



OPEN ACCESS

EDITED BY

Hande Erkut,
Social Science Research Center
Berlin, Germany

REVIEWED BY

Yiming Liu,
Humboldt University of Berlin, Germany
Zvonimir Basic,
University of Glasgow, United Kingdom

*CORRESPONDENCE

Reuben Kline
✉ reuben.kline@stonybrook.edu

RECEIVED 06 September 2024

ACCEPTED 13 January 2025

PUBLISHED 21 February 2025

CITATION

Kline R, Galeotti F and Orsini R (2025)
Meritocracy or malfeasance: violations of
meritocracy erode civic rule following.
Front. Behav. Econ. 4:1492421.
doi: 10.3389/frbhe.2025.1492421

COPYRIGHT

© 2025 Kline, Galeotti and Orsini. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Meritocracy or malfeasance: violations of meritocracy erode civic rule following

Reuben Kline^{1*}, Fabio Galeotti² and Raimondello Orsini³

¹Department of Political Science, Stony Brook University, Stony Brook, NY, United States, ²CNRS, Université Lumière Lyon 2, Université Jean Monnet Saint-Etienne, emlyon business school, GATE, Lyon, France, ³Department of Economics, University of Bologna, Bologna, Emilia-Romagna, Italy

The perceived legitimacy of institutions, along with the voluntary compliance with authority that it undergirds, is crucial for stable governance and economic development. Legitimacy varies greatly across individuals and societies. We introduce a simple model of meritocratic equity—the notion that in a social exchange, individuals should receive greater compensation if their contributions exceed those of others. We argue that violations of meritocratic equity undermine the legitimacy of authority, leading to breaking rules, laws and civic norms—behaviors we refer to as justified malfeasance—in an effort to reduce perceived inequity. Using data from an incentivized laboratory experiment conducted in the United States and Italy and complemented by data from the World Values Survey, we investigate the effect of meritocratic violations on malfeasance. We find convergent evidence that meritocratic inequity explains variation in justified malfeasance across individuals and across countries. We conclude by discussing the implications of our results for multiple equilibria in societal levels of malfeasance and voluntary compliance with authority.

KEYWORDS

malfeasance, corruption, meritocracy, equity theory, behavioral ethics, multi-level modeling

Legitimacy, meritocracy, and malfeasance

The legitimacy of political and organizational authority is an important topic going back to Plato and Aristotle. Recently, research across the social sciences has brought this topic to the forefront of academic discourse once again, underscoring its perennial importance. This recent, more empirical, research in political science (Dickson et al., 2015, 2009), behavioral economics (Akerlof, 2017; Xiao, 2013; Frey et al., 2004), and social psychology (Tyler, 2011, 2006) demonstrates that rules, authorities, and institutions are considered legitimate if they are regarded as just, fair, and proper (Zelditch, 2001).

An influential conception of legitimate authority comes from the psychological literature on procedural justice, which defines legitimacy as “judgment by group members that they ought to voluntarily obey social rules and authorities irrespective of the likelihood of reward or punishment” (Tyler, 1997, p. 323). When authority is legitimate, individuals will be predisposed to obey the authority voluntarily rather than out of any concern for either reward or punishment. Voluntary obedience is important for the proper functioning of institutional authority because enforcement costs will likely outweigh the beneficial aspects of institutions if individuals do not generally voluntarily comply with authority (Tyler, 2006).

It is not merely specific rules, norms, or institutions that may or may not be considered legitimate; perceptions of legitimacy also apply to the societal distribution of resources. As Tyler (2006, p. 384) notes,

People also judge the legitimacy of social arrangements such as economic markets and/or the social or economic standing of people or groups. Whenever there are differences in social or economic standing between people or between groups, issues are raised about the legitimacy of those differences and of the processes through which they arise.

Although all societies make judgments regarding the legitimacy of the resource distribution, they differ on which procedures constitute legitimate processes for the societal allocation of wealth. Some tend to prefer more egalitarian arrangements, and others meritocratic ones (Reynolds and Xian, 2014; Kunovich and Slomczynski, 2007; Osberg and Smeeding, 2006). Although there are a number of potential sources of legitimacy in the societal income distribution (Tyler, 2011), here we are interested in meritocracy, or allocating rewards as a function of merit (Almås et al., 2010; Sen, 2000; Young, 1994).

We investigate how legitimacy is affected when a society's income distribution is perceived as unmeritocratic. We argue that when individuals view income as disconnected from merit, it undermines the legitimacy of authority, leading them to feel justified in violating rules and norms to address perceived inequities. We refer to this behavior of breaking laws and civic rules as *justified malfeasance*, which serves as a proxy for legitimacy.¹ According to the dictionary, malfeasance is “wrongdoing, especially by a public official.” Here we have a broader concept in mind, one for which there is no extant term that perfectly fits. We revise this definition to define malfeasance as “self-serving wrongdoing, especially within a civil, legal, or public sphere.” Note that the broader definition we adopt here is also in line with some usages that fall outside this dictionary definition, e.g., *corporate malfeasance*. Essentially, we operationalize malfeasance as the polar opposite of civic-minded behavior. For example, it is civic-minded to pay your taxes and it is malfeasant to evade taxes; It is civic-minded to be an uncorrupted public official and it is malfeasant to take bribes; it is civic-minded to avail oneself only of public benefits to which one is entitled and it

is malfeasant to fraudulently obtain government benefits, and so on.

Our focus is on meritocratic equity. But first it is useful to first differentiate meritocratic equity from the more familiar concept of egalitarianism. Egalitarianism simply prescribes that equity be defined by equality of outcomes, *irrespective of inputs*. Recent work suggests that individuals also care about other types of equity, including merit-based equity (Starmans et al., 2017). These non-egalitarian equity motives are often referred to with the more general term of “fairness” (Szynger et al., 2017; Starmans et al., 2017). Starmans et al. (2017) argue that in general, people have a preference for “fairness,” including meritocracy, rather than strict equality. They argue that because the subjects in laboratory studies enter the lab as equals and are typically un-differentiated within the experiment, what often appears to be a preference for equality is actually confounded with fairness—the equal outcome is also the (artificially) fair one.

This potential confounding is important for the question at hand because many societies that are unequal also tend to be characterized by corruption, fraud, cronyism, nepotism and a lack of the rule of law (Alesina and Angeletos, 2005a,b). These factors contribute to a perception that the distribution of income is un-meritocratic, and therefore inequitable. To the extent that this is true, then inequality and meritocratic inequity are likely to be confounded in cross-national observational data as well. Even if most existing experimental designs are typically not able to differentiate between egalitarianism and meritocratic equity, observational data on its own cannot fully uncover the causal mechanisms such a relationship.

To overcome the limitations of a single type of data collection we leverage two distinct but ultimately convergent sources of evidence. First, we present results from an incentivized, real-effort laboratory experiment conducted in the United States and Italy in which we directly manipulate the equality and equity of the income distribution. Second, we complement this data with country level estimates from multilevel regression analysis (MLM) of four items measuring malfeasance from World Values Survey (WVS) data, covering 77 countries, including the United States and Italy.

To ground our empirics in theory, we develop a simple formal model based on an influential psychological theory of equity, simply called equity theory (Adams, 1963). We draw upon this theory because its conception of equity is a simple comparison of inputs and outcomes across individuals in a social exchange. Following Adams, we define meritocratic equity as a characteristic of a rule or process prescribing that individual B should receive a greater outcome (reward) than individual A if B's inputs (contribution) are higher than A's in a given social exchange (Adams, 1963). This is a simple and intuitive way to capture meritocratic equity, and it can easily be operationalized, controlled, and manipulated in an incentivized laboratory experiment. Moreover, there is a widely used survey measure of meritocracy in the form of an item from the WVS that asks respondents to express their agreement with the statement “hard work brings success, or success is more a matter of luck and connections.” Adams' equity theory provides an intuitive formulation of meritocracy that can be operationalized using multiple sources of data.

¹ Engaging in malfeasance indicates a breakdown in the legitimacy of authority, even when such actions are intended to rectify perceived injustices (e.g., taking money from someone who has benefited from violations of meritocracy). This is because the individual disobeys the rules established by that authority. This dynamic is captured in our experiment. However, it is also possible that violations of meritocracy undermine the legitimacy of authority, which may lead to malfeasance, even if these actions do not contribute to reducing inequity. For instance, one might engage in malfeasance after experiencing unmeritocratic treatment as a way to harm the authority or organization responsible for that treatment. In our experiment, taking money from the experimenter (as the authority figure) rather than from another player could better illustrate this dynamic (we thank an anonymous reviewer for this suggestion). Unfortunately, our experiment does not address this specific aspect, which may limit our findings; we suggest this as a direction for future research. Nonetheless, as we will see later, this limitation is somewhat mitigated by including a separate study based on survey data from the World Value Survey, which supports our experimental findings.

Meritocracy across countries and in the lab

There is extensive evidence that individuals care not only about the final distribution of income but also about the equity of the process that gives rise to it (e.g., [Starmans et al., 2017](#); [Duru-Bellat and Tenret, 2012](#); [Bolton et al., 2005](#); [Konow, 2000](#)). Recent evidence indicates that a preference for meritocracy, at least under certain circumstances, is deeply rooted in human psychology, and tends to occur naturally and regularly as part of general childhood development ([Kanngiesser and Warneken, 2012](#); [Almås et al., 2010](#)). Nonetheless, there is variation in the extent of both the belief in meritocracy across societies and actual differences across societies. What causes this variation?

While the belief in the extent to which meritocracy should play a role in the distribution of income within a society varies across individuals—partly as a function of political ideology ([Mitchell et al., 2003](#)) and context ([Scott and Bornstein, 2009](#))—most individuals in advanced industrial societies support the principle of merit as one criteria, among several, in decisions regarding distributive justice ([Kunovich and Slomczynski, 2007](#); [Scott et al., 2001](#)). Still, even among advanced industrialized market economies, support for meritocratic equity varies substantially ([Almås et al., 2020](#); [Reynolds and Xian, 2014](#); [Kunovich and Slomczynski, 2007](#); [Osberg and Smeeding, 2006](#)).

In natural settings, merit, or one's inputs, can be assessed in many different ways, as can one's outcomes/rewards. Moreover, in many such contexts, both inputs and outcomes are at least partly unobservable. Experimental economics is based on *induced value theory* ([Smith, 1976](#)), which requires that the material rewards in an experiment are salient—meaning that the mapping from participants' actions to the resulting payoffs is clear and transparent—and monotonic—meaning that, all else equal, participants prefer more of the reward to less. Because of these foundational features of the approach, it is well suited to investigate preferences for various ethical principles, including distributive fairness ([Fehr and Schmidt, 1999](#); [Bolton and Ockenfels, 2000](#)) and procedural fairness principles ([Dold and Khadjavi, 2017](#); [Bolton et al., 2005](#); [Hoffman et al., 1994](#)).

