribution depends on whether welfare benefit recipients are believed to be responsible for their income situation (van Oorschot 2006, 26). In a vignette survey experiment, Reetskens and van der Meer (2019) show that personal responsibility is the most important factor explaining deservingness of unemployment support. The political framing of deservingness is in turn a useful tool to legitimize welfare retrenchment (Esmark and Schoop 2017). Those who believe in success through effort and talent tend to oppose redistribution more than those who see luck and other uncontrollable factors as the main reasons for success.

Experimental research shows that support for redistribution is lower when endowments are earned through work or effort than when they are randomly assigned (Cabrales, Nagel, and Mora 2012; Durante and Puttermann 2014; Gey, Migueis, and Parsa 2017; Gill and Stone 2010, 2015; Lorenz, Paetzel, and Tepe 2017).

This article, by contrast, treats not the endowments for performance but rather the rules of reward that translate these endowments into unequal pre-tax incomes. Only few studies have addressed this topic by manipulating merit through a flat tax experimental vehicle. Lefgren, Sims, and Stoddard (2016), who do not study redistributive voting, use Amazons’ MTurk to randomly assign subjects to either a high or low payrate and either a high or low effort level in performing an online task. They find that about one-quarter of both high and low payrate subjects vote in favor of between-group redistribution when rewards are not related to effort. Bokemer, DeScioli, and Kline (2019) manipulate whether unequal pay for a text transcription task was either arbitrary or based on stable characteristics such as political party or eye color (non-meritocratic). They find that participants judge unequal pay to be less fair when it is based on a stable characteristic. Implementing a redistribution regime similar to the one utilized in this study, Esarey, Salmon, and Barrilleaux (2012b) use a within-subject design to test how reward rules affect the willingness to redistribute. They show that subjects’ decisions on redistribution are driven by rational self-interest only: the reward rule has no effect.
Our study design aims to advance this line of experimental research in a number of ways. First, since experimental studies show that support for redistribution is generally affected by whether endowments are earned or randomly assigned, we use a novel real-effort counting task (Ryvkin and Semykina 2017) to create pre-tax income. Moreover, Esarey, Salmon, and Barrilleaux (2012b, 611) proxy effort with a twenty-question multiple-choice spelling test which subjects have ninety seconds to complete. They use subjects’ performance in the first two periods of that task to allocate a high or low tax rate. Here we extend this design in two ways. First, we place subjects with either a low, medium, or high payrate into one group. And importantly, we actually use two different real-effort tasks—one to proxy merit (the quiz) and another to proxy performance (counting task). We believe this better mirrors the fact that the societal processes rewarding work through payrates are distinct from the actual contributions to society (work effort).

Second, Lefgren, Sims, and Stoddard (2016, 90) fix from the start both the amount of work to be done (the number of words each subject has to encode correctly) and the payrate subjects receive for completing a word encoding task. While this design allows them to bypass the loss of control that tournaments imply, subjects themselves have no control whatsoever over how their effort affects their payrate. By contrast, the present study relies on a tournament-type design to both allocate payrates and elicit differing levels of effort. Rather than just assigning payrates, we allow voluntary prior preparation for the learning-knowledge-quiz (Rauhut and Winter 2010) to induce a real sense of entitlement for one’s payrate (subjective feeling of merit).

### Voting on Redistribution: Theories and Hypotheses

#### Simple Model of Reward and Redistribution

We adopt a simple model of voting on redistribution (e.g., Esarey, Salmon, and Barrilleaux 2012a, 2012b; Krawczyk 2010; Lefgren, Sims, and Stoddard 2016), which imposes a flat tax on all incomes. The redistribution mechanism for N individuals with pre-tax incomes xi, . . ., x_N can be defined as: yi = (1−τ)xi + τx̄, where yi is the pre-tax income of individual i under redistributive tax rate τ and x̄ is the group’s average income. The tax revenues are distributed in equal shares among all group members. Thus, individuals with endowments exactly at x̄ are indifferent about tax levels in the self-interested standard model because their post-tax and pre-tax income remains the same regardless of the collectively agreed tax rate. The single-dimensional conflict over the size of the proportional rate is decided through majority rule, meaning the median voters preferred tax rate is implemented.

The societal reward rule is represented by the payrate factor pi. Individuals’ pre-tax income is defined as xi = pi × ci, where ci is the subject’s number of correct answers in the real-effort counting task and pi is a subject’s payrate. The payrate factor pi is either low (p_l), medium (p_m), or high (p_h). In each three-person group, there is one subject with pl, one with pm, and one with ph. Thus, even if all group members put the same effort into the real-effort task, the payrate factors pi will cause inequality in the distribution of pre-tax incomes xi.

### Merit Reasoning about Redistributive Taxation

Support for redistribution tends to be smaller when income is determined by factors perceived as justified, such as skill and effort (Krawczyk 2010, 132). The extent to which individuals are perceived as accountable for differences in economic performance is crucial for the “fairness-legitimacy” linkage (Konow 2000, 2003) Equality of opportunity is key here (Dworkin 2002; Roemer 1998).

