Are Elites Meritocratic and Efficiency-Seeking? Evidence from MBA Students*

Marcel Preuss Germán Reyes Jason Somerville Joy Wu

June 2026

Abstract

Elites disproportionately influence policymaking, yet little is known about their fairness and efficiency preferences—key determinants of support for redistributive policies. We investigate these preferences in an incentivized lab experiment with future elites: Ivy League MBA students. We find that MBA students implement substantially more unequal earnings distributions than the average American, regardless of whether inequality stems from luck or merit. Their redistributive choices are also far more responsive to efficiency costs than the near-zero response found in representative U.S. samples. These patterns partly reflect distinct fairness ideals: a large share of MBA students falls outside standard classifications, instead displaying "weak meritocratic" tendencies that tolerate inequality even when it stems from luck. These findings identify a channel through which elite preferences may sustain U.S. inequality.

*Preuss: Cornell University, email: mp2222@cornell.edu; Reyes (corresponding author): Middlebury College and IZA, email: greyes@middlebury.edu; Somerville: University of California, Santa Barbara, email: jasonsomerville@ucsb.edu; Wu: University of British Columbia, email: joy.wu@sauder.ubc.ca. For helpful comments and suggestions, we thank Peter Andre, Alexander Cappelen, Peter Matthews, Ricardo Pérez-Truglia, Erik Sørensen, Bertil Tungodden, and participants in FAIR NHH and the Cornell Behavioral Economics Research Group. The experiment reported in this paper was registered in the American Economic Association's registry for randomized controlled trials under ID AEARCTR-0015448. The experiment was reviewed and granted an exemption by the Institutional Review Board at Cornell University (Protocol #0147786).

1 Introduction

Support for redistributive policies depends on how individuals weigh fairness considerations against efficiency costs (e.g., Alesina and Angeletos, 2005; Bénabou and Tirole, 2006). Studies of representative U.S. samples find that individuals prefer to redistribute income when inequality stems from luck rather than effort, but show little sensitivity to efficiency costs (Almås et al., 2020, 2025; Cohn et al., 2023).

Policy outcomes, however, are not determined by the preferences of the general public alone. Theories of "elite control" in political science argue that elites—individuals with substantial economic, political, or social capital—largely determine which policies are adopted.[1] Consistent with these theories, empirical evidence shows that elites have a disproportionate influence on policymaking, while the median voter has little or no influence. For example, Gilens and Page (2014) show that policy adoption in the U.S. is strongly related to support from business groups and the richest ten percent, and is uncorrelated with the preferences of the average American.[2] Understanding U.S. inequality therefore requires measuring elites' fairness and efficiency preferences.

Yet little work has estimated these preferences, partly because elites are difficult to reach. We address this gap by eliciting incentivized redistribution choices from two cohorts of Cornell MBA students—*future* elites. Students from elite universities are disproportionately likely to reach the top one percent (Chetty et al., 2020), and graduates from these programs are overrepresented in U.S. leadership positions (Wai et al., 2024; Chetty et al., 2026). Alumni of the Cornell MBA program include CEOs of Fortune 500 companies, members of Congress, board members, and founders of major companies (Cornell University, 2026).

To estimate MBA students' redistributive preferences, we follow the impartial-spectator paradigm (Cappelen et al., 2013). This experimental design involves "workers" and "impartial spectators" across three stages. In the production stage, workers complete a real-effort task. In the earnings stage, we randomly pair workers and determine their earnings based

[1] See, among others, Mills (1956), Burch (1980), Ferguson (1995), Winters and Page (2009), Winters (2011), Beard (2012), and Domhoff (2013, 2017).

[2] For further evidence in the U.S., see Hacker and Pierson (2010), Gilens (2012), Rigby and Wright (2013), Gilens and Page (2014), Bartels (2016), Page and Gilens (2017), Page et al. (2018), and Hertel-Fernandez (2019). Evidence on the influence of elites on policymaking from other countries is mixed. Single-country studies from Germany, Norway, and the Netherlands show that elites exert disproportionate influence on policymaking (Elsässer et al., 2021; Schakel, 2021; Mathisen, 2023). However, in cross-country regressions, the preferences of low-income individuals are more predictive of redistribution than those of high-income individuals (Maréchal et al., 2025).

on either task performance or chance. Thus, earnings inequality is either due to merit or luck. In the redistribution stage, MBA students acting as impartial spectators choose the final earnings allocations for three worker pairs. Each worker pair differs in the source of inequality (merit vs. luck) and the cost of redistributing earnings (no cost vs. costly redistribution). Our main analysis uses spectators' first redistributive choice, the standard between-subject design in the literature. We complement this with a within-subject analysis that uses all three decisions per spectator.

We report two main findings. First, MBA students implement more unequal earnings distributions than those chosen by the average American, regardless of whether worker earnings are due to luck or merit. MBA students implement a Gini coefficient of 0.43 when worker earnings reflect luck and 0.61 when they reflect task performance. These levels exceed the corresponding pooled estimates of 0.29 and 0.57 across four representative U.S. samples (Almås et al., 2020; Cohn et al., 2023; Preuss et al., 2025; Harrs and Sterba, 2025).

Second, MBA students' redistributive choices respond to the efficiency cost of redistribution. When worker earnings are randomly assigned, introducing an efficiency cost increases the implemented Gini coefficient by 0.21 points. This response is much larger than in representative U.S. samples, where estimated responses to efficiency costs are typically statistically indistinguishable from zero (Almås et al., 2020, 2025). We replicate these findings with Cornell undergraduate business students—another sample of future elites.

To understand why future elites implement more unequal distributions than the average American, we classify them according to established fairness ideals in the literature. Following the standard methodology of Almås et al. (2020), we estimate that 9.9 percent of MBA students are egalitarians, 23.2 percent are libertarians, 22.6 percent are meritocrats, and 44.3 percent do not fit any of these three fairness ideals. Relative to representative U.S. samples, the fraction of egalitarians is similar, the fractions of meritocrats and libertarians are lower, and the unclassified fraction is much larger (Almås et al., 2020, 2025; Cohn et al., 2023; Preuss et al., 2025; Harrs and Sterba, 2025). We use our repeated-measures design and find that many of these unclassified spectators display weak meritocratic tendencies—they redistribute more when inequality stems from luck than from performance, but still allow luck-based winners to retain an earnings premium. We refer to these unclassified weak meritocrats as "moderates." Open-ended responses on the drivers of U.S. inequality show that moderates attribute inequality to multiple forces—education, unequal opportunities, and discrimination—a profile distinct from any standard fairness ideal.

Our findings suggest that U.S. redistributive policies remain more limited than the

average citizen would prefer because elites, exerting disproportionate influence on policy, tolerate more inequality. This mismatch fits a broader puzzle: the U.S. remains a highly unequal country despite widespread support for a more egalitarian distribution (Norton and Ariely, 2011, 2013). This preference gap is also consistent with recent evidence that the welfare weights implied by actual U.S. tax and transfer policies are less progressive than those of the general population, with high-income individuals receiving disproportionate weight (Capozza and Srinivasan, 2024).

We contribute to the literature on the relationship between fairness views and income redistribution. Empirical evidence shows that the source of inequality—effort versus luck—shapes individuals' attitudes toward redistribution.[3] Recent work estimates how fairness ideals are distributed in representative population samples (Almås et al., 2020, 2025; Müller and Renes, 2021; Cohn et al., 2023; Harrs and Sterba, 2025). We estimate fairness ideals in a sample of *future* elites who will likely influence policymaking and business decisions that directly affect inequality, such as worker compensation schemes.

Two related papers also study elites' distributional preferences: Fisman et al. (2015) and Cohn et al. (2023). Fisman et al. (2015) study Yale Law students' distributional preferences using modified dictator games in which subjects divide an endowment between themselves and another participant. They find that elites place more weight on their own payoff and prioritize efficiency over *altruism*—i.e., reducing their own payoff to help others. By contrast, we find that elites prioritize efficiency over concerns for *equality*. Moreover, when subjects have a personal stake in the allocation as in Fisman et al. (2015), separating self-interest from fairness preferences requires strong functional-form assumptions.[4] Further, Nax et al. (2021) report a failed replication and argue that the observed elite–non-elite differences may reflect protocol variations. Our impartial-spectator design addresses both concerns: by removing subjects' stake in the outcome, we measure fairness and efficiency preferences directly without structural assumptions, and by using a consistent protocol across populations, we ensure comparability with existing studies. Cohn et al. (2023) find that high-income or high-wealth Americans—households in the top 5 percent by income or wealth, identified using self-reported survey data—are less inequality-averse than the

[3]Empirical work using lab experiments includes Cappelen et al. (2007), Cappelen et al. (2010), Cappelen et al. (2013), Cappelen et al. (2022), Almås et al. (2010), Almås et al. (2011), Durante et al. (2014), Mollerstrom et al. (2015), Andre (2025), Dong et al. (2025), Bhattacharya and Mollerstrom (2025), Preuss et al. (2025), Harrs and Sterba (2025), and Yusof and Sartor (2025).

[4]Individuals use uncertainty or inefficiencies as "moral wiggle room" to excuse selfishness (Dana et al., 2007; Exley, 2016, 2020). The functional form used to estimate efficiency concerns neglects this possible confound, which may inflate efficiency estimates for more selfish subjects.

general population. Beyond studying a different population, our paper complements and is distinct from their work in two main ways. First, we examine how efficiency concerns shape redistribution—a parameter that prior representative-sample studies estimate to be near zero, a puzzling null result. Second, our in-class recruitment mitigates the self-selection bias of survey-based samples, in which individuals with particular social preferences are more likely to participate (Levitt and List, 2007).

Finally, we contribute to the literature on the determinants of labor market inequality. While firm wage-setting policies drive substantial earnings inequality (e.g., Card et al., 2018), why firms pay different wages to observationally equivalent workers remains poorly understood.[5] One potential explanation centers on managers' role in designing compensation schemes (Frank et al., 2015; Acemoglu et al., 2025; He and le Maire, 2025). If managers' fairness preferences shape wage-setting decisions, then heterogeneity in these preferences could generate firm-level wage dispersion. We address this hypothesis by having future managers allocate real earnings between observationally equivalent workers, varying only the stated source of worker earnings across decisions. We document substantial heterogeneity in implemented earnings inequality across participants, suggesting that variation in managers' fairness preferences may contribute to wage differentials among otherwise similar workers.

2 Experimental Design

We follow the standard impartial-spectator paradigm (Cappelen et al., 2013). The experiment has two types of participants—workers and impartial spectators—across three stages: a production stage, an earnings stage, and a redistribution stage (Appendix Figure A1 shows the flow of the experiment).

2.1 Stages of the Experiment

In the production stage, workers have five minutes to complete a real-effort encryption task (Appendix Figure A2 shows an example). We inform workers that their earnings depend on their relative performance and a third party's decision, but we do not describe the redistribution mechanism.

[5] Recent work studying firm wage-setting includes Hall and Krueger (2012), Caldwell and Harmon (2019), Lachowska et al. (2022), Cullen et al. (2025), Derenoncourt and Weil (2025), Dube et al. (2025), Reyes (2025), Hazell et al. (2026), and Hjort et al. (2026).

In the earnings stage, workers are randomly paired and compete in a winner-take-all format. We determine the winner of each pair either by the number of encryptions completed ("performance") or by a coin flip ("luck"). Winners receive an initial allocation of \$6, while losers receive \$0. Workers never learn the outcome of their competition—they observe only their final payment after redistribution.

In the redistribution stage, spectators (MBA students) choose final earnings allocations for three worker pairs. Each pair differs in how the winner is determined and whether redistribution is costly. Spectators know which worker won each competition but not how many encryptions workers completed. Spectators can redistribute any amount ranging from \$0 to \$6 in \$0.50 increments. We present each decision as an adjustment schedule. We incentivize spectators by randomly selecting one of their decisions to implement.[6]

After the redistribution stage, spectators complete an exit survey covering demographics, socioeconomic background, career aspirations, social views, and an attention check. The second MBA cohort also answers an open-ended question on the main drivers of income inequality in the U.S.

2.2 Treatment Conditions

We study three conditions, presented to spectators in a randomized order (see Appendix Figure A3 for screenshots of each treatment). In the *luck* treatment, the winner is determined by a coin flip. In the *performance* treatment, the winner is the worker who completed more encryptions. In the *efficiency cost* treatment, the winner is determined by a coin flip, and redistribution incurs an "adjustment cost" that reduces total participant earnings. For every \$1.00 reduction in the winner's earnings, the loser's earnings increase by only \$0.50, resulting in a net loss of total income.[7]

[6] To mitigate the impact of anchoring effects on redistribution decisions, we tell spectators that workers were not informed about whether they won or lost their match, but would only be told their final earnings.

[7] One potential concern is that spectators might try to minimize cognitive effort by selecting default options, which could bias our results. Our experimental design addresses this in two ways: no option was pre-selected, requiring spectators to actively choose, and we presented the options horizontally rather than vertically to prevent defaulting to "top" choices. No spectator clicked through all redistributive decisions without making at least one active choice, suggesting effort-minimizing concerns are unlikely to bias our results.

3 Data and Summary Statistics

3.1 Recruitment of Workers and Spectators

We recruited 880 U.S.-based individuals on Prolific to work on the encryption task.[8] We paid workers a $1.50 participation fee for completing the task, plus a bonus of up to $6 based on a randomly chosen spectator's decision. The average bonus was $2.86, resulting in average total compensation of $4.36 per worker. (See Appendix B.1 for additional details on the worker sample.)

We collected MBA students' redistributive decisions from two consecutive cohorts (2023 and 2024) during mandatory classes. This approach mitigates the self-selection concerns that can arise in experiments (Levitt and List, 2007). Of the 338 MBA students who started the experiment, we exclude 67 who quit before reaching the first redistribution screen, leaving an analysis sample of 271 students.[9]

3.2 Summary Statistics and Balance

Appendix Table A1 presents summary statistics for the spectators, both overall (column 1) and by the first treatment shown (columns 2–4). Most MBA students are aged 24–31 (83 percent), male (55 percent), and U.S.-born (56 percent). They strongly prefer private-sector careers (84 percent) and managerial roles (91 percent), with most planning to work in the U.S. after graduation (94 percent). The majority voted in recent elections (69 percent) and believe that hard work leads to a better life (81 percent). Observable characteristics are balanced across the first treatment shown: F-tests of joint equality fail to reject for any condition.

To benchmark our results, we compare MBA students' redistributive choices with those in four studies that estimate the distribution of fairness ideals in the U.S. general population: Almås et al. (2020), Cohn et al. (2023), Harrs and Sterba (2025), and the Survey of Consumer Expectations (SCE) collected by Preuss et al. (2025). For studies that report the share of earnings redistributed, we use that measure to construct the corresponding

[8] These were paired into 440 worker pairs, which were matched to redistributive decisions from our 271 MBA students and 167 undergraduate business students.

[9] The experiment was conducted at the end of class sessions with voluntary participation clearly stated. MBA students often have commitments immediately following class, including case interview practice sessions, recruiting calls, and networking meetings, which may have led some students to leave before completing the experiment.

Gini coefficient.[10] We also report a pooled estimate across all representative samples using a DerSimonian–Laird random-effects meta-analysis (DerSimonian and Laird, 1986). Appendix C compares the experimental design, recruitment, and sample characteristics across benchmark studies.[11]

4 Origin of Income Differences and Implemented Inequality

4.1 Effects of Merit and Efficiency on Redistribution

We estimate linear models of the form:

$$\text{Gini}_{ip} = \beta_0 + \beta_1 \text{Performance}_p + \beta_2 \text{EfficiencyCost}_p + \varepsilon_{ip}, \tag{1}$$

where Gini_{ip} is the Gini coefficient of the final earnings allocation in worker pair p implemented by spectator i, Performance_p and EfficiencyCost_p are indicators that equal one if pair p competed in the performance or efficiency cost treatments and zero if they competed in the luck treatment. β_0 measures the average Gini when worker earnings are randomly assigned and there is no redistribution cost. β_1 measures the causal impact of assigning worker earnings based on performance on implemented inequality. β_2 measures the effect of introducing an efficiency cost on the Gini.[12]

Table 1 presents regression estimates of equation (1). Columns 1 and 2 use all redistributive decisions from spectators, with column 1 including no controls (thus, identification is based on both between- and within-subject variation) and column 2 including spectator fixed effects (identification based on within-subject variation). Columns 3 and 4 use only spectators' first decisions (identification based on between-subject variation), though the order in which we presented the treatments has no statistically significant effect on redistributive choices (Appendix Table A2). We cluster standard errors at the spectator

[10] While these comparison studies were conducted at different times, research shows that distributional preferences remain stable over time (Fisman et al., 2023; Harrs and Sterba, 2026), making these cross-study comparisons informative.

[11] Another relevant comparison is Fisman et al. (2015), who also study elites' distributional and efficiency preferences. Because their experimental design, data, and identification strategy differ substantially from ours—they use a dictator game with structural estimation of a CES utility function—we do not include them in the main benchmarking exercise. In Appendix C.4, we compare the implied efficiency parameters in our paper using their methodology.

[12] While β_2 represents the effect of switching EfficiencyCost_p from 0 to 1, we can derive the implied arc elasticity with respect to the efficiency cost c. With $c = 0$ in the no-cost condition and $c = 0.5$ in the efficiency cost condition, the arc elasticity is approximately $\epsilon \approx (\beta_2/\bar{G})/(\Delta c/\bar{c})$, where $\bar{G}$ is the mean Gini coefficient, $\Delta c = 0.5$, and $\bar{c} = 0.25$ is the midpoint of the two cost levels.

level throughout. Appendix B.2 shows descriptive evidence on implemented inequality and redistribution choices across experimental conditions.

Merit-based earnings inequality and efficiency costs increase implemented inequality. When worker earnings are randomly assigned, spectators implement a Gini coefficient of $\hat{\beta}_0 = 0.420$ (columns 1 and 2). Assigning worker earnings based on their performance increases the implemented Gini coefficient by $\hat{\beta}_1 = 0.240$ Gini points ($p < 0.01$), or 57 percent of the luck condition Gini. Similarly, introducing a redistribution cost increases the Gini coefficient by $\hat{\beta}_2 = 0.188$ Gini points ($p < 0.01$), or 45 percent of the luck condition Gini. This corresponds to an implied elasticity of approximately 0.17: a one percent increase in the cost is associated with roughly a 0.17 percent increase in the Gini.

The coefficients are similar when using only the first decision of each spectator, albeit with a slightly smaller impact of the performance condition. Column 3 shows that assigning worker earnings based on performance (relative to luck) increases the Gini coefficient by $\hat{\beta}_1 = 0.191$ Gini points ($p < 0.01$), or 44 percent of the luck condition Gini. Still, given the standard errors, we cannot reject equality of coefficients between columns 1 and 3 at conventional levels.