At the same time, evidence from experimental economics has also made it clear that people bring their culture into the lab with them. Therefore, although researchers are not able to fully control subjects' preferences, it also means that they can nonetheless infer something about cultural variation based on their behavior in the experiment ([Barr and Serra, 2010](#); [Smith, 2007](#); [Cárdenas and Ostrom, 2004](#)). This feature makes the paradigm especially useful for assessing preferences for meritocracy across societies: through careful calibration of the material incentives as a function of performance on some sort of task, researchers can, in a very transparent way, make the "societies" in their experiment more or less meritocratic and then directly observe responses to these manipulations. Although the tasks differ—and range from envelope-stuffing (e.g., [Konow, 2000](#)) to solving mathematical problems (e.g., [Sutter and Weck-Hannemann, 2003](#)), to word puzzles (e.g., [Burrows and Loomes, 1994](#)) to manipulating computerized slider bars (e.g., [Gill and Prowse, 2011](#)) to letter counting (e.g., [Galeotti et al., 2017](#))—they have been shown to have

effects on the beliefs and behaviors of participants in a way that seems consistent with preferences for meritocracy.

There is also an extensive literature that uses formal theory and survey data to study distributional preferences in a cross-national setting, and suggests that there are substantial within- and between-nation differences in the extent to which individuals perceive the income distribution in their society as being meritocratically equitable ([Fong, 2001](#); [Alesina and Angeletos, 2005a,b](#); [Alesina and La Ferrara, 2005](#)). This literature suggests that cross-national variation in these perceptions explains substantial variation in aggregate preferences for redistribution. In these models, self-fulfilling beliefs sustain multiple equilibria in the relationship between meritocratic equity and attitudes toward redistribution. [Alesina and Angeletos \(2005a\)](#) derive a model in which individuals support greater income redistribution when they perceive their society's income distribution to be determined by corruption, connections and fraud rather than merit. Increased redistribution, in turn, increases the size of government, and opportunities for corruption while reducing incentives for hard, honest work. In [Bénabou and Tirole \(2006\)](#), belief in meritocracy increases hard work, reinforcing the meritocracy of the income distribution and reducing demand for redistribution.

This interdisciplinary experimental literature makes clear that meritocratic equity is an important (though far from exclusive) determinant of distributive preferences, that it can be meaningfully operationalized, and that it can be successfully manipulated in the lab. Moreover, related cross-national survey-based research demonstrates that there is likely to be endogeneity between perceptions of meritocracy and certain behaviors and preferences. Here we take advantage of the causal inference afforded by controlled experiments, precluding such endogeneity, while at the same time exploiting the cross-cultural variation and generalizability that is associated with cross-national survey research. This complementary research design brings together the two related, but mostly separate, lines of research discussed in this section, leading to more robust inference and increased construct validity.

To study the effect of meritocracy violations on individual malfeasance, it is crucial to obtain reliable indicators of malfeasance. For cross-national survey data, we rely on several measures of acceptability for malfeasant activities (e.g., accepting bribes, claiming undeserved government benefits, avoiding public transport fares, and tax evasion) reported in the World Values Survey (WVS), which covers a large number of countries, including Italy and the United States—the two primary countries of focus in our research.

For the laboratory experiment, we draw on the extensive experimental literature in economics and psychology that uses simple tasks, such as asking subjects to report the outcome of a random draw, to measure lying behavior as a proxy for dishonesty (see, e.g., [Shalvi et al., 2011](#); [Fischbacher and Föllmi-Heusi, 2013](#); [Abeler et al., 2014](#); [Gneezy et al., 2018](#)). This literature shows that many individuals are reluctant to lie maximally, even if there is no detection mechanism or risk of being punished, suggesting a general aversion to lying. In the context of our research question, lying aversion is a potential enabler of the voluntary obedience necessary for legitimate authority. Related studies have investigated

how this aversion to lying is shaped by both internal factors, such as self and social image concerns (e.g., Bašić and Quercia, 2022), motivated forgetting (e.g., Galeotti et al., 2020) or loss aversion (Garbarino et al., 2019), and environmental factors, such as the nature and size of the incentives at stake (e.g., Kajackaite and Gneezy, 2017; Benistant et al., 2022), or information about others (e.g., Dimant et al., 2024), among others. Multiple studies have also demonstrated that lying behavior observed in these laboratory tasks correlates with various unethical activities in real-world settings (e.g., Potters and Stoop, 2016; Hanna and Wang, 2017; Cohn and Maréchal, 2018; Dai et al., 2018), providing evidence that these tasks are useful for investigating malfeasance.

Within the experimental literature on lying, several studies that link inequality and fairness concerns to lying are particularly relevant to this paper. A common feature of these studies is the manipulation of subjects' wealth levels in various ways before administering the lying task: as a function of luck (Gino and Pierce, 2009, 2010; Galeotti et al., 2017), subjectively (Gino and Pierce, 2010), according to decisions made in a previous game (Houser et al., 2012; Alempaki et al., 2019), or based on merit (Galeotti et al., 2017; Grundmann, 2020).² Our previous study (Galeotti et al., 2017) where we also explore the link between meritocracy and dishonesty. We extend their work by examining the generality and scope of this relationship across different countries and malfeasant behaviors, using a combination of survey and experimental evidence.

A theory of meritocracy and malfeasance

We adopt a simple formalization of meritocracy based on equity theory (Adams, 1963). According to this theory, an individual evaluates the equity of a given context by comparing the ratio of their own outcomes (or rewards) and inputs (or contributions) with the ratio of outcomes to inputs of another individual who serves as a reference point for interpersonal comparisons of equity. Unequal ratios cause the individual mental distress or tension, a phenomenon known in psychology as cognitive dissonance. This dissonance can be attenuated through a variety of mechanisms, including the modification of inputs and/or outcomes. In our setup, one way to resolve this dissonance is by rebalancing the outcomes through malfeasance.³ In Adams' theory, inputs are broadly construed, and consist of "the participant's

contributions to the exchange, which are seen as entitling him to rewards or costs." These rewards or costs in turn constitute the outcomes. Outcomes are typically defined in monetary terms, though they are often construed more broadly.

We carefully distinguish the type of inequity that we are using here from the more familiar concept of inequality. The equality/inequality dimension is exclusively concerned with the distribution of outcomes. However, the equity/inequity dimension takes into account both inputs and outcomes. A meritocratically equitable scenario is defined as one in which the ratio of inputs to outcomes between two individuals is equal. An unequal scenario, then, is not necessarily inequitable. For example, a more productive employee may be paid more than a less productive colleague. Assuming that the increased productivity is fairly remunerated—thus ensuring equality of the ratio of outcomes to inputs—then such a situation would be, under our classification, *unequal but equitable*. Likewise, scenarios in which outcomes are equal are not necessarily equitable. If two workers get the same salary, but one is twice as productive as the other, according to our logic this scenario is *equal but inequitable*, as the ratio of outcomes to inputs for the less productive worker would be twice that of the more productive one.

According to equity theory, an equitable scenario is one in which the ratios of outcomes to inputs for two persons are equal. For persons A and B we formally define it as:

$$O_A/I_A = O_B/I_B \quad (1)$$

where O_i and I_i are, respectively, the outcomes and inputs of individual $i = \{A, B\}$.

Conversely, inequity arises when these ratios are *not* equal, i.e., when $O_A/I_A > O_B/I_B$ or $O_A/I_A < O_B/I_B$. In our experiment, we manipulate the values of O_A and O_B across treatments by varying payments as a function of relative performance in a real effort task.

Assume that B is the individual who may take a malfeasant action to reduce her distress caused by inequity. We can normalize her output such that $O_B = 1$, and then solve for the value of O_A that equalizes the two ratios. We will denote the value that equalizes the equity ratios as O^* . Since we normalize $O_B = 1$, then $O^* = I_A/I_B$. Thus, when $O_A = O^*$ the scenario is equitable. Conversely, when O_A is not equal to O^* the scenario is inequitable. When $O_A > O^*$ the scenario is characterized by negative inequity for B (or positive inequity for A) and when $O_A < O^*$ the scenario is characterized by positive inequity for B (negative inequity for A). There are different combinations of inputs and outcomes that produce scenarios differentiated by both their inequity and their inequality. These combinations are summarized in Table 1.

We focus on scenarios where the input of Person B (the decision maker) is greater than or equal to the input of Person A—the first four scenarios. Why? Negative or disadvantageous inequity (which

initially acts ethically then feels justified in engaging in less ethical behaviors. Individuals who experience unfairness may be particularly susceptible to moral licensing (Bobek et al., 2024). In our context, an individual who works hard but perceives the system as unfair may be more likely to feel a "moral license" to lie or cheat. While our study was not designed to explore these alternative mechanisms, we acknowledge their relevance and suggest that they warrant further research.

² A related study is Birkelund and Cherry (2020) who manipulate whether subjects receive equal and unequal advantages in a real-effort task where cheating is possible.

³ Adams (1963) discusses cognitive dissonance as the principal mechanism through which inequity operates. However, there may be other explanations for the effect of inequity on malfeasance. For instance, experiencing inequity may diminish an agent's intrinsic motivations (see, e.g., Frey and Jegen, 2001, for a review on motivation crowding effects). Specifically, if an agent feels that the system has exploited them (e.g., by demanding high effort without fair compensation), they may be inclined to disregard any additional costly rules imposed by the system, even if this leads to increased malfeasance. Another possible mechanism is moral licensing (we thank a reviewer for suggesting this alternative). Moral licensing occurs when a person who

TABLE 1 Equitable and inequitable scenarios.

Scenario	Inputs	Outcomes	Inequity	Inequality
1	$I_B = I_A$	$O_A = O_B$	None	No
2	$I_B > I_A$	$O_A = O_B$	Negative	No
3	$I_B > I_A$	$O_A < O_B$	Either	Yes
4	$I_B \geq I_A$	$O_A > O_B$	Negative	Yes
5	$I_B < I_A$	$O_A > O_B$	Either	Yes
6	$I_B < I_A$	$O_A = O_B$	Positive	No
7	$I_B \leq I_A$	$O_A < O_B$	Positive	Yes

“Inequity” occurs when $O_A/I_A = O_B/I_B$, and in this table is characterized by the perspective of individual, *B*. “None” means that $O_A/I_A = O_B/I_B$. “Negative” inequity is disadvantageous to *B* because $O_A/I_A > O_B/I_B$, implying that *B*’s ratio of outcomes to inputs is lower than that of *A*. Positive inequity is advantageous to *B* because $O_A/I_A < O_B/I_B$, implying that *B*’s ratio of outcomes to inputs is higher than *A*’s. “Either” indicates that the inequalities between the outcomes and the inputs are in opposing directions, and therefore in these cases the type of inequity depends on the relative magnitudes of these inequalities. In Scenario 4 (as well as Scenario 7), we group together situations where I_B is strictly greater (smaller) than I_A with situations where I_B is equal to I_A . Within the equity theory framework, these theoretical predictions go in the same direction. “Inequality” is defined exclusively with respect to the distribution of outcomes between individuals *A* and *B*.

characterizes Scenarios 2, 4 and potentially 3, depending on the relative size of the inequality) has been found to have a stronger effect than advantageous inequity on behavior (Bloom, 1999; Fehr and Schmidt, 1999). Moreover, our theory of justified malfeasance only applies to cases of disadvantageous inequity. Scenario 1 (or any other in which $O_A/I_A = O_B/I_B$) where $O^* = O_A = 1$, is the only necessarily equitable scenario.