Starting conditions ought to be equal, and success derived from effort and skills, not luck. Yet, effort and skill might be misleading indicators of entitlement to market income in the presence of unfair rules of reward that translate effort into actual income.

Under the premise of self-interested reasoning, subjects’ tax vote will be based solely on whether their pre-tax income is above or below x̄. Subjects’ payrate factor pi, subjects’ effort ci, and the mechanisms through which the payrate factor pi is allocated are irrelevant for rational and fully egoistic subjects. Moral reasoning, in contrast, predicts that subjects take into account the cause of inequality in the distribution of pre-tax incomes xi. If pre-tax income inequality stems from unfairness in the pre-tax income-generating process, redistributive taxation provides a means to compensate for unfairness.

To test the compensation argument, we manipulate the allocation of pi. The payrate factor pi is either allocated based on subjects’ performance in a learning and knowledge task or allocated randomly. The random allocation of pi created unfairness in the pre-tax income-generating process. It is a matter of luck which pi a group member receives. Allocating pi on the basis of subjects’ performance in a learning and knowledge task with voluntary preparation opportunity may make differences in the income-generating process justifiable. If the different payrate factors that cause inequality are distributed based on performance in a learning and knowledge task (earned), there is no violation of fairness norms in the income-generating process and therefore no need for compensation through redistribution. If the different payrate factors, however, are randomly distributed, redistribution is considered to be a mechanism with which to
compensate for a low pre-tax income due to a low payrate. Our main treatment hypothesis therefore stipulates:

**Hypothesis 1a (H1a)—Direct merit effect (full information treatment):** If the assignment of \( p_i \) is based on merit, this will lead to a lower voted tax rate \( \tau \) than if there is a random assignment of \( p_i \).

H1a is based on the presumption that subjects are fully aware of who in their group got \( p_{i*} \), \( p_m \), and \( p_h \). We expect such “full merit knowledge” to be the mechanism that triggers “merit reasoning” as a salient, possibly overriding, driver of tax thinking. By the same token, however, in the absence of full merit knowledge, we expect the salience of merit reasoning to be much weaker. As part of the robustness analysis, we therefore conducted the random and earned treatment again, this time without giving subjects any information on \( p_i \) or \( c_i \). That is, subjects here do not know who in their group had which payrate and who was lazy or hardworking in the real-effort task.

In other words, in this no-information treatment, any group member knows how well he or she performed himself or herself and observes pre-tax income of the other members, but he or she is in the dark as to how the other group members earned their pre-tax income. We expect this information manipulation to help illuminate the mechanisms (clear signals of deservingness) that produce the “merit reasoning.” In the case of self-interest, knowing that one’s pre-tax income is also likely to matter for tax voting. However, knowledge about how other people obtained their pre-tax income is also likely to matter for tax voting. When such “merit knowledge” is available, self-interested and ideological reasoning could then be conditioned by “merit reasoning.” In the case of self-interest, knowing that others made efforts to merit their income (rather than being lucky) may represent an additional normative price (or tax) on selfishness. When the individual stakes from redistribution are high (a lot to be gained or lost), such merit reasoning is unlikely to outweigh or cancel out self-interest reasoning. But with more moderate stakes, merit reasoning might lower support for redistribution. We therefore hypothesize:

**Hypothesis 2a (H2a)—Direct self-interest effect:** The relative gain from redistribution has a positive effect on the voted tax rate \( \tau \).

**Hypothesis 3a (H3a)—Direct political ideology effect:** More leftist self-reported ideology will lead to higher voted tax rates \( \tau \).

However, knowledge about how other people obtained their pre-tax income is also likely to matter for tax voting. When such “merit knowledge” is available, self-interested and ideological reasoning could then be conditioned by “merit reasoning.” In the case of self-interest, knowing that others made efforts to merit their income (rather than being lucky) may represent an additional normative price (or tax) on selfishness. When the individual stakes from redistribution are high (a lot to be gained or lost), such merit reasoning is unlikely to outweigh or cancel out self-interest reasoning. But with more moderate stakes, merit reasoning might lower support for redistribution. We therefore posit:

**Hypothesis 2b (H2b)—Conditional self-interest effect:** Self-interest motives for voting on tax rate \( \tau \) will be weakened (but not neutralized) when stakes are moderate, but only in the merit condition and not in the random condition.

**Standard Drivers of Redistributive Beliefs: Self–Interest and Ideology**

Self-interest and political ideology are widely seen as strong drivers of support for income redistribution. Self-interested reasoning predicts that the more subjects’ pre-tax income is above \( \bar{x} \), the lower will be the \( \tau \) they vote for, and vice versa. Political ideology, as conceptualized on a left-right dimension, posits that leftwing voters prefer higher taxes and more equality and rightwingers greater individual responsibility and economic freedom and lower taxes (Bobbio 1996; Feldman and Zaller 1992; Jacoby 1994; Mair 2007). There are few experimental studies of the behavioral consequences of ideology on redistributive preferences. Esarey, Salmon, and Barrilleaux (2012b) show that votes for a redistributive tax are almost entirely in accordance with self-interest, not partisan ideology (also see Esarey, Salmon, and Barrilleaux 2012a). Barber, Beramendi, and Wibbels (2013) find that self-reported partisanship does not predict voting behavior in a simplified version of the MR model. These non-findings appear to be puzzling, given the overwhelming evidence from survey-based studies showing that self-reported leftwing ideology is a robust predictor of support for income redistribution (e.g., Jaeger 2006, 2008, 363; Scheepers and Grotenhuis 2005).

