Defaults or Fairness Norms? Explaining Differences in Redistributive Preferences

Justin Valasek[1], Pauline Vorjohann[2]
and Weijia Wang[1*†]

October 3, 2025

Abstract

A large literature uses impartial spectators to elicit preferences for redistribution. We investigate whether differences in redistributive choices reflect underlying differences in fairness norms or default effects. Our model illustrates that identification of relative fairness norms requires an assumption of equal default effects. We test this assumption in an experiment with representative US and Swedish samples, and find that merit-luck differences reflect genuine fairness norms, while cross-country differences largely arise from heterogeneous default effects. Our results caution against equating redistributive choices with underlying fairness norms, and we provide a simple method for distinguishing differences in fairness norms from heterogeneity in default effects.

Keywords: fairness norms, redistributive preferences, reference dependence, default effects, cross-country comparison, experiment.

JEL Classification Codes: C91, D63, H23.

*Thanks to Ingvild Almås, Yves Breitmoser, Alexander Cappelen, Thomas de Haan, Simone Valerie Häckl-Schermer, Ranveig Falch, Ingar Haaland, Steffen Huck, Vincent Somville, Erik Sørensen, Heidi Thyssen and Christian Zehnder. This study was funded by the NHH Smaaforsk Fund and by FAIR, the Centre for Experimental Research on Fairness, Inequality and Rationality through grants from Telenor. The project is approved by the Institutional Review Board at the Norwegian School of Economics.

†[1]NHH Norwegian School of Economics, [2]University of Exeter. Contact e-mails: justin.valasek@nhh.no, p.l.vorjohann@exeter.ac.uk,weijia.wang@nhh.no.

1 Introduction

In the face of rising global inequality, understanding how individuals form judgments about fairness, inequality and redistribution is a central concern. A growing body of experimental research uses redistributive choices made by impartial third-party "spectators" to elicit preferences for redistribution, aiming to uncover the normative principles that guide judgments about how income ought to be allocated. However, recent evidence suggests that these redistributive choices are highly sensitive to contextual features of the decision environment, such as information about the default allocation that spectators observe prior to making a choice over redistribution (see Charité et al., 2022 for evidence on third-party reference dependence). This sensitivity to contextual factors implies that default allocations may shift behavior independently of underlying fairness norms, and raises a fundamental identification problem: when observed redistributive choices vary across treatments or countries, do these differences reflect variation in fairness norms, or do they simply reflect differences in the strength of default effects?

In this paper, we show that without accounting for default effects, standard methods risk misinterpreting heterogeneity in default effects as differences in fairness norms. We develop a simple formal framework in which spectator choices are a weighted average of an underlying fairness norm—a distributional preference—and the default allocation. This framework clarifies that comparisons across contexts or populations require strong assumptions about the absence or equality of default effects. We also show how these assumptions can be tested using data from multiple defaults.

We then implement this approach in a pre-registered experiment that modifies the influential design of Almås et al. (2025); Almås et al. (2020). Similar to the original design, we vary whether intermediate payoffs are based on merit or luck and whether the sample population is drawn from the U.S. or Sweden. Additionally, we vary whether the default allocation is equal, unequal, or absent. This design allows us to disentangle differences in fairness norms from differences in default effects.

Our findings are clear. Defaults substantially shape spectator choices in both the U.S. and Sweden, and in both merit and luck contexts. However, our results indicate that differences between merit and luck reflect genuine differences in fairness norms, while differences between U.S. and Swedish spectators are driven largely by heterogeneous default effects. Importantly, cross-country differences appear only under an unequal default and vanish under an equal or no default.

Taken together, our findings carry important methodological and substantive implications for the study of redistributive preferences. Methodologically, due to default effects, the redistributive choices of third-party spectators do not always reflect the underlying fairness norms. Accordingly, we establish that comparisons of fairness norms cannot be inferred using standard approaches without making strong and potentially unwarranted assumptions about the absence or equality of default effects across contexts and populations. However, our model provides a simple diagnostic: varying the default is sufficient to test whether cross-context or cross-country comparisons capture fairness norms or default effects.

Substantively, our analysis suggests that widely cited differences in redistributive preferences may be overstated if default effects are disregarded. More broadly, our results suggest that understanding variation in redistributive choices requires carefully accounting for how decision-makers respond to elements of the choice architecture, such as the status quo. By studying a formal model of third-party reference dependence and providing new experimental evidence on default effects, our results offer a clear basis for interpreting variation in redistributive choices across contexts, treatments, and populations.

Literature

The importance of default allocations for spectators has been highlighted in Charité et al. (2022), who show that spectators equalize earnings less frequently when workers are informed about the intermediate distribution of payoffs, relative to the case where the workers are not informed about the intermediate distribution. Their work confirms that defaults impact spectator's preferences through the channel of beliefs about worker expectations, an effect we capture in our formal framework.[1] Our paper builds

[1]Telle and Tjøtta (2023) also show that spectators are responsive to priming effects regarding the salience of labeling a change of the default as "redistribution," suggesting

on this result by formally illustrating the challenge default effects introduce for identifying differences in fairness norms from spectators' redistributive choices. We further provide a simple method for testing for the equality of default effects across treatments and populations.

There are a large number of papers that use redistributive choices of third-party spectators as a metric to compare implemented inequality across treatments and populations, many of which use either an exogenous or endogenous intermediate distribution of earnings—a default allocation. For example, articles and working papers that use spectator choices to measure preferences for redistribution published since 2020 include: Almås et al. (2025); Almås et al. (2020); Andre (2024); Bhattacharya and Mollerstrom (2024, 2025); Bortolotti et al. (2025a,b); Cappelen et al. (2024a, 2025, 2022a, 2024b, 2023, 2022b); Charité et al. (2022); Cohen et al. (2025); Cohn et al. (2023); Dong et al. (2022); Drouvelis and Gavassa-Pérez (2024); Dwenger et al. (2024); Espinosa et al. (2020); Fischbacher et al. (2023); Freyer and Guenther (2025); Haeckl et al. (2025); Hufe and Weishaar (2025); Konow et al. (2020); Lu (2025); Mo et al. (2023); Müller and Renes (2021); Piraino and Ryan (2024); Preuss et al. (2025); Santamaría et al. (2023); Sartor (2025); Sterba (2022); Strang and Schaube (2025); Telle and Tjøtta (2023); Valasek et al. (2024); Wang (2024); Yusof and Sartor (2025) (also see Cappelen et al., 2020 and Almås et al., 2024 for reviews of the literature). Importantly, our results highlight that the use of different contexts, populations and default distributions limits the comparability *across different studies*: that is, without testing for equal default effects, differences in elicited fairness norms across studies may actually reflect differences in default effects across studies.