To formalize our theory about the relationship between meritocratic inequity and malfeasance, we model the degree to which an individual will justify malfeasance as a simple linear function of the extent to which the individual perceives the situation to be characterized by disadvantageous meritocratic inequity:

$$D_B = \alpha + \beta(O_A - O^*) \tag{2}$$

where D_B is malfeasance by person *B*. α captures an individual’s baseline tendency to commit malfeasance, regardless of any recent unfair treatment, while β reflects how an individual reacts specifically to experiencing unfair treatment.

This implies that the larger the unfavorable equity violation experienced by person *B* the more likely she is to engage in malfeasance. In other words, $O_A > O^*$ implies that that person *B* experiences meritocratic inequity with respect to person *A*. We re-state Equation (2) as our first hypothesis:

Hypothesis 1 *Violations of meritocratic equity will increase justified malfeasance at the individual level.*

The basic model in Equation (2) can capture the individual-level reaction from experiencing a violation of meritocratic equity. However, as previously discussed, there is considerable variation in the degree of meritocracy across countries, even among advanced industrial democracies. This country-level variation in meritocracy may influence the general propensity to engage in malfeasance (α), regardless of whether an individual has recently received

unfair treatment. Specifically, we expect that individuals living in countries where meritocratic principles are frequently violated may be conditioned to believe that dishonesty is a necessary strategy to get ahead in life. As a result, they may be more prone to commit malfeasance. Additionally, the degree of meritocracy in a country may affect how individuals respond to unfair treatment (β). In countries with lower levels of meritocracy, people are more accustomed to violations of these principles and, therefore, may react less strongly to a new instance of unfair treatment.⁴

We can capture these notions with a simple extension of the basic model:

$$D_B = \alpha_j + \beta_j(O_A - O^*) \tag{3}$$

where D_B , O_A^* , and O_A are defined as above, and the function’s parameters, α_j and β_j , are allowed now to vary across countries/societies, representing country-specific levels of justified malfeasance and its sensitivity to violations of meritocratic equity respectively.

We conjecture that the tendency toward malfeasance will also depend on α_j —which one can think of as the country-specific baseline level of malfeasance—and on β_j —a society-specific coefficient measuring the strength with which citizens respond to violations of meritocratic equity with malfeasance. In other words, it is the marginal effect of inequity on malfeasance. Citizens in an un-meritocratic society are confronted with a relatively high α_j and should generally not expect meritocratic treatment. This should in turn make them less sensitive to a given instance of meritocratic inequity, implying a lower β_j . We estimate these parameters for the United States and Italy using our experimental data. The advantage of this approach is that the data is derived from a tightly controlled and incentivized experiment, conducted in an identical manner across the two countries. Still, inferences from such data are limited because we cannot exogenously manipulate norms, culture or citizenship, and in any case we have only two country-level data points. To complement this limited cross-national comparison, we turn to WVS data, giving us data on 77 countries. We cannot lend any causal interpretation to the observational results directly, but we can determine whether and how α_j and β_j vary across countries.

We conjecture that meritocratic equity is correlated cross-nationally with malfeasance. But, aside from Kunovich and Slomczynski (2007)—who analyze just 14 countries—we are not aware of any cross-national measure of meritocratic inequity. Therefore, we cannot determine whether meritocratic inequity is cross-nationally associated with malfeasance across a large number of countries. Fortunately, however, two of the countries included in Kunovich and Slomczynski (2007) are the United States and Italy. Thus, we can compare our results to this benchmark, which indicates that the United States’ income distribution is in fact perceived as more meritocratic than Italy’s. With

⁴ While meritocratic inequity is central to these arguments, cultural, historical, and institutional factors may also contribute to cross-country differences. These factors can also influence the extent to which meritocracy is regarded as a desired ideal. In the current paper, we focus on meritocratic inequity and do not address the effects of these other factors, directing the reader to previous literature for accounts of these influences (see, e.g., Gächter and Schulz, 2016; Cohn et al., 2019).

this difference in mind, and to complement our experimental results, we use WVS data to estimate country-specific parameters for Italy and the U.S. Further, this provides some evidence regarding our conjectures about cross-national variation in the relationship between meritocracy and malfeasance. If the qualitative relationship between these parameters for the two countries are similar across our two methods of data collection, this enhances construct validity and lends more credibility to our findings. We elaborate these ideas with the following conjectures.

Conjecture 1 *We expect to observe cross-national variation in baseline rates of justified malfeasance. This variation should be associated with the degree of perceived meritocracy in the country, with individuals in countries characterized by more meritocratic inequity exhibiting higher rates of malfeasance. Specifically, we expect that $\alpha_{0,Italy} > \alpha_{0,US}$.*

Conjecture 2 *We expect to observe cross-national variation in the sensitivity of justifying malfeasance when encountering violations of meritocratic equity. This variation should be associated with the degree of perceived meritocracy in the country, with individuals in countries characterized by less meritocratic inequity exhibiting a greater sensitivity to violations of meritocratic inequity. Specifically, we expect that $\beta_{US} > \beta_{Italy}$.*

Together, Conjectures 1 and 2 imply that malfeasance is a concave function of exposure to violations of meritocracy, with diminishing effects on sensitivity as one encounters additional violations.

Experimental design

Experimental sessions were conducted at university laboratories in the Northeastern United States and Northern Italy, using a sample of undergraduate students at each site. The experimental sessions were conducted according to standard practices in experimental economics research, further details of which can be found in the [Appendix](#). The experiment is divided into two main stages. First, we use a real-effort task to create a binary categorization of subjects based on their relative performance on the task. This is followed by a custodial stage in which one subject in each pair is put in a privileged position in which she can misreport private information in such a way so as to benefit herself at the expense of their partner. Subjects were compensated for both the real-effort task and their decisions in the custodial stage. The identities of the partners remained anonymous throughout the experiment. We use a between-subjects design; each subject participated in exactly one treatment.

The real-effort task

We use a real-effort task which has been shown to “induce a sense of ownership” over the outputs in previous experimental studies (e.g., [Hoffman et al., 1994](#); [Banerjee, 2016](#)). This is important for our equity theory model of malfeasance. There have been various real-effort tasks employed in experimental economics, including mathematical tasks, quizzes and envelope-stuffing tasks. For a good review of the various tasks used in the literature, see [Gill and Prowse \(2011\)](#) and [Charness et al. \(2018\)](#).

We choose a “letter-counting” task where we require subjects to tally the occurrences of the letters “c” and “e” in each line of a 15-line text written in the German language (more details about the text are found in the [Appendix](#)). Advantages of this task include: it is easy for the subjects to understand; there is a fair amount of variation in performance; and it does not require specialized knowledge or skill. Most importantly, perhaps, is that it is tedious and mentally costly for the subjects, and therefore relative performance on it can plausibly serve as a legitimate merit-based distinction. Performance was measured by absolute difference from the true number of letters “c” and “e” in each line, with lower scores indicating higher performance. It was public knowledge that subjects were ranked according to a median split, with those above the median designated as “high performers” (HPs) and those below the median “low-performers” (LPs). Before the task, subjects are informed that they will be assigned to one of four scenarios after the task, and that these scenarios are differentiated by the relative payments (outcomes) of the HPs and LPs. The instructions state “in three of the four scenarios the payment to the high performers will be greater than or equal to the payment to the low performers. In one out of the four scenarios the payment to the high performers will be less than the payment to the low performers.” The design is incentive compatible because it pays to be a HP in a majority of the scenarios. In the language of equity theory, the task stage represents a social exchange in which the participants’ inputs—their (relative) performance on the task—are mapped into outcomes, reflecting our design’s foundation in our theoretical model.

The custodial stage

The second stage is called the “custodial” stage. In each treatment subjects are randomly paired as players A and B, and one is assigned to be the “owner” of an “escrow” account valued at \$15 (€10).⁵ B does not receive any additional income and plays the role of the “custodian” of the owner’s account. Only the subjects assigned the role of Player B, the custodian, make decisions in this stage, while subjects assigned the role of Player A, the owner, simply wait for their counterpart’s decisions.⁶ Which performance type is assigned to which role, as well as the payments associated with each role, differ by treatment (see next section). Each custodian receives a series of 25 binary signals (red or green), that are generated from a random process with a 50% chance for each color. The value of each signal is known to the custodian but not to the owner. The custodian must then report each signal as green or red. Subjects are instructed that the value recorded by the custodian determines the relative payment between the custodian and the owner. The subjects are told that a green signal means that an amount equal to $1/25$ of the escrow account (\$0.60/0.40€) will be transferred to herself or, in the case of a red signal, remain in the owner’s account. The design allows the custodians to increase or decrease their own

⁵ At the time these amounts were approximately equivalent in terms of purchasing power parity.

⁶ While waiting, we asked them to predict how many red signals their counterpart would report. This prediction was not incentivized.

payoff at the expense or benefit of their partner *by misreporting the signal*.

By virtue of their performance, *B* is given a privileged position, officially and explicitly entrusted as the Custodian of *A*'s account. Reporting a red signal as green increases their own payoff at the expense of the Owner of the account. Although we make it clear that they cannot otherwise be punished (or rewarded) for misreporting a signal, it is a direct violation of the rules that we gave the participants, and thus represents malfeasance.⁷

Treatments

Our experimental manipulations relate to the equity and the equality of the remuneration (the “outcomes”) of the first stage as a function of the relative performance (the “inputs”) of the participants in the task.⁸ Our 2×2 design yields four treatments, described below.

Treatment EE: equal and equitable

In EE, the first stage incomes of HPs and LPs are equal. It is also equitable: in the second stage subjects are randomly paired within performance types (i.e., HPs are paired with HPs and LPs are paired with LPs). In EE, the income distribution among partners is equal *and* equitable because both subjects within each pair are compensated with the same amount for an identical category of performance. To control for income effects we conduct two sub-treatments, in which both HPs and LPs receive \$10 (€7) in the *high income* sub-treatment or \$3 (€2) in the *low income* sub-treatment. In both of the income sub-treatments the payments are equal ($O_A = O_B$), and because the pairings are homogeneous with respect to performance, then the inputs must be equal ($I_A = I_B$). Therefore, $O_A/I_A = O_B/I_B$. This situation is equivalent to Scenario 1 in Table 1.