This linear egoistic prediction based on individual gaps above and below the group average and subjects’ political ideology will serve as reference points against which the explanatory power of moral reasoning about redistribution can be tested. We therefore hypothesize:

**Hypothesis 3b (H3b)—Conditional political ideology effect:** More leftist self-reported ideology will lead to higher voted tax rates \( \tau \).
with normative implications about politics and society. As such, they contain larger generic beliefs about the relative roles of individual contributions and societal reward rules in bringing about success in life (Sabbagh and Vanhuysse 2006; Skitka and Tetlock 1993). For instance, conservatives tend to endorse luck as influential to success considerably less than liberals (Fong 2001; Gromet, Hartson, and Sherman 2015). These strong a priori ideological disagreements over how success is achieved, which relate to the random element in life, may be at the heart of seemingly unbridgeable ideological divides about income redistribution (Gromet, Hartson, and Sherman 2015). We therefore posit:

Hypothesis 3b (H3b)—Conditional political ideology effect: Leftist self-reported ideology will increase support for tax rates $\tau$ in the random but not the merit assignment of $pi$.

The full set of hypotheses is summarized in Figure 1.

### Experimental Design and Analysis

#### Experimental Procedures

All sessions followed the same procedures. At the beginning of each session, after randomly handing out place cards, the experimenter read out the instructions and answered questions (see Online Appendix B for Instructions). As part of the instructions, subjects also answered comprehension questions to ensure that they were fully informed about the experimental procedure. The experimental procedure consisted of three stages. In the first stage, subjects were assigned to one of five treatment conditions (Table 1).

1. In the equal payrate treatment, each subject received three Euro cents for a correctly solved counting task (“equal pay”).
2. In the two unequal payrate treatments, subjects were assigned a payrate of either three (low), six (medium), or nine (high) Euro cents for a correctly solved task. In one unequal treatment, the assignment of payrates was random (“luck rule”). In the other, it was based on subjects’ performance in a quiz with voluntary prior week-long preparation effort (“merit rule”). We use a knowledge quiz to assign payrates in the unequal “merit” treatment. Five days before the experiment, subjects in this treatment received the text of a Wikipedia entry on the White House (official residence and workplace of the U.S. President) via email. Subjects were informed that their knowledge on this topic was going to influence their possible earnings in the experiment (Rauhut and Winter 2010, 1185). Following Rauhut and Winter (2010, 1185), we chose a specific topic to ensure that everybody had to learn the text and nobody could benefit from his or her respective field of studies. On the day of the experiment, subjects had to answer

### Table 1. Treatment Conditions.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Subjects</th>
<th>Rounds</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal pay</td>
<td>39</td>
<td>5</td>
<td>195</td>
</tr>
<tr>
<td>Merit rule with full information</td>
<td>93</td>
<td>5</td>
<td>465</td>
</tr>
<tr>
<td>Luck rule with full information</td>
<td>93</td>
<td>5</td>
<td>465</td>
</tr>
<tr>
<td>Merit rule without information</td>
<td>63</td>
<td>5</td>
<td>315</td>
</tr>
<tr>
<td>Luck rule without information</td>
<td>69</td>
<td>5</td>
<td>345</td>
</tr>
<tr>
<td>Overall</td>
<td>357</td>
<td></td>
<td>1,785</td>
</tr>
</tbody>
</table>

Figure 1. Theoretical framework.
knowledge questions in a quiz on the White House. Payrate factors were then assigned according to performance on the quiz (Online Appendix C Figure C1). The third with the most correct answers in the session was assigned the high payrate factor (nine Euro cents), the second-best third the medium rate (six Euro cents), and the lowest third the low payrate (three Euro cents). We implemented a between-subject design to test the treatment effect. Thus, subjects’ payrate did not change over the course of the experiment.

The second stage of the experiment was the real-effort counting task. In each treatment, subjects earned their pre-tax income through a task designed by Ryvkin and Semykina (2017) that required them to count numbers in random sequences of numbers and letters. Correct answers were rewarded with Euro cents according to subjects’ assigned payrate. The counting task was played for five minutes, during which time subjects had to solve as many tasks as possible (Online Appendix C Figure C2).

The third stage of the experiment was the decision stage. After subjects played the counting task, they voted on a linear (flat) tax rate on their pre-tax earnings in groups consisting of a high, medium, and low payrate member (except for the equal treatment). Each subject casted a vote for his or her preferred tax rate (in the range of 0%-100%) and the median vote was then implemented, thus determining subjects’ final post-tax income (Esarey, Salmon, and Barrilleaux 2012b). Each subject entered their preferred tax rate, without talk or negotiation with other group members. To ensure perfect information about consequences of a possible tax rate, subjects in the information conditions were informed about the payrate, number of correctly solved tasks, and pre-tax income of all three group members. In addition, subjects were equipped with an on-screen calculator that displayed the earnings of each group member before and after redistributive taxation for any given tax rate (Online Appendix C Figure C3).