2 A Model of Reference-Dependent Fairness Preferences

In this section we develop a theoretical framework that incorporates both redistributive preferences—a fairness norm—and default effects. Conceptually, a default effect may represent a tendency to leave a suggested division

that more mechanical default effects may also be at play.

of payoffs unchanged (see Choi et al., 2003), or the default may impact preferences by, say, influencing spectators' beliefs regarding workers' expectations about payoffs (see Charité et al., 2022). The framework we present here is based on Breitmoser and Vorjohann, 2024 and allows for preferences to depend on both a fairness norm and a default, regardless of the behavioral mechanism driving the observed default effect.

The setting we model is of an agent—a "spectator"—i who chooses a distribution of payoffs for two workers. The workers, $j \in \{A, B\}$, complete a real effort task and each have a set of characteristics $\kappa_j \in \mathbf{K}$. For example, κ_A could indicate that worker A was "lucky," or that they were the most productive in the real effort task.

After completing the task, an intermediate payoff distribution is determined by a decision rule that maps worker characteristics into "intermediate" payoffs $\{\tilde{y}_A, \tilde{y}_B\}$. We normalize these intermediate payoffs as follows:

$$d = \frac{\tilde{y}_A}{\tilde{y}_A + \tilde{y}_B}, \tag{1}$$

and refer to $d, 1-d$ as the "default".

Next, the spectator observes the worker characteristics $\kappa \equiv \{\kappa_A, \kappa_B\}$, the decision rule mapping characteristics into the default, and the default $d, 1-d$. The spectator is then given the opportunity to change the payoffs to any final payoffs $\{y_A = y_i, y_B = (1-y_i)\}$, with $y_i \in [0, 1]$. This framework is general enough to capture most studies on redistributive preferences cited above.

In the theoretical appendix (Appendix A.2), we use the axiomatic framework of Breitmoser and Vorjohann, 2024 to show that the spectator's preferences can be represented by a reference-dependent utility function where the reference point reflects both default effects and distributional preferences. More precisely, the reference point can be expressed as a linear function of the default d and the spectator's "fairness norm", $m_i(\kappa)$, specifying the payoff distribution $\{y_A = m_i(\kappa), y_B = 1 - m_i(\kappa)\}$ deemed fair by the spectator given worker characteristics κ.

Given this representation, the model's prediction is that the spectator will choose an allocation that is a weighted average of their fairness norm and the default:

$$y_i = \beta_i m_i(\kappa) + (1 - \beta_i)d, \tag{2}$$

where β_i captures the weight spectators place on the fairness norm relative to the default.

This representation considers a single individual in a specific context. Our goal, however, is to compare choices across populations or contexts. Therefore, we must account for the fact that the relative weight on the fairness norm may differ between settings. For example, Swedish spectators may have a different weight on their fairness norm compared to US spectators ($\beta^{SWE} \neq \beta^{US}$), and spectators may have different weights on their fairness norm depending on whether intermediate payoffs are determined by luck or merit ($\beta^{Luck} \neq \beta^{Merit}$).

2.1 Identifying Differences in Fairness Norms

As discussed above, we are concerned with comparing fairness norms across different contexts and populations. Accordingly, we focus on detailing sufficient conditions to identify *comparisons* of fairness norms given unobserved default effects. To formalize comparisons, consider $x, x' \in \mathbf{X}$ to be two different contexts or populations. Take $(y^x(d), y^{x'}(d))$ to be the observed choices in the respective contexts given default d, β^x and $\beta^{x'}$ to be the respective weights on the fairness norm, and $m(\kappa^x)$ and $m(\kappa^{x'})$ to be the respective fairness norms.

To summarize, in this section our focus is on characterizing sufficient conditions for identifying differences in fairness norms, $m(\kappa^x) - m(\kappa^{x'})$, from observed choices, $(y^x(d), y^{x'}(d))$, given default effects that are unobserved and that may be heterogeneous across contexts.

First, we use Equation 2 to solve for $m(\kappa^x)$ as a function of $y^x(d)$ and β^x:

$$m(\kappa^x) = \frac{y^x(d) - (1 - \beta^x)d}{\beta^x}, \tag{3}$$

which gives us the following equation for the difference in fairness norms:

$$m(\kappa^x) - m(\kappa^{x'}) = \frac{\beta^{x'}(y^x(d) - (1 - \beta^x)d) - \beta^x(y^{x'}(d) - (1 - \beta^{x'})d)}{\beta^x \beta^{x'}}. \tag{4}$$

This equation shows that given $(y^x(d), y^{x'}(d))$, data on β^x and $\beta^{x'}$ are required to identify $m(\kappa^x) - m(\kappa^{x'})$. Still, partial identification of $m(\kappa^x) - m(\kappa^{x'})$ is possible in some special cases. In the following, we consider three different cases that are relevant for identifying *relative differences* in fairness norms.

Case 1: $\beta^x \neq \beta^{x'}$. When default effects are not equal across x and x', Equation 4 shows that we can only draw precise inference about the relative size of $m(\kappa^x)$ and $m(\kappa^{x'})$ from $(y^x(d), y^{x'}(d))$ when $y^x(d)$ and $y^{x'}(d)$ are on opposite sides of the default. This is because in that case, $m(\kappa^x)$ is on the same side of the default as $y^x(d)$ and $m(\kappa^{x'})$ is on the same side of the default as $y^{x'}(d)$.

Result 1 *If $y^x(d)$ and $y^{x'}(d)$ are interior and $y^x(d), y^{x'}(d) \leq (\geq) d$, then the relative position of $m(\kappa^x)$ and $m(\kappa^{x'})$ is not identified by $(y^x(d), y^{x'}(d))$. That is, given $= y^x(d), y^{x'}(d)$ there exists $\beta^x, \beta^{x'}$ such that $m(\kappa^x)$ could be smaller or greater than $m(\kappa^{x'})$.*

We illustrate this identification problem with an example from Almås et al. (2020), who compare the fairness preferences of Norwegian and US spectators. In their experiment, spectators are faced with an intermediate payoff distribution of 6 USD to one worker, and 0 USD to the other worker, which corresponds to $d = 1$. On average, US spectators choose $y^{US} = 0.71$ and Norwegian spectators choose $y^{NOR} = 0.61$.