Treatment EI: equal and inequitable

HPs are (randomly) paired with LPs. The income distribution is equal, but meritocratically inequitable because HPs do not receive greater compensation for their performance than LPs. This situation is equivalent to Scenario 2 in Table 1, where $O_A = O_B$ and $I_B > I_A$. Because the outcome for each subject in the pair is equal ($O_A = O_B$), the scenario would only be equitable if $I_A = I_B$,

⁷ The custodial stage reflects tasks commonly used in the experimental literature on lying behavior to measure dishonesty, as previously mentioned. In designing this stage, we rely particularly on a protocol where lying is observable by the experimenter, following methods used in Galeotti et al. (2017), Kocher et al. (2018), Gneezy et al. (2018), Gneezy and Kajackaite (2020), and Hakimov and Kajackaite (2024), among others. This approach allows us to measure dishonesty at the individual level. While overall lying may be lower compared to tasks where lying is unobservable by the experimenter (Gneezy et al., 2018; Abeler et al., 2019; Fries et al., 2021), this should not affect the differences between treatments.

⁸ In our experiment, the only information about inputs that is available to subjects is the binary performance ranking. Given that the ranking is coarse, subjects might have exhibited unobserved heterogeneity in beliefs regarding the magnitude or ultimate source of any such performance differences.

TABLE 2 Experimental design: payments by treatment and performance rank.

Performance rank	Equitable		Inequitable	
	High	Low	High	Low
Equal*	\$10/€7	\$10/€7	\$10/€7	\$10/€7
Unequal	\$10/€7	\$3/€2	\$3/€2	\$10/€7

*For the sake of simplicity, the payments listed for the Equal conditions are those for the “high” income condition. In the “low” income conditions, the payments were \$3/€2. This controls for the absolute effects of income.

implying an identical valuation of both low and high performance. If one has even the slimmest belief in meritocratic equity then this is an inequitable scenario for *B* (the HP), because she receives compensation equivalent to her LP partner, despite the common knowledge of her superior performance. Just as in the EE treatment, to control for income effects, we conduct two sub-treatments.

Treatment UE: unequal and equitable

HPs receive \$10 (€7) after the first stage of the experiment, while LPs receive \$3 (€2). As in EI, each HP was randomly paired with a LP. Hence, the income distribution between partners could be viewed as equitable, depending on the relative valuations placed on high and low performance. More precisely, $O_A/O_B = 3/10$ (or $2/7$ in the Italian sessions), or, in other words low performance (I_A) must be discounted by a factor of 0.3 (0.29) relative to high performance (I_B) in order for the scenario to be equitable. This treatment is one parameterization of Scenario 3 in Table 1.

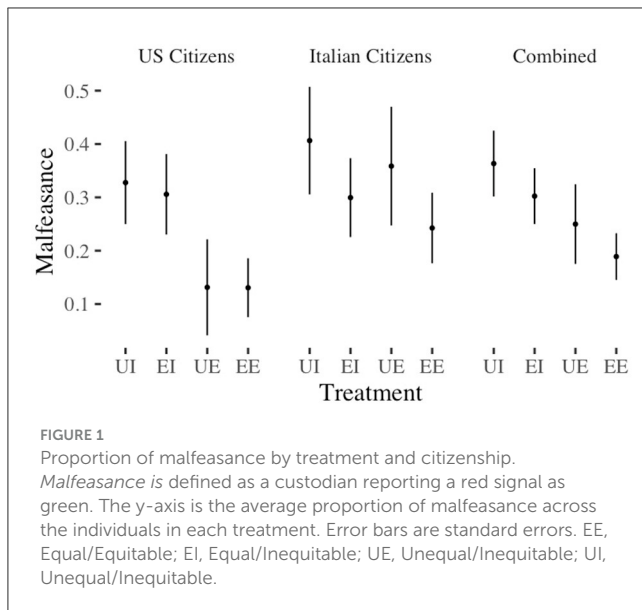
Treatment UI: unequal and inequitable

Here the HPs receive a compensation of \$3 (€2) after the first stage, while the LPs receive \$10 (€7). Hence, the income distribution between partners is both unequal and inequitable, since those who perform better in the real effort task are paid *less* than their partners who perform *worse*. In this treatment, $O_A/O_B = 10/3$ ($7/2$). In order for the scenario to be equitable, one must value low performance by a ratio of $10/3$ ($7/2$) compared to high performance. Thus, even if one valued low performance over high performance at a ratio of two-to-one, the scenario would still be inequitable toward the HP, leaving considerable room for justified malfeasance aimed at reducing the inequity. This treatment corresponds to what we refer as Scenario 4 in Table 1.

Our four between-subjects treatments are summarized in Table 2.

Experimental results

In the context of our experiment, each red signal reported as green represents a case in which a Custodian breaks the rules of the game in order to enrich him/herself at the expense of the Owner. This variable provides a normalized measure of the rate of malfeasance. Figure 1 displays the rates of malfeasance for each experimental treatment, separately displaying results from the U.S. and Italian sessions, as well as the combined results. There is no significant difference, both in aggregate and within each



sample pool, between the high- and low-income sub-treatments of either the EE or the EI treatments ($p > 0.1$; Mann-Whitney [MW] test), so we pool the data of the sub-treatments together. We also exclude the 15% of the U.S. based sample that are not U.S. citizens (all of the subjects in Italy are Italian citizens).⁹ Results with all participants (citizens and non-citizens) can be found in the [Appendix](#).¹⁰ Excluding the non-citizens increases the differences between the two samples (potentially bolstering our claims about cross-cultural differences) but does not change the qualitative nature of the results.

The behavior of the American participants in the inequitable treatments (EI and UI) were characterized by a much greater degree of malfeasance than the equitable treatments (UE and EE). On the other hand, Italians seem relatively unaffected by inequity, but somewhat sensitive to inequality. If we look at the mean rate of malfeasance across treatments of the combined sessions, the relative magnitude of these rates is exactly as predicted by the O^* parameter for each Treatment: $D(UI) = 0.36 > D(EI) = 0.30 > D(UE) = 0.25 > D(EE) = 0.19$. Generally, the larger the value of O^* , the greater amount of malfeasance we would expect to observe. The relationship between $D(UE)$ and $D(EE)$ might need further clarification. In EE, the performance ranking is homogeneous within pairs, and, therefore, it cannot be used as a basis of merit: from the standpoint of the rules and information available, the Custodian—a HP (LP)—cannot differentiate himself from his partner—another HP (LP). Thus, $O_A/I_A = O_B/I_B$ and any malfeasance should not be a result of feelings of inequity.

⁹ Non-citizens were excluded because their exposure to different cultural norms and experiences with meritocracy could introduce variability that may confound our cross-country comparison between the U.S. and Italy.

¹⁰ We failed to record citizenship data in the UI sessions conducted in the U.S. Citizenship in UI is assigned as the average of 1000 permutations of the expected number of citizens in UI based on the actual number in the other sessions. The [Appendix](#) contains additional information about this procedure.

On the other hand, in UE, even if HPs are compensated for their higher performance compared to their partner (LP), the difference may be small enough that they believe the scenario still to be inequitable, and therefore equity theory does allow for some degree of malfeasance in this case, depending on the beliefs of the Custodian as to *how much* higher performance should be differentially rewarded. We thus expect at least as much malfeasance in UE compared to EE.

Principally, we want to determine whether our manipulation of meritocratic equity (the “equity” factor) has an effect on malfeasance in our experiment. Using our factorial design, we can simultaneously test whether our manipulation of equality of the initial incomes has an independent effect on malfeasance. Given the binary nature of the decision, we use logistic regression. To reduce noise resulting from stochastically different numbers of red signals that a given subject observes, we also control for the total number of red signals each subject receives. We did not detect any statistically significant difference across treatments or countries in the number of red signals received by the Custodians ($p > 0.1$, MW test), so our results are unlikely to be explained by differences in the number of red signals—which, recall, are randomly generated by the computer for each Custodian. To get the most efficient parameter estimates out of our data given our data generating process, we employ a MLM framework, in this case with the 25 decisions of each Custodian (level 1) being nested within each individual Custodian (level 2). Though intuitively it would also seem to make sense to nest the individual Custodians within their country (Italy or the U.S.), a dichotomous variable cannot be employed as a grouping variable in a MLM framework, so we treat it as a simple dummy variable. [Table 3](#) displays results from three multilevel logistic regressions. All coefficients here are reported as odds ratios. The regressions include up to five terms (not including the intercept): *Inequity* is a dummy variable for the treatments equal to 1 for the inequitable treatments (EI, UI) and 0 for the equitable treatments (EE, UE); *Inequality* represents the second manipulated factor and is coded as 1 for the unequal treatments (UE, UI) and 0 for the equal treatments (EE, EI); *Red Signals* is the number of red signals received by each Custodian, and takes the values 0–25; *U.S.* is a dummy variable taking value 1 for a U.S.-based Custodian and 0 for an Italy-based Custodian; and *U.S. × Inequity* is a multiplicative interaction term interaction the variables U.S. and Inequity. In column (1) of [Table 3](#), just as in column (1) of [Table 4](#) we can investigate our basic theoretical model, $D_B = O_A - O^*$. The results, which report behavior aggregated across both the U.S. and Italy, show that a Custodian in an inequitable treatment was 2.37 times more likely to be malfeasant (i.e., report a red signal as green) than those in the equitable treatments. This effect is significant at 5% level ($p = 0.017$) whereas inequality has no significant effect on malfeasance ($p = 0.230$). Also, as in all of our specifications, the number of red signals received significantly increases a Custodian’s odds of being malfeasant. A possible interpretation of this finding is that individuals who experience a higher number of red signals may perceive this as unfair, given that the signals are generated with equal probability for each color. This perceived unfairness could motivate them to misreport more frequently. In our experiment, inequity significantly increases malfeasance, providing us with evidence consistent with Hypothesis 1. At the

TABLE 3 Mixed-effects logistic regressions: determinants of malfeasance.

	(1) Baseline	(2) US dummy	(3) US × inequity	(4) US × inequality
Inequity	2.37	2.44	1.38	1.37
<i>p</i> - value	0.017	0.013	0.500	0.501
Inequality	1.57	1.59	1.53	1.89
<i>p</i> - value	0.230	0.212	0.245	0.199
Red Signals	1.26	1.26	1.26	1.25
<i>p</i> - value	0.004	0.003	0.003	0.004
U.S.	–	0.50	0.25	0.29
<i>p</i> - value	–	0.051	0.009	0.036
U.S. × Inequity	–	–	3.54	3.63
<i>p</i> - value	–	–	0.078	0.072
U.S. × Inequality	–	–	–	0.635
<i>p</i> - value	–	–	–	0.539
Intercept	0.002	0.002	0.003	0.003
<i>p</i> - value	0.000	0.000	0.000	0.000
<i>N</i> (# obs)	3,488	3,488	3,488	3,488
<i>M</i> (# Custodians)	140	140	140	140

All models were estimated using Stata's `melogit` command. All coefficients reported as odds ratios. Numbers in bold relate to our theoretical quantities of interest. *Variables*: DV: Malfeasance (reporting red signal as green). *Inequity* refers to the two “inequitable” treatments, EI and UI. *Inequality* refers to the two “unequal” conditions, UI and UE; *Red signals* is the total number of red signals received by the individual; *U.S.* is a dummy for whether the individual is a U.S. citizen; *USxInequity* and *USxInequality* are the multiplicative interactions of the those sets of variables respectively.

same time, our results do not indicate any significant effect of our inequality manipulation.