The results are robust to several sample restrictions, specification checks, and alternative dependent variables. Appendix B.3 shows robustness to (i) excluding spectators who failed the attention check (Appendix Table B2); (ii) excluding spectators who rushed through the experiment (Appendix Table B3); (iii) excluding spectators who allocated strictly more earnings to the loser than to the winner (Appendix Table B4); (iv) excluding spectators whom we classify as having non-standard fairness ideals (Appendix Table B5); (v) excluding spectators who redistributed more when there was an efficiency cost (Appendix Table B6); (vi) estimating the regression separately for each cohort (Appendix Table B7); (vii) including round fixed effects (Appendix Table B8); and (viii) using the net-of-efficiency-cost share of earnings redistributed as the outcome (Appendix Table B9).

4.2 Are Elite Redistributive Choices Different from the Average American's?

To compare MBA students with U.S. benchmark samples, Figure 1 plots mean Gini coefficients by treatment condition and Table 2 presents side-by-side regression estimates of equation (1), with Panel A comparing MBA students to the general population and Panel B to higher-income subsamples. Column 1 reproduces the first-decision specification from column 4 of Table 1, which is the design most comparable to the benchmark studies.

Future elites are more tolerant of inequality when earnings differences stem from luck.

In the luck condition, MBA students implement a Gini of 0.425 (Table 2, Panel A, column 1), compared with a pooled estimate of 0.292 across the four representative U.S. samples (column 6). The MBA Gini significantly exceeds those of Cohn et al. (2023) ($p < 0.001$) and Harrs and Sterba (2025) ($p = 0.003$), and is statistically indistinguishable from the higher-inequality Almås et al. (2020) and Preuss et al. (2025) samples ($p = 0.410$ and $p = 0.663$, respectively). The MBA–pooled gap amounts to about 68 percent of the U.S.–Norway implemented-inequality gap in Almås et al. (2020). The conclusion extends to higher-income benchmark samples (Panel B): the pooled luck-condition Gini is 0.281. The gap persists even against Cohn et al. (2023)'s top-5%-by-income-or-wealth sample, which implements a luck-condition Gini of 0.271—about 36 percent smaller than the MBA estimate.[13]

Future elites are less responsive to the source of inequality than the average American. The performance condition increases the implemented Gini by 0.184 points among MBA students, compared with a pooled estimate of 0.279 across the representative samples (Table 2, column 6). Yet MBA students start from a much higher baseline in the luck condition, so they still implement more inequality in absolute terms: their performance-condition Gini is 0.609, compared with a pooled estimate of 0.571 across representative samples. The pattern is similar for higher-income subsamples: the pooled performance-condition Gini is 0.540 (Panel B), still below the MBA estimate. The same pattern holds against Cohn et al. (2023)'s top-5%-by-income-or-wealth sample: although their performance effect (0.253 points) exceeds the MBA estimate, their lower luck-condition baseline (0.271) leaves their performance-condition Gini at 0.524—still below the MBA level.[14]

The largest gap between MBA students and the general population is in their responsiveness to efficiency costs. The efficiency-cost condition increases the Gini by 0.211 points among MBA students, compared with a statistically insignificant 0.011 points in the representative U.S. sample of Almås et al. (2020)—the only other study with an equivalent

[13] Cohn et al. (2023) find that business owners are the most tolerant of inequality. Inequality tolerance in this subgroup is also the closest to our sample, although a gap remains in the luck treatment where business owners still implement less inequality than MBA students (a Gini of 0.350 vs. 0.432).

[14] Some political economy models suggest that the median voter's preferences, rather than the mean, determine policy outcomes (e.g., Downs, 1957). To assess whether our comparisons are sensitive to this distinction, we compute the median and mean implemented Gini in each of the representative samples and take a simple unweighted average across studies (distinct from the random-effects pooled estimate of 0.292 reported above). In the performance condition, the mean and median Gini are similar (0.59 vs. 0.54). In the luck condition, however, the median Gini (0.05) is substantially lower than the mean (0.30). Therefore, comparing future elites to the median citizen yields an even larger preference gap than comparing to the mean citizen.

efficiency-cost treatment (same 50 percent cost of redistribution). The difference between these two efficiency responses is statistically significant ($p = 0.004$). In a re-analysis of Almås et al. (2020)'s data, we find that respondents with household income above $100,000 also respond significantly to efficiency costs (0.174 Gini points, $p < 0.05$), though MBA students' response remains larger.[15]

4.3 Replication with Elite Undergraduate Business Students

To assess the robustness of these results, we replicate our analysis with *undergraduate* Cornell business students from the Dyson School of Applied Economics and Management (henceforth "Dyson students"; see Appendix D for details). Dyson students are another future-elite sample: Ivy League universities show high mobility to the top one percent (Chetty et al., 2020) and disproportionate representation in leadership (Chetty et al., 2026).

Dyson students implement earnings distributions as unequal as those of MBA students: a Gini coefficient of 0.416 when worker earnings are determined by luck and 0.727 when determined by performance—substantially higher than the Ginis in representative U.S. samples. Dyson students' sensitivity to efficiency costs also mirrors that of MBA students: redistribution costs increase the Gini by 0.198 points, nearly identical to the MBA estimate and substantially larger than the near-zero response in representative samples. This consistency across undergraduate and graduate business populations supports our conclusion that future elites tolerate more inequality and respond more strongly to efficiency costs than the general population.

5 The Fairness Ideals of the Elite

5.1 Measuring Fairness Ideals

To understand why future elites implement more inequality than the general population, we classify MBA students by their fairness ideal. We develop a statistical framework that models fairness ideals as mappings from initial earnings distributions to final allocations (see Appendix E). Following the literature, we focus on three fairness ideals: egalitarian,

[15] Recent work by Almås et al. (2025), for which microdata are not available, estimates that the efficiency-cost condition increases the Gini by about 0.08 points in their U.S. sample (see their Figure 5, panel b), a coefficient that is also statistically indistinguishable from zero. The MBA efficiency response exceeds that of all countries in their worldwide sample, confirming that future elites are substantially more responsive to efficiency costs than representative populations.

libertarian, and meritocratic. Egalitarians equalize workers' earnings regardless of how the earnings were generated. Libertarians leave the initial earnings unchanged regardless of how earnings differences arose. Meritocrats condition their redistributive decisions on the source of inequality: they equalize earnings when inequality is entirely luck-based (i.e., there is no merit to the earnings allocation), but redistribute less when inequality reflects performance differences. These fairness ideals predict individuals' social preferences for redistribution (Harrs and Sterba, 2025).

We use two complementary designs to measure fairness ideals. The first approach relies on between-subject comparisons, using only each spectator's first redistributive decision and excluding those who saw the efficiency cost environment first. This is the standard design used in the literature (e.g., Almås et al., 2020, 2025; Cohn et al., 2023) and lets us benchmark our results against prior work. The second approach uses within-subject variation from multiple decisions per spectator across the luck and performance treatments. This approach, also used by Harrs and Sterba (2025), lets us identify each spectator's fairness ideal, not just aggregate shares. We extend the standard identification assumptions to accommodate these multiple observations per spectator (see Appendix E).

The first-choice design provides clean comparisons across individuals and avoids biases arising from sequential decisions, but is more vulnerable to measurement error and choice noise. The repeated-measures design reduces measurement error by incorporating more information about each spectator's behavior, but requires assuming that initial choices do not influence subsequent decisions.[16] We present results from both approaches to establish robustness and to ease comparison with existing research.

5.2 Estimates of Fairness Ideals

Based on the spectators' first choices, we estimate that 9.9 percent of MBA students are egalitarians, 23.2 percent are libertarians, 22.6 percent are meritocrats, and 44.3 percent follow other fairness ideals (Table 3, Panel A; Figure 2 plots these estimates alongside the pooled random-effects benchmark). These estimates are similar to those from the repeated-measures design, which classifies 6.5 percent as egalitarians, 15.3 percent as libertarians, 28.7 percent as meritocrats, and 49.4 percent as following other fairness ideals. Differences in proportions across the two designs are not statistically significant at conventional levels (Appendix Figure A4). A chi-squared Wald test fails to reject equality of distributions

[16] Recent evidence shows that individual-level fairness type classifications are highly stable over time (Harrs and Sterba, 2026), further supporting the use of within-subject variation to identify fairness ideals.

across the two designs ($p = 0.27$), suggesting that our fairness-ideal estimates are robust to the elicitation method.[17]

The distribution of fairness ideals among MBA students differs markedly from that of the general population (Table 3, Panel B). Relative to the pooled benchmark across all four representative U.S. samples, the shares of libertarians and meritocrats are smaller among MBA students. The largest difference is in the "other" category: 44.3 percent of future elites are unclassified, compared with a pooled estimate of 19 percent across the four benchmark samples.[18] A chi-squared Wald test rejects the equality of distributions between the MBA sample and each of the four representative samples individually: $p = 0.005$ for the SCE, $p = 0.002$ for Almås et al. (2020), $p < 0.001$ for Cohn et al. (2023), and $p < 0.001$ for Harrs and Sterba (2025).

MBA students' fairness preferences differ systematically from those of higher-income Americans (Table 3, Panel C). MBA students are less likely to be meritocrats, and a larger proportion falls outside the three conventional fairness ideal classifications. This contrasts with prior work showing that higher-income Americans are less frequently egalitarian and more frequently meritocratic than the average American, though findings vary across studies. A chi-squared Wald test rejects equal distributions between the MBA sample and the higher-income samples in Cohn et al. (2023) ($p < 0.001$) and Harrs and Sterba (2025) ($p = 0.004$), and yields marginal rejections for the SCE ($p = 0.060$) and Almås et al. (2020) ($p = 0.081$).

5.3 Understanding the Fairness Views of Future Elites

Who are the spectators who do not conform to standard fairness ideals? To answer this, we first assess whether unclassified fairness views reflect deliberate choices. Then, we identify systematic patterns in redistribution decisions using the repeated-measures design.

Are Non-Standard Fairness Views Deliberate or Random? Spectators classified as "other" make deliberate choices rather than randomly clicking through the survey. Four pieces of evidence support this. First, unclassified spectators spend just as much time making their redistribution decisions as classified spectators (Appendix Table A3). Second,

[17] Some cross-design differences are sizable despite being statistically insignificant, particularly for libertarians (23.2 vs. 15.3 percent).

[18] The high proportion of individuals in the "other" category is not an artifact of our definitions of fairness ideals—we employ the same classification approach as Almås et al. (2020) in our between-subject analysis, yet still find a substantially larger share of unclassified individuals than in prior studies.

they pass the embedded attention check at the same rate as other spectators.[19] Third, the joint distribution of their redistribution choices across the luck and performance conditions shows systematic patterns rather than the uniform distribution we would expect from random clicking (discussed below). Fourth, our replication study with another future-elite sample—the Dyson students—finds similar rates of unclassified spectators (Section 4.3).

Redistributive Patterns Leading to Unclassified Elites. To identify the patterns behind unclassified fairness ideals, we examine the joint distribution of redistribution choices in the performance and luck conditions (Appendix Table A4). The most common unclassified pattern is to redistribute more in the luck condition than in the performance condition, but without fully equalizing earnings under luck. Specifically, 11.1 percent of spectators redistribute \$1 or \$2 in the luck condition while redistributing \$0 in the performance condition, and 8.4 percent redistribute \$2 in the luck condition and \$1 in the performance condition. These spectators behave as "weak meritocrats": they reward effort differences but still let luck-based winners retain a premium by not fully equalizing earnings in the luck condition. Together, these choices account for 19.5 percent of all spectators (or 39.5 percent of unclassified spectators). We refer to these spectators as "moderates."[20]

Moderates implement more unequal earnings distributions than meritocrats. On average, moderates redistribute 18.8 percent of earnings and implement a Gini coefficient of 0.624, whereas meritocrats redistribute 31.7 percent of earnings and implement a Gini coefficient of 0.367. These differences are statistically significant at $p < 0.01$. The key distinction between moderates and meritocrats is that moderates preserve some inequality even when it results purely from chance, contradicting the meritocratic principle that only merit-based differences should be rewarded.[21]

[19] In our data, 6.7 percent of all spectators fail the attention check. A bivariate regression of an indicator for failing the attention check on the "other ideal" dummy yields a statistically insignificant $\hat{\beta} = -0.005$ ($p = 0.88$).

[20] The remaining 60 percent of unclassified spectators show no single dominant pattern—the largest subgroup redistributes the same intermediate amount regardless of the source of inequality, and the rest are scattered across the joint distribution (Appendix Table A4)—so they fit none of the standard fairness ideals and are best understood as a heterogeneous residual rather than a coherent type. The only characteristic that predicts membership is a marginally lower propensity to have voted in recent elections (Appendix Table A5; $p < 0.10$).

[21] Several individual characteristics correlate significantly with fairness type (Appendix Table A5). Men are more likely than women to be libertarians ($p < 0.01$), respondents who believe that working long hours is necessary are more likely to be moderates ($p < 0.05$), and politically active students—measured by voting in recent elections—are more likely to be meritocrats ($p < 0.05$).

Inequality Narratives. To further characterize moderates relative to other fairness types, we analyze the second MBA cohort's open-ended responses to the question: "What do you believe is the main driver of income inequality in the United States?" We use large language models (LLMs) to classify the responses into eight categories developed inductively from the data: education, unequal opportunities, discrimination, government policy, corporate/elite power, historical legacy, behavioral/cultural explanations, and other. For each response, the LLMs produced a single-best label and multi-label coding. For example, responses mentioning "school," "college," or "studying" are classified under education, while mentions of race, gender, or bias are classified under discrimination. Appendix F details the methodology, robustness checks, and representative quotes.

The classification shows systematic differences across fairness types (Appendix Table F2 and Figure F1). Meritocrats emphasize education (38.9 percent), libertarians emphasize behavioral and cultural explanations (26.7 percent), and egalitarians emphasize barriers—unequal opportunities (62.5 percent) and discrimination (25.0 percent). Moderates share meritocrats' emphasis on education but cite barriers—discrimination (23.1 vs. 11.1 percent) and unequal opportunities (26.9 vs. 19.4 percent)—more often. Unlike meritocrats, who frequently cite historical and systemic forces (25.0 percent), moderates almost never do (3.8 percent), instead attributing inequality to multiple proximate causes. Moderates differ from libertarians in their low emphasis on behavioral explanations (11.5 vs. 26.7 percent) and high emphasis on barriers. Taken together, these patterns indicate that moderates hold a distinct view of why inequality exists, one not captured by standard fairness type classifications.[22]

6 Conclusion

We examine which types of inequality future elites consider worthy of redistribution and how much weight they give to efficiency. We find that MBA students implement substantially more unequal earnings distributions than the average American and are far more responsive to efficiency costs.

Our results speak to a puzzle in the political economy of inequality: while most Amer-

[22] As an exploratory accounting exercise, we ask how much of the moderate–meritocrat gap in redistribution can be explained by their differing narratives. We regress the share of earnings redistributed on indicators for each narrative category, pooling moderates and meritocrats, and multiply each coefficient by the moderate–meritocrat difference in how often each category is cited. Summing across categories, narratives explain about 20 percent of the gap, driven mainly by meritocrats citing historical and systemic causes more often, which is associated with more redistribution.

icans express strong preferences for more equal income distributions (Norton and Ariely, 2011, 2013), the U.S. remains a highly unequal nation (Chancel et al., 2022; Saez and Zucman, 2022). Our results suggest one explanation: future elites, who are likely to enter positions with disproportionate influence over policy and organizational decisions, tolerate substantially more unequal earnings distributions than the average American, and may therefore be more reluctant to adopt redistributive policies. This aligns with recent evidence that the general population's welfare weights are substantially more progressive than those implied by actual tax and transfer policies (Capozza and Srinivasan, 2024). The prevalence of moderates among future elites further suggests that elites do not just tolerate more inequality—they think about its causes differently, in ways existing fairness ideal classifications miss.

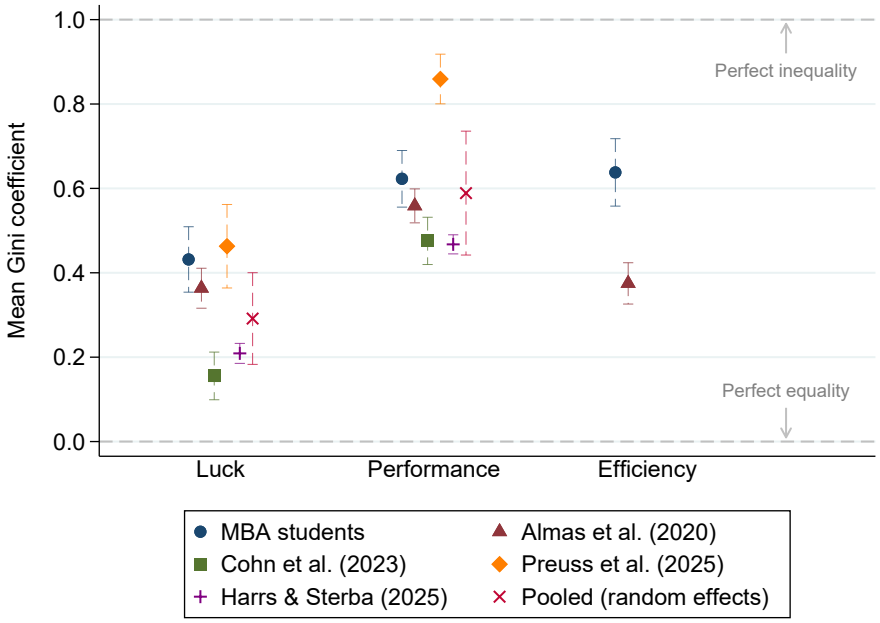
Although we document differences in redistributive preferences between future elites and the general population, we cannot pin down why they differ. Observable characteristics such as education, socioeconomic background, or career aspirations may explain these differences, but the direction of causality remains unclear. Individuals with efficiency-seeking preferences may self-select into elite business programs, or these programs may shape their preferences through training and socialization (Bauman and Rose, 2011; Sundemo and Löfgren, 2025). While these causal mechanisms warrant future research, our core finding is unchanged: a sizable preference gap exists between the general population and those who will likely shape policy.

Whether these preferences persist as MBA students advance into positions of power is an open question. Three pieces of evidence suggest these preferences may endure. First, MBA and undergraduate business students yield similar results, indicating that future elites' redistributive preferences are stable across educational stages. Second, distributional preferences tend to be stable over time (Fisman et al., 2023; Harrs and Sterba, 2026). Third, Almås et al. (2020) find that age has no economically significant effect on redistributive preferences, suggesting that preference gaps between future elites and the general population are unlikely to narrow with age alone.