As illustrated by Result 1, given $d = 1, y^{US} = 0.71, y^{NOR} = 0.61$ it is not possible to identify whether the observed ordering, $y^{US} > y^{NOR}$ is due to cross-country differences in default effects, or due to differences in fairness norms. That is, given y^{US}, y^{NOR} are both interior and smaller than the default, it is feasible that $m(\kappa^{US}) < m(\kappa^{NOR})$ and $\beta^{US} > \beta^{NOR}$. Determining the relationship between $m(\kappa^{US})$ and $m(\kappa^{NOR})$ would, therefore, require additional information on β^{US} and β^{NOR}.

Next, we focus on the two cases in which it is possible to identity the relative difference of fairness norms across contexts from data without variation in defaults.

Case 2: No default effects ($\beta^x = \beta^{x'} = 1$). Identification of the difference in fairness norms is straightforward if there are no default effects, since by Equation 2, spectators will choose $y^x(d)$ and $y^{x'}(d)$ equal to their fairness norm.

Result 2 *If $\beta^x = \beta^{x'} = 1$, then $y^x(d) = m(\kappa^x)$ and $y^{x'}(d) = m(\kappa^{x'})$ for all d, and $m(\kappa^x) - m(\kappa^{x'}) = y^x(d) - y^{x'}(d)$.*

Case 3: Equal default effects $\beta^x = \beta^{x'} < 1$. When default effects exist, $y^x(d)$ is not sufficient to identify $m(\kappa^x)$. However, when $\beta^x = \beta^{x'}$, it follows by Equation 4 that the difference in fairness norms is represented by:

$$m(\kappa^x) - m(\kappa^{x'}) = \frac{y^x(d) - y^{x'}(d)}{\beta^x}, \tag{5}$$

Result 3 *Given equal relative weights on the fairness norm, $\beta^x = \beta^{x'} < 1$, spectator choices only identify the difference of the fairness norms weighted by a scalar of $1/\beta^x > 1$.*

Result 3 shows that if $\beta^x = \beta^{x'} < 1$, then $y^x(d) - y^{x'}(d)$ serves as a lower-bound estimate of $m(\kappa^x) - m(\kappa^{x'})$.

2.2 Testing for existence and equality of default effects:

Results 1-3 show that the ability to draw inference on fairness norms based on data on spectator choices depends on whether default effects exist, and whether they are equal. Here, we show that using choice data for two different defaults, d and d', for both x and x' we can derive testable predictions for each case.

First, note that Result 2 gives the straightforward and testable prediction that, absent default effects (Case 2), spectators will select the same allocation of earnings regardless of the default.

Prediction 1 (Case 2: $\beta^x = \beta^{x'} = 1$) *If agents put no weight on the default, then $y^x(d) = y^x(d')$ and $y^{x'}(d) = y^{x'}(d')$.*

If default effects are present, however, then data from two different defaults also allows the researcher to test whether the default effects are equal across x and x' (Case 3). That is, given two data points for different defaults $\{y^x(d), y^x(d')\}$, Equation 5 provides two different expressions for $m(\kappa^x)$—one for each default. These expressions can be combined to give the following equation for β^x as a function of $\{y^x(d), y^x(d')\}$:

$$\beta^x = \frac{y^x(d) - y^x(d')}{d - d'}, \tag{6}$$

This gives the testable prediction that if $\beta^x = \beta^{x'}$, then the difference between spectator choices under different defaults are the same across x and x':

Prediction 2 (Case 3: $\beta^x = \beta^{x'}$) *If agents put an equal weight on the default in A and B, then $y^x(d) - y^x(d') = y^{x'}(d) - y^{x'}(d')$.*

Prediction 2 also implies that if $y^x(d) - y^x(d') \neq y^{x'}(d) - y^{x'}(d')$, then $\beta^x \neq \beta^{x'}$ (Case 1).

Next, we detail an experiment designed to test Predictions 1 and 2 across both contexts and populations. This experiment will provide information on whether observed differences in redistributive choice are due to differences in fairness norms, or heterogeneous default effects.

3 Experiment and Empirical Analysis

To test the assumptions on default effects required to identify fairness norm comparisons across contexts and countries, we implement versions of the "Luck" and "Merit" treatments from Almås et al. (2025); Almås et al. (2020), applied to two populations: a representative U.S. sample and a representative Swedish sample. In addition to being highly influential articles that compare impartial third-party preferences for redistribution, their design allows for a straightforward modification to vary the default allocation. Specifically, the design allocates intermediate worker payoffs based either on a random draw or on which worker was the most productive. Spectators are then informed of the allocation rule and asked whether they want to redistribute earnings from one worker to the other. Using this design, we implement two different defaults, $d = 1$ and $d = 0.5$.

Additionally, we implement a "No Default" treatment for the Merit context, in which spectators are not informed about an intermediate allocation of worker payoffs. Arguably, the choice data from the No Default treatment captures the underlying fairness norm since there is no default effect by design.

For each treatment arm, we collected a representative sample of at least 250 U.S. spectators and 250 Swedish spectators, for a total of 2,873 spectators. The data were collected by a professional survey firm, and the English-language questionnaires used in the treatments are provided in Section B of the Appendix. All details of data collection and our empirical analysis were pre-registered in the AEA registry (AEARCTR-0012990).

In each treatment, spectators made a decision about the distribution of earnings for two workers who had completed a real-effort task. In all treatments, spectators could choose any distribution of earnings from the set $(3, 3), (4, 2), (5, 1), (6, 0)$ (workers also received a fixed participation fee of 2 USD). For example, if the spectator chose $(4, 2)$, then one randomly selected worker would earn 4 USD for the assignment, and the other would earn 2 USD.

In the "Unequal Default" treatment, spectators were informed that the workers had been told about an intermediate allocation in which one worker would be randomly selected to earn 6 USD, while the other would earn 0 USD ($d = 1$). Spectators were also told that the workers had been informed that a third person would have the opportunity to change this payment plan.

In the "Equal Default" treatment, spectators were informed that the workers had been told about an intermediate allocation in which both would earn 3 USD each ($d = 0.5$). As in the Unequal Default treatment, spectators were told that the workers had been informed that a third person could change this payment plan.

Finally, in the "No Default" treatment, run only for the Merit setting, spectators were asked to make a distributional decision without the workers having been informed of any intermediate allocation of earnings.