Finally, we implemented two additional treatments in which we manipulated subjects’ information set. The aim was to better illuminate how clear information signals about merit may be the mechanisms that bring “merit reasoning” to the foreground and produce the differences between the “merit rule” and “luck rule” treatments. Our fourth (“merit rule without information”) and fifth (“luck rule without information”) treatment conditions were equivalent to treatment conditions two (“merit rule with full information”) and three (“luck rule with full information”) except that on the decision screen, we now hid the information on fellow group members’ payrate and performance in the counting task.

The second and third stages of the experiment were repeated five times using random matching for the group assignment. To avoid learning and sequence effects, subjects were not informed about the implemented tax rates until after the last period was played. The final pay-off in Euros was the sum of post-tax incomes in all five periods. Subjects completed a post-experimental questionnaire including socio-demographics and a measure of their ideological orientation. Finally, subjects were individually paid in private and in cash. Subjects earned approximately twenty-one Euros on average and sessions lasted approximately ninety minutes.

Sample and Analysis

The experiment was programmed and conducted with the laboratory experimental software zTree (Fischbacher 2007). Subjects were recruited with the online registration platform hroot (Bock, Baetge, and Nicklisch 2014). The experimental sessions were conducted at two universities in Northern Germany. A total of 357 subjects participated. Descriptive statistics by treatment condition are presented in Appendix Table 1 and results of balancing tests in Online Appendix A Table A1.

The dependent variables measure subjects’ voted tax rate. Since this is a proportion between 0 and 1, a fractional logistic model is used for the regression analysis (Papke and Wooldridge 1996). Since subjects make repeated tax choices, the standard errors are adjusted for two-way clustering on subjects (N = 357) and sessions (N = 19) (Cameron, Gelbach, and Miller 2008, 2011). As part of the robustness analysis, Online Appendix A Table A4 presents the result from fractional logistic regression and generalized least squares (GLS) random-effects with one-way clustered standard errors. The three independent variables of main theoretical interest are the treatment condition, the self-interest prediction, and self-reported political ideology. The treatment conditions are measured with a set of dummies. Subjects’ self-interested preference is measured as the difference between their pre-tax income and the mean income of the group. Political orientation is measured on a standard one to ten left-right scale. Online Appendix A Figure A1 compares the distribution of this left-right measure for our sample of university students with a general German population sample restricted to ages eighteen to thirty years. Although the student sample is slightly skewed to the left side of the ideological continuum (student sample: mean = 0.40, SD = 0.17) compared with the population sample (ALLBUS: mean = 0.47, SD = 0.16), both scales are normally distributed showing substantial variation on ideological self-placement. There is long debate on the appropriateness of the left-right scale to measure people’s ideological orientation and whether it captures the economic dimension (e.g., attitudes toward redistribution) or the societal dimension (e.g., attitudes on privacy and the reach of the state). For our purposes, it is especially important to
know whether and to which extent ideological self-placement reflects specific views on social justice. Therefore, we tested to which extent this variable reflects justice beliefs as captured by the four sub-dimensions from the Basic Social Justice Orientations (BSJO) scale (Liebig, Hülle, and May 2016, 28). These measure support for four different justice principles: merit, equality, need, and entitlement. Online Appendix A Table A3 shows that the left–right scale adequately captures normative justice beliefs, as it is positive correlated with merit and entitlement and negatively correlated with equality and need.

The regression models control for subjects’ age, gender, field of study, and a set of dummies for the five times stages two and three of the experiment were repeated. In addition, each model includes a dummy variable accounting for subject-pool effects as the experiments were conducted at two different laboratories.

**Empirical Results**

**Design Check**

Figure 2 reports the descriptive evidence on the functioning of the experimental manipulation of payrates. The left-hand side reports the distribution of correctly solved tasks in the real-effort counting task for the three different payrates. The right-hand side shows the resulting pre-tax income across the three payrates. The rules of reward that are in place do not much change actual performance on the counting task. The vast majority of subjects correctly solve about fifty-four tasks, regardless of their payrate. Rules of reward obviously strongly affect how performance is translated into actual pre-tax earned income. Subjects who were assigned a high payrate were able to earn within a wide range from very high to low pre-tax incomes. Those assigned a medium or low payrate were limited to a much narrower range of smaller pre-tax incomes. Clearly, the unequal payrates that translate efforts levels into earnings, not effort levels themselves, are what cause pre-tax income inequality in the experimental data.

Probing deeper into the functioning of the experimental manipulation, Figure 3 reports the conditional means for the number of correctly solved counting tasks by payrate and treatment condition (Online Appendix A Table A2). Having a high payrate based on their prior good performance in the White House quiz treatment (merit) leads subjects to perform worse on the subsequent counting task than those with a randomly assigned high payrate. Possibly, the former group, aware that they had already earned a beneficial payrate, now felt free to coast along and rest on their laurels. Alternatively, it might be that high earners anticipated that other group members would implement high tax rates, which would reduce their incentive to perform well. Moreover, their subsequent low performance may admittedly have caused low and medium types to propose lower tax rates, simply because there was not as much to be gained from the high types. Other than that, the overlapping confidence intervals show that all other treatments generally have no effect on subjects’ performance in the real-effort task.