Another promising direction is exploring preference heterogeneity across elite groups. Demographic and ideological differences exist across Silicon Valley, Wall Street, and Capitol Hill elites (Broockman et al., 2019; Bühlmann et al., 2025). Business leaders may have different preferences from political, academic, or cultural elites. Measuring these preferences across elite groups and tracing how they translate into policy influence are natural extensions.

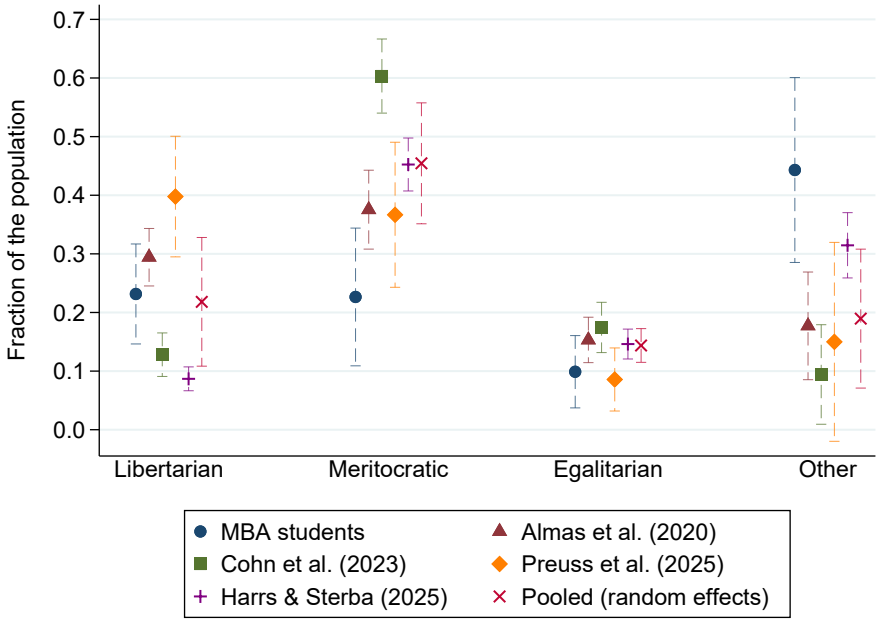
Figures and Tables

Figure 1: Implemented Inequality Across Treatments: Elites vs. Benchmark Studies



Notes: This figure plots mean Gini coefficients with 95 percent confidence intervals for each treatment condition across studies. The Gini coefficient measures implemented inequality in the final earnings allocation, ranging from zero (perfect equality) to one (maximum inequality). For Cohn et al. (2023), we construct the Gini as $1 - 2 \times$ (share redistributed); observations are weighted to be representative of the U.S. population. For Harrs and Sterba (2025), we construct the Gini as $|1 - \text{transfer}/2|$ using each respondent's first decision only. The efficiency treatment is available only in our experiment and in Almås et al. (2020). The pooled estimate is a DerSimonian–Laird random-effects meta-analysis across the four benchmark studies (DerSimonian and Laird, 1986).

Figure 2: The Distribution of Fairness Ideals Among MBA Students and in the U.S.



Notes: This figure shows estimated fractions of egalitarians, libertarians, meritocrats, and unclassified ("other") spectators across studies, based on a between-subject design using one redistributive choice per spectator. Egalitarians are spectators who equalize earnings when the winner is determined by performance. Libertarians are spectators who do not redistribute when earnings are randomly assigned. Meritocrats are identified as the difference between the share of spectators allocating more to the winner in the performance condition and the share allocating more to the winner in the luck condition (following Almås et al., 2020). Remaining spectators are classified as having "other" ideals. Vertical dashed lines represent 95 percent confidence intervals from robust standard errors clustered at the spectator level.

We estimate the distribution of fairness ideals in the Preuss et al. (2025) data using only spectators' first decision in the "lucky outcomes" condition, in which the winner is determined by a coin flip with some probability and by performance otherwise. We keep only decisions in which the coin flip determined the outcome with probability one (equivalent to our "luck" condition) or performance did so with probability one (equivalent to our "performance" condition). For Harrs and Sterba (2025), we use their U.S. Prolific sample, in which spectators make redistribution decisions in both luck and merit conditions, presented in randomized order. For the between-subject estimates shown here, we use only the condition shown first. The winner receives \$4 and the loser \$0; spectators can transfer \$0–\$4. We classify fairness ideals using the same definitions as in the other studies.

We also report a pooled estimate across all representative samples using a DerSimonian–Laird random-effects meta-analysis (DerSimonian and Laird, 1986), which accounts for within-study sampling error and between-study heterogeneity by weighting each estimate by $1/(\text{se}_i^2+\hat{\tau}^2)$, where $\hat{\tau}^2$ is the estimated between-study variance.

Table 1: Implemented Gini Coefficient Across Environments

	Dependent variable: Gini coefficient			
	All redistributive decisions		First decision only	
	(1)	(2)	(3)	(4)
Performance condition	0.240***	0.242***	0.191***	0.184***
	(0.029)	(0.035)	(0.052)	(0.054)
Efficiency condition	0.188***	0.187***	0.206***	0.211***
	(0.024)	(0.029)	(0.057)	(0.060)
Constant (Mean Gini Luck)	0.420***	0.420***	0.432***	0.425***
	(0.024)	(0.018)	(0.040)	(0.071)
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	793	793	271	271
N (individuals)	271	271	271	271

Notes: This table displays estimates of β_0, β_1, and β_2 from equation (1). The omitted treatment is the luck condition. The outcome is the Gini coefficient of the final earnings allocation in worker pair p implemented by spectator i, Gini_{ip}. The Gini coefficient takes a value between zero and one, with zero denoting perfect equality (both workers have the same post-redistribution earnings) and one denoting perfect inequality (one worker holds all earnings). We calculate the Gini coefficient as:

$$\text{Gini}_{ip} = \frac{|\text{Income Worker 1} - \text{Income Worker 2}|}{\text{Income Worker 1} + \text{Income Worker 2}}.$$

The specifications in columns 2 and 4 include additional controls. In column 2, we include spectator fixed effects. In column 4, we control for gender, age bins, parental financial situation while growing up, and a dummy for being born in the U.S. Heteroskedasticity-robust standard errors clustered at the spectator level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Implemented Gini Coefficient: Elites vs. Benchmark Studies

	Dependent variable: Gini coefficient					
		Representative U.S. samples				
	MBA students	Almås et al. (2020)	Cohn et al. (2023)	Preuss et al. (2025)	Harrs and Sterba (2025)	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. General population samples						
Performance	0.184***	0.195***	0.320***	0.396***	0.258***	0.279***
	(0.054)	(0.032)	(0.040)	(0.059)	(0.017)	(0.033)
Efficiency	0.211***	0.011				0.011
	(0.060)	(0.035)				(0.035)
Constant	0.425***	0.363***	0.156***	0.463***	0.209***	0.292***
	(0.071)	(0.024)	(0.029)	(0.050)	(0.012)	(0.055)
N	271	1,000	614	197	1,476	3,287
Panel B. Higher-income subsamples						
Performance	0.184***	0.271***	0.253***	0.467***	0.244***	0.259***
	(0.054)	(0.065)	(0.045)	(0.124)	(0.033)	(0.025)
Efficiency	0.211***	0.174**				0.174**
	(0.060)	(0.077)				(0.077)
Constant	0.425***	0.316***	0.271***	0.456***	0.268***	0.281***
	(0.071)	(0.048)	(0.033)	(0.118)	(0.024)	(0.019)
N	271	213	316	49	398	976

Notes: This table displays estimates of equation (1) across studies. The dependent variable is the Gini coefficient of the final earnings allocation. The omitted treatment is the luck condition, so the constant estimates the average Gini when earnings are randomly assigned. "Performance condition" and "Efficiency condition" display the treatment effects of switching from the luck condition to the performance and efficiency conditions, respectively. Column 1 reports estimates from our MBA sample using spectators' first decisions and controlling for gender, age bins, parental financial situation while growing up, and a dummy for being born in the U.S. Columns 2–5 report estimates from representative U.S. samples in the indicated studies, without additional controls. Column 6 reports pooled estimates from a DerSimonian–Laird random-effects meta-analysis across the four benchmark studies (DerSimonian and Laird, 1986). The efficiency-cost condition is available in our experiment and in Almås et al. (2020), but not in the other studies; the pooled efficiency estimate therefore reflects a single study. For Preuss et al. (2025), we compute the Gini directly from each spectator's allocation between the two workers, using the first decision per respondent. For Cohn et al. (2023), we construct the Gini coefficient as $1 - 2 \times$ (share redistributed); observations are weighted to be representative of the U.S. population. For Harrs and Sterba (2025), we construct the Gini as $|1 - \text{transfer}/2|$, where the transfer ranges from \$0 to \$4; we use each respondent's first decision only. Panel A uses the full sample from each study. Panel B restricts to higher-income subsamples: individuals with household income $\geq$\$100,000 in Almås et al. (2020), the SCE, and Harrs and Sterba (2025), and the top 5% by income or wealth in Cohn et al. (2023). The MBA column is identical across panels. Heteroskedasticity-robust standard errors in parentheses; pooled standard errors account for between-study heterogeneity. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Distribution of Fairness Ideals: Elites vs. Average Citizens

	Percentage of...			
	Libertarians (1)	Meritocrats (2)	Egalitarians (3)	Other ideals (4)
Panel A. MBA students				
MBA students (spectators' first choices)	23.16	22.65	9.89	44.30
	(4.35)	(5.99)	(3.15)	(8.05)
MBA students (repeated-measures design)	15.33	28.74	6.51	49.43
	(1.58)	(1.98)	(1.08)	(2.76)
Panel B. Average American				
SCE – Average American	39.77	36.67	8.57	14.99
	(5.25)	(6.31)	(2.75)	(8.66)
Almås et al. (2020) – Average American	29.43	37.54	15.32	17.72
	(2.50)	(3.43)	(1.98)	(4.69)
Cohn et al. (2023) – Average American	12.80	60.33	17.45	9.42
	(1.89)	(3.22)	(2.19)	(4.33)
Harrs and Sterba (2025) – Average American	8.68	45.24	14.61	31.46
	(1.04)	(2.30)	(1.30)	(2.84)
Panel C. Higher-income samples				
SCE – USA income over 100k	44.44	40.32	3.23	12.01
	(12.05)	(13.27)	(3.23)	(18.22)
Almås et al. (2020) – USA income over 100k	22.67	41.92	14.08	21.33
	(4.87)	(7.11)	(4.16)	(9.57)
Cohn et al. (2023) – USA top 5 percent	25.15	59.14	8.97	6.75
	(3.33)	(4.37)	(2.38)	(5.99)
Harrs and Sterba (2025) – USA income over 100k	8.96	41.44	7.61	41.99
	(2.02)	(4.42)	(1.89)	(5.21)

Notes: This table shows estimates of the fraction of egalitarians, libertarians, meritocrats, and unclassified ("other") spectators in various studies. Estimates in Panels B and C are based on a between-subject design that uses one redistributive choice per spectator. See Appendix E for the definition and estimation of each fairness ideal.

We estimate the distribution of fairness ideals in the Survey of Consumer Expectations (SCE), collected by Preuss et al. (2025), using only data from the first decision of spectators in the "lucky outcomes" condition, in which the winner of a worker pair is determined by a coin flip with some probability and by performance with the remaining probability. We keep only decisions in which the coin flip determined the outcome with probability one (equivalent to our "luck" condition) or performance determined the outcome with probability one (equivalent to our "performance" condition).

For Harrs and Sterba (2025), we use their U.S. Prolific sample, in which spectators redistribute earnings in both a luck and a merit condition. Each respondent sees both conditions in randomized order; we use only the condition shown first. The winner receives \$4 and the loser \$0; spectators can transfer \$0–\$4 in \$0.10 increments. Income is reported in seven categories; we define "high-income" as categories 6 (\$100–150k) and 7 (\$150k+).

Panel C shows estimates for higher-income samples. In Almås et al. (2020), the SCE data, and Harrs and Sterba (2025), these are individuals with annual household income above \$100,000. In Cohn et al. (2023), these are households in the top 5 percent by income or wealth (i.e., annual household income above \$250,000 or gross liquid assets of at least \$1 million).

References

Acemoglu, D., He, A., and le Maire, D. (2025). Eclipse of rent-sharing: The effects of managers' business education on wages and the labor share in the US and Denmark. *American Economic Review (conditionally accepted)*.

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

Almås, I., Cappelen, A. W., Lind, J. T., Sørensen, E. Ø., and Tungodden, B. (2011). Measuring unfair (in)equality. *Journal of Public Economics*, 95(7-8):488–499.

Almås, I., Cappelen, A. W., Sorensen, E. O., and Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science*, 328(5982):1176–1178.

Almås, I., Cappelen, A. W., Sørensen, E. Ø., and Tungodden, B. (2025). Fairness across the world. Working Paper 06/2025, NHH Norwegian School of Economics, Department of Economics.

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. (2025). Shallow meritocracy. *Review of Economic Studies*, 92(2):772–807.

Bartels, L. M. (2016). *Unequal Democracy: The Political Economy of the New Gilded Age*. Princeton University Press, Princeton, NJ, 2 edition.

Bauman, Y. and Rose, E. (2011). Selection or indoctrination: Why do economics students donate less than the rest? *Journal of Economic Behavior & Organization*, 79(3):318–327.

Beard, C. A. (2012). *An Economic Interpretation of the Constitution of the United States*. Simon and Schuster, New York.

Bellavance, F., Dionne, G., and Lebeau, M. (2009). The value of a statistical life: A meta-analysis with a mixed effects regression model. *Journal of Health Economics*, 28(2):444–464.

Bénabou, R. and Tirole, J. (2006). Belief in a just world and redistributive politics. *The Quarterly Journal of Economics*, 121(2):699–746.

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

Broockman, D. E., Ferenstein, G., and Malhotra, N. (2019). Predispositions and the political behavior of American economic elites: Evidence from technology entrepreneurs. *American Journal of Political Science*, 63(1):212–233.

Bühlmann, F., Christesen, C. A., Cousin, B., Denord, F., Ellersgaard, C. H., Lagneau-Ymonet, P., Larsen, A. G., Savage, M., Thine, S., Young, K., et al. (2025). Varieties of economic elites? preliminary results from the World Elite Database (WED). *The British Journal of Sociology*, 76(3):663–673.

Burch, P. H. (1980). *Elites in American History*. Holmes and Meier, New York.

Caldwell, S. and Harmon, N. (2019). Outside options, bargaining and wages: Evidence from coworker networks. *AEA Papers and Proceedings*, 109:203–207.

Capozza, F. and Srinivasan, K. (2024). Who should get money? estimating welfare weights in the US. URPP Equality of Opportunity Discussion Paper Series 50, University of Zurich.

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

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

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

Cappelen, A. W., Sørensen, E. Ø., and Tungodden, B. (2010). Responsibility for what? fairness and individual responsibility. *European Economic Review*, 54(3):429–441.

Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):S13–S70.

Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2022). *World Inequality Report 2022*. Harvard University Press.

Chetty, R., Deming, D. J., and Friedman, J. N. (2026). Diversifying society's leaders? the determinants and causal effects of admission to highly selective private colleges. *Quarterly Journal of Economics*, 141(1):51–145.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., and Yagan, D. (2020). Income segregation and intergenerational mobility across colleges in the United States. *The Quarterly Journal of Economics*, 135(3):1567–1633.

Cohn, A., Jessen, L. J., Klašnja, M., and Smeets, P. (2023). Wealthy Americans and redistribution: The role of fairness preferences. *Journal of Public Economics*, 225:104977.

Cornell University (2026). Notable alumni. https://www.johnson.cornell.edu/about/75th/alumni/. Accessed: May 30, 2026.

Cullen, Z. B., Li, S., and Perez-Truglia, R. (2025). What's my employee worth? the effects of salary benchmarking. *Review of Economic Studies*.

Dana, J., Weber, R. A., and Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1):67–80.

Derenoncourt, E. and Weil, D. (2025). Voluntary minimum wages: The local labor market effects of national retailer policies. *Quarterly Journal of Economics*, 140(3):1901–1958.

DerSimonian, R. and Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3):177–188.

Domhoff, G. W. (2013). *Who Rules America? The Triumph of the Corporate Rich*. McGraw-Hill, New York, 7 edition.

Domhoff, G. W. (2017). *The Power Elite and the State: How Policy is Made in America*. Routledge.

Dong, L., Huang, L., and Lien, J. W. (2025). 'they never had a chance': Unequal opportunities and fair redistributions. *The Economic Journal*, 135(667):914–942.

Downs, A. (1957). *An Economic Theory of Democracy*. Harper and Row, New York.

Dube, A., Manning, A., and Naidu, S. (2025). Monopsony and employer misoptimization explain why wages bunch at round numbers. *American Economic Review*, 115(8):2689–2721.

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

Elsässer, L., Hense, S., and Schäfer, A. (2021). Not just money: Unequal responsiveness in egalitarian democracies. *Journal of European Public Policy*, 28(12):1890–1908.

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

Exley, C. L. (2020). Using charity performance metrics as an excuse not to give. *Management Science*, 66(2):553–563.

Ferguson, T. (1995). *Golden Rule: The Investment Theory of Party Competition and the Logic of Money-Driven Political Systems*. University of Chicago Press.

Fisman, R., Jakiela, P., Kariv, S., and Markovits, D. (2015). The distributional preferences of an elite. *Science*, 349(6254):aab0096.

Fisman, R., Jakiela, P., Kariv, S., and Vannutelli, S. (2023). The distributional preferences of Americans, 2013–2016. *Experimental Economics*, 26(4):727–748.

Frank, D. H., Wertenbroch, K., and Maddux, W. W. (2015). Performance pay or redistribution? cultural differences in just-world beliefs and preferences for wage inequality. *Organizational Behavior and Human Decision Processes*, 130:160–170.

Gilens, M. (2012). *Affluence and Influence: Economic Inequality and Political Power in America*. Princeton University Press.

Gilens, M. and Page, B. I. (2014). Testing theories of American politics: Elites, interest groups, and average citizens. *Perspectives on Politics*, 12(3):564–581.

Hacker, J. S. and Pierson, P. (2010). *Winner-Take-All Politics: How Washington Made the Rich Richer—and Turned Its Back on the Middle Class*. Simon & Schuster, New York, NY.

Hall, R. E. and Krueger, A. B. (2012). Evidence on the incidence of wage posting, wage bargaining, and on-the-job search. *American Economic Journal: Macroeconomics*, 4(4):56–67.

Harrs, S. and Sterba, M.-B. (2025). Fairness and support for redistribution: The role of preferences and beliefs. Working Paper 47, Cluster of Excellence "The Politics of Inequality", University of Konstanz.