Hypothesis and Empirical Strategy: We test the following hypotheses that, based on the predictions of the simple theoretical framework introduced above, will provide evidence about the existence and equality of default effects across contexts and populations.

Hypothesis 1 *Spectators implement the same level of inequality in the Unequal Default and Equal Default treatments.*

Hypothesis 2 *The difference between implemented inequality in the Unequal Default and Equal Default treatments is the same in the Merit and Luck treatments.*[2]

Hypothesis 3 *The difference between implemented inequality in the Unequal Default and Equal Default treatments is the same in the US and Sweden.*

Hypothesis 4 *The difference between implemented inequality in the Unequal Default and No Default treatments is the same in the US and Sweden.*

To test these hypotheses, we use a normalized measure of the inequality implemented by spectator i that is analogous to the measure introduced in our model:

$$g_i = \frac{|Income\ Worker\ A_i - Income\ Worker\ B_i|}{Total\ Income} \in [0, 1], \tag{7}$$

This inequality measure is is equal to one if the spectator chooses an allocation of $(6, 0)$ and zero if the spectator chooses an allocation of $(3, 3)$, and is equivalent to the Gini coefficient.

Since the data for the Luck treatments and the Merit treatments were gathered in different rounds (Jan 24' and June 25', respectively), we test Hypothesis 1 and 3 separately for the two treatments. However, as preregistered, we also report the joint estimation, and we also use this joint estimation to compare default effects across the Merit and Luck contexts (Hypothesis 2).

[2]Our preanalysis plan for the Merit treatments specifies an empirical estimation comparing the Merit and Luck data. Due to an oversight, the plan does not include a hypothesis related to this empirical estimation. However, we believe that Hypothesis 2 follows naturally from our research design (see the Appendix A.1 for futher discussion).

The baseline empirical specification we use to study the treatment effects on implemented inequality is:

$$g_i = \alpha + \delta_0 EqualDefault_i + \delta_1 Sweden_i + \delta_2 EqualDefault_i \times \ Sweden_i + \gamma \mathbf{C}_i + \epsilon_i, \tag{8}$$

where $EqualDefault_i$ is an indicator variables for spectator i being in the Equal Default treatment (as opposed to the Unequal Default treatment), $Sweden_i$ is an indicator variable for spectator i being from Sweden, and $\mathbf{C}_i$ is a vector of control variables. For the Merit treatment, we also include a indicator variable for the No Default treatment, $\delta_3 NoDefault_i$, and an interaction with the No Default and Sweden indicators, $\delta_4 NoDefault_i \times Sweden_i$ in our estimation. Additionally, for the joint estimation of Luck and Merit treatments, we include a indicator variable for the Merit treatments, $\delta_5 Merit_i$, and an interaction with Equal Default, $\delta_6 EqualDefault_i \times Merit_i$.

The estimated value of δ_0 tests for default effects (Hypothesis 1), δ_2 and δ_6 tests for equal default effects $(1-\beta)$ between, respectively, the Luck and Merit treatments and the US and Sweden samples (Hypothesis 2 and 3), and δ_4 tests whether the difference in implemented inequality between US and Sweden samples is the same for the Unequal Default and No Default treatments (Hypothesis 4).

Earlier studies have shown that age, gender, education and income are important control variables when studying fairness preferences (Almås et al., 2020). Therefore, we also report the results both with and without these control variables, where education is a binary indicator for whether the spectator has the equivalent of a bachelor's degree and income is a binary indicator for whether the spectator has an individual income that is higher or lower than the country median.

3.1 Empirical Analysis

We begin with a descriptive analysis, focusing first on the average implemented inequality illustrated in Figure 1. While we report the formal hypothesis testing below, note that there is a clear impact of the default on implemented inequality, with the average Gini coefficient decreasing by close to one half for the Merit treatment and two thirds for the Luck treatment. However, the level of the decrease is similar for both Merit and Luck.

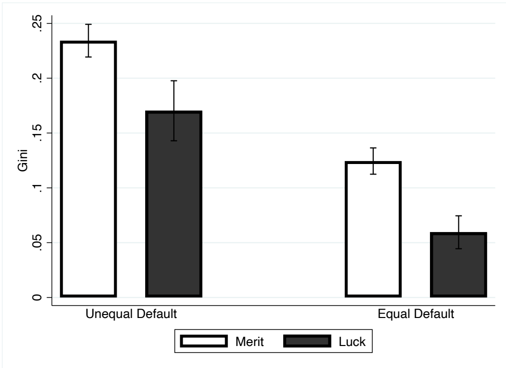
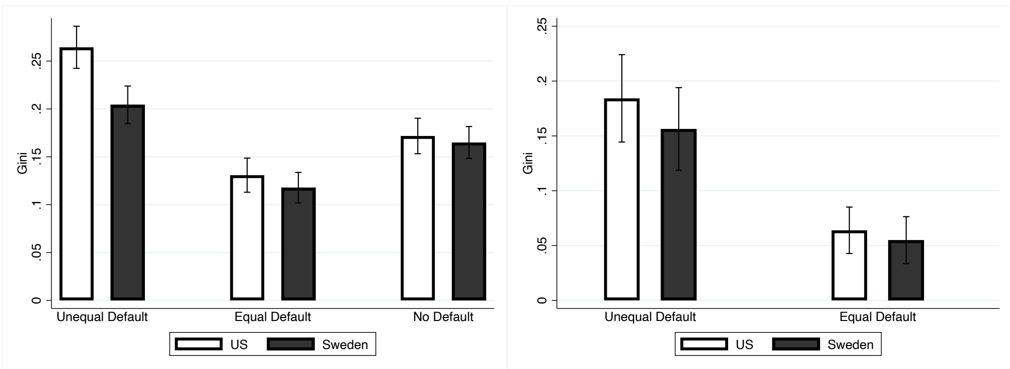


Figure 1: Average implemented inequality by treatment

Next, we consider the implemented inequality disaggregated by treatment and population (US and Sweden) in Figure 2. For the Merit treatment, we observe a clear difference in the implemented inequality between the US and Sweden for the Unequal Default. However, a similar difference is not observed for the Equal Default treatment. Additionally, we included data for the No Default/Merit treatment, with also shows almost no difference between the US and Sweden. The level differences in the Luck treatments follow a similar pattern, albeit with a nosier measure (we did not gather data for a No Default/Luck treatment).