**Main Effect**

Table 2 presents three different model specifications. Model 1 serves to test our main theoretical hypotheses
H1a, H1b, H2a, and H3a. Models 2 and 3 in turn test the conditional hypotheses for self-interest (H2b) and ideology (H3b). The control variables, including gender and field of study, generally do not affect preferred tax levels. There is a lab effect and a period effect (not reported), as support for redistribution increases over the course of the experiment. Below we use predicted effect plots to ease the interpretation of estimation results on our main variables of interest. As these plots take into account the effect of control variables (some of which have an effect on support of redistribution), they provide a clearer representation of the treatment effect.

Figure 4 compares the predicted conditional mean tax rate over the five treatment conditions. Comparing the “merit rule with full information” and the “luck rule with full information,” conditions show the expected treatment effect: merit rule leads to lower tax rates (H1a). Compared to the “equal pay” reference treatment, the preferred to tax rate is higher but not statistically different from under the “merit rule with full information.” Comparing the “merit rule with full information” condition with the two no-information treatments shows that preferred tax rates are higher under the no-information conditions. However, this relationship is not significant for the “luck rule without information” condition. In line with H1b, the voted tax rates in the merit and random assignments of $p_i$ are not significantly different from each other in the no-information condition.

Figure 5 displays the effects of self-interest and subjects’ self-reported political ideology on the conditional mean preferred tax rate. The left panel shows the single most important predictor of subjects’ tax choice: self-interest. The more the pre-tax income of a subject is below the mean group income, the higher the tax rate he or she chooses. Likewise, the more his or her pre-tax income is above the mean group income, the lower his or her chosen tax rate (H2a). Equally in line with the self-interest hypothesis, subjects whose pre-tax income equals the group’s choose a 50-percent tax rate. Note in this context that in our experimental design each subject was able to calculate his or her post-tax income for a given tax rate on screen. The right panel of Figure 5 shows the negative effect of subjects’ ideological left–right placement on their preferred tax rate (H3a). This effect, however, turns out to be substantively rather small and statistically weak and sensitive to model specification (see Table A6).

Figure 6 is based on Model 2 from Table 2. It illustrates the effect of self-interest on preferred tax rates as

---

**Figure 3.** Correctly solved tasks by payrate and treatment. See Online Appendix A Table A2 for full estimation results ($N = 1,785$). 95% confidence intervals.
Table 2. Determinants of Average Tax Decision.