Harrs, S. and Sterba, M.-B. (2026). How stable is meritocratic ideology in times of crises? Work in progress.

Hazell, J., Patterson, C., Sarsons, H., and Taska, B. (2026). National wage setting. *American Economic Review (forthcoming)*.

He, A. X. and le Maire, D. (2025). Managing inequality: Manager-specific wage premiums and selection in the managerial labor market. *Review of Economics and Statistics (forthcoming)*.

Hertel-Fernandez, A. (2019). *State Capture: How Conservative Activists, Big Businesses, and Wealthy Donors Reshaped the American States—and the Nation*. Oxford University Press, New York.

Hjort, J., Li, X., and Sarsons, H. (2026). Across-country wage compression in multinationals. *American Economic Review*, 116(1):52–87.

Lachowska, M., Mas, A., Saggio, R., and Woodbury, S. A. (2022). Wage posting or wage bargaining? a test using dual jobholders. *Journal of Labor Economics*, 40(S1):S469–S493.

Levitt, S. D. and List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, 21(2):153–174.

Maréchal, M. A., Cohn, A., Yusof, J., and Fisman, R. (2025). Whose preferences matter for redistribution? cross-country evidence. *Journal of Political Economy Microeconomics*, 3(1):1–24.

Mathisen, R. B. (2023). Affluence and influence in a social democracy. *American Political Science Review*, 117(2):751–758.

Meager, R. (2019). Understanding the average impact of microcredit expansions: A Bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics*, 11(1):57–91.

Meager, R. (2022). Aggregating distributional treatment effects: A Bayesian hierarchical analysis of the microcredit literature. *American Economic Review*, 112(6):1818–1847.

Mills, C. W. (1956). *The Power Elite*. Oxford University Press, New York.

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

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

Nax, H. H., Depoorter, B., Grech, P. D., Halten, J., Hoeppner, S., Newton, J., Pradelski, B. S. R., Soos, A., Wehrli, S., Ikica, B., et al. (2021). Are elites really less fair-minded? Legal Studies Research Paper Series 3766728, UC College of the Law San Francisco.

Norton, M. I. and Ariely, D. (2011). Building a better America—one wealth quintile at a time. *Perspectives on Psychological Science*, 6(1):9–12.

Norton, M. I. and Ariely, D. (2013). American's desire for less wealth inequality does not depend on how you ask them. *Judgment and Decision Making*, 8(3):393–394.

Page, B. I. and Gilens, M. (2017). *Democracy in America?: What Has Gone Wrong and What We Can Do About It*. University of Chicago Press, Chicago, IL.

Page, B. I., Seawright, J., and Lacombe, M. J. (2018). *Billionaires and Stealth Politics*. University of Chicago Press, Chicago.

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

Reyes, G. (2025). Coarse wage-setting and behavioral firms. *Review of Economics and Statistics (accepted)*.

Rigby, E. and Wright, G. C. (2013). Political parties and representation of the poor in the American states. *American Journal of Political Science*, 57(3):552–565.

Saez, E. and Zucman, G. (2022). Top wealth in America: A reexamination. Working Paper 30396, National Bureau of Economic Research.

Schakel, W. (2021). Unequal policy responsiveness in the Netherlands. *Socio-Economic Review*, 19(1):37–57.

Sundemo, M. and Löfgren, Å. (2025). Do business and economics studies erode prosocial values? *Southern Economic Journal*, 92(2):504–526.

Vivalt, E. (2020). How much can we generalize from impact evaluations? *Journal of the European Economic Association*, 18(6):3045–3089.

Wai, J., Anderson, S. M., Perina, K., Worrell, F. C., and Chabris, C. F. (2024). The most successful and influential Americans come from a surprisingly narrow range of 'elite' educational backgrounds. *Humanities and Social Sciences Communications*, 11(1):1129.

Winters, J. A. (2011). *Oligarchy*. Cambridge University Press, New York.

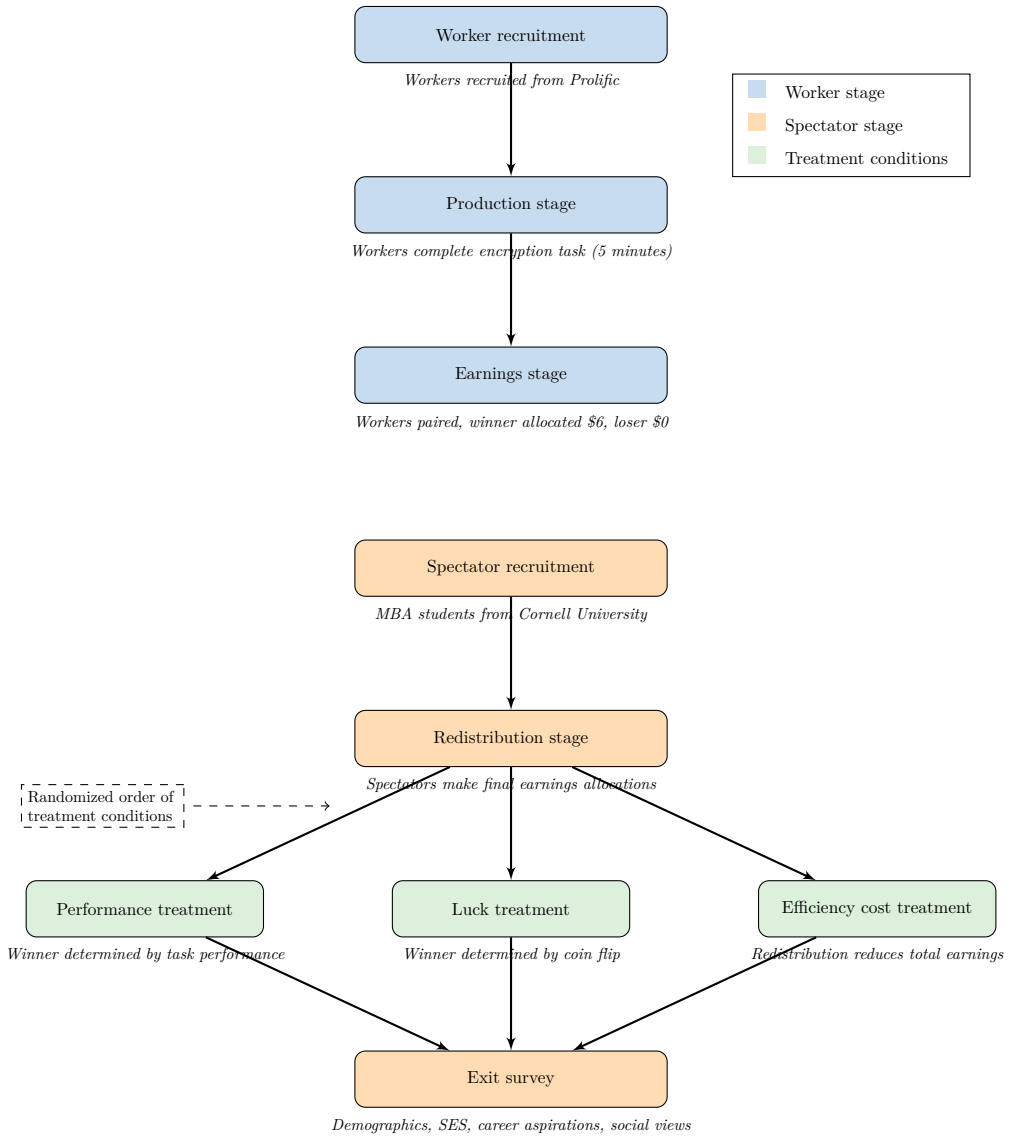
Winters, J. A. and Page, B. I. (2009). Oligarchy in the United States? *Perspectives on Politics*, 7(4):731–751.

Yusof, J. and Sartor, S. (2025). Market luck: Skill-biased inequality and redistributive preferences. Working Paper 475, University of Zurich, Department of Economics.

Appendix

A Appendix Figures and Tables

Figure A1: Flow of the Experiment



Note: This figure illustrates the three-stage experimental design. In the production stage, workers complete an encryption task. In the earnings stage, workers are paired and initial earnings are assigned based on either performance or luck. In the redistribution stage, MBA student spectators determine final earnings allocations across three treatment conditions presented in randomized order.

Figure A2: Example of the Worker Encryption Task

Q	**X**	**D**	**A**	**C**	**V**	**U**	**R**	**P**	**W**	**L**	**Y**	**G**
754	579	860	708	344	725	950	314	532	595	654	838	327
Z	**F**	**M**	**N**	**T**	**B**	**K**	**O**	**H**	**S**	**E**	**I**	**J**
190	776	627	980	830	803	603	673	536	490	545	445	925

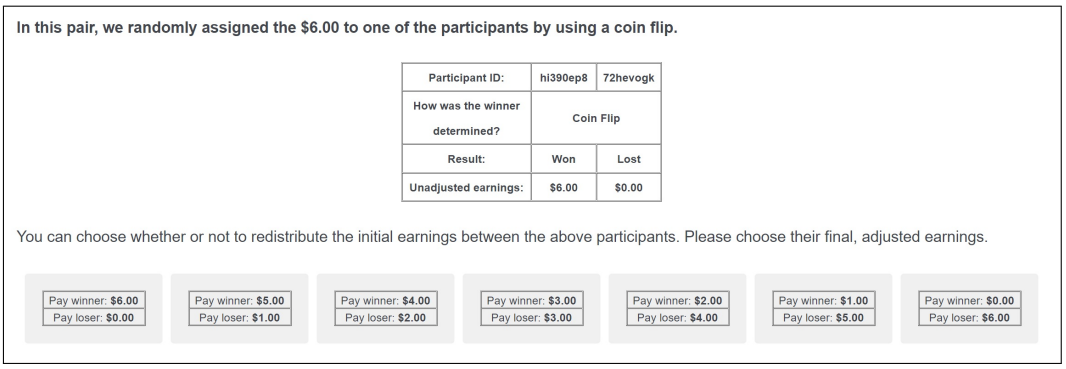
Please translate the following word into code:

RPZ: ☐ ☐ ☐

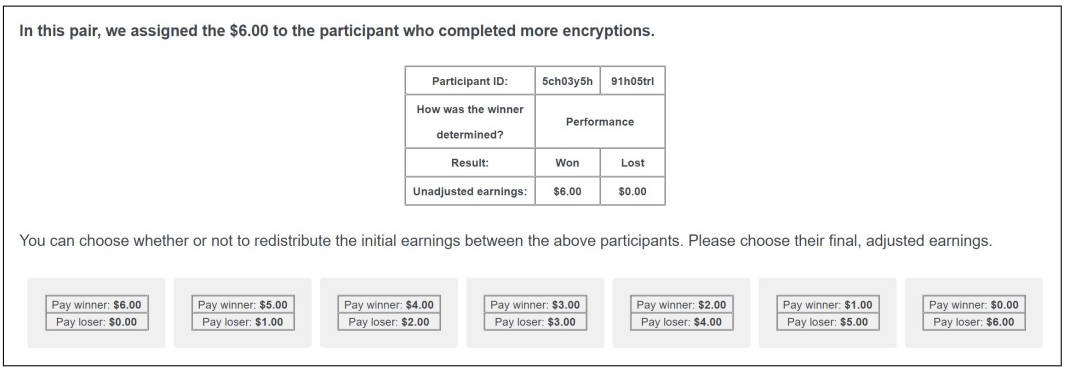
Notes: This figure shows an example of an encryption completed by workers. For each three-letter "word," workers receive a codebook that maps letters to three-digit numbers. After encrypting one word, a new word appears along with a new codebook. The words, codes, and the sequence of letters in the codebook are randomized for each new word. Feedback on the correctness of encryptions is not provided. Workers have five minutes to complete as many encryptions as possible.

Figure A3: Screenshots of Redistributive Decision Screens

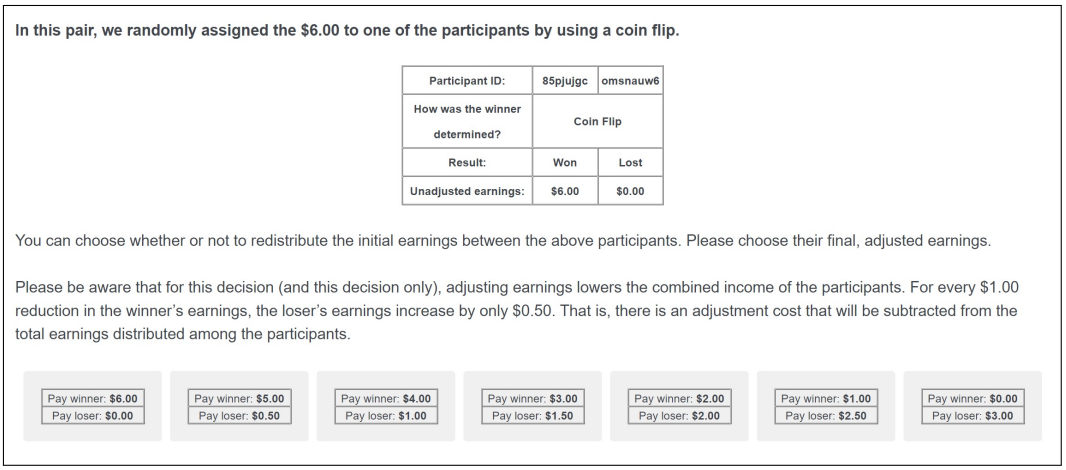
Panel A. Luck condition



Panel B. Performance condition

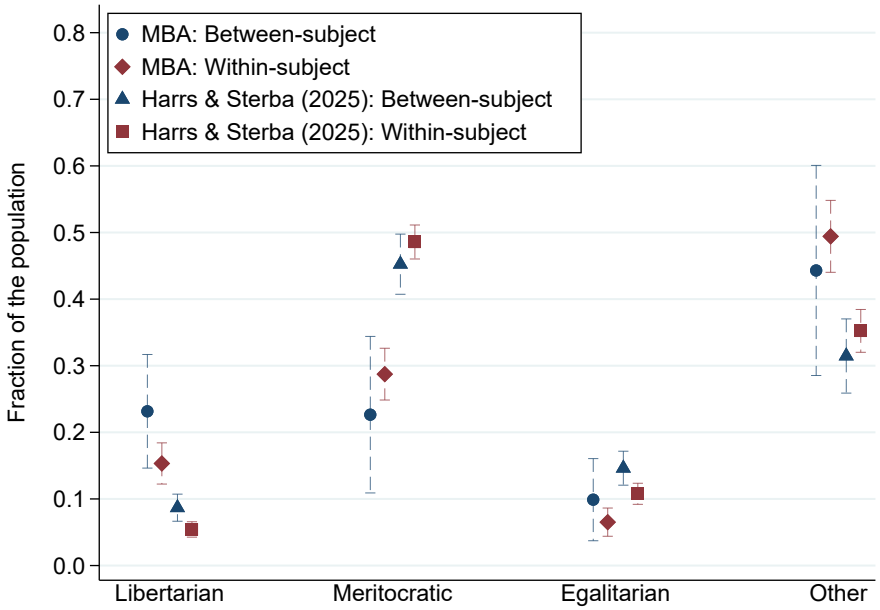


Panel C. Efficiency-cost condition



Notes: This figure shows screenshots of the redistribution screen shown to spectators for each worker pair condition.

Figure A4: Fairness Ideals: Between- vs. Within-Subject Identification



Notes: This figure shows estimates of the fraction of libertarians, meritocrats, and egalitarians using two different research designs for MBA students and the Harrs and Sterba (2025) Prolific sample. See Section E for the definition and estimation of each fairness ideal. For MBA students, between-subject estimates (circles) use one redistributive choice per spectator and within-subject estimates (diamonds) use two choices. For Harrs and Sterba (2025), between-subject estimates (triangles) use only the condition shown first, while within-subject estimates (squares) use both the luck and merit conditions per respondent. Vertical dashed lines represent 95 percent confidence intervals calculated using robust standard errors clustered at the spectator level.

Table A1: Summary Statistics of the Sample by First Treatment Shown

		First condition shown		
	All	Luck	Performance	Efficiency
	(1)	(2)	(3)	(4)
Panel A. Demographic characteristics				
Age ≤23	0.031	0.043	0.034	0.013
Age 24–27	0.380	0.419	0.315	0.408
Age 28–31	0.453	0.387	0.483	0.500
Age ≥32	0.136	0.151	0.169	0.079
Male	0.554	0.521	0.589	0.553
Cohort 2023	0.509	0.537	0.451	0.541
Cohort 2024	0.491	0.463	0.549	0.459
Born in the USA	0.562	0.606	0.545	0.527
Panel B. Financial situation while growing up				
Always had enough money for necessities and luxuries	0.198	0.191	0.169	0.240
Always had enough for necessities and occasional luxuries	0.422	0.447	0.438	0.373
Usually had enough for necessities, rarely for luxuries	0.279	0.277	0.281	0.280
Sometimes did not have enough money for necessities	0.081	0.074	0.090	0.080
Frequently did not have enough money for necessities	0.019	0.011	0.022	0.027
Panel C. Employment preferences				
Planning on getting a job in the USA	0.941	0.978	0.899	0.946
Working long hours is sometimes necessary	0.811	0.806	0.851	0.770
Work-life balance is more important than a high salary	0.524	0.505	0.448	0.635
Wants career in private sector	0.839	0.806	0.851	0.865
Wants career in a non-profit	0.173	0.204	0.138	0.176
Wants to become an entrepreneur	0.484	0.430	0.483	0.554
Aspire to be a manager	0.913	0.903	0.920	0.919
Panel D. Voting behavior and social views				
Voted in the last elections	0.693	0.714	0.701	0.658
Success is luck and connections	0.362	0.355	0.414	0.311
Hard work brings a better life	0.815	0.828	0.816	0.797
Passed attention check	0.933	0.968	0.908	0.919
Minutes spent in survey	4.553	4.560	4.640	4.454
Completed all three redistributive decisions	0.959	0.989	0.989	0.894
Number of individuals	271	95	91	85
Number of redistributive decisions	793	284	271	238
F-statistic	–	1.042	1.049	1.034
p-value F-statistic	–	0.410	0.404	0.417

Notes: This table shows summary statistics of MBA students in our sample. All variables are based on data self-reported by MBA students in the exit survey of our study. Employment preferences and social views are based on MBA students' agreement with several statements in a five-point Likert scale grid. For each statement, we define an indicator variable that equals one if the student selects "strongly agree" or "agree," and zero otherwise. Column 1 reports overall summary statistics. Columns 2–4 split the sample by the first condition shown: luck (earnings randomly assigned), performance (earnings determined by performance on the encryption task), or efficiency cost (earnings randomly assigned with costly redistribution).