Figure 2: Average implemented inequality by treatment and country



Merit Luck

While the descriptive analysis is suggestive, we report the results of our model estimations in Table 1 and present the formal tests our hypotheses below (Table 1 omits the coefficient estimates of the control variables, see Appendix A.3 for these results).

Table 1: Estimated treatment effects on Gini Coefficients

	Merit	Merit	Luck	Luck	Both	Both
Equal Default (δ_0)	-0.134***	-0.134***	-0.120***	-0.124***	-0.128***	-0.129***
	(0.0133)	(0.0132)	(0.0227)	(0.0224)	(0.0161)	(0.0160)
Sweden (δ_1)	-0.0601***	-0.0632***	-0.0279	-0.0324	-0.0446***	-0.0502***
	(0.0133)	(0.0133)	(0.0225)	(0.0222)	(0.0129)	(0.0128)
Eq.Def.×Swe. (δ_2)	0.0471**	0.0482**	0.0189	0.0235	0.0334*	0.0361**
	(0.0189)	(0.0188)	(0.0323)	(0.0318)	(0.0184)	(0.0182)
No Default (δ_3)	-0.0927***	-0.0934***				
	(0.0133)	(0.0133)				
No.Def.×Swe. (δ_4)	0.0534***	0.0557***				
	(0.0189)	(0.0189)				
Merit (δ_5)					0.0640***	0.0607***
					(0.0129)	(0.0129)
Eq.Def.×Merit (δ_6)					0.00089	0.00081
					(0.0184)	(0.0183)
Controls	-	✓	-	✓	-	✓
Constant	0.264***	0.256***	0.184***	0.249***	0.193***	0.207***
	(0.00943)	(0.0201)	(0.0159)	(0.0385)	(0.0113)	(0.0236)
N	1768	1767	1071	1071	2251	2250

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

First, our formal model establishes that if there are no default effects, then spectator choices will not vary based on the default (Prediction 1). Since the coefficient estimate on the indicator variable for the Equal Default treatment (δ_0) is statistically and economically significant across all specifications, we can reject the assumption that spectator choices are constant under different defaults.

Result 1 *We reject the null hypothesis that spectators implement the same level of inequality in the Unequal Default and Equal Default treatments for both the Merit and Luck treatments, and therefore reject the assumption of no default effects.*

Having established the existence of default effects, we next consider whether default effects are equal across treatments and populations using the test developed in our theoretical model. Prediction 2 states that if default effects are equal across treatments (or populations), then the difference in spectator choices between the Equal Default and Unequal Default will be the same across the Merit and Luck (or US and Sweden).

Looking first between treatments, the coefficient estimate on the interaction between the indicator variable for Equal Default and Merit (δ_6) indicates a precisely estimated null effect ($\delta_6 = 0.000892$), indicating that there is no evidence of a difference in the default effect between Merit and Luck.

Result 2 *We do not reject the null hypothesis that the difference in spectator choices between the Equal Default and the Unequal Default is equal between the Merit and Luck treatments. Therefore, we do not reject the assumption that default effects are equal across treatments.*

Results 1 and 2 establish that while default effects are present, we do not reject the assumption that they are equal across treatments, implying that the difference in spectator choices between treatments reflects differences in fairness norms.

Next, we test whether default effects are equal across the US and Sweden. First, note that the test detailed in Prediction 2 aims to distinguish whether cross-population differences are due to differences in fairness norms. That is, it only applies in cases where cross-population differences *exist*. Our model estimates only show evidence of cross-country differences in the Merit treatment, and not in the Luck treatment—the coefficient estimate on the indicator variable for Sweden is statistically significant for Merit, and statistically insignificant for Luck. Therefore, we are only able to use the data from the Merit treatment to test for equal default effects across the US and Sweden.

In the Merit treatment, the coefficient estimate on the interaction of the indicator variables for Equal Default and Sweden (δ_2) is positive and statistically significant at the five percent level.

Result 3 *We reject the null hypothesis that the difference in spectator choices across defaults is equal for the US and Sweden populations. Therefore, we reject the assumption that default effects are equal across these populations.*

Result 3 implies that the difference in implemented inequality between the U.S. and Sweden samples is at least partly due to differences in default effects across populations. Interestingly, the linear combination of Sweden and Eq.Def.×Swe. is not statistically significant ($p = 0.328$ without controls, $p = 0.262$ with controls), which shows that there is no evidence of cross-country differences in the Equal Default treatment. In contrast, there is a highly statistically significant cross-country difference in the Unequal Default treatment ($p = 0.000$, with and without controls).

To further explore whether cross-country differences in the Unequal Default treatment are due to different fairness norms or heterogeneous default effects, we compare the Unequal Default treatment to the No Default treatment. Again, since spectators are not informed about an intermediate allocation of worker payoffs in the No Default treatment, it arguably provides a clean measure of fairness norms.

The comparison between the Unequal and No Default treatments shows a positive and significant interaction of No Default and Sweden (δ_4 is positive and significant at the one percent level). Additionally, there is no evidence of cross-country differences in the No Default treatment: the linear combination of Sweden and No.Def.×Swe. is close to zero and is not statistically significant ($p = 0.611$ without controls, $p = 0.576$ with controls).

Result 4 *We reject the null hypothesis that the difference in spectator choices between the Unequal Default and No Default is equal for the US and Sweden populations. We find no evidence of cross-country differences in the No Default treatment.*

Overall, Results 3 and 4 provide evidence that the difference between the U.S. and Swedish populations observed in the Unequal Default treatment is driven by differences in average default effects, rather than differences in fairness norms.

4 Conclusion

This paper demonstrates that redistributive choices made by impartial spectators cannot be straightforwardly interpreted as fairness norms without accounting for default effects. We develop a simple formal framework that clarifies the identification problem and show experimentally how the presence and heterogeneity of default effects shape observed behavior.

Our findings carry three broader lessons. First, comparisons of redistributive preferences across experimental treatments are relatively robust: the difference between merit- and luck-based allocations appears to reflect genuine variation in fairness norms. Second, cross-country comparisons are far less robust: the difference between U.S. and Swedish spectators is driven by heterogeneous default effects rather than divergent fairness norms. Third, and most important, the widespread use of single-default designs in the literature risks conflating fairness norms with default effects, limiting the comparability of results across studies.