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merit rule with full information</td>
<td>−0.202</td>
<td>−0.136</td>
<td>−0.404</td>
</tr>
<tr>
<td></td>
<td>[0.16]</td>
<td>[0.16]</td>
<td>[0.40]</td>
</tr>
<tr>
<td>Luck rule with full information</td>
<td>0.447***</td>
<td>0.512***</td>
<td>1.033***</td>
</tr>
<tr>
<td></td>
<td>[0.12]</td>
<td>[0.14]</td>
<td>[0.42]</td>
</tr>
<tr>
<td>Merit rule without information</td>
<td>0.365**</td>
<td>0.410**</td>
<td>1.283***</td>
</tr>
<tr>
<td></td>
<td>[0.14]</td>
<td>[0.16]</td>
<td>[0.42]</td>
</tr>
<tr>
<td>Luck rule without information</td>
<td>0.132</td>
<td>0.193</td>
<td>−0.0670</td>
</tr>
<tr>
<td></td>
<td>[0.16]</td>
<td>[0.16]</td>
<td>[0.66]</td>
</tr>
<tr>
<td>Egoistic vote (cont.)</td>
<td>−0.0118***</td>
<td>−0.0247****</td>
<td>−0.0119***</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.00]</td>
</tr>
<tr>
<td>Left Right</td>
<td>−0.693*</td>
<td>−0.657*</td>
<td>−0.381</td>
</tr>
<tr>
<td></td>
<td>[0.39]</td>
<td>[0.39]</td>
<td>[0.77]</td>
</tr>
<tr>
<td>Age</td>
<td>0.001</td>
<td>0.001</td>
<td>−0.002</td>
</tr>
<tr>
<td></td>
<td>[0.02]</td>
<td>[0.02]</td>
<td>[0.02]</td>
</tr>
<tr>
<td>Female</td>
<td>0.0875</td>
<td>0.0503</td>
<td>0.112</td>
</tr>
<tr>
<td></td>
<td>[0.19]</td>
<td>[0.18]</td>
<td>[0.18]</td>
</tr>
<tr>
<td>Semester</td>
<td>−0.011</td>
<td>−0.013</td>
<td>−0.010</td>
</tr>
<tr>
<td></td>
<td>[0.02]</td>
<td>[0.02]</td>
<td>[0.02]</td>
</tr>
<tr>
<td>Natural science</td>
<td>0.152</td>
<td>0.114</td>
<td>0.0776</td>
</tr>
<tr>
<td></td>
<td>[0.22]</td>
<td>[0.22]</td>
<td>[0.22]</td>
</tr>
<tr>
<td>Social science</td>
<td>0.122</td>
<td>0.0970</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td>[0.20]</td>
<td>[0.20]</td>
<td>[0.22]</td>
</tr>
<tr>
<td>Arts and humanities</td>
<td>−0.0871</td>
<td>−0.134</td>
<td>−0.118</td>
</tr>
<tr>
<td></td>
<td>[0.28]</td>
<td>[0.28]</td>
<td>[0.27]</td>
</tr>
<tr>
<td>Business science</td>
<td>−0.0785</td>
<td>−0.0900</td>
<td>−0.0910</td>
</tr>
<tr>
<td></td>
<td>[0.16]</td>
<td>[0.17]</td>
<td>[0.16]</td>
</tr>
<tr>
<td>Lab.</td>
<td>−0.384**</td>
<td>−0.391***</td>
<td>−0.424***</td>
</tr>
<tr>
<td></td>
<td>[0.15]</td>
<td>[0.15]</td>
<td>[0.16]</td>
</tr>
<tr>
<td>Merit rule with full information # Egoistic vote (cont.)</td>
<td>0.0128***</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>Luck rule with full information # Egoistic vote (cont.)</td>
<td>0.0143***</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>Merit rule without information # Egoistic vote (cont.)</td>
<td>0.0102****</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>Luck rule without information # Egoistic vote (cont.)</td>
<td>0.0139****</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>Merit rule with full information # Left Right</td>
<td>0.512</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.81]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luck rule with full information # Left Right</td>
<td>−1.475*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.87]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merit rule without information # Left Right</td>
<td>−2.321**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[1.08]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luck rule without information # Left Right</td>
<td>0.440</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[1.30]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,785</td>
<td>1,785</td>
<td>1,785</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>.299</td>
<td>.306</td>
<td>.302</td>
</tr>
<tr>
<td>AIC</td>
<td>1,771.7</td>
<td>1,754.4</td>
<td>1,762.8</td>
</tr>
<tr>
<td>BIC</td>
<td>1,870.5</td>
<td>1,853.1</td>
<td>1,861.5</td>
</tr>
</tbody>
</table>

Fractional logistic regression with standard errors cluster by subject (N = 357) and session (N = 19). Dependent variable: tax decision (0–1). Robust standard errors in brackets. Constant and Period Dummies include but not reported. AIC = Akaike information criterion; BIC = Bayesian information criterion.

*p ≤ .1, **p ≤ .05, ***p ≤ .01.
conditioned by treatment. Each of the three panels presents the predicted tax rate for the “merit rule with full information” condition and compares it to the other treatment conditions. The top left panel is theoretically of particular interest, as it compares the effect of self-interest in the merit and luck conditions. Clearly, the “merit rule with full information” treatment does not fully suspend or override self-interested reasoning, least of all when the gains (resp. losses) from redistribution are high. But it does shift support for higher taxes downward among those just below, at, and above the group’s mean income. Consistent with H2b, compared with the “luck rule with full information” treatment, the “merit rule with full information” treatment leads to lower taxes among these low-stake middle and upper middle-income voters. The top right panel shows that fully equal payrates make voters even more egoistic than in the “merit rule with full information” treatment: even small deviations from an indifferent pre-tax income position (0 on the x-axis) lead to rather steep increases (resp. declines) in the preferred tax rate. Otherwise said, the merit treatment flattens, or softens, the effect of self-interest compared with pure equality. The bottom panel of Figure 5 in turn compares the “merit rule” and the “luck rule” treatments under the “no information” conditions. In the absence of clear information signals about merit, tax voting is not significantly different in merit versus luck circumstances.

**Conclusion and Discussion**

Redistributive taxation is generally perceived to be both more acceptable and more just if and when it equalizes privileges that are based on luck rather than merit (Dworkin 2002; Konow 2000; Roemer 1998). To test this argument, our experimental design manipulated subjects’ payrates in a real-effort task. In other words, we treated not the endowments resulting from individual performance, but rather the “societal” rules of reward (merit, luck, or equality) that translate these endowments into actual pre-tax incomes. To flesh out how “merit reasoning” might affect tax voting, we also explored how the absence of “full merit information” affects tax voting. Our main experimental results can be summarized as follows. First, voted tax rates are highest in the unequal random treatment (“luck rule with full information”); about seven to ten percent higher than in the equal and the unequal earned (“merit rule with full information”) treatments. In other words, voters want to redistribute less when incomes are truly merited by effort. But this only holds when subjects have information about how others obtained their incomes: this information signal appears to be the mechanism that activates merit reasoning. When no such merit knowledge is available, tax rates do not differ in the merit and luck conditions. Second, merit rule does not suspend or override self-interested reasoning, least of all when the gains (resp. losses) from redistribution are high. At best, it lowers support for taxation somewhat among those who stand to either gain or lose modestly from it. Third, merit overrides political ideology—again only when the information needed for engaging in merit reasoning is available. We have surmised that
the presence of such a merit signal is needed to activate or make salient such merit reasoning. Only when more left-leaning subjects know that payrates are random do they seem to activate their ideological beliefs and vote for higher taxation, regardless of their income position.