In column (2), we introduce a country dummy into the specification in (1), allowing us to estimate the α parameter of our theoretical model. The results show that U.S. Custodians are two times less likely to engage in malfeasance, and the effect of inequity remains essentially unchanged from (1). These results are consistent with Conjecture 1.

In column (3), we add a multiplicative interaction term (*U.S. × Inequity*) to our specification in (2). This term represents the β term in our theoretical model and allows us to determine whether American Custodians are more sensitive to inequity than their Italian counterparts, as suggested in Conjecture 2. It is generally a challenge to interpret the estimated coefficients for such terms when using non-linear models such as logistic regression. Some, like Norton et al. (2004), have developed tools for more sophisticated analysis of these estimates. This is even more of a challenge when dealing with multilevel non-linear models, such as our multilevel logistic regressions (Ai and Norton, 2003). In light of these challenges, we report all estimated coefficients as odds ratios, and to avoid confusion, do not report marginal effects for any of the estimated coefficients. With those limitations in mind, we find suggestive support for Conjecture 2 with the coefficient demonstrating that, compared to Italians, Americans are 3.54 times more likely ($p = 0.078$) to report a red signal as green

in the inequitable treatments. In this estimation the main effect of inequity on malfeasance is substantially weakened ($p = 0.50$). Taken together, these results imply that to the extent that the effect of the inequity manipulations is being driven by the American participants, and once nationality is accounted for, the significance of the main effect fades. All of these results are consistent with the graphical display of this data in Figure 1.

What do our experimental results say about the roles that meritocratic inequity and inequality respectively play in driving malfeasance? Violations of meritocracy are an important factor in the behavior of American subjects, but less so in Italy. On the other hand, there is some indication that inequality was more important in driving the decision of the Italian subjects. Thus, though overall all Italians exhibit a higher level of malfeasance than Americans, the American subjects are relatively more likely to respond with malfeasance to violations of inequity than inequality, while the reverse seems true for the Italians. To test this idea, we added another interaction term (*U.S. × Inequality*) to our model with the results in column (4). This coefficient has a *p*-value of 0.54, and the rest of the results remain nearly identical. Overall, we see stronger evidence in favor of conjecture 1—that baseline rates of justified malfeasance (a) vary across the two countries—than for conjecture 2—that the sensitivity of justifying malfeasance (b) varies across the two countries. We don't find any evidence that inequality has any systematic effect on justified malfeasance in our experiment. Of course, any similarities or differences between our Italian sample and our American sample we find in our experimental data might be due to the particular samples rather than any true country-level features. To complement these results and search for convergent evidence, we now turn to cross-national survey evidence to see if we observe similar patterns.

Cross-national survey evidence

In our experiments, subjects respond to violations of meritocratic equity with justified malfeasance. However, this tendency appears stronger among American citizens than Italians. We now turn to a second source of evidence—data from the World Values Survey—to fully leverage the cross-national variation to obtain an efficient estimate of the association between variables at an individual level (Gelman and Hill, 2007). Moreover, the use of MLM allows us to estimate country-specific parameters while simultaneously controlling for country-level and individual-level factors. To do so, we employ both a random-intercepts only model and a random-intercepts with random-slopes model.

Like Bénabou and Tirole (2006); Alesina and Glaeser (2004); Alesina and Angeletos (2005b); Alesina and La Ferrara (2005) we capture perceptions of meritocratic equity with a WVS item that asks subjects to express their agreement with the statement “hard work brings success” on a scale from 1–10, in which [1] represents close agreement (“Hard work usually brings a better life.”), and [10] represents strong disagreement (“Hard work doesn't generally bring success—it's more a matter of luck and connections.”). Because higher scores indicate a stronger perception of meritocratic inequity, we call this variable *Inequity*, and use it as our chief explanatory variable. We use four outcome measures which for four distinct malfeasant activities, asks respondents to rate, on

TABLE 4 WVS data: effect of meritocratic inequity on four measures of justifying malfeasance, mixed-effects regressions.

Outcome measure		(1) <i>Baseline</i>	(2) <i>Intercept only</i>	(3) <i>Intercept and slope</i>
Bribe-taking	coef. (s.e.)	0.062 (0.002)	0.064 (0.002)	0.058 (0.007)
	<i>p</i> -value	0.000	0.000	0.000
Benefits fraud	coef. (s.e.)	0.046 (0.004)	0.045 (0.004)	0.047 (0.010)
	<i>p</i> -value	0.000	0.000	0.000
Avoiding transit fare	coef. (s.e.)	0.080 (0.004)	0.075 (0.004)	0.073 (0.008)
	<i>p</i> -value	0.000	0.000	0.000
Cheating on taxes	coef. (s.e.)	0.062 (0.004)	0.064 (0.004)	0.063 (0.009)
	<i>p</i> -value	0.000	0.000	0.000

All outcomes measured in terms of the activities “justifiability” (10 = “always justifiable”; 1 = “never justifiable”). All models were estimated using Stata’s xtmixed command.

a scale of 1–10, whether the activity is “never justifiable” [1] or “always justifiable” [10]. Here we use four distinct but conceptually related outcome measures from the WVS that are available for a large number of countries, including Italy and the United States. In particular, these survey items ask to what extent “accepting a bribe in the course of one’s duties;” “claiming undeserved government benefits;” “avoiding a fare on public transport;” and “cheating on taxes if you have a chance” are justifiable activities. Using four distinct outcome measures lends greater construct validity to our study and enhances its complementarity with our experimental results.

We include three other relevant survey items (individual level variables) as controls. First we include a variable (*Inequality*) which measures respondents’ views on income inequality by asking on a range of 1 to 10, whether “income should be made more equal” [1] or “we need larger income differences as incentives” [10]. Second is a measure of trust (*Trust*) which takes values of [1], corresponding to “Most people can be trusted” and [2] corresponding to “Can’t be too careful.” The third gives a self-reported measure of the respondent’s household income (*Income*), measured in ten brackets (but these brackets are not deciles).

For country-level predictors, we include GDP per capita in purchasing power parity terms (*GDP*) and a measure of the country’s Gini coefficient (*Gini*—a number from 0–100 with higher levels representing greater inequality) as a measure of income inequality, both taken from World Bank, and a measure of corruption at the country level taken from Transparency International’s Corruption Perceptions Index (*Corruption*), taking values from 0–10, with higher values indicating *less* pervasive corruption. Given differences in the CPI scale, we had to make some adjustments. More details are in the [Appendix](#). Because this measure begins only in 1995, our sampling begins then and extends until 2015. The years are pooled in order to maximize the number of countries we can include. In addition to the limited availability of the country level data, the inequity variable is only intermittently available. Here we include WVS data from waves 1–6, encompassing the years 1981–2014. See [Appendix](#) for further details about the MLM specification. Based on the definition of malfeasance we have adopted here, corruption as measured by this index is just a subset of malfeasance (albeit a particularly salient one).

In our multilevel model, there are two levels of analysis: individual respondent (level 1) and the country in which the respondent resides (level 2).

We estimate the following mixed model:

$$\begin{aligned} \text{Malfeasance}_{ij} = & \alpha_{00} + \beta_{10} * \text{inequity}_{ij} + \beta_{20} * \text{inequality}_{ij} \\ & + \beta_{30} * \text{trust}_{ij} + \beta_{40} * \text{income}_{ij} + \alpha_{01} * \text{gini}_{ij} + \alpha_{02} * \text{corruption}_{ij} \\ & + \alpha_{03} * \text{gdp}_j + \delta_{0j} + c_{ij} \end{aligned} \quad (4)$$

with the random intercept for country j estimated by the following equation:

$$\alpha_{0j} = \alpha_{00} + \alpha_{01} * \text{gini}_{ij} + \alpha_{02} * \text{cpi}_j + \alpha_{03} * \text{gdp}_j + \delta_{0j} \quad (5)$$

and the random slope for country j on the Inequity variable estimated by:

$$\beta_{1j} = \beta_{10} + \delta_{1j} \quad (6)$$

[Table 4](#) displays the results of three MLM specifications: column (1) contains the baseline results for an estimation without random slopes or intercepts ([Equation 4](#)); column (2) has the results from a random intercepts specification (i.e. it embeds [Equation 5](#) in [Equation 4](#)); and column (3) is a specification with both random intercepts and random slopes (which embeds [Equations 5, 6](#) into [Equation 4](#)). [Table 4](#) shows only the results for our variable of interest—inequity. The full results of regressions for each of the four outcome measures can be found in the [Appendix](#).

Note the consistency: The results across all four specifications are consistent with Hypothesis 1. At the individual level the correlational effects of inequity on justified malfeasance are in all cases significant even with the inclusion of the included individual and country-level controls. Together with the experimental results presented above, the results in [Table 4](#) provide convergent evidence that meritocratic inequity is associated with an increase in justified malfeasance across a variety of measures. These aggregate effects, however, give us little insight into cross national variation. In particular we want to know whether any such variation exhibits the same qualitative relationships that we observed in experimental data across Italy and the US.

To test our two conjectures, we investigate cross-national variation in the α_{0j} and β_{1j} parameters in our theoretical model. The

results of the likelihood ratio test indicate cross-national variation sufficient to justify the use of MLM. α_{0j} and β_{1j} are, respectively, the country-level random intercepts and random slopes parameters in the MLM for our cross-national sample. Column (2) of Table 4 displays results from a random-intercepts MLM model, allowing us to more directly investigate the α_{0j} term in our theoretical model. Likewise, column (3) displays the estimates from a model which combines random-intercepts (α_{0j}) with country-specific random slope (β_{1j}) parameters for the effect of the inequity variable on all four measures of malfeasance. The full results of the specifications are available in the Appendix. Recall that in our experimental data, the Italian participants had higher baseline rates of malfeasance than the American subjects overall ($\alpha_{0,Italy} > \alpha_{0,US}$) but the American subjects were more sensitive to the effects of inequity ($\beta_{1,US} > \beta_{1,Italy}$).