One implication of our research is methodological: inference about fairness norms either requires a no-default condition or explicit testing for equality of default effects by varying the default. For the broader economics literature, the implication is substantive: cultural variation in fairness judgments may be overstated unless default effects are properly accounted for. Our results suggest that what often looks like cross-country heterogeneity in fairness norms may instead be differences in how strongly individuals anchor on the status quo. These results suggest that future work should apply this approach to other domains where fairness and redistribution are studied to assess how much of the observed variation in redistributive choices across populations reflects genuine differences in fairness norms, and how much reflects the pervasive influence of defaults.

References

Almås, I., Cappelen, A., Sørensen, E., and Tungodden, B. (2025). Fairness Across the World. Working Paper No. 6/2025, Discussion Paper Series in Economics, Norwegian School of Economics, Department of Economics.

Almås, I., Cappelen, A. W., Sørensen, E. Ø., and Tungodden, B. (2024). Attitudes to inequality: preferences and beliefs. *Oxford Open Economics*, 3:64–79.

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

Andre, P. (2024). Shallow meritocracy. *Review of Economic Studies*, 89(6):2958–2991.

Bhattacharya, P. and Mollerstrom, J. (2024). Fairness preferences over parental wealth transfer. Available at SSRN: `https://ssrn.com/abstract=4934746`.

Bhattacharya, P. and Mollerstrom, J. (2025). Lucky to work. *The Review of Economics and Statistics*. Forthcoming.

Bortolotti, S., Kölle, F., Soraperra, I., and Sutter, M. (2025a). Social risk, fairness types, and redistribution. Technical Report DP No. 18122, IZA – Institute of Labor Economics.

Bortolotti, S., Soraperra, I., Sutter, M., and Zoller, C. (2025b). Too lucky to be true: Fairness views under the shadow of cheating. *The Review of Economics and Statistics*, 107(3):771–785.

Breitmoser, Y. and Vorjohann, P. (2024). Fairness-based Altruism. *Mimeo*.

Cappelen, A. W., De Haan, T., and Tungodden, B. (2024a). Fairness and limited information: Are people bayesian meritocrats? *Journal of Public Economics*, 233.

Cappelen, A. W., Falch, R., Huang, Z., and Tungodden, B. (2025). Acceptance of inequality between children: Large-scale experimental evidence from china and norway. *The Economic Journal*, 135(667):999–1020.

Cappelen, A. W., Falch, R., and Tungodden, B. (2020). Fair and unfair income inequality. In *Handbook of Labor, Human Resources and Population Economics*, pages 1–25. Springer.

Cappelen, A. W., Fest, S., Sørensen, E. Ø., and Tungodden, B. (2022a). Choice and personal responsibility: What is a morally relevant choice? *Review of Economics and Statistics*, 104(5):1110–1119.

Cappelen, A. W., Liu, Y., Nielsen, H., and Tungodden, B. (2024b). Fairness in a society of unequal opportunities. Technical Report No. 17/2024, Norwegian School of Economics, Department of Economics.

Cappelen, A. W., Moene, K. O., Skjelbred, S.-E., and Tungodden, B. (2023). The merit primacy effect. *The Economic Journal*, 133(651):951–970.

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

Charité, J., Fisman, R., Kuziemko, I., and Zhang, K. (2022). Reference points and redistributive preferences: Experimental evidence. *Journal of Public Economics*, 216.

Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2003). Optimal defaults. *American Economic Review*, 93(2):180–185.

Cohen, R., Maltz, A., and Ofek-Shanny, Y. (2025). Social preferences over taste-based inequality. *International Economic Review*. First published online: 15 April 2025.

Cohn, A., Jessen, L. J., Klašnja, M., and Smeets, P. (2023). Wealthy americans and redistribution: The role of fairness preferences. *Journal of Public Economics*, 225.

Dong, L., Huang, L., and Lien, J. W. (2022). They never had a chance: Unequal opportunities and fair redistributions. *CeDEx Discussion Paper Series*.

Drouvelis, M. and Gavassa-Pérez, E. (2024). Redistribution over gains and losses: Social preferences and moral rules. Mimeo.

Dwenger, N., Hoem Sjursen, I., and Vietz, J. (2024). What is fair? experimental evidence on fair equality vs fair inequality. Technical Report 11289, CESifo. Available at SSRN.

Espinosa, R., Deffains, B., and Thöni, C. (2020). Debiasing preferences over redistribution: an experiment. *Social Choice and Welfare*, 55:823–843.

Fischbacher, U., Grammling, D., Hausfeld, J., and Zíka, V. (2023). Identity breeds inequality: Evidence from a laboratory experiment on redistribution. *Journal of Public Economics*, 222:104866.

Freyer, T. and Guenther, L. (2025). Passive inequality and the dilemma of meritocracy. *Experimental Economics*, 28(2):379–400.

Haeckl, S., Moeller, J., and Zednik, A. (2025). Can being competitive but unsuccessful harm you, even more so if you are a woman? *Journal of Economic Behavior & Organization*, 236:107108.

Hufe, P. and Weishaar, D. (2025). Just cheap talk? investigating fairness preferences in hypothetical scenarios. *Journal of Economic Inequality*. Forthcoming.

Konow, J., Saijo, T., and Akai, K. (2020). Equity versus equality: Spectators, stakeholders and groups. *Journal of Economic Psychology*, 77:102171. Decision Making for Others.

Lu, K. (2025). Silver spoons and scales of justice: The fairness preference over unequal intergenerational wealth transfers between americans and chinese. *Journal of Economic Behavior & Organization*, 236:107093.

Mo, Z., Huang, Y., and Belguise, M. (2023). Non-meritocrats or conformist meritocrats? a redistribution experiment in china and france. Available at SSRN: `https://ssrn.com/abstract=4598219`.

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

Piraino, P. and Ryan, J. (2024). The price of fairness: Experimental evidence on the limits to demand for redistribution. *Journal of Economic Behavior & Organization*, 227:106737.

Preuss, M., Reyes, G., Somerville, J., and Wu, J. (2025). Inequality of opportunity and income redistribution. *Journal of Political Economy: Microeconomics*. Forthcoming.

Santamaría, D., Miller, L., Ubeda, P., and Padilla-Angulo, L. (2023). Determinants of inequality acceptability in a representative sample of the spanish population. *Journal of Behavioral and Experimental Economics*, 106.

Sartor, S. (2025). Meritocracy and birthrights. *Mimeo*.

Sterba, M. (2022). The fairness of inequality due to risk and effort choices. Technical Report 2022/5, Max Planck Institute for Research on Collective Goods.

Strang, L. and Schaube, S. (2025). (not) everyone can be a winner â the role of payoff interdependence for redistribution. *Journal of Public Economics*, 243:105320.