We find that the presence of clear information signals about whether unequal income distributions are created by merit reduces the role of political ideology (also see Brown-Iannuzzi et al. 2015; Esarey, Salmon, and Barrilleaux 2012a; Gromet, Hartson, and Sherman 2015a).

Figure 5. Effect of self-interest and Left Right ideology on tax decision. Based on Model 1 from Table 2. 95% confidence intervals. Bars represent distribution in percentage points.

Figure 6. Self-interest conditional on treatment. Based on Model 3 from Table 2. 95% confidence intervals.
Our study design allowed us to instill an unusually clear distinction in subjects’ minds, before they voted on taxes, about the key moral question at hand. Of course, in real-world settings, the causal role of merit and luck in unequally rewarding effort is rarely as clear-cut. Indeed, we have shown that muddying the information waters by hiding signals on how other group members earned their incomes significantly alters how subject vote about redistribution. This also implies that fuller real-world transparency about the precise pathways of success (earnings) might well lead to stronger calls for tax justice and more urgent debates about whether society rewards contributions or luck.

The observation that our subjects generally (and leftwing subjects particularly) voted for redistribution only when inequalities were based on luck behaviorally corroborates the normative arguments of luck egalitarianism (Dworkin 2002). Leftwing political parties in turn should take note that since their voter base may not be uniformly susceptible to tax-the-rich policies, they might be better advised to frame their redistributive proposals in terms of procedural justice and merit reasoning. Future experimental research might take advantage of online crowd-working platforms that enable us to test the effect of reward rules on support for redistribution on larger samples (e.g., Bokemper, DeScioli, and Kline 2019), even if that means that the collective vote choice, which is central to this study, moves into the background.

In sum, our experimental findings support the conclusion that egoistic preferences are the most important determinant of voting behavior on redistributive taxation, but they also show that deviations from self-interested reasoning can at least partially be attributed to merit reasoning and to political ideology when reward rules are arbitrary. Voters’ perception of the causes of income inequality and the specifics of the income inequality-generating process within their society are key to whether or not distribution conflicts are fought out ideologically. Taken together, these findings do not provide the final word on merit, luck, and taxes, or on how these notions relate to larger societal debates about meritocracy, the role of the state, and the drivers of privilege. But they do indicate that voters may buy into the idea that luck is an undeserved reason for riches. What is more: voters may be willing to put their money where their mouths are, and to significantly redistribute (only) luck-driven riches.

**Figure 7.** Political ideology conditional on treatment. Based on Model 3 from Table 2. 95% confidence intervals.
## Appendix

### Table 1. Descriptive Statistics by Treatment Condition.

<table>
<thead>
<tr>
<th></th>
<th>Equal pay</th>
<th>Merit rule with full information</th>
<th>Luck rule with full information</th>
<th>Merit rule without information</th>
<th>Luck rule without information</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M SD</td>
<td>M SD</td>
<td>M SD</td>
<td>M SD</td>
<td>M SD</td>
<td>M SD</td>
</tr>
<tr>
<td>Left Right</td>
<td>0.42 0.16</td>
<td>0.40 0.21</td>
<td>0.40 0.17</td>
<td>0.39 0.15</td>
<td>0.42 0.15</td>
<td>0.40 0.17</td>
</tr>
<tr>
<td>Merit (equity)*</td>
<td>0.69 0.20</td>
<td>0.62 0.19</td>
<td>0.61 0.19</td>
<td>0.63 0.18</td>
<td>0.66 0.20</td>
<td>0.63 0.19</td>
</tr>
<tr>
<td>Equality</td>
<td>0.55 0.23</td>
<td>0.51 0.26</td>
<td>0.51 0.23</td>
<td>0.56 0.22</td>
<td>0.52 0.22</td>
<td>0.52 0.23</td>
</tr>
<tr>
<td>Need</td>
<td>0.84 0.15</td>
<td>0.86 0.13</td>
<td>0.85 0.15</td>
<td>0.85 0.15</td>
<td>0.85 0.11</td>
<td>0.85 0.14</td>
</tr>
<tr>
<td>Entitlement</td>
<td>0.49 0.16</td>
<td>0.46 0.18</td>
<td>0.42 0.17</td>
<td>0.42 0.17</td>
<td>0.48 0.17</td>
<td>0.45 0.17</td>
</tr>
<tr>
<td>Age</td>
<td>25.5 4.22</td>
<td>24.6 5.22</td>
<td>26.1 6.08</td>
<td>26.6 5.09</td>
<td>24.8 6.79</td>
<td>25.5 5.72</td>
</tr>
<tr>
<td>Female</td>
<td>0.44 0.50</td>
<td>0.52 0.50</td>
<td>0.56 0.50</td>
<td>0.57 0.50</td>
<td>0.62 0.49</td>
<td>0.55 0.50</td>
</tr>
<tr>
<td>Semester</td>
<td>6.03 3.41</td>
<td>5.56 3.43</td>
<td>6.65 4.51</td>
<td>5.83 3.52</td>
<td>5.00 3.51</td>
<td>5.83 3.81</td>
</tr>
<tr>
<td>Natural sciences</td>
<td>0.18 0.18</td>
<td>0.13</td>
<td>0.14</td>
<td>0.19</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>0.26 0.15</td>
<td>0.24</td>
<td>0.29</td>
<td>0.25</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Social science</td>
<td>0.18 0.26</td>
<td>0.18</td>
<td>0.11</td>
<td>0.20</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Arts and humanities</td>
<td>0.05 0.12</td>
<td>0.15</td>
<td>0.05</td>
<td>0.12</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Business science</td>
<td>0.33 0.29</td>
<td>0.30</td>
<td>0.41</td>
<td>0.25</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>Lab</td>
<td>0.62 0.48</td>
<td>0.48</td>
<td>1.00</td>
<td>0.30</td>
<td>0.55</td>
<td></td>
</tr>
</tbody>
</table>