With the descriptive statistics, we observe that Italians are more likely than Americans to perceive their country as exhibiting an inequitable income distribution. For the *inequity* variable, the average response is 5.11 (s.d. = 2.75) in the Italian sample compared to 3.65 (s.d. = 2.41) in the U.S. sample. These descriptive statistics are also consistent with the data reported by Kunovich and Slomczynski (2007) who measure observed levels of meritocracy in the income distribution and find that the U.S. has higher levels than Italy. According to our theory, then, Italians should exhibit a greater tendency to justify malfeasance. To determine whether that is the case, we directly examine the country-specific α_{0j} and β_{1j} estimates corresponding to column (3) of Table 4 and then compare these estimates to the analogous estimates from our experimental data above.

We are unaware of a canonical way to formally test for differences between two random effects parameters estimated by a MLM. In the absence of any consensus method for such comparisons, we employ two simple empirical strategies for assessing differences in country specific random coefficients. The results of the two methods are virtually identical. First, we employ simple Wald tests. We tested for a difference by subtracting the Italy coefficient from the US one. Our conjectures would therefore imply that the differences in the slopes should be negative and the differences in the intercepts should be positive. The results of each of the eight tests are displayed in Table 5. Each of the differences is in the hypothesized direction. All but two (β_1 for *fare* and *benefits*) are significant at the 95% level, and only one other is not significant at the 99% level. Overall, these results are consistent with our conjectures, and with the experimental data.

Second, to offer additional evidence in favor of our conjectures, we take a graphical approach. Figure 2 displays the estimated country-specific random intercepts for both the U.S. ($\alpha_{0,US}$) and Italy ($\alpha_{0,Italy}$) for each of our four outcome measures, along with their 83% confidence intervals. Austin and Hux (2002) and Cumming and Finch (2005) show that the existence of overlap between 83% confidence intervals is approximately equivalent to a two-tailed test with $\alpha = 0.05$ (For comparison purposes we also include the global intercept parameter, α_0). Again our analysis yields results that are consistent across five separate outcome measures. As in the experimental data, for each of our four WVS outcome measures, the baseline rate (the intercept) is significantly larger for the Italian sample—where

TABLE 5 Wald tests for differences between country level random effects coefficients, United States and Italy.

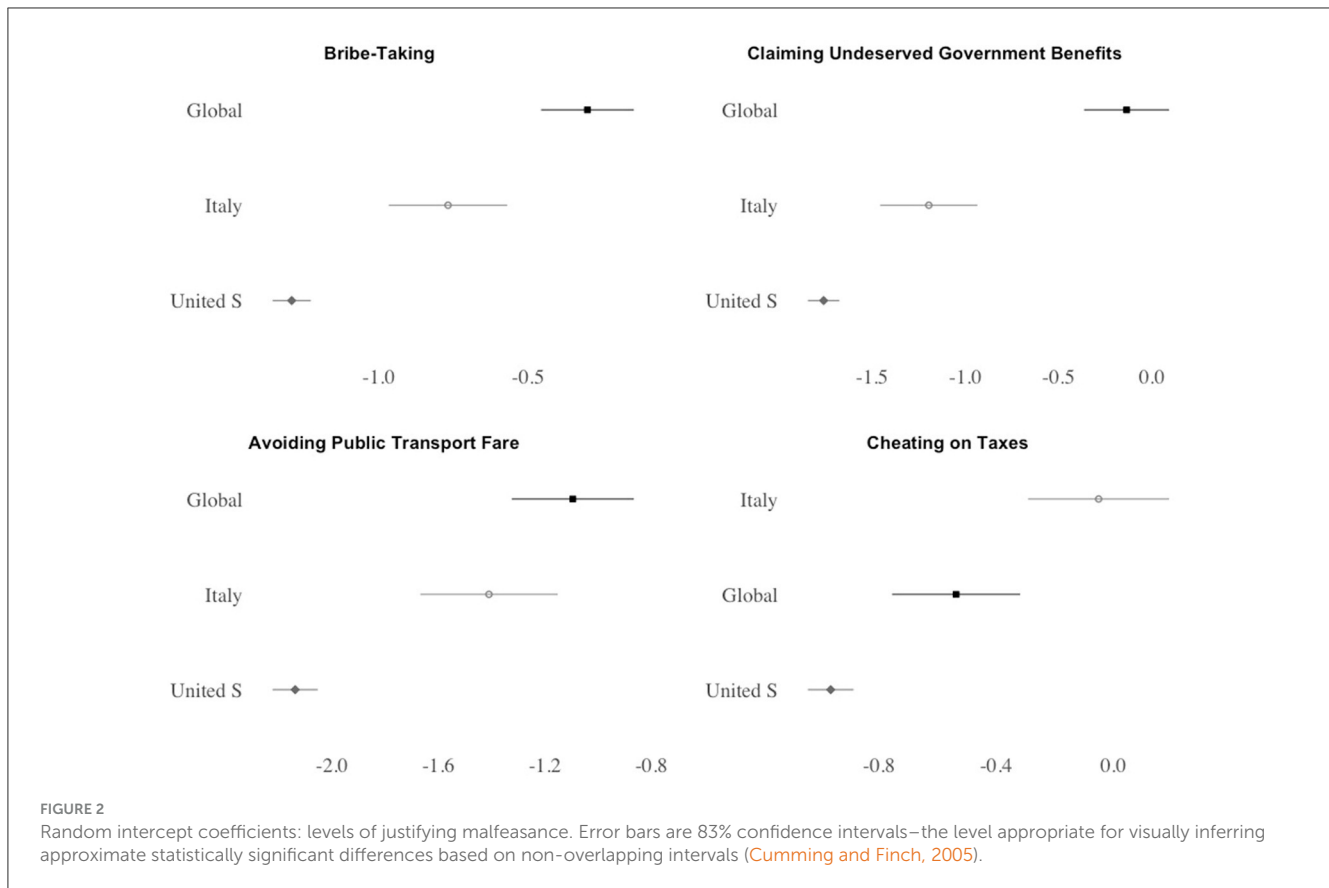
Outcome measure	Coefficient	Difference	Wald-z	p-value
Bribe-taking	Slope	-0.08	-2.89	0.004
	Intercept	0.52	3.44	0.0006
Benefits fraud	Slope	-0.06	-1.79	0.073
	Intercept	0.56	2.82	0.005
Avoiding transit fare	Slope	-0.07	-1.95	0.051
	Intercept	0.73	3.67	0.0002
Cheating on taxes	Slope	-0.07	-2.30	0.021
	Intercept	0.91	4.95	0.0000007

All differences calculated as Italy - U.S. All outcomes measured in terms of the activities "justifiability" (10 = "always justifiable"; 1 = "never justifiable").

perceptions of meritocratic inequity are also greater—than the American one ($\alpha_{0,Italy} > \alpha_{0,US}$). These results are consistent with Conjecture 1, the experimental data, and the results from the Wald tests. The Appendix contains more information about cross-national variation in α_{0j} and a figure displaying the estimated parameters for each of the 77 countries in our sample.

Now we turn to estimates of the country specific slope parameter, or β_{1j} . Our theory predicts that Americans, because their society is characterized by a greater degree of meritocracy, will be more sensitive to violations of it, and therefore, at the margin, respond with a greater degree of justified malfeasance when exposed to a meritocratic equity violation. In other words, we expect that $\beta_{1,US} > \beta_{1,Italy}$, and indeed that is what we observe in the experimental data. Figure 3 shows the estimates and 83% confidence intervals for the U.S. ($\beta_{1,US}$), Italy ($\beta_{1,Italy}$) and for the global slope parameter (β_1). In this case we again find consistency: across all four WVS outcome measures, and the experimental data, $\beta_{1,US} > \beta_{1,Italy}$, though there is considerably greater variation in the Italian sample. Country differences estimated for the benefits and the fare outcome variables are only slightly overlapping. This indicates, as was the case with the Wald tests, that those differences fall just slightly above the 0.05 cutoff.

To summarize, we find consistent and robust evidence to support our chief claim that violations of meritocratic equity are associated with malfeasance (Hypothesis 1). Moreover, we observe consistent cross-national variation in both $\alpha_{0,j}$ and β_{1j} . Across all five measures of malfeasance we use $\alpha_{0,Italy} > \alpha_{0,US}$ and $\beta_{1,US} > \beta_{1,Italy}$. Violations of meritocracy do seem to be associated with malfeasance. Again we find evidence that the baseline tendency and the sensitivity of this effect varies across countries in a way that is consistent with our theorized conjectures, though again this evidence is stronger in the case of $\alpha_{0,j}$ compared to differences in β_{1j} . Still, there is clearly a lot of cross-national variation that is not obviously explained by our theory, and further study is necessary to take other factors into account.



Discussion

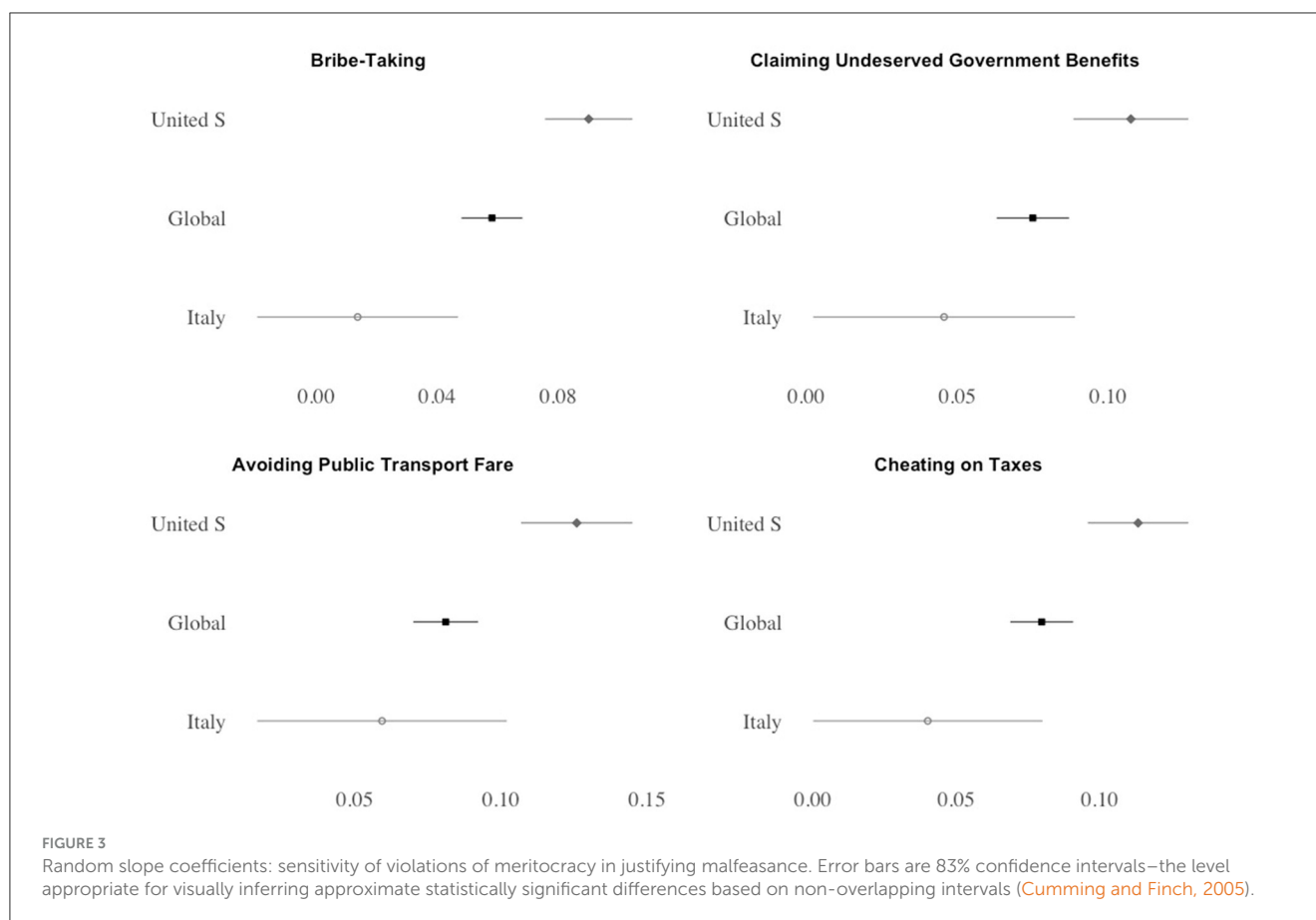
In this study we investigate the role that meritocratic inequity plays in determining levels of malfeasance across individuals and societies. Voluntary compliance is a necessary and crucial aspect of an authority's legitimacy. Income distributions that are perceived to be meritocratically inequitable undermine legitimacy, causing individuals to engage in rule-breaking—what we call justified malfeasance—as a corrective to this inequity. Such a dynamic, if widespread, is detrimental because it undermines civic-mindedness, trust, the rule of law, and the provision of public goods.