Telle, I. O. and Tjøtta, S. (2023). Context and preferences for equality in the spectator game. *Journal of the Economic Science Association*, 9(2):227–238.

Valasek, J., Vorjohann, P., and Wang, W. (2024). Fairness preferences, inequality acceptance and default effects. *Mimeo.*

Wang, W. (2024). Incentives vs. fairness: Navigating trade-offs in managerial decision making. *Mimeo.*

Yusof, J. and Sartor, S. (2025). Market luck: Skill-biased inequality and redistributive preferences. *Mimeo.*

A Appendix

A.1 Discussion of deviations from the PAP

Here we discuss two deviations from our Pre-analysis Plan (PAP). First, since we gathered data for the Luck and Merit treatments in two different rounds, we submitted separate PAPs for each round, and specified that we would analyze the data separately. However, in our second PAP, we do state that we will also run a model estimation on the combined data with "with the appropriate indicator variables for "Merit" (i.e. a baseline of Luck)." Due to an oversight, we did not register any experimental hypotheses related to this model estimation.

However, given that our experimental design is clearly suited to testing for differences in default effects across treatments, we find it justifiable to state and test Hypothesis 2 using the same method that we use to test for differences in default effects across populations.

Second, our PAP states that we will use our data to run additional tests to directly estimate β, and measure the impact of the treatments on the number of individuals who select an interior allocation of earnings. We have decided to omit these tests due to the short-paper format, and since they will not provide much additional insight beyond the results of our main analysis.

A.2 Additional theoretical results

Here we present a model for spectator preferences that allows us to formally distinguish between a spectator's conditional distributional preferences and default effects, and which rationalizes the representation of spectator preferences we derive our predictions from. To do so, we rely on the work of Breitmoser and Vorjohann (2024). The authors provide an axiomatic representation result that clarifies how contextual factors such as worker characteristics and defaults determine an altruistic decision-maker's (the spectator's) distributional preferences.

Applied to our framework, their result implies that the spectator's utility is determined by the sum of worker "well-beings" as captured by prospect theoretical value functions. That is, each worker's payoff is evaluated in comparison to a reference point with a value function that exhibits loss aversion. Importantly, Breitmoser and Vorjohann (2024)'s analysis further implies that the spectator's reference point for each worker is a linear function of the default and other observable contextual factors. In line with the literature, we subsume other contextual factors, specifically worker characteristics, by the spectator's "fairness norm"—i.e. the underlying conditional distributional preferences—as defined below.

Formally, the spectator's preferences are represented by the utility function:

$$u_i(y_i, \kappa, d) = v_i(y_{iA} - r_{iA}(\kappa, d)) + v_i(1 - y_i - r_{iB}(\kappa, d))$$

where v_i corresponds to the value function capturing workers' well-beings and $r_{ij}(\kappa, d)$ captures the reference point for worker j.[3] To simplify the analysis, we restrict attention to linear value functions given by

$$v_i(p) = \begin{cases} p & \text{if } p \geq 0 \\ -\delta_i(-p) & \text{if } p < 0 \end{cases}$$

with $\delta_i > 1$ implying loss-aversion.[4]

[3]Since all elements entering the spectator's utility function are to be interpreted from the spectator's viewpoint, value functions hold spectator indices and reference points hold both spectator and worker indices.

[4]Breitmoser and Vorjohann (2024)'s result allows for S-shaped value functions in the spirit of prospect theory. However, to avoid corner solutions, which are less relevant for our setting, we consider linear value functions.

In line with the literature, we assume worker characteristics give rise to a fairness norm $m_i(\kappa)$ specifying the payoff distribution $\{y_A = m_i(\kappa), y_B = 1 - m_i(\kappa)\}$ deemed fair by the spectator. That is, the "fairness norm" is the distribution preferred by the agent, absent any default effect, given the available information on contextual factors. As part of their representation result, Breitmoser and Vorjohann (2024) show that reference points are linear functions of observables, which in our case are the fairness norm given the worker characteristics and their default payoffs.

Formally, we express reference points as weighted averages

$$r_{iA}(\kappa, d) = \beta_i m_i(\kappa) + (1 - \beta_i)d$$

$$r_{iB}(\kappa, d) = \beta_i(1 - m_i(\kappa)) + (1 - \beta_i)(1 - d)$$

where $\beta_i \in [0, 1]$ captures the weight attached to the fairness norm. Note that this representation implies that reference points exactly exhaust the spectator's distribution budget, i.e. $r_{iA}(\kappa, d) = 1 - r_{iB}(\kappa, d) =: r_i(\kappa, d)$. The spectator's utility function then reduces to:

$$u_i(y_i, \kappa, d) = v_i(y_i - r_i(\kappa, d)) + v_i(r_i(\kappa, d) - y_i). \tag{9}$$

Given this representation of the spectator's preferences, the model's prediction is that the chosen division of payoffs is equal to the reference point division; i.e. the spectator selects:

$$y_i = \beta_i m_i(\kappa) + (1 - \beta_i)d. \tag{10}$$

A.3 Additional empirical results

Here we report on the coefficients on the control variables for Models 1,3 and 5 from Table 1.

Table 2: Coefficients on control variables in Table 1

	(1) Merit	(3) Luck	(5) Both
Age	-0.000627**	-0.00318***	-0.00181***
	(0.000297)	(0.000589)	(0.000345)
Female	-0.00306	0.0243	0.0104
	(0.00785)	(0.0161)	(0.00915)
Other	-0.0200	-0.116	-0.0737
	(0.0814)	(0.151)	(0.108)
Above Median Income	-0.00688	-0.0126	-0.00242
	(0.00842)	(0.0175)	(0.00995)
Political Orientation	0.0128***	0.0239***	0.0207***
	(0.00400)	(0.00839)	(0.00476)
Higher Education	0.00428	-0.0162	-0.00544
	(0.00955)	(0.0176)	(0.0108)
N	1767	1071	2250

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Online Appendix: "Spectator" Questionnaires

Treatment 1: Merit × Unequal Default

In contrast to traditional survey questions that are about hypothetical situations, we now ask you to make a choice that has consequences for a real life situation. A few days ago two individuals, let us call them worker A and worker B, were recruited via an international online market place to conduct an assignment.