Table A2: The Impact of the Order in Which Spectators Saw Each Condition

	Dependent variable: Gini coefficient			
	Separately by condition			
	Luck	Performance	Efficiency	Pooled
	(1)	(2)	(3)	(4)
Luck × Screen 2	−0.014			−0.014
	(0.059)			(0.059)
Luck × Screen 3	−0.022			−0.022
	(0.059)			(0.060)
Performance × Screen 2		0.028		0.028
		(0.050)		(0.050)
Performance × Screen 3		0.087*		0.087*
		(0.052)		(0.052)
Efficiency × Screen 2			0.017	0.017
			(0.057)	(0.057)
Efficiency × Screen 3			−0.105*	−0.105*
			(0.058)	(0.058)
Constant	0.432***	0.623***	0.638***	0.565***
	(0.040)	(0.034)	(0.041)	(0.022)
N (individuals)	261	263	269	271

Notes: This table displays estimates of order effects on the Gini coefficient. In columns 1–3, we regress the Gini on dummies for seeing a given condition on the second and third screen. The constant represents the average Gini coefficient for spectators who saw a given condition on the first screen. In column 4, we interact all conditions with the order dummies. Heteroskedasticity-robust standard errors clustered at the spectator level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Time Spent on Redistribution Screens by Spectators With Unclassified Ideals

	Dependent variable: Time spent (seconds)			
		By redistribution condition		
	All screens	Performance	Luck	Efficiency Cost
	(1)	(2)	(3)	(4)
Other ideal	0.382	−1.280	2.311	−2.770
	(4.467)	(1.983)	(1.674)	(2.974)
Constant	73.822***	21.266***	19.336***	35.338***
	(3.489)	(1.467)	(1.225)	(2.104)
N (individuals)	260	263	261	269

Notes: This table presents regression estimates of the relationship between having an unclassified fairness ideal and time spent on redistribution decision screens. The outcome is the total time spent across all redistribution screens. Standard errors clustered at the spectator level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Joint Distribution of Redistribution in Luck and Performance Conditions

Luck Condition	Earnings Redistributed in Performance Condition: \$0	\$1	\$2	\$3	\$4	\$5	\$6
\$0	0.153	0.031	0.027	0.011	0.000	0.000	0.008
\$1	0.034	0.057	0.023	0.008	0.004	0.000	0.000
\$2	0.077	0.084	0.080	0.008	0.000	0.000	0.000
\$3	0.138	0.069	0.080	0.065	0.000	0.000	0.004
\$4	0.000	0.004	0.015	0.004	0.000	0.000	0.000
\$5	0.000	0.000	0.000	0.000	0.000	0.000	0.000
\$6	0.000	0.000	0.000	0.004	0.004	0.004	0.004

Libertarians
Meritocrats
Egalitarians

Notes: This table shows the joint distribution of redistribution choices across luck and performance conditions, excluding the efficiency cost treatment. Each cell represents the proportion of spectators who chose to redistribute \$$x$ dollars in the luck condition (rows) and \$$y$ dollars in the performance condition (columns). Choices were elicited in \$0.50 increments and aggregated to \$1 bins for this table. For example, 15.3 percent of spectators chose to redistribute \$0 in both conditions, while 6.5 percent chose to redistribute \$3 (i.e., equalize earnings) in both conditions.

The green shaded cell indicates spectators classified as libertarians (no redistribution in either condition), the blue shaded cells indicate spectators classified as meritocratic (earnings equalization in the luck condition and more redistribution in luck than performance), and the orange shaded cell indicates spectators classified as egalitarians (equal redistribution in both conditions). Spectators whose redistribution choices fall outside the shaded areas are classified as having "other" fairness ideals.

Table A5: Correlates of Fairness Ideals

	Dependent Variable:				
	Libertarian	Moderate	Meritocrat	Egalitarian	Other
	(1)	(2)	(3)	(4)	(5)
Male	0.146***	−0.082	−0.102	−0.038	0.075
	(0.042)	(0.050)	(0.057)	(0.032)	(0.057)
Aged 28 or over	0.016	0.031	0.029	−0.016	−0.060
	(0.045)	(0.050)	(0.057)	(0.032)	(0.058)
Born USA	0.064	−0.027	−0.026	0.007	−0.019
	(0.044)	(0.051)	(0.058)	(0.031)	(0.058)
Stay USA	0.080	0.002	−0.260	0.001	0.178
	(0.069)	(0.107)	(0.133)	(0.067)	(0.093)
Always enough money	0.030	0.055	−0.015	−0.009	−0.062
	(0.045)	(0.050)	(0.058)	(0.032)	(0.060)
Voted last elections	0.007	−0.044	0.159**	0.041	−0.163*
	(0.048)	(0.057)	(0.056)	(0.030)	(0.065)
Success luck network	0.015	−0.109	−0.048	0.065	0.076
	(0.048)	(0.047)	(0.059)	(0.036)	(0.061)
Hard work better life	0.032	0.075	−0.034	0.030	−0.103
	(0.055)	(0.056)	(0.075)	(0.035)	(0.078)
Long hours necessary	−0.042	0.130**	−0.103	0.031	−0.016
	(0.062)	(0.049)	(0.077)	(0.034)	(0.074)
Work−life balance	−0.038	0.093	−0.043	0.033	−0.044
	(0.046)	(0.049)	(0.057)	(0.031)	(0.058)
Career in priv. sector	0.067	−0.095	0.086	−0.037	−0.021
	(0.053)	(0.074)	(0.072)	(0.049)	(0.079)
Career in a non−profit	−0.021	0.019	−0.023	0.056	−0.032
	(0.058)	(0.067)	(0.074)	(0.051)	(0.075)
Wants entrepreneur	0.049	−0.020	−0.029	0.044	−0.044
	(0.046)	(0.049)	(0.057)	(0.032)	(0.058)
Wants manager	0.019	0.008	−0.079	−0.026	0.079
	(0.077)	(0.087)	(0.107)	(0.064)	(0.095)
Mean Dep. Var.	0.181	0.188	0.277	0.063	0.292
N (Spectators)	271	271	271	271	271

Notes: This table shows estimates of the correlates of having a given fairness ideal. The regression equation is:

$$Y_i = \alpha + \gamma X_i + \nu_i,$$

where Y_i is a fairness ideal and X_i is an individual-level characteristic.

Each column shows the result for a different dependent variable. In column 1, the outcome equals one if the student does not redistribute earnings in either the luck or the performance environments. In column 2, the outcome equals one if the student is classified as a moderate (as defined in Section 5.3). In column 3, the outcome equals one if the student equalizes earnings in the luck environment but gives strictly more to the winner in the performance environment. In column 4, the outcome equals one if the student divides earnings equally in the luck and performance environments. In column 5, the outcome equals one if the student has "other" fairness ideals but is not classified as a moderate.

Heteroskedasticity-robust standard errors clustered at the spectator level in parentheses. Thresholds * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ are Bonferroni corrected for five tests.

B Empirical Appendix

B.1 Worker Sample

This appendix provides details on the characteristics of the worker sample.

Appendix Table B1 provides summary statistics on workers. The sample is diverse in terms of age, gender, and racial background. The average worker is approximately 40 years old, with a standard deviation of 13.8 years. About 49 percent of workers are male, and 94 percent were born in the U.S. The majority identify as White (70 percent), followed by Black (12 percent), Asian (9 percent), and individuals of mixed or other racial backgrounds.

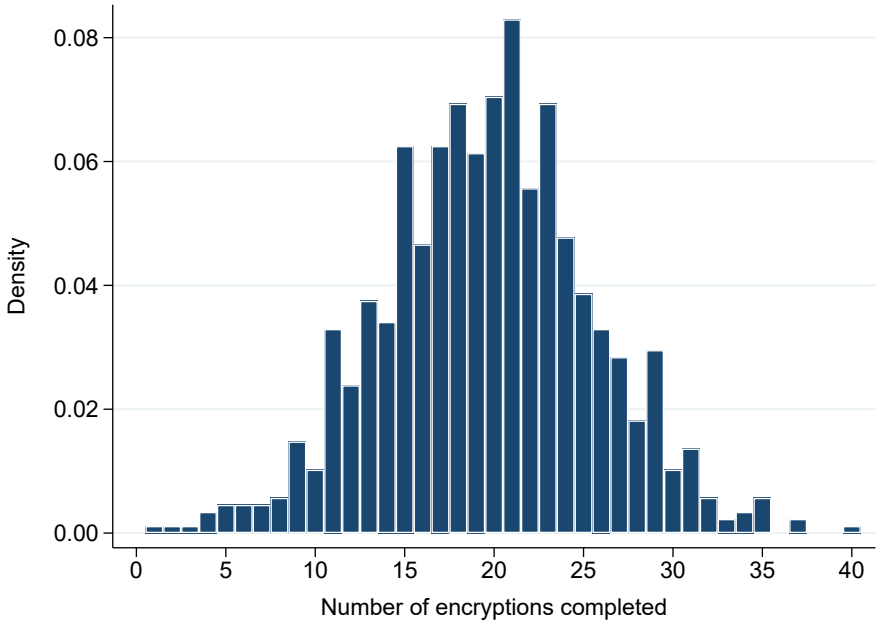
Appendix Figure B1 shows the distribution of encryption task performance. Workers completed an average of 19.67 encryptions, with a standard deviation of 5.91. The average number of encryption attempts was 20.92. The distribution is approximately normal, with most workers completing between 10 and 30 encryptions.

Table B1: Summary Statistics of the Worker Sample

	Mean (1)	SD (2)	N (3)
Panel A. Demographic characteristics			
Age	40.403	13.829	863
Male	0.487	0.500	871
White	0.697	0.460	855
Black	0.116	0.320	855
Asian	0.090	0.286	855
Mixed race	0.065	0.248	855
Other race	0.032	0.175	855
Born in the USA	0.945	0.229	869
Language is English	0.947	0.223	875
Panel B. Employment status			
Is a student	0.136	0.343	589
Works full-time	0.503	0.501	475
Works part-time	0.162	0.369	475
Unemployed	0.112	0.315	475
Panel C. Performance on the encryption task			
Tasks attempted	20.919	5.821	880
Tasks completed	19.670	5.910	880

Notes: This table shows summary statistics of our worker sample. Variables in Panels A and B are based on demographic and employment data provided by Prolific rather than collected through our survey. Sample sizes vary across variables as this information is only available for workers who have provided it to Prolific. Variables in Panel C are based on performance on the encryption task and are available for all workers.

Figure B1: Distribution of Tasks Completed in the Worker Task



Notes: This figure shows the distribution of the total number of correct three-letter encryptions completed by workers. The mean number of encryptions completed is 19.7 and the standard deviation is 5.9.

B.2 Descriptive Evidence

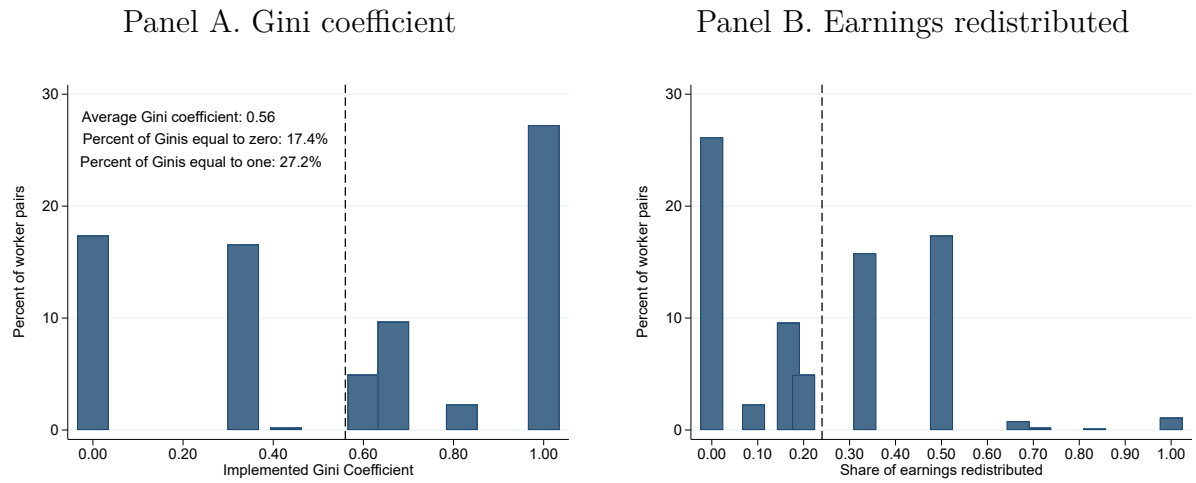
This appendix presents descriptive evidence on implemented inequality and redistribution choices across experimental conditions.

Appendix Figure B2 shows a histogram of the implemented Gini coefficients (Panel A) and share of earnings redistributed (Panel B). Implemented inequality varies substantially across spectators. On average across worker pairs, spectators implemented a Gini coefficient of 0.56 and redistributed \$1.30 of the winner's earnings to the loser. The two modal Gini coefficients are one (perfect inequality, 34.8 percent of worker pairs) and zero (perfect equality, 22.2 percent).

Appendix Figure B3 presents the average Gini coefficient (Panel A) and the share of earnings redistributed (Panel B) across different experimental conditions. Panel A shows that implemented inequality varies significantly by condition. In the luck treatment, where earnings are randomly assigned, the average Gini coefficient is 0.43. In contrast, inequality is substantially higher in the performance and efficiency cost conditions. The efficiency cost condition generates the highest inequality, with an average Gini coefficient of 0.64. Panel B

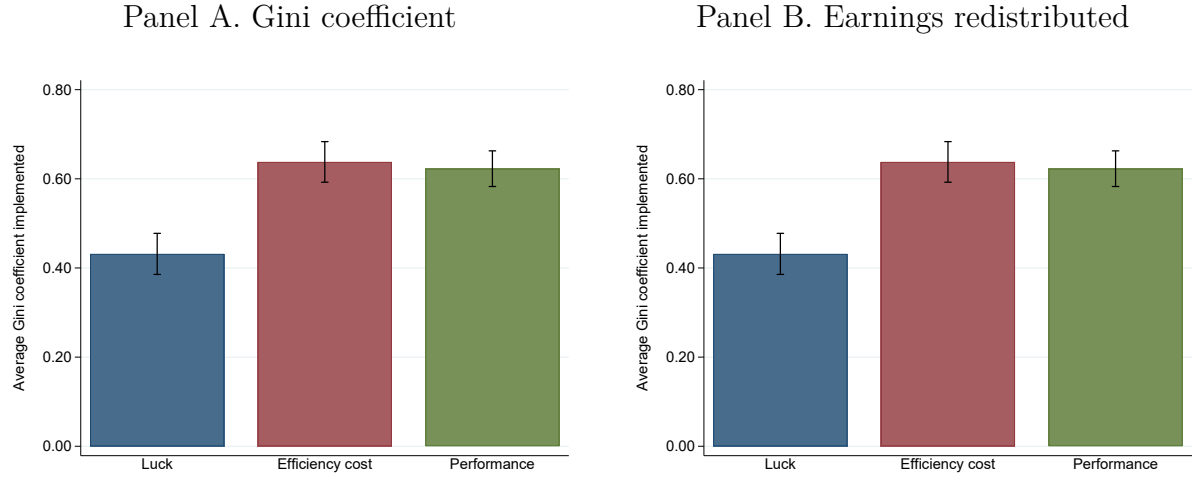
illustrates the fraction of earnings redistributed across conditions. Spectators redistribute the highest share of earnings in the luck condition, while redistribution decreases sharply in the performance and efficiency cost conditions.

Figure B2: Histogram of the Gini Coefficient and Earnings Redistributed



Notes: This figure shows a histogram of the implemented Gini coefficients (Panel A) and share of earnings redistributed (Panel B). To produce this figure, we use data on all spectator redistributive decisions.

Figure B3: Average Gini and Earnings Redistributed by Condition



Notes: This figure shows the average implemented Gini coefficients (Panel A) and share of earnings redistributed (Panel B) across each condition. To produce this figure, we only use the first redistributive decision of each spectator.

B.3 Robustness Checks

This appendix presents several robustness checks of our main results. We examine the sensitivity of our findings to different sample restrictions and alternative outcome measures. All robustness checks follow the specification in equation (1) and present results using both within-subject and between-subject variation.

First, we verify that our results are not driven by inattentive participants. Appendix Table B2 reproduces our main analysis excluding spectators who failed an embedded attention check in our survey. The attention check asked participants to select a specific response option within a matrix of survey questions. The estimates remain virtually unchanged after excluding inattentive respondents.

Second, we examine whether our results are affected by participants who rushed through the experiment. Appendix Table B3 presents estimates excluding spectators who completed the survey in less than 2.20 minutes, corresponding to the bottom five percent of the completion time distribution. The magnitude and statistical significance of our main effects remain stable in this restricted sample.

Third, we assess the robustness of our findings to potentially inconsistent response patterns. Appendix Table B4 shows results after excluding spectators who allocated strictly more earnings to the loser than to the winner in at least one worker pair—a pattern that could indicate confusion about the task. Our main conclusions about the source-of-inequality effect and efficiency concerns remain unchanged in this restricted sample.

Fourth, we examine whether our results are driven by spectators whose redistributive choices do not conform to standard fairness ideals. Appendix Table B5 presents estimates excluding spectators classified as having "other" fairness ideals—those who do not fit the egalitarian, libertarian, or meritocratic classifications. The performance condition continues to increase the Gini coefficient by 0.296 points in the between-subject specification without controls, and the efficiency-cost condition increases it by 0.199 points (both $p < 0.01$).

Fifth, we examine whether our results are driven by spectators who may have misunderstood the efficiency cost treatment. Appendix Table B6 presents estimates excluding spectators who redistributed more earnings when there was an efficiency cost than when there was no efficiency cost—a pattern inconsistent with standard economic theory. This restriction excludes 46 spectators (17 percent of our sample). Despite this exclusion, our main findings remain robust. The performance condition continues to increase the Gini coefficient by 0.234 points in the within-subject specification without controls (column 1),

nearly identical to our baseline estimate of 0.240. The efficiency-cost condition increases the Gini by 0.243 points, larger than our baseline estimate of 0.188. The baseline inequality level increases slightly from 0.420 to 0.447.

Sixth, we examine whether our results differ systematically across the two MBA cohorts in our sample. Appendix Table B7 presents estimates from equation (1) estimated separately for MBA students from the 2023 and 2024 cohorts. The effect of the performance condition is stable across cohorts, ranging from 0.235 to 0.245 in the within-subject specifications without controls. Similarly, the efficiency-cost condition effects are consistent across cohorts, with estimates of 0.182 and 0.195 for the 2023 and 2024 cohorts, respectively. The baseline levels of inequality, captured by the constants, are also similar across cohorts (0.419 and 0.421).