To formalize our theory, we create a simple model based on equity theory (Adams, 1963). We derive several implications from our model. First, individuals, when confronted with meritocratic inequity, will tend to respond with malfeasance. Moreover, any such responses and the baseline rates of malfeasance will vary across countries. We test these hypotheses using two datasets: data from an incentivized experiment conducted in the United States and Italy, and an MLM model using WVS data from 77 countries in which we estimate country-specific slopes (β_j) and intercepts (α_j) for the US and Italy. Our findings are consistent across our two data sources: we find that meritocratic inequity increases malfeasance. Results from both datasets show that there is significant variation in baseline tendencies toward malfeasance (α_j), and, in particular, $\alpha_{Italy} > \alpha_{US}$ indicating that Italian subjects and respondents are more likely overall to engage in malfeasance in the experiment and justify malfeasance in the survey.

We also expect that there will be variation in β_j . In both sets of data, we find that Americans are more sensitive to violations of meritocratic equity, i.e., that $\beta_{US} > \beta_{Italy}$. Though the evidence in favor of differences in sensitivity is overall weaker than that for the differences in the baseline rates. Still, these results are robust not only to the mode of data collection, but using four different items from the WVS as dependent variables. Our convergent results reveal a strong and consistent relationship that is robust to different conceptualizations of malfeasance. In other words, our study exhibits a high degree of construct validity.

The robustness of our results lends credence to our principal claim—that meritocratic equity violations lead to greater malfeasance. If true, this suggests many additional avenues of investigation. Given the importance of legitimacy in fostering trust, the rule of law, and the provision of public goods, combined with the capacity of malfeasance to undermine them, institutions that are widely perceived to be meritocratic are likely to be an important factor in political development and stability.

In sum, we find that Italians are more likely to engage in malfeasance overall, while simultaneously being less likely to engage in malfeasance at the margin, i.e., all else equal, they are less likely to respond to a particular instance of meritocratic violation. Kunovich and Slomczynski (2007) provide evidence that the levels of both actual and perceived meritocracy are lower in Italy than the United States. Based on this evidence and informed by our theory, we offer a plausible, though speculative, explanation for these robust but seemingly contradictory findings. Given that Italians experience less meritocracy in their daily lives, they are in a sense “pre-treated” with meritocratic inequity, and



therefore more likely than Americans, all else equal, to engage in malfeasance as a corrective mechanism to the more widespread inequity they are accustomed to experiencing. At the same time, because of this experience, they generally do not *expect* meritocratic treatment, making them less sensitive to a specific violation of meritocratic equity and therefore unlikely to increase their baseline tendency toward malfeasance as a result of a given violation. Another potential explanation could be that Italians have a shared understanding of low expectations in terms of effort put into work (Gambetta and Origgi, 2013). As a result, they may feel that meritocracy is not particularly relevant if the expectation is that compatriots will most likely shirk (though causality here would most likely run in both directions).

Though our study offers insight into the behavioral drivers of malfeasance, more research needs to be done in order to explain cross-national variation in malfeasance, to uncover the mechanisms behind this relationship, and more generally understand how normative factors shape behavior and in turn affect economic and political development.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Stony Brook University Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

RK: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. FG: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. RO: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the Einaudi Institute for Economics and Finance, the Russell Sage Foundation's small grants in behavioral economics, and the Project Decision of the French National Research Agency (ANR-19-CE26-0019).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Abeler, J., Becker, A., and Falk, A. (2014). Representative evidence on lying costs. *J. Public Econ.* 113, 96–104. doi: 10.1016/j.jpubeco.2014.01.005
- Abeler, J., Nosenzo, D., and Raymond, C. (2019). Preferences for truth-telling. *Econometrica* 87, 1115–1153. doi: 10.3982/ECTA14673
- Adams, J. S. (1963). Towards an understanding of inequity. *J. Abnormal Soc. Psychol.* 67:422. doi: 10.1037/h0040968
- Ai, C., and Norton, E. C. (2003). Interaction terms in logit and probit models. *Econ. Lett.* 80, 123–129. doi: 10.1016/S0165-1765(03)00032-6
- Akerlof, R. (2017). The importance of legitimacy. *World Bank Econ. Rev.* 30, S157–S165. doi: 10.1093/wber/lhw009
- Alempaki, D., Dogan, G., and Saccardo, S. (2019). Deception and reciprocity. *Exper. Econ.* 22, 980–1001. doi: 10.1007/s10683-018-09599-3
- Alesina, A., and Angeletos, G. -M. (2005a). Corruption, inequality, and fairness. *J. Monet. Econ.* 52, 1227–1244. doi: 10.1016/j.jmoneco.2005.05.003
- Alesina, A., and Angeletos, G. -M. (2005b). Fairness and redistribution. *Am. Econ. Rev.* 95, 960–980. doi: 10.1257/0002828054825655
- Alesina, A., and Glaeser, E. (2004). *Fighting Poverty in the US and Europe: A World of Difference*. Oxford: Oxford University Press. doi: 10.1093/0199267669.001.0001
- Alesina, A., and La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. *J. Public Econ.* 89, 897–931. doi: 10.1016/j.jpubeco.2004.05.009
- Almás, I., Cappelen, A. W., Sørensen, E. Ø., and Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science* 328, 1176–1178. doi: 10.1126/science.1187300
- Almás, I., Cappelen, A. W., and Tungodden, B. (2020). Cutthroat capitalism versus cuddly socialism: are Americans more meritocratic and efficiency-seeking than Scandinavians? *J. Polit. Econ.* 128, 1753–1788. doi: 10.1086/705551
- Austin, P. C., and Hux, J. E. (2002). A brief note on overlapping confidence intervals. *J. Vasc. Surg.* 36, 194–195. doi: 10.1067/mva.2002.125015
- Banerjee, R. (2016). On the interpretation of bribery in a laboratory corruption game: moral frames and social norms. *Exper. Econ.* 19, 240–267. doi: 10.1007/s10683-015-9436-1
- Barr, A., and Serra, D. (2010). Corruption and culture: an experimental analysis. *J. Public Econ.* 94, 862–869. doi: 10.1016/j.jpubeco.2010.07.006
- Bašić, Z., and Quercia, S. (2022). The influence of self and social image concerns on lying. *Games Econ. Behav.* 133, 162–169. doi: 10.1016/j.geb.2022.02.006
- Bénabou, R., and Tirole, J. (2006). Belief in a just world and redistributive politics. *Q. J. Econ.* 121, 699–746. doi: 10.1162/qjec.2006.121.2.699
- Benistant, J., Galeotti, F., and Villeval, M. C. (2022). Competition, information, and the erosion of morals. *J. Econ. Behav. Organ.* 204, 148–163. doi: 10.1016/j.jebo.2022.10.008
- Birkelund, J., and Cherry, T. L. (2020). Institutional inequality and individual preferences for honesty and generosity. *J. Econ. Behav. Organ.* 170, 355–361. doi: 10.1016/j.jebo.2019.12.014
- Bloom, M. (1999). The performance effects of pay dispersion on individuals and organizations. *Acad. Manag. J.* 42, 25–40. doi: 10.2307/256872
- Bobek, D., Hageman, A., and Hausserman, C. (2024). Are fairness perceptions related to moral licensing behavior? Evidence from tax compliance. *J. Bus. Ethics* 1–23. doi: 10.1007/s10551-024-05784-y
- Bolton, G. E., Brandts, J., and Ockenfels, A. (2005). Fair procedures: evidence from games involving lotteries. *Econ. J.* 115, 1054–1076. doi: 10.1111/j.1468-0297.2005.01032.x
- Bolton, G. E., and Ockenfels, A. (2000). ERC: a theory of equity, reciprocity, and competition. *Am. Econ. Rev.* 90, 166–193. doi: 10.1257/aer.90.1.166
- Burrows, P., and Loomes, G. (1994). “The impact of fairness on bargaining behaviour,” in *Experimental Economics*, ed. J. D. Hey (Heidelberg: Physica), 201–221. doi: 10.1007/BF01175872
- Cárdenas, J.-C., and Ostrom, E. (2004). What do people bring into the game? Experiments in the field about cooperation in the commons. *Agric. Syst.* 82, 307–326. doi: 10.1016/j.agry.2004.07.008
- Charness, G., Gneezy, U., and Henderson, A. (2018). Experimental methods: measuring effort in economics experiments. *J. Econ. Behav. Organ.* 149, 74–87. doi: 10.1016/j.jebo.2018.02.024
- Cohn, A., and Maréchal, M. A. (2018). Laboratory measure of cheating predicts school misconduct. *Econ. J.* 128, 2743–2754. doi: 10.1111/eoj.12572
- Cohn, A., Maréchal, M. A., Tannenbaum, D., and Zünd, C. L. (2019). Civic honesty around the globe. *Science* 365, 70–73. doi: 10.1126/science.aau8712
- Cumming, G., and Finch, S. (2005). Inference by eye: confidence intervals and how to read pictures of data. *Am. Psychol.* 60:170. doi: 10.1037/0003-066X.60.2.170
- Dai, Z., Galeotti, F., and Villeval, M. C. (2018). Cheating in the lab predicts fraud in the field: an experiment in public transportation. *Manage. Sci.* 64, 1081–1100. doi: 10.1287/mnsc.2016.2616
- Dickson, E. S., Gordon, S. C., and Huber, G. A. (2009). Enforcement and compliance in an uncertain world: an experimental investigation. *J. Polit.* 71, 1357–1378. doi: 10.1017/S0022381609990235
- Dickson, E. S., Gordon, S. C., and Huber, G. A. (2015). Institutional sources of legitimate authority: an experimental investigation. *Am. J. Pol. Sci.* 59, 109–127. doi: 10.1111/ajps.12139
- Dimant, E., Galeotti, F., and Villeval, M. C. (2024). Motivated information acquisition and social norm formation. *Eur. Econ. Rev.* 167:104778. doi: 10.1016/j.eurocorev.2024.104778
- Dold, M., and Khadjavi, M. (2017). Jumping the queue: An experiment on procedural preferences. *Games Econ. Behav.* 102:127–137. doi: 10.1016/j.geb.2016.12.002
- Duru-Bellat, M., and Tenret, E. (2012). Who's for meritocracy? Individual and contextual variations in the faith. *Compar. Educ. Rev.* 56, 223–247. doi: 10.1086/661290
- Fehr, E., and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Q. J. Econ.* 114, 817–868. doi: 10.1162/003353599556151
- Fischbacher, U., and Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *J. Eur. Econ. Assoc.* 11, 525–547. doi: 10.1111/jeea.12014
- Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *J. Public Econ.* 82, 225–246. doi: 10.1016/S0047-2727(00)00141-9