Worker A and worker B were each offered a participation compensation of 2 USD regardless of what they were paid for completing the assignment. After completing the assignment, they were told that their productivity may determine their earnings from the assignment. They were not informed about who was the most productive worker. The most productive worker would earn 6 USD for the assignment and the other worker would earn nothing for the assignment. However, they were told that a third person would be informed about the assignment and who was the most productive worker, and would be given the opportunity to redistribute the earnings and thus determine how much they were paid for the assignment.

You are the third person and we now want you to choose whether to change the earnings for the assignment between worker A and worker B. Your decision is completely anonymous. The workers will receive the payment that you choose for the assignment within a few days, but will not receive any further information.

Worker A was most productive and earned 6 USD for the assignment, thus worker B earned nothing for the assignment. Please state which of the following alternatives you choose:

I do not change the earnings:

- worker A is paid 6 USD and worker B is paid 0 USD.

I do change the earnings:

- worker A is paid 5 USD and worker B is paid 1 USD.
- worker A is paid 4 USD and worker B is paid 2 USD.
- worker A is paid 3 USD and worker B is paid 3 USD.
- worker A is paid 2 USD and worker B is paid 4 USD.
- worker A is paid 1 USD and worker B is paid 5 USD.
- worker A is paid 0 USD and worker B is paid 6 USD.

Treatment 2: Merit × Equal Default

In contrast to traditional survey questions that are about hypothetical situations, we now ask you to make a choice that has consequences for a real life situation. A few days ago two individuals, let us call them worker A and worker B, were recruited via an international online market place to conduct an assignment.

Worker A and worker B were each offered a participation compensation of 2 USD regardless of what they were paid for completing the assignment. After completing the assignment, they were told that their productivity may determine their earnings from the assignment. They were not informed about who was the most productive worker. Both the most productive worker and the other worker would earn 3 USD for the assignment. However, they were told that a third person would be informed about the assignment and who was the most productive worker, and would be given the opportunity to redistribute the earnings and thus determine how much they were paid for the assignment.

You are the third person and we now want you to choose whether to change the earnings for the assignment between worker A and worker B. Your decision is completely anonymous. The workers will receive the payment that you choose for the assignment within a few days, but will not receive any further information.

Worker A was most productive and earned 3 USD for the assignment, and worker B also earned 3 USD for the assignment. Please state which of the following alternatives you choose:

I do not change the earnings:

- worker A is paid 3 USD and worker B is paid 3 USD.

I do change the earnings:

- worker A is paid 6 USD and worker B is paid 0 USD.
- worker A is paid 5 USD and worker B is paid 1 USD.
- worker A is paid 4 USD and worker B is paid 2 USD.
- worker A is paid 2 USD and worker B is paid 4 USD.
- worker A is paid 1 USD and worker B is paid 5 USD.
- worker A is paid 0 USD and worker B is paid 6 USD.

Treatment 3: Merit $\times$ No Default

In contrast to traditional survey questions that are about hypothetical situations, we now ask you to make a choice that has consequences for a real life situation. A few days ago two individuals, let us call them worker A and worker B, were recruited via an international online market place to conduct an assignment.

Worker A and worker B were each offered a participation compensation of 2 USD regardless of what they were paid for completing the assignment. After completing the assignment, they were told that their productivity may determine their earnings from the assignment. They were not informed about who was the most productive worker. However, they were told that a third person would be informed about the assignment and who was the most productive worker, and would choose how the total earnings for completing the assignment, $6, would be divided between the two of them and thus determine how much they were paid for the assignment.

You are the third person and we now want you to choose how to divide the earnings for the assignment between worker A and worker B. Your decision is completely anonymous. The workers will receive the payment that you choose for the assignment within a few days, but will not receive any further information.

Worker A was most productive. Please state which of the following alternatives you choose:

- worker A is paid 6 USD and worker B is paid 0 USD.
- worker A is paid 5 USD and worker B is paid 1 USD.
- worker A is paid 4 USD and worker B is paid 2 USD.
- worker A is paid 3 USD and worker B is paid 3 USD.
- worker A is paid 2 USD and worker B is paid 4 USD.
- worker A is paid 1 USD and worker B is paid 5 USD.
- worker A is paid 0 USD and worker B is paid 6 USD.

Treatment 4: Luck × Unequal Default

In contrast to traditional survey questions that are about hypothetical situations, we now ask you to make a choice that has consequences for a real life situation. A few days ago two individuals, let us call them worker A and worker B, were recruited via an international online market place to conduct an assignment.

Worker A and worker B were each offered a participation compensation of 2 USD regardless of what they were paid for completing the assignment. After they had completed the assignment, they were told that it was randomly decided that one of them would earn an additional 6 USD for the work on the assignment while the other would not earn anything additional for the work on the assignment. However, they were also told that a third person could change how the additional earnings would be divided between the two of them and thus determine how much they were paid for the assignment.

You are the third person and we now want you to choose whether to change the earnings for the assignment between worker A and worker B. Your decision is completely anonymous. The workers will receive the payment that you choose for the assignment within a few days, but will not receive any further information.

Worker A was randomly selected to earn 6 USD for the assignment, thus worker B earned nothing for the assignment. Please state which of the following alternatives you choose:

I do not change the earnings:

- worker A is paid 6 USD and worker B is paid 0 USD.

I do change the earnings:

- worker A is paid 5 USD and worker B is paid 1 USD.
- worker A is paid 4 USD and worker B is paid 2 USD.
- worker A is paid 3 USD and worker B is paid 3 USD.

Treatment 5: Luck × Equal Default

In contrast to traditional survey questions that are about hypothetical situations, we now ask you to make a choice that has consequences for a real life situation. A few days ago two individuals, let us call them worker A and worker B, were recruited via an international online market place to conduct an assignment.

Worker A and worker B were each offered a participation compensation of 2 USD regardless of what they were paid for completing the assignment. After they had completed the assignment, they were told that it was randomly decided that both of them would earn an additional 3 USD for the work on the assignment. However, they were also told that a third person could change how the additional earnings would be divided between the two of them and thus determine how much they were paid for the assignment.

You are the third person and we now want you to choose whether to change the earnings for the assignment between worker A and worker B. Your decision is completely anonymous. The workers will receive the payment that you choose for the assignment within a few days, but will not receive any further information.

Please state which of the following alternatives you choose:

I do not change the earnings:

- worker A is paid 3 USD and worker B is paid 3 USD.

I do change the earnings:

- worker A is paid 6 USD and worker B is paid 0 USD.
- worker A is paid 5 USD and worker B is paid 1 USD.
- worker A is paid 4 USD and worker B is paid 2 USD.