Seventh, we account for potential order effects by including round fixed effects in our estimation. Appendix Table B8 presents estimates from our main specification with additional controls for the order in which spectators encountered each treatment condition. The results remain consistent with our baseline findings. When controlling for round fixed effects, the performance condition increases the Gini coefficient by 0.240 points in the within-subject specification without additional controls, nearly identical to our main estimate. Similarly, the efficiency cost condition increases the Gini coefficient by 0.188 points, matching our baseline result. This consistency indicates that our findings are not driven by the sequential nature of the experimental design or by potential learning effects across redistributive decisions.

Finally, we verify that our results are not sensitive to our choice of dependent variable. Appendix Table B9 presents estimates using the net-of-efficiency-cost share of earnings redistributed as the outcome variable instead of the Gini coefficient. This alternative specification directly measures redistribution behavior while accounting for efficiency costs. The results remain qualitatively similar, confirming robustness to alternative outcome measures.

Across all these robustness checks, the key patterns in our data—MBA students' greater tolerance for inequality, weaker sensitivity to the source of inequality than the general population, and stronger response to efficiency costs—remain stable and statistically significant.

Table B2: Robustness Check of Main Effects to Excluding Inattentive Spectators

	Dependent variable: Gini coefficient			
	All redistributive decisions		First decision only	
	(1)	(2)	(3)	(4)
Performance condition	0.242***	0.242***	0.186***	0.186***
	(0.030)	(0.036)	(0.054)	(0.057)
Efficiency condition	0.186***	0.186***	0.194***	0.202***
	(0.025)	(0.030)	(0.062)	(0.063)
Constant	0.415***	0.415***	0.430***	0.437***
	(0.026)	(0.019)	(0.040)	(0.073)
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	711	711	237	237
N (individuals)	237	237	237	237

Notes: This table is analogous to Table 1, but excludes students who failed the attention check. See notes to Table 1 for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B3: Robustness Check of Main Effects to Excluding Spectators Who Rushed Through the Experiment

	Dependent variable: Gini coefficient			
	All redistributive decisions		First decision only	
	(1)	(2)	(3)	(4)
Performance condition	0.249***	0.248***	0.206***	0.202***
	(0.029)	(0.036)	(0.053)	(0.055)
Efficiency condition	0.181***	0.180***	0.186***	0.194***
	(0.024)	(0.030)	(0.059)	(0.061)
Constant	0.421***	0.421***	0.430***	0.430***
	(0.025)	(0.019)	(0.040)	(0.073)
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	762	762	257	257
N (individuals)	257	257	257	257

Notes: This table is analogous to Table 1, but excludes students who completed the survey in less than 2.20 minutes (corresponding to the bottom five percent of the total completion time distribution). See notes to Table 1 for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B4: Robustness Check of Main Effects to Excluding Spectators Who Allocated Strictly More Earnings to the Loser

	Dependent variable: Gini coefficient			
	All redistributive decisions		First decision only	
	(1)	(2)	(3)	(4)
Performance condition	0.257***	0.259***	0.212***	0.214***
	(0.029)	(0.036)	(0.053)	(0.056)
Efficiency condition	0.186***	0.186***	0.221***	0.236***
	(0.025)	(0.030)	(0.059)	(0.062)
Constant	0.410***	0.409***	0.402***	0.423***
	(0.025)	(0.019)	(0.040)	(0.072)
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	744	744	254	254
N (individuals)	254	254	254	254

Notes: This table is analogous to Table 1, but excludes students who allocated strictly more earnings to the loser than to the winner in at least one worker pair. See notes to Table 1 for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B5: Robustness Check of Main Effects to Excluding Spectators with Non-Standard Fairness Ideals

	Dependent variable: Gini coefficient			
	All redistributive decisions		First decision only	
	(1)	(2)	(3)	(4)
Performance condition	0.419***	0.418***	0.296***	0.325***
	(0.037)	(0.046)	(0.089)	(0.092)
Efficiency condition	0.253***	0.248***	0.199**	0.174*
	(0.038)	(0.047)	(0.098)	(0.098)
Constant	0.303***	0.305***	0.378***	0.439***
	(0.040)	(0.026)	(0.073)	(0.117)
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	405	405	141	141
N (individuals)	141	141	141	141

Notes: This table is analogous to Table 1, but excludes spectators classified as having "other" fairness ideals (i.e., those who do not fit the standard egalitarian, libertarian, or meritocratic classifications). See notes to Table 1 for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B6: Robustness Check of Main Effects to Excluding Spectators Who Redistributed More Under Efficiency Costs

	Dependent variable: Gini coefficient			
	All redistributive decisions		First decision only	
	(1)	(2)	(3)	(4)
Performance condition	0.234***	0.237***	0.187***	0.187***
	(0.034)	(0.041)	(0.060)	(0.062)
Efficiency condition	0.243***	0.245***	0.250***	0.252***
	(0.027)	(0.033)	(0.061)	(0.066)
Constant	0.447***	0.445***	0.456***	0.437***
	(0.028)	(0.021)	(0.043)	(0.081)
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	655	655	225	225
N (individuals)	225	225	225	225

Notes: This table is analogous to Table 1, but excludes spectators who redistributed more earnings when there was an efficiency cost than when there was no efficiency cost. This exclusion addresses concerns that some spectators may have misunderstood the efficiency cost treatment. See notes to Table 1 for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B7: Robustness Check of Main Effects to Estimating the Regression Separately for Each Cohort

	Dependent variable: Gini coefficient			
	All redistributive decisions		First decision only	
	(1)	(2)	(3)	(4)
Panel A. 2023 Cohort				
Performance condition	0.235***	0.234***	0.180**	0.140*
	(0.042)	(0.052)	(0.075)	(0.079)
Efficiency condition	0.182***	0.184***	0.195**	0.187**
	(0.034)	(0.042)	(0.078)	(0.086)
Constant	0.419***	0.419***	0.438***	0.439***
	(0.035)	(0.027)	(0.053)	(0.114)
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	403	403	138	138
N (individuals)	138	138	138	138
Panel B. 2024 Cohort				
Performance condition	0.245***	0.251***	0.202***	0.218***
	(0.039)	(0.048)	(0.075)	(0.078)
Efficiency condition	0.195***	0.191***	0.220***	0.246***
	(0.034)	(0.041)	(0.084)	(0.088)
Constant	0.421***	0.421***	0.424***	0.422***
	(0.035)	(0.025)	(0.060)	(0.095)
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	390	390	133	133
N (individuals)	133	133	133	133

Notes: This table is analogous to Table 1, but estimated separately for MBA students of the 2023 cohort (Panel A) and 2024 cohort (Panel B). See notes to Table 1 for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B8: Robustness Check of Main Effects to Including Round Fixed Effects

	Dependent variable: Gini coefficient			
	All redistributive decisions		First decision only	
	(1)	(2)	(3)	(4)
Performance condition	0.240***	0.242***	0.191***	0.184***
	(0.029)	(0.029)	(0.052)	(0.054)
Efficiency condition	0.188***	0.187***	0.206***	0.211***
	(0.024)	(0.024)	(0.057)	(0.060)
Constant	0.420***	0.420***	0.432***	0.425***
	(0.024)	(0.015)	(0.040)	(0.071)
Round fixed effects?	Yes	Yes	Yes	Yes
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	793	784	271	271
N (individuals)	271	262	271	271

Notes: This table is analogous to Table 1, but the regression equation also controls for round fixed effects (only relevant for the columns 1 and 2). See notes to Table 1 for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B9: Robustness Check of Main Effects to Using the Share of Earnings Redistributed as the Outcome

	Dependent variable: Share of earnings redistributed			
	All redistributive decisions		First decision only	
	(1)	(2)	(3)	(4)
Performance condition	−0.123***	−0.124***	−0.116***	−0.121***
	(0.016)	(0.019)	(0.031)	(0.031)
Efficiency condition	−0.103***	−0.106***	−0.125***	−0.145***
	(0.014)	(0.017)	(0.033)	(0.033)
Constant	0.313***	0.314***	0.323***	0.305***
	(0.014)	(0.010)	(0.024)	(0.037)
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	793	793	271	271
N (individuals)	271	271	271	271

Notes: This table is analogous to Table 1, but the dependent variable is the net-of-efficiency-cost share of earnings redistributed. See notes to Table 1 for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C Comparison of Benchmark Studies

This appendix compares experimental designs, recruitment strategies, and sample characteristics across the five studies (Tables 2–3). Appendix Table C1 summarizes the key design features and spectator demographics.

C.1 Benchmark Studies

Almås et al. (2020). Almås et al. (2020) conducted a spectator experiment across the United States and Norway using nationally representative samples recruited through Norstat and its U.S. collaborator Research Now. Workers completed assignments on Amazon Mechanical Turk. Each spectator was randomized into one of three conditions—luck, merit, or efficiency cost—and made a single redistribution decision (between-subject design), choosing how to divide \$6 between a winning and a losing worker. In the luck and merit treatments, spectators chose from a discrete menu in \$1 increments; in the efficiency treatment, redistribution incurred a 50 percent cost. We use only the U.S. sample for our comparison.

Cohn et al. (2023). Cohn et al. (2023) recruited a large sample of U.S. respondents through YouGov between December 2016 and April 2017, oversampling households in the top 5 percent by income or wealth (annual household income above \$250,000 or gross liquid assets of at least \$1 million). Workers completed a real-effort task (checking and correcting digitized identification entries) on Amazon Mechanical Turk. Each spectator made a single redistribution decision (between-subject design), choosing how to divide the winner's \$6 earnings using a discrete menu in \$1 increments. The study includes three treatments—luck, mixed, and merit—but does not include an efficiency cost treatment.

Preuss et al. (2025) (SCE). Preuss et al. (2025) embedded a spectator experiment within the New York Federal Reserve's Survey of Consumer Expectations (SCE), yielding a nationally representative U.S. sample. Workers completed an encryption task on Amazon Mechanical Turk. Each spectator made 12 redistribution decisions (within-subject design), allocating \$5 between workers in \$0.50 increments. Decisions varied in the probability that earnings were determined by a coin flip versus worker performance; we use only the pure luck and pure performance conditions.

Harrs and Sterba (2025). Harrs and Sterba (2025) recruited a broadly representative U.S. sample of spectators through Prolific. Workers were recruited separately on Amazon Mechanical Turk. Each spectator made two redistribution decisions in randomized order—

one in a luck condition and one in a merit condition (within-subject design). The total earnings at stake are $4 and spectators can transfer amounts in $0.10 increments—lower stakes and finer increments than the other studies ($5–$6 and $0.50–$1, respectively).

C.2 Design Differences

All five studies share the impartial-spectator paradigm (Cappelen et al., 2013) in which spectators allocate income between other workers. The workers whose income spectators redistribute are U.S.-based online workers, recruited from Prolific in our study and from Amazon Mechanical Turk in the benchmark studies. Nonetheless, some differences between the cited studies exist.

First, our MBA spectators were recruited during mandatory class sessions, reducing the self-selection concerns that arise in online experiments (Levitt and List, 2007). The benchmark studies use nationally representative panels (Norstat/Research Now for Almås et al.; YouGov for Cohn et al.; the NY Fed SCE for Preuss et al.) or online platforms (Prolific for Harrs and Sterba).

Second, the number of decisions per spectator differs across studies. Our design and those of Preuss et al. (2025) and Harrs and Sterba (2025) are within-subject, meaning each spectator makes multiple redistribution decisions across conditions. In contrast, Almås et al. (2020) and Cohn et al. (2023) use between-subject designs in which each spectator makes a single decision. When comparing across studies in our main analysis, we use between-subject estimates—restricting to spectators' first decisions in within-subject designs—to ensure comparability.

Third, the transfer mechanisms and stakes differ. Almås et al. (2020) and Cohn et al. (2023) use discrete transfer menus in $1 increments from a $6 total. Preuss et al. (2025) uses $0.50 increments from a $5 total. Harrs and Sterba (2025) uses finer increments ($0.10) and lower stakes ($4). Despite these differences, all studies allow us to compute a standardized Gini coefficient as the outcome measure.

Fourth, the treatment set differs. Our experiment and Almås et al. (2020) include an efficiency cost treatment alongside the luck and performance conditions. Cohn et al. (2023) includes luck, mixed, and merit treatments but no efficiency cost. The remaining benchmark studies include only luck and performance (or merit) conditions.

Table C1: Summary of Study Designs and Spectator Demographics

	This study	Almås et al. (2020)	Cohn et al. (2023)	Preuss et al. (2025)	Harrs and Sterba (2025)
Panel A. Study design					
Year of data collection	2023–2024	2014	2016–2017	2021	2020–2022
Spectator recruitment	In class	Research Now	YouGov	NY Fed SCE	Prolific
Worker recruitment	Prolific	MTurk	MTurk	MTurk	MTurk
Stakes	\$6	\$6	\$6	\$5	\$4
Decisions per spectator	3	1	1	12	2
Transfer increments	\$0.50	\$1	\$1	\$0.50	\$0.10
Treatments	L, P, E	L, P, E	L, Mx, P	L, P	L, P
Design	Within	Between	Between	Within	Within
Panel B. Spectator demographics					
Male (%)	55.4	48.9	44.0	50.3	48.2
Mean age (years)	28.3	44.5	47.9	49.5	43.5
College educated (%)	100.0	58.2	44.8	62.4	69.9
HH income $\geq$ \$100k (%)	—	24.5	18.9	25.4	27.0
White (%)	—	—	74.6	82.2	75.5
N	271	1,000	882	197	1,476

Notes: This table compares study designs and spectator demographics across the five studies used in our benchmark analysis. Panel A describes key design features; "—" indicates the variable is not available in that dataset. Panel B reports spectator demographics. For Cohn et al. (2023), demographics in Panel B are population-weighted to account for the oversampling of households in the top 5% by income or wealth; household income $\geq$\$100k is computed from self-reported family annual income categories. For Almås et al. (2020), annual household income is computed from monthly income. All incentivized studies implement one randomly selected spectator decision. L = Luck, P = Performance/Merit, Mx = Mixed, E = Efficiency cost.

C.3 Pooling the Benchmark Studies

We summarize four benchmark estimates from comparable U.S. impartial-spectator designs ((Almås et al., 2020; Cohn et al., 2023; Preuss et al., 2025; Harrs and Sterba, 2025)) with a single pooled estimate, computed using a DerSimonian and Laird (1986) random-effects meta-analysis. The DerSimonian and Laird estimator is a widely used conventional random-effects estimator in the medical and social sciences, and random-effects or hierarchical pooling of study-level estimates is also common in economics—for example, in Bellavance et al. (2009)'s meta-analysis of the value of a statistical life and in the hierarchical analyses of Meager (2019, 2022) and Vivalt (2020).

We combine K study-level estimates $\hat{\theta}_i$, each with squared standard error s_i^2. Let the inverse-variance (fixed-effect) weights be $w_i = 1/s_i^2$, and let $Q = \sum_i w_i(\hat{\theta}_i - \bar{\theta}_{\text{FE}})^2$, where

$\bar{\theta}_{\mathrm{FE}} = \sum_i w_i \hat{\theta}_i / \sum_i w_i$, denote Cochran's heterogeneity statistic. The DerSimonian and Laird method-of-moments estimator of the between-study variance τ^2 is

$$\hat{\tau}^2 = \max\left\{0,\ \frac{Q-(K-1)}{\sum_i w_i - \sum_i w_i^2 / \sum_i w_i}\right\}. \tag{C1}$$

The pooled estimate weights each study by the inverse of its total variance, $w_i^* = 1/(s_i^2+\hat{\tau}^2)$:

$$\hat{\theta}_{\mathrm{RE}} = \frac{\sum_i w_i^* \hat{\theta}_i}{\sum_i w_i^*}, \qquad \widehat{\mathrm{Var}}(\hat{\theta}_{\mathrm{RE}}) = \frac{1}{\sum_i w_i^*}, \tag{C2}$$

and the 95 percent confidence intervals we report for the pooled estimate (e.g., Figure 1 and the "Pooled" column of Table 2) are $\hat{\theta}_{\mathrm{RE}} \pm 1.96\,\widehat{\mathrm{Var}}(\hat{\theta}_{\mathrm{RE}})^{1/2}$. The estimator nests inverse-variance fixed effects as the special case $\hat{\tau}^2 = 0$, and the share of cross-study variation reflecting genuine heterogeneity rather than sampling error is $I^2 = \max\{0, (Q-(K-1))/Q\}$.

C.4 Comparison with Fisman et al. (2015)

Fisman et al. (2015) study elites' efficiency preferences using a modified dictator game in which subjects divide an endowment between themselves and another participant. To compare efficiency concerns across designs, we estimate the analogue of their efficiency parameter ρ in our data. Specifically, Fisman et al. (2015) assume that dictators maximize a CES utility function:

$$\left(\alpha\, y_S^{\rho} + (1-\alpha)\, y_O^{\rho}\right)^{1/\rho}, \tag{C3}$$

where y_S is the income the dictator assigns to themselves, y_O is the income assigned to the other participant, α captures the weight on own payoff, and ρ governs the equality–efficiency tradeoff. For our impartial spectators, we assume the analogous function with equal weights ($\alpha = 1/2$), since spectators have no stake in the allocation:

$$\left(\tfrac{1}{2}\, y_H^{\rho} + \tfrac{1}{2}\, y_L^{\rho}\right)^{1/\rho}, \tag{C4}$$

where y_H and y_L denote the earnings assigned to the winner and loser, respectively. In the efficiency-cost condition, the spectator's budget constraint is $y_H + 2y_L = M$, where the price of redistribution is 2 (each dollar taken from the winner generates only 50 cents for

the loser). The first-order condition yields $(y_L/y_H)^{\rho-1} = 2$, which we invert to recover the implied ρ for each spectator:

$$\rho_i = 1 - \frac{\ln(2)}{\ln(y_H/y_L)}. \tag{C5}$$

Spectators who do not redistribute ($y_L = 0$) are assigned $\rho = 1$ (utilitarian); those who fully equalize ($y_H = y_L$) are assigned $\rho = -20$, matching the effective floor in the Fisman et al. (2015) replication data. Since ρ is unbounded below and can take arbitrarily negative values, we compare medians to limit the influence of outliers.

A key challenge is that ρ has different interpretations across designs. In our setting, ρ measures how quickly spectators become more tolerant of inequality as redistribution becomes costlier. In Fisman et al. (2015), by contrast, dictators allocate their own income, so ρ measures how quickly they become more tolerant of advantageous inequality as the cost of redistribution rises. Because 78 percent of YLS students are estimated to prioritize their own income over others' (that is, have $\alpha > 0.6$), the efficiency parameter in their setting partly reflects self-interest rather than a pure preference for efficiency over equality.