*306 subjects, otherwise 357.

### Authors’ Note

Experimental protocol, data set, and syntax will be made publicly available at the Harvard Dataverse (https://doi.org/10.7910/DVN/DLWEUY).

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received financial support for the research, authorship, and/or publication of this article: This study has been funded by the German Research Foundation (Grant Number TE1022/2-1).

### ORCID iD

Markus Tepe [https://orcid.org/0000-0002-6679-0346](https://orcid.org/0000-0002-6679-0346)

### Notes

1. See, for example, Article 4 of the European Social Charter, Article 7 of the UN International Covenant on Economic, Social and Cultural Rights or the Constitution of the International Labor Organization.

2. A large body of research has documented persistent forms of gender wage gaps (Blau and Kahn 2007; Weichselbaumer and Winter-Ebmer 2005), public sector wage gaps (Lucifora and Meurs 2006), and ethnic wage gaps (Blackaby et al. 1998; Johnston and Lordan 2016).

3. Note that all our subjects perform the same real-effort task. Yet, those with the same number of correct counts can still be unequally remunerated financially because of the rules of reward, which in our design are either luck-based (sheer randomness) or merit-based (scores on a quiz with voluntary preparation).

4. See Alesina and La Ferrara (2005), Fong (2001), and Sabbagh and Vanhuyse (2006). Alesina and La Ferrara (2005) add that beliefs about social mobility can further decrease the level of redistribution. People who believe that a society offers equal possibilities to everybody are averse to redistribution when mobility increases.

5. Important related experimental studies not use a flat tax voting vehicle. Almas, Cappelen, and Tungodden (2017) allow subjects to redistribute a bonus between a high and a low performer. Using a modified version of the dictator game, Almas et al. (2010) study the development of inequality acceptance in adolescence. Cappelen et al. (2013) let subjects take a risky choice in the first stage and then let them redistribute in the second stage, to study fairness views about risk taking.

6. We cannot disguise $p_i$ and $c_i$ separately because subjects always need to know $x_i$ in order to vote and with one piece of additional information ($p_i$ or $c_i$) they can infer the third missing parameter.

7. We are grateful to a reviewer for pointing this out.

8. See Rauhut and Winter (2010). A potential limitation of this design is that this White House quiz has only been conducted in the unequal earned treatments (both with and without merit information). This task could also have been conducted in each treatment, but used to allocate payrates in the unequal earned (merit) and unequal earned (no-information) treatment. We have deliberately chosen to conduct the quiz when needed to allocate payrates, as otherwise subjects might have falsely assumed a (non-existing) linkage between their quiz performance and the
subsequent counting task. Moreover, having a quiz that is irrelevant for the subsequent treatment should still be incentivized, which would have created more problematic differences between treatments.

9. To make sure that the attitudinal measures of theoretical interests, specifically subjects’ self-reported ideological orientation and self-reported support for the merit principle, are not affected by the treatment condition, Online Appendix A Figure A3 reports a set of box-whisker-plots by treatment condition.

10. The zTree code, the data set, and syntax will be made available at the Harvard Dataverse for replication purposes.

11. For the balancing test, we regress the set of control variables on the treatment conditions using a multinomial regression model. Results show that the controls are not perfectly balanced for the no-information treatments (see Online Appendix A Table A1). Against this background, it is important to account for these variables in the statistical analysis and not to dive into analysis in which samples are separated by treatment condition.

12. We are grateful to a reviewer for pointing this out.

13. Furthermore, the estimation results presented in Online Appendix A Table A2 show that while subjects’ ideological self-placement has no effect on their effort, we find that females and elder students perform poorer in the counting tasks. These auxiliary findings shed some light in the limitations of the real-effort task, which is age and gender sensitive.

14. Also see Online Appendix A Figure A2 for the effect of self-reported preference for the merit principle on the preferred tax rate conditioned by treatments.

Supplemental Material

Supplemental materials for this article are available with the manuscript on the Political Research Quarterly (PRQ) website.

References