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frbhe.2025.1492421/full#supplementary-material>

- Frey, B. S., Benz, M., and Stutzer, A. (2004). Introducing procedural utility: not only what, but also how matters. *J. Instit. Theor. Econ.* 160, 377–401. doi: 10.1628/0932456041960560
- Frey, B. S., and Jegen, R. (2001). Motivation crowding theory. *J. Econ. Surv.* 15, 589–611. doi: 10.1111/1467-6419.00150
- Fries, T., Gneezy, U., Kajackaite, A., and Parra, D. (2021). Observability and lying. *J. Econ. Behav. Organ.* 189, 132–149. doi: 10.1016/j.jebo.2021.06.038
- Gächter, S., and Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature* 531, 496–499. doi: 10.1038/nature17160
- Galeotti, F., Kline, R., and Orsini, R. (2017). When foul play seems fair: exploring the link between just deserts and honesty. *J. Econ. Behav. Organ.* 142, 451–467. doi: 10.1016/j.jebo.2017.08.007
- Galeotti, F., Saucet, C., and Villeval, M. C. (2020). Unethical amnesia responds more to instrumental than to hedonic motives. *Proc. Nat. Acad. Sci.* 117, 25423–25428. doi: 10.1073/pnas.2011291117
- Gambetta, D., and Origgi, G. (2013). The ll game: the curious preference for low quality and its norms. *Polit. Philos. Econ.* 12, 2–23. doi: 10.1177/1470594X11433740
- Garbarino, E., Slonim, R., and Villeval, M. C. (2019). Loss aversion and lying behavior. *J. Econ. Behav. Organ.* 158, 379–393. doi: 10.1016/j.jebo.2018.12.008
- Gelman, A., and Hill, J. (2007). *Data Analysis Using Regression and Hierarchical/Multilevel Models*. Cambridge, UK: Cambridge University Press. doi: 10.32614/CRAN.package.arm
- Gill, D., and Prowse, V. L. (2011). *A novel computerized real effort task based on sliders*. Available at: SSRN 1732324. doi: 10.2139/ssrn.1732324
- Gino, F., and Pierce, L. (2009). Dishonesty in the name of equity. *Psychol. Sci.* 20, 1153–1160. doi: 10.1111/j.1467-9280.2009.02421.x
- Gino, F., and Pierce, L. (2010). Lying to level the playing field: why people may dishonestly help or hurt others to create equity. *J. Business Ethics* 95, 89–103. doi: 10.1007/s10551-011-0792-2
- Gneezy, U., and Kajackaite, A. (2020). Externalities, stakes, and lying. *J. Econ. Behav. Organ.* 178, 629–643. doi: 10.1016/j.jebo.2020.08.020
- Gneezy, U., Kajackaite, A., and Sobel, J. (2018). Lying aversion and the size of the lie. *Am. Econ. Rev.* 108, 419–453. doi: 10.1257/aer.20161553
- Grundmann, S. (2020). “Do just deserts and competition shape patterns of cheating?” in *Passauer Diskussionspapiere-Volkswirtschaftliche Reihe*.
- Hakimov, R., and Kajackaite, A. (2024). Breaking bad: Malfunctioning control institutions erode good behavior in a cheating game. *Games Econ. Behav.* 148, 162–178. doi: 10.1016/j.jeb.2024.09.005
- Hanna, R., and Wang, S. Y. (2017). Dishonesty and selection into public service: evidence from India. *Am. Econ. J.* 9, 262–290. doi: 10.1257/pol.20150029
- Hoffman, E., McCabe, K., Shachat, K., and Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games Econ. Behav.* 7:346–380. doi: 10.1006/game.1994.1056
- Houser, D., Vetter, S., and Winter, J. (2012). Fairness and cheating. *Eur. Econ. Rev.* 56, 1645–1655. doi: 10.1016/j.eurocorev.2012.08.001
- Kajackaite, A., and Gneezy, U. (2017). Incentives and cheating. *Games Econ. Behav.* 102, 433–444. doi: 10.1016/j.jeb.2017.01.015
- Kanngiesser, P., and Warneken, F. (2012). Young children consider merit when sharing resources with others. *PLoS ONE* 7:e43979. doi: 10.1371/journal.pone.0043979
- Kocher, M. G., Schudy, S., and Spantig, L. (2018). I lie? We lie! Why? Experimental evidence on a dishonesty shift in groups. *Manage. Sci.* 64, 3995–4008. doi: 10.1287/mnsc.2017.2800
- Konow, J. (2000). Fair shares: accountability and cognitive dissonance in allocation decisions. *Am. Econ. Rev.* 90, 1072–1091. doi: 10.1257/aer.90.4.1072
- Kunovich, S., and Slomczynski, K. M. (2007). Systems of distribution and a sense of equity: a multilevel analysis of meritocratic attitudes in post-industrial societies. *Eur. Sociol. Rev.* 23, 649–663. doi: 10.1093/esr/jcm026
- Mitchell, G., Tetlock, P. E., Newman, D. G., and Lerner, J. S. (2003). Experiments behind the veil: structural influences on judgments of social justice. *Polit. Psychol.* 24, 519–547. doi: 10.1111/0162-895X.00339
- Norton, E. C., Wang, H., and Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *Stata J.* 4, 154–167. doi: 10.1177/1536867X0400400206
- Osberg, L., and Smeeding, T. (2006). “Fair” inequality? Attitudes toward pay differentials: the United States in comparative perspective. *Am. Sociol. Rev.* 71, 450–473. doi: 10.1177/000312240607100305
- Potters, J., and Stoop, J. (2016). Do cheaters in the lab also cheat in the field? *Eur. Econ. Rev.* 87, 26–33. doi: 10.1016/j.eurocorev.2016.03.004
- Reynolds, J., and Xian, H. (2014). Perceptions of meritocracy in the land of opportunity. *Res. Soc. Stratif. Mobil.* 36:121–137. doi: 10.1016/j.rssm.2014.03.001
- Scott, J. T., and Bornstein, B. H. (2009). What’s fair in foul weather and fair? Distributive justice across different allocation contexts and goods. *J. Polit.* 71, 831–846. doi: 10.1017/S0022381609090744
- Scott, J. T., Matland, R. E., Michelbach, P. A., and Bornstein, B. H. (2001). Just deserts: an experimental study of distributive justice norms. *Am. J. Polit. Sci.* 45, 749–767. doi: 10.2307/2669322
- Sen, A. (2000). “Merit and justice,” in *Meritocracy and Economic Inequality*, eds. K. Arrow, S. Bowles, S. Durlauf (Princeton: Princeton University Press), 5–16. doi: 10.1515/9780691190334-003
- Shalvi, S., Dana, J., Handgraaf, M. J., and De Dreu, C. K. (2011). Justified ethicality: observing desired counterfactuals modifies ethical perceptions and behavior. *Organ. Behav. Hum. Decis. Process.* 115, 181–190. doi: 10.1016/j.obhdp.2011.02.001
- Smith, V. L. (1976). Experimental economics: induced value theory. *Am. Econ. Rev.* 66, 274–279.
- Smith, V. L. (2007). *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge: Cambridge University Press.
- Starmans, C., Sheskin, M., and Bloom, P. (2017). Why people prefer unequal societies. *Nat. Hum. Behav.* 1:0082. doi: 10.1038/s41562-017-0082
- Sutter, M., and Weck-Hannemann, H. (2003). Taxation and the veil of ignorance—a real effort experiment on the laffer curve. *Public Choice* 115, 217–240. doi: 10.1023/A:1022873709156
- Szycer, D., Seal, M. F. L., Sell, A., Lim, J., Porat, R., Shalvi, S., et al. (2017). Support for redistribution is shaped by compassion, envy, and self-interest, but not a taste for fairness. *Proc. Nat. Acad. Sci.* 114, 8420–8425. doi: 10.1073/pnas.1703801114
- Tyler, T. (2011). Procedural justice shapes evaluations of income inequality: Commentary on Norton and Ariely (2011). *Perspect. Psychol. Sci.* 6, 15–16. doi: 10.1177/1745691610393981
- Tyler, T. R. (1997). The psychology of legitimacy: a relational perspective on voluntary deference to authorities. *Person. Soc. Psychol. Rev.* 1, 323–345. doi: 10.1207/s15327957pspr0104_4
- Tyler, T. R. (2006). Psychological perspectives on legitimacy and legitimation. *Annu. Rev. Psychol.* 57, 375–400. doi: 10.1146/annurev.psych.57.102904.190038
- Xiao, E. (2013). Profit-seeking punishment corrupts norm obedience. *Games Econ. Behav.* 77, 321–344. doi: 10.1016/j.jeb.2012.10.010
- Young, M. D. (1994). *The Rise of the Meritocracy*. Piscataway: Transaction Publishers.
- Zelditch, M. (2001). “Theories of legitimacy,” in *The psychology of legitimacy: Emerging perspectives on ideology, justice, and intergroup relations*, eds. J. T. Jost & B. Major (Cambridge: Cambridge University Press), 33–53.