We present two comparisons (Appendix Table C2). The first comparison uses the full YLS sample ($N = 208$): the median ρ is 0.699 for YLS students and 0.500 for MBA students—statistically indistinguishable ($p = 0.432$, rank-sum test). This comparison includes all subjects but is harder to interpret because the dictator design conflates efficiency preferences with selfishness for subjects with high $\hat{\alpha}$. The second comparison restricts the Fisman et al. (2015) data to the 15 percent of YLS students they classify as "fair-minded" ($0.45 \leq \hat{\alpha} \leq 0.55$, $N = 30$), who assign roughly equal weight to their own and the other participant's payoff. This removes the selfishness contamination, yielding a median ρ of 0.533—close to our MBA estimate of 0.500.

Both comparisons point to the same conclusion: MBA students' efficiency preferences are similar to those of YLS elites. Both elite groups are substantially more efficiency-seeking than the general population, whose fair-minded subsample has a median ρ of 0.276 (Table C2, Panel B). Our Dyson replication confirms this pattern, with Dyson students exhibiting a median ρ of 0.500—identical to MBA students. However, only 30 of the 208 YLS students are classified as fair-minded. For the remaining 85 percent, the estimated ρ confounds efficiency preferences with self-interest, so the Fisman et al. (2015) data cannot tell us how the majority of elites would trade off efficiency against inequality between others. Our impartial-spectator design closes this gap by measuring efficiency preferences

for the full sample, free of self-interest confounds.

Table C2: Efficiency Preferences: MBA Spectators vs. Fisman et al. (2015)

	Median ρ	Mean ρ	p25	p75	N
Panel A. Spectator design (self-interest removed)					
MBA students	0.500	-3.834	0.000	1.000	266
Dyson students	0.500	-3.665	0.000	1.000	167
Almås et al. — U.S. sample	-20.000	-11.327	-20.000	1.000	332
Panel B. Dictator design, fair-minded subjects ($0.45 \leq \hat{\alpha} \leq 0.55$)					
Fisman et al. — YLS students	0.533	-1.521	0.024	0.904	30
Fisman et al. — General population	0.276	-1.314	-0.114	0.466	227
Panel C. Dictator design, all subjects					
Fisman et al. — YLS students	0.699	-0.182	0.214	0.826	208
Fisman et al. — General population	0.034	-1.243	-0.731	0.477	717

Notes: This table compares the CES efficiency parameter ρ across studies. Higher ρ indicates stronger efficiency preferences ($\rho = 1$: utilitarian; $\rho = 0$: Cobb–Douglas; $\rho \to -\infty$: Rawlsian). Panel A reports implied ρ for our spectator samples, recovered from efficiency-cost treatment choices using $\rho_i = 1 - \ln(2)/\ln(y_H/y_L)$. Panel B restricts the Fisman et al. (2015) sample to "fair-minded" subjects, for whom the selfishness–efficiency confound is minimal. Panel C shows all Fisman et al. (2015) subjects. The spectator design in Panel A removes the self-interest confound by construction.

D Replication: Dyson Business Students

This appendix presents our replication with undergraduate business students from Cornell University's Dyson School of Applied Economics and Management.

D.1 Sample Characteristics

Appendix Table D1 presents summary statistics of the Dyson sample. The sample consists primarily of young students (mean age 19.5 years), with a roughly equal gender split (54.3 percent male). The majority were born in the U.S. (74.7 percent) and report relatively high socioeconomic status during childhood—71.8 percent had enough money for necessities and at least occasional luxuries while growing up. Most aspire to managerial positions (80.7 percent) and private-sector careers (72.7 percent), and about half plan to pursue an MBA (51.6 percent). Only 20.4 percent report voting in recent elections, likely reflecting their young age.

D.2 Implemented Inequality

Appendix Table D2 presents estimates of equation (1) for the Dyson sample. When worker earnings are randomly assigned, Dyson students implement a Gini coefficient of 0.416 (column 1), similar to the MBA estimate. The performance condition increases the implemented Gini coefficient by 0.311 Gini points ($p < 0.01$), or 74.8 percent of the baseline Gini. The efficiency cost condition increases the Gini coefficient by 0.198 Gini points ($p < 0.01$), or 47.6 percent of the baseline. These effects are robust to including spectator fixed effects (column 2) and using only spectators' first choices (columns 3 and 4).

D.3 Distribution of Fairness Ideals

Appendix Tables D3 and D4 present the distribution of fairness ideals in the Dyson sample. Using the between-subject design, we find that 1.9 percent are egalitarians, 23.9 percent are libertarians, 37.6 percent are meritocrats, and 36.6 percent have other fairness ideals. The within-subject design yields similar estimates. Compared to MBA students, Dyson students show lower rates of egalitarianism (1.9 vs. 9.9 percent), higher rates of meritocracy (37.6 vs. 22.6 percent), and similar rates of libertarianism (23.9 vs. 23.2 percent). Like MBA students, they are less likely to conform to standard fairness ideals than the general population. About a quarter of Dyson students (24.5 percent) behave as "moderates."

Table D1: Summary Statistics of the Dyson Sample

	Mean (1)	SD (2)	N (3)
Panel A. Demographic characteristics			
Average age	19.534	0.689	161
Male	0.543	0.500	162
Born in the USA	0.747	0.436	162
Panel B. Financial situation while growing up			
Always had enough money for necessities and luxuries	0.244	0.431	156
Always had enough for necessities and occasional luxuries	0.474	0.501	156
Usually had enough for necessities, rarely for luxuries	0.231	0.423	156
Sometimes did not have enough money for necessities	0.038	0.193	156
Frequently did not have enough money for necessities	0.013	0.113	156
Panel C. Employment preferences			
Working long hours is sometimes necessary	0.776	0.418	161
Wants career in a non-profit	0.199	0.400	161
Wants to become an entrepreneur	0.516	0.501	161
Wants career in private sector	0.727	0.447	161
Work-life balance is more important than a high salary	0.596	0.492	161
Aspire to be a manager	0.807	0.396	161
Wants to do an MBA	0.516	0.501	161
Panel D. Attitudes toward pay determination			
How much responsibility should decide pay	0.994	0.079	161
How well someone works should decide pay	1.000	0.000	161
How hard someone works should decide pay	0.957	0.205	161
Panel E. Voting behavior and social views			
Voted in the last elections	0.204	0.404	152
Success is luck and connections	0.280	0.450	161
Hard work brings a better life	0.863	0.345	161
Completed all three redistributive decisions	0.970	0.171	167
Passed attention check	0.938	0.242	161
Minutes spent in survey	5.674	2.849	167

Notes: This table shows summary statistics of Dyson students in our sample. All variables are based on data self-reported by Dyson students in the exit survey of our study. Employment preferences and social views are based on Dyson students' agreement with several statements in a five-point Likert scale grid. For each statement, we define an indicator variable that equals one if the student selects "strongly agree" or "agree," and zero if the student selects "neither agree nor disagree," "disagree," or "strongly disagree." For the attitudes toward pay determination, we define an indicator variable that equals one if the student selects "essential," "very important," or "fairly important" and zero if the student selects "not very important" or "not important at all."

Table D2: Implemented Gini Coefficient Across Environments (Dyson Sample)

	Dependent variable: Gini coefficient			
	All redistributive decisions		First decision only	
	(1)	(2)	(3)	(4)
Performance condition	0.311***	0.304***	0.334***	0.330***
	(0.036)	(0.044)	(0.071)	(0.072)
Efficiency condition	0.198***	0.194***	0.157**	0.152*
	(0.028)	(0.035)	(0.078)	(0.080)
Constant	0.416***	0.420***	0.413***	0.441***
	(0.031)	(0.023)	(0.060)	(0.087)
Additional controls?	No	Yes	No	Yes
N (redistributive decisions)	494	494	167	167
N (individuals)	167	167	167	167

Notes: This table displays estimates of β_0, β_1, and β_2 from equation (1) estimated on the Dyson sample. The omitted treatment category is the luck condition. The specifications in columns 2 and 4 include additional controls. In column 2, we include spectator fixed effects. In column 4, we control for gender, parental financial situation while growing up, and a dummy for being born in the U.S. Heteroskedasticity-robust standard errors clustered at the spectator level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D3: Joint Distribution of Redistribution in Luck and Performance Treatments (Dyson Sample)

		Earnings Redistributed in Performance Condition						
		\$0	\$1	\$2	\$3	\$4	\$5	\$6
Luck Condition	\$0	0.160	0.061	0.018	0.000	0.000	0.000	0.000
	\$1	0.061	0.037	0.025	0.006	0.006	0.000	0.000
	\$2	0.117	0.067	0.049	0.006	0.000	0.000	0.000
	\$3	0.166	0.067	0.074	0.061	0.000	0.000	0.000
	\$4	0.000	0.000	0.000	0.000	0.000	0.000	0.006
	\$5	0.000	0.006	0.000	0.000	0.000	0.000	0.000
	\$6	0.000	0.000	0.000	0.006	0.000	0.000	0.000

Libertarians
Meritocrats
Egalitarians

Notes: This table shows the joint distribution of redistribution choices across luck and performance conditions in the Dyson sample, excluding the efficiency cost treatment. See notes to Table A4 for details.

Table D4: Distribution of Fairness Ideals: Elites vs. Average Citizens (Including Dyson Sample)

	Percentage of...			
	Libertarians (1)	Meritocrats (2)	Egalitarians (3)	Other ideals (4)
Panel A. MBA and undergraduate business students				
MBA students (spectators' first choices)	23.16	22.65	9.89	44.30
	(4.35)	(5.99)	(3.15)	(8.05)
MBA students (repeated-measures design)	15.33	28.74	6.51	49.43
	(1.58)	(1.98)	(1.08)	(2.76)
Dyson students (spectators' first choices)	23.91	37.60	1.85	36.63
	(6.36)	(7.78)	(1.85)	(10.22)
Dyson students (repeated-measures design)	15.95	30.67	6.13	47.24
	(2.03)	(2.56)	(1.33)	(3.53)
Panel B. Average American				
SCE – Average American	39.77	36.67	8.57	14.99
	(5.25)	(6.31)	(2.75)	(8.66)
Almås et al. (2020) – Average American	29.43	37.54	15.32	17.72
	(2.50)	(3.43)	(1.98)	(4.69)
Cohn et al. (2023) – Average American	12.80	60.33	17.45	9.42
	(1.89)	(3.22)	(2.19)	(4.33)
Harrs and Sterba (2025) – Average American	8.68	45.24	14.61	31.46
	(1.04)	(2.30)	(1.30)	(2.84)
Panel C. Higher-income samples				
SCE – USA income over 100k	44.44	40.32	3.23	12.01
	(12.05)	(13.27)	(3.23)	(18.22)
Almås et al. (2020) – USA income over 100k	22.67	41.92	14.08	21.33
	(4.87)	(7.11)	(4.16)	(9.57)
Cohn et al. (2023) – USA top 5 percent	25.15	59.14	8.97	6.75
	(3.33)	(4.37)	(2.38)	(5.99)
Harrs and Sterba (2025) – USA income over 100k	8.96	41.44	7.61	41.99
	(2.02)	(4.42)	(1.89)	(5.21)
Panel D. Low-income samples				
SCE – USA income below 100k	38.57	35.14	10.81	15.48
	(5.86)	(7.32)	(3.63)	(10.06)
Almås et al. (2020) – USA income below 100k	31.40	36.33	15.65	16.63
	(2.89)	(3.92)	(2.25)	(5.37)
Cohn et al. (2023) – USA bottom 95 percent	12.06	60.54	17.83	9.57
	(2.75)	(4.62)	(3.06)	(6.19)
Harrs and Sterba (2025) – USA income below 100k	8.58	46.71	17.16	27.55
	(1.21)	(2.69)	(1.62)	(3.36)

Notes: This table shows estimates of the fraction of egalitarians, libertarians, and meritocrats in various studies and samples. See notes to Table 3 for details.

E Empirical Framework

This appendix presents a statistical framework of income redistribution that mirrors our experimental setup. We use the framework to define each type of fairness ideal and the identification assumptions.

E.1 Defining Fairness Ideals

We begin by formalizing how spectators make redistribution decisions based on their fairness ideals. Consider a spectator who observes the initial earnings distribution $(w_H, w_L) \in \mathbb{R}_+ \times \mathbb{R}_+$ of two workers who competed for a fixed prize, where $w_H > w_L$ represents the earnings of the worker who won the competition (the "high-earnings" worker) and w_L the earnings of the worker who lost (the "low-earnings" worker).

We study an environment in which worker earnings are allocated based on some combination of performance and luck. Let $z_p \in [0, 1]$ denote the probability that the outcomes of a worker pair p are determined by chance. z_p can be interpreted as the role that luck plays in determining income inequality. When $z_p = 0$, the initial earnings are determined by workers' performance. The worker with a better performance earns w_H, and the other worker earns w_L; thus, inequality reflects differences in performance. When $z_p = 1$, earnings are randomly assigned to workers, and inequality therefore reflects luck alone. Spectators observe w_H, w_L, and z_p but do not observe worker performance.

Let Y_{ip} be the earnings redistributed from the high- to the low-earnings worker in worker pair p by spectator i. We assume that a spectator's choice of how much earnings to redistribute, Y_{ip}, is guided by their *fairness ideal*, that is, their preferences for what constitutes a fair earnings allocation given the role played by luck in determining worker earnings, z_p. We model fairness ideals as a mapping from initial distributions to final earnings distributions. Following the literature, we focus on three main types of fairness ideals:

Definition 1 (Fairness ideals)**.**

- *Spectator i is an* egalitarian *if $Y_{ip}(z_p) = \frac{w_H - w_L}{2}$ for all z_p.*
- *Spectator i is a* libertarian *if $Y_{ip}(z_p) = 0$ for all z_p.*
- *Spectator i is a* meritocrat *if $Y_{ip}(z_p)$ is strictly increasing in z_p and $Y_{ip}(1) = \frac{w_H - w_L}{2}$.*

Egalitarian spectators equalize the workers' earnings regardless of how the earnings were generated. Libertarian spectators leave the initial earnings unchanged regardless of how earnings differences arose. Meritocratic spectators condition their redistributive decisions on the source of inequality. Specifically, as the role of chance in determining earnings z_p increases, they redistribute more earnings and, when inequality is entirely luck-based (i.e., there is no merit to the earnings allocation), they equalize earnings.

The aggregate level of redistribution in a population of spectators, $\mathbb{E}[Y_{ip}]$, links directly to the distribution of fairness ideals. Let α_E, α_L, and α_M be the proportions of egalitarians, libertarians, and meritocrats in a population. By the law of iterated expectations

$$\mathbb{E}[Y_{ip}] = \alpha_E \frac{w_H - w_L}{2} + \alpha_M Y_{ip}^M + (1 - \alpha_E - \alpha_L - \alpha_M) Y_{ip}^O, \tag{E1}$$

where $Y_{ip}^M \equiv \mathbb{E}[Y_{ip}|\text{Meritocrat}]$ and $Y_{ip}^O \equiv \mathbb{E}[Y_{ip}|\text{Other ideal}]$ are the average earnings redistributed conditional on being a meritocrat and having other fairness ideals, respectively (the libertarian term is omitted because libertarians never redistribute).

Equation (E1) highlights two reasons why elites may have different redistributive choices than non-elites. First, the distribution of fairness ideals among elites (the α's) may differ from those among non-elites. Second, Y_{ip}^M and Y_{ip}^O may differ for elites and non-elites. In other words, elites who follow a meritocratic fairness ideal or fall outside the three classifications above may redistribute different amounts than the corresponding non-elites.

E.2 Identification of Fairness Ideals

Research design. Our first research design relies on comparisons *across* subjects, using only each spectator's first redistributive decision and excluding those who saw the efficiency cost environment first. We identify the fraction of egalitarians by the fraction of individuals who divide earnings equally when the winner is determined by performance. We identify the fraction of libertarians by the fraction of individuals who redistribute no earnings when the winner is determined by luck. Finally, we identify the fraction of meritocrats by the difference between (i) the fraction of individuals who allocate more to the winner when worker outcomes are determined by performance and (ii) the fraction of individuals who allocate more to the winner when worker outcomes are determined by luck. We refer to the remainder of the population as having "other" fairness ideals.

Our second research design relies on *within*-subject comparisons across worker pairs, using all redistributive decisions except those from the efficiency cost environment. We

classify a spectator as an egalitarian if they divide earnings equally in both the luck and performance environments, as a libertarian if they do not redistribute earnings in either environment, and as a meritocrat if they equalize earnings in the luck environment but give strictly more to the winner in the performance environment. This within-subject definition of meritocrats coincides with that of Harrs and Sterba (2025). We refer to individuals who do not fit any of these classifications as having "other" fairness ideals.

Identification assumptions. For the between-subject design, we follow the standard assumptions in the literature (e.g., Almås et al., 2020; Cohn et al., 2023), which we formalize through the lens of our statistical framework. For the within-subject design, we extend these assumptions to account for multiple observations per spectator.

To identify egalitarians in the between-subject design, our identification assumption is that individuals who divide earnings equally when worker earnings are assigned based on performance would have also divided earnings equally under any other environment. Formally, if $Y_{ip}(z_p) = \frac{w_H - w_L}{2}$ for $z_p = 0$, then $Y_{ip}(z_p) = \frac{w_H - w_L}{2}$ for all z_p. In the within-subject design, our identification assumption is that individuals who divide equally in the luck *and* performance environments would have divided equally under any other environment. Formally, if $Y_{ip}(z_p) = \frac{w_H - w_L}{2}$ for $z_p = 0$ *and* $Y_{ip'}(z_{p'}) = \frac{w_H - w_L}{2}$ for $z_{p'} = 1$, then $Y_{ip}(z_p) = \frac{w_H - w_L}{2}$ for all z_p.

To identify libertarians in the between-subject design, our identification assumption is that individuals who do not redistribute earnings in the luck environment would not redistribute earnings under any other environment. Formally, if $Y_{ip}(z_p) = 0$ for $z_p = 1$, then $Y_{ip}(z_p) = 0$ for all z_p. In the within-subject design, our identification assumption is that individuals who do not redistribute earnings in either the luck or performance environments would not redistribute earnings under any other environment. Formally, if $Y_{ip}(z_p) = 0$ for $z_p = 0$ *and* $Y_{ip'}(z_{p'}) = 0$ for $z_{p'} = 1$, then $Y_{ip}(z_p) = 0$ for all z_p.

To identify meritocrats in the between-subject design, we follow Almås et al. (2020). The share of spectators who allocate more to the winner in the performance condition equals the combined share of libertarians (who never redistribute) and meritocrats (who reward performance differences), while the share who allocate more to the winner in the luck condition equals the share of libertarians alone (since meritocrats equalize under luck). The difference between these two shares therefore identifies the meritocrat fraction. In the within-subject design, our identification assumption is that those who equalize earnings in the luck environment but give strictly more to the winner in the performance environment

would strictly redistribute more earnings as the role of luck increases. Formally, if $Y_{ip}(0) < Y_{ip}(1) = \frac{w_H - w_L}{2}$, then $Y_{ip}(z_p) > Y_{ip}(z'_p)$ whenever $z_p > z'_p$.

Both the between- and within-subject designs have advantages and disadvantages for identification. To understand these, recall that a fairness ideal is a mapping from circumstances to choices. Reconstructing this mapping requires observing an infinite number of counterfactual choices. The research designs "extrapolate" spectator behavior based on a limited set of observed choices: one in the between-subject design and two (sequential) choices in the within-subject design. The between-subject design uses a single choice, so it avoids assumptions about sequential-decision bias but is more susceptible to measurement error and choice noise. In contrast, the within-subject design, using two choices, mitigates the measurement error concerns by incorporating more information about spectators' behavior, at the cost of assuming that the first choice does not influence the second.

E.3 A Statistical Model of Redistribution and Inequality Source

Recall Y_{ip} represents the earnings redistributed from the high- to the low-earnings worker in worker pair p by spectator i. We model Y_{ip} as a function of the role played by luck in determining worker earnings, z_p.[23] For the remainder of this subsection, we restrict attention to $z_p \in \{0, 1\}$ and consider the two environments studied by most experimental work: worker earnings are determined by performance ($z_p = 0$) or by chance ($z_p = 1$). Let $Y_{ip}(0)$ be the amount redistributed if earnings are determined by performance and $Y_{ip}(1)$ the amount redistributed if outcomes are determined by chance. These two redistribution levels denote potential outcomes for different levels of luck, but only one of the two outcomes is observed for a given worker pair. The observed redistribution, $Y_{ip}(z_p)$, can be written in terms of these potential outcomes as

$$Y_{ip}(z_p) = Y_{ip}(0) + \underbrace{\Big(Y_{ip}(1) - Y_{ip}(0)\Big)}_{\text{"Source-of-inequality effect"}(\beta_i)} z_p, \tag{E2}$$

where $Y_{ip}(1) - Y_{ip}(0) \equiv \beta_i$ measures the effect of changing the income-generating process from performance to chance, or "source-of-inequality effect," for short.

Suppose that we observed spectator choices for two worker pairs, p and p', with earnings in pair p determined by performance ($z_p = 0$) and earnings in pair p' randomly assigned

[23] We abstract from modeling Y_{ip} as a function of w_H and w_L for simplicity. In our experiment, w_H and w_L are constant across worker pairs, and the only variable that changes across worker pairs is z_p.

($z_{p'} = 1$). One could then compare average redistribution in the random-assignment pair, $\mathbb{E}[Y_{ip'}|z_{p'} = 1]$, with average redistribution in the performance pair, $\mathbb{E}[Y_{ip}|z_p = 0]$. This comparison can be written as

$$\mathbb{E}[Y_{ip'}|z_{p'} = 1] - \mathbb{E}[Y_{ip}|z_p = 0] = \underbrace{\mathbb{E}[Y_{ip'}(1)|z_{p'} = 1] - \mathbb{E}[Y_{ip'}(0)|z_{p'} = 1]}_{\text{Term 1: Source-of-inequality effect}} + \underbrace{\mathbb{E}[Y_{ip'}(0)|z_{p'} = 1] - \mathbb{E}[Y_{ip}(0)|z_p = 0]}_{\text{Term 2: Sequential-decision bias}}. \tag{E3}$$

Equation (E3) shows that comparing the redistributive behavior of spectators for different worker pairs yields the sum of two terms. The first one is the effect of the source of inequality on earnings redistributed. The second term is a potential bias that arises when comparing sequential decisions. This term captures mechanisms such as anchoring effects, contrast effects, and learning effects. To address this potential bias, most literature uses a between-subject design in which spectators make only one decision. We provide both between-subject estimates (using only the first decision) and within-subject estimates (using the decisions in the luck and performance conditions) for comparison.

In addition to measuring the source-of-inequality effect, we also examine how introducing an efficiency cost affects redistribution (Y_{ip}). We analyze this using the same framework, holding constant the source of inequality and re-interpreting z_p as an indicator for redistribution having a cost. In this case, β_i in equation (E2) and Term 1 in equation (E3) measure the responsiveness of redistribution to the efficiency cost of redistribution.

F Narratives of Inequality by Fairness Ideal

This appendix classifies MBA students' open-ended responses about the drivers of income inequality in the United States using large language models (LLMs). We describe the classification methodology, report the distribution of responses by fairness ideal, and provide qualitative evidence.

F.1 Classification Methodology

We classify the 116 open-ended responses collected from the second MBA cohort to the question: *"What do you believe is the main driver of income inequality in the United States?"* Responses are typically short (most are 1–10 words), so many are ambiguous between adjacent categories.

We developed the codebook inductively, providing all responses to ChatGPT 5.2 Pro and instructing it to derive 6–8 categories from the data without a predefined scheme. The procedure yielded eight categories (see Appendix Table F4): education (C01), unequal opportunities (C02), discrimination (C03), government policy (C04), corporate/elite power (C05), historical legacy (C06), behavioral/cultural (C07), and other/non-substantive (C08).

Each response was then classified using OpenAI's gpt-5-mini with structured outputs, producing both a multi-label coding (all relevant categories) and a single-best primary label. To establish robustness, we re-ran the classification with three additional models from two providers (gpt-5-nano, claude-sonnet-4-6, claude-haiku-4-5). Agreement across all four models is 82.8 percent, and most disagreements involve substantively adjacent categories or short, ambiguous responses (see Appendix Tables F3–F5).

F.2 Overall Distribution and Cross-Tabulation

Appendix Table F1 reports the multi-label distribution across all 116 responses. Because the classifier assigns every relevant category to each response, percentages sum to more than 100 percent. Education and skill formation is the most frequently cited driver (31.9 percent), followed by unequal opportunities (24.1 percent), government policy (17.2 percent), and discrimination (16.4 percent). Overall, 21.6 percent of responses mention two or more categories.

Table F1: Overall Distribution of Inequality Drivers, Multi-Label

Code	Category	N	%
C01	Education and skill formation	37	31.9
C02	Unequal opportunities, mobility barriers, starting conditions	28	24.1
C03	Discrimination and identity-based bias	19	16.4
C04	Government policy, politics, institutional governance	20	17.2
C05	Corporate/elite power, wealth concentration	13	11.2
C06	Historical legacy, broad systemic/structural inequity	13	11.2
C07	Behavioral, cultural, and ideological explanations	12	10.3
C08	No clear driver / rejects premise / non-substantive	3	2.6

Notes: Multi-label distribution of 116 MBA students' open-ended responses across eight LLM-identified categories. Each response may be coded into multiple categories; percentages sum to more than 100%. Classifications are based on gpt-5-mini.

F.3 Qualitative Evidence by Fairness Ideal

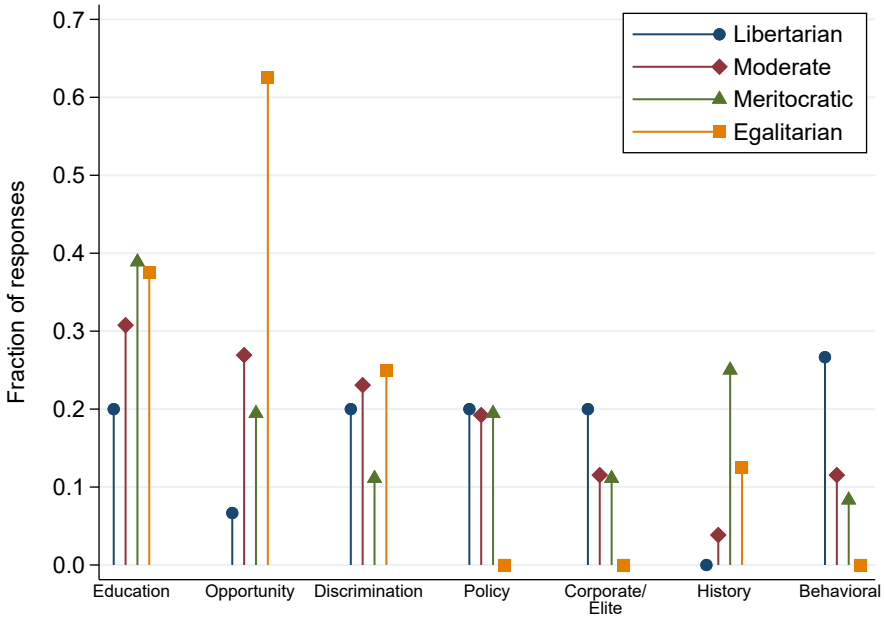
Appendix Table F2 reports the multi-label category distribution by fairness ideal for the 85 classified respondents, and Appendix Figure F1 visualizes these shares. The qualitative evidence below illustrates each type's distinctive profile with representative quotes.

Table F2: Inequality Drivers by Fairness Ideal, Multi-Label (%)

	Libertarian	Moderate	Meritocratic	Egalitarian
C01: Education	20.0	30.8	38.9	37.5
C02: Opportunity	6.7	26.9	19.4	62.5
C03: Discrimination	20.0	23.1	11.1	25.0
C04: Policy	20.0	19.2	19.4	0.0
C05: Corporate/elite	20.0	11.5	11.1	0.0
C06: History	0.0	3.8	25.0	12.5
C07: Behavioral	26.7	11.5	8.3	0.0
C08: Non-substantive	6.7	0.0	0.0	12.5
N	15	26	36	8

Notes: Each cell reports the share of respondents within a fairness ideal whose response mentions the given category. Because responses may be coded into multiple categories, columns sum to more than 100 percent. The 31 respondents not classified as libertarian, moderate, meritocratic, or egalitarian are omitted.

Figure F1: Inequality Drivers by Fairness Ideal



Notes: Each marker shows the fraction of respondents within a fairness ideal whose response mentions the given category. Because responses may be coded into multiple categories, fractions can sum to more than one within each fairness type. Classifications are based on gpt-5-mini.

Libertarian Narratives. Libertarians are the only fairness type whose modal category is behavioral and cultural explanations (26.7 percent). They attribute inequality to personal choices, work ethic, and cultural differences—responses such as "willingness to work hard," "poor choices," and "cultural differences." This individualistic framing is consistent with their preference not to redistribute: if inequality reflects personal agency, there is less justification for corrective transfers.

Beyond the behavioral emphasis, libertarian responses are spread evenly across education, discrimination, corporate power, and policy (each cited by 20.0 percent). They are the least likely of any type to cite barriers such as unequal opportunities. Some express outright skepticism about inequality claims: "Women demanding more money than what Men make, there is no such thing as income inequality in the U.S. unless you are brainwashed by false statistics, causation without correlation."

Moderate Narratives. Moderates identify multiple distinct mechanisms behind inequality. Their top categories are education (30.8 percent), unequal opportunities (26.9 percent),

discrimination (23.1 percent), and policy (19.2 percent). 50 percent of moderates cite at least one barrier category, compared to only 27 percent of libertarians.

Many point to starting conditions: "Circumstance," "Socioeconomic status at birth," "How you grew up," and "access to opportunities." This focus on initial endowments aligns with their experimental pattern of redistributing more when inequality stems from luck while still allowing some luck-based inequality to persist. Others emphasize educational factors—"difference in accessibility of education" and "intellect and education inequality"—without the systemic framing of egalitarians.

The most distinctive feature of the moderate profile is what it *excludes*: near-zero historical or systemic framing (3.8 percent), in stark contrast to meritocrats (25.0 percent). Identity factors appear through mentions of "white supremacy," "racism, sexism, greed," and "race," but typically as proximate mechanisms rather than deep structural forces. This pattern suggests that moderates are not well characterized as soft meritocrats or soft libertarians, but instead hold a distinct worldview.

Meritocratic Narratives. Meritocrats concentrate on education (38.9 percent), the highest share of any type. They describe "access to high quality and relevant education," "lack of equitable access to quality education," and education being "too expensive for more people to access." This framing positions unequal educational access—rather than education per se—as the primary barrier to mobility, consistent with a worldview that values merit but acknowledges an uneven playing field.

Meritocrats also cite historical and systemic factors at the highest rate of any type (25.0 percent). One student references "historical issues that continue to affect current socio-economic stratification"; another cites "history (residual injustice and cultural difference) exacerbated by a financial system that gives the 'further ahead' exponential resources to get ahead." This willingness to acknowledge structural forces—while still operating within a framework that values merit-based outcomes—distinguishes meritocrats from libertarians, who never invoke historical framing.

Egalitarian Narratives. Egalitarians emphasize barriers: 62.5 percent cite unequal opportunities and 25.0 percent cite discrimination—the highest barrier rates of any type. Rather than viewing education as a meritocratic sorting mechanism, egalitarians frame it as reinforcing existing disparities: "Access to education (cost of studying/school/etc)" and "barriers to opportunity/education." Identity-based discrimination features prominently,

with one student citing "sexism and racism" as primary drivers and another pointing to "historical systems around identity and location."

This framing—inequality as the product of institutional barriers rather than individual failings—is consistent with egalitarians' strong preference for redistribution. No egalitarian cites behavioral or cultural explanations, the sharpest contrast with libertarians (26.7 percent).

F.4 Classification Robustness and Codebook

This section provides details on the codebook and inter-model agreement.

Table F3: Inter-Model Agreement Rates

Comparison	Agree / N	Agreement (%)
Within-provider (OpenAI)		
gpt-5-mini vs. gpt-5-nano	107 / 116	92.2
Within-provider (Anthropic)		
claude-sonnet-4-6 vs. claude-haiku-4-5	107 / 116	92.2
Cross-provider		
gpt-5-mini vs. claude-sonnet-4-6	102 / 116	87.9
gpt-5-mini vs. claude-haiku-4-5	99 / 116	85.3
gpt-5-nano vs. claude-sonnet-4-6	104 / 116	89.7
gpt-5-nano vs. claude-haiku-4-5	105 / 116	90.5
All four models agree	96 / 116	82.8

Notes: Each row reports the share of 116 responses where the listed models assign the same single-best primary category. Cross-provider comparisons (OpenAI vs. Anthropic) are especially informative as the models differ in training data, architecture, and fine-tuning.

Table F4: Codebook for Open-Ended Inequality Response Classification

Code	Category	Include if	Exclude if
C01	Education and skill formation	Mentions education (K–12, college) as key driver; emphasizes affordability/cost of schooling; points to skill gaps, credential requirements, or knowledge deficits	Focuses on unequal opportunity in general without education being central → C02; focuses on race/-gender discrimination → C03
C02	Unequal opportunities, mobility barriers, starting conditions	Mentions lack of opportunities, "uneven playing field"; emphasizes parental income, upbringing, geography; highlights connections/social capital	Core argument is education access/quality → C01; identity-based discrimination → C03; specific government lever → C04
C03	Discrimination and identity-based bias	References race/racism, gender/-sexism, white supremacy, or discrimination; mentions bias tied to identity	Favoritism via networks without identity framing → C02; historical injustice without present-day discrimination → C06
C04	Government policy, politics, institutional governance	Mentions tax policy, regulation, laws, "the government"; references politics, money in politics, or named political leaders; mentions social welfare/healthcare policy	Blames corporations/elites without emphasizing government levers → C05; only broad "systemic" language → C06
C05	Corporate/elite power, wealth concentration	References economic elite, corporations, or wealth concentration; mentions "greed" as elite capture; points to executive pay dynamics or capitalism	Focuses primarily on taxes/regulation/politics → C04; on starting conditions without elite capture → C02
C06	Historical legacy, broad systemic/structural inequity	Mentions history, residual injustice, long-run stratification; uses "systemic challenges," "social inequity," "structural inequity"	Specifies a primary mechanism fitting another category → code that; gives concrete opportunity pathway without historical framing → C02
C07	Behavioral, cultural, ideological	Mentions work ethic, willingness to work hard, poor choices, confidence, luck; mentions culture/cultural differences; frames inequality as driven by beliefs/perceptions	Emphasizes structural barriers to education → C01; opportunity constraints → C02; identity discrimination → C03
C08	No clear driver / rejects premise	Expresses uncertainty or no answer; denies inequality is real; says "no one factor" without naming any	Any specific driver is named (even if hedged) → code relevant category

Notes: Codebook developed inductively by providing all 116 responses to ChatGPT 5.2 Pro with instructions to derive categories from the data without predefined schemes. Categories are defined to be mutually exclusive at the definition level; individual responses may receive multiple labels.

Table F5: Responses with Inter-Model Disagreement

Response text	gpt-5 mini	gpt-5 nano	claude sonnet	claude haiku
Barriers to opportunity/education	C02	C02	**C01**	**C01**
Current laws and inequality of race and gender in the decision room	**C04**	C03	C03	C03
Intellect and education inequality	**C07**	C01	C01	C01
Lack of centralized knowledge and access to resources	**C01**	C02	C02	C02
Lack of environment that encourages competition	**C05**	**C02**	C07	C07
Lack of parental guidance and inadequate education	**C02**	C01	C01	C01
Laws, access to opportunities	C04	C04	**C02**	C04
Leftist ideas	C07	C07	C07	**C08**
Opportunity and education	C02	C02	C02	**C01**
People think investing is gambling. Basic knowledge about markets is absent	C01	C01	**C07**	**C07**
Structural inequity, unequal opportunities	C06	C06	**C02**	C06
Systemic bias	C03	C03	**C06**	**C06**
Systemic issues in regards to race, gender and overall socioeconomic situations	C03	C03	C03	**C06**
systemic poverty and capitalism	C06	C06	**C05**	**C05**
Transparency and Opportunity	**C04**	C02	C02	C02
Trump	C04	**C08**	C04	**C08**
Wage stagnation relative to inflation	C05	**C06**	C05	**C04**
Women demanding more money than what Men make, there is no such thing as income inequality. . .	C03	C03	**C08**	**C08**
Wrong people at the wrong place, playing a perception game on general population	C05	**C02**	C05	**C07**
Wrongly enforced initiative for diversity initiatives	C04	C04	**C03**	C04

Notes: Table lists all 20 responses (of 116) where at least one of four models assigns a different primary category. Responses are sorted alphabetically. **Bold** entries indicate the model(s) that deviate from the modal classification. Of the 20 disagreements, 11 are 3-vs-1 splits and 9 are 2-vs-2 (or 2-vs-1-vs-1) splits. Category codes: C01 = Education; C02 = Opportunities; C03 = Discrimination; C04 = Policy; C05 = Corporate/elite power; C06 = Systemic; C07 = Behavioral/cultural; C08 = Non-substantive. See Appendix Table F4 for full definitions.