An Analysis of Moral Motives in Economic and Social Decisions

Marc Fleurbaey, Ravi Kanbur and Dennis J. Snower

15 October 2021

Abstract

Abstract

Abstract

We present an analysis of moral motives underlying economic and social decision making. The underlying idea is that moral motives are psychological instruments that induce people to cooperate in pursuit of collective goals and to suppress destructive competition. Different social contexts are associated with different collective action problems, which call for different cooperative relationships. These different cooperative relationships are associated with different moral motives. Cooperative contexts, associated with positive externalities, call for moral motives that reward people for promoting the common good, thereby enabling them to internalize the positive externalities. Competitive contexts, associated with negative externalities, call for moral motives that discourage people from harming one another, thereby promoting the internalization of the negative externalities. In this framework of analysis, people are subject to multiple, context-dependent moral motives. The diverse moral principles underlying the diverse moral motives are to be understood not as mutually exclusive alternatives, but as components to be applied in combination with one another.

1 Introduction

This paper analyzes the role of moral motives in economic and social decisions. Our point of departure is a well-known insight from anthropology, sociology and evolutionary science: one of the main drivers of human evolutionary success has been our capacity to cooperate beyond the bounds of enlightened self-interest.[1] This capacity has been shaped by moral motives throughout human development.[2] Our analysis provides an account of moral motives in economic decision-making and we show that this account can also be applied to social decision-making (social exchanges that do not involve economic transactions).

---

[1] For example, Henrich (2016, 2020) and Sober and Wilson (1998), and Wilson (2015).
[2] For example, Tomasello (2016).

The inclusion of moral motives extends our analysis of economic decisions beyond individualistic purposes (described by payoffs to the individual) to include collective purposes (described by collective payoffs, including but extending beyond payoffs to the individual) as well. There is a large body of evidence that humans are driven by both sets of purposes,[3] whose relative salience differs across individualist versus collectivist cultures[4] and across competitive versus cooperative economic and social settings.[5] We provide a functionalist explanation of moral motives: Moral motives induce people to cooperate in the pursuit of common goals and to suppress destructive competition.

That economic behaviour is driven by collective, as well as individualistic, goals is not surprising, since homo sapiens and their ancestors have been living in social groups for at least seven million years, and thus may be expected to have developed a variety of psychological[6] and cultural[7] mechanisms for promoting group cohesion, suppressing selfishness, and organizing collective action. Thereby humans have found ways of overcoming the free rider problem, with varying degrees of success. Our analysis views moral motives as psychological devices that induce people to cooperate when self-interest stands in the way of achieving collective goals. These motives provide intrinsic rewards and punishments, working alongside the extrinsic mechanisms[8] in the formation and maintenance of cooperation.

Given that moral motives are recognized as playing an important role in people's private lives, as well as their political allegiances and activities, it would be very surprising to find that they were absent in their economic decisions. After all, economic decisions are associated with many of the same moral problems that are present in the private and political domains. Our analysis indicates that moral motives make a difference to our economic decisions and thus that there is no Occam's Rasor case for omitting them from economic analysis.

There is a large body of evidence that people are motivated to pursue multiple moral values in practice, with different values becoming salient under different circumstances.[9] The context-specificity of multiple moral values follows from the recognition that different collective goals call for different forms of interpersonal relationships and multiple values drive such multiple relationships.

---

[3] See, for example, Shweder and Bourne (1984), distinguishing between societies where individuals are meant primarily to serve the needs of society and ones where society is meant to serve the individual.

[4] In individualistic cultures, people's individual goals are salient and one is expected to achieve these goals primarily on one's own; whereas in collectivist cultures, collective goals are salient and people support one another in pursuing these goals (e.g, Hofstede (2001)). This divide affects people's self-concept and personality: people in individualistic cultures define themselves in terms of their own preferences and abilities, whereas those in collectivist cultures define themselves as aspects of their social groups (e.g., Triandis (2002)). Furthermore, the divide affects behavior patterns in the presence of social dilemmas (e.g., Parks and Vu (1994)).

[5] For example, McGuire et al. (2018).

[6] For example, Kurzban, Tooby and Cosmides (2001).

[7] For example, Henrich (2004), Richerson and Boyd (2005), Turchin (2006) and Wilson (2002a).

[8] Ffor example, Fehr and Gächter (2002).

[9] For example, Fiske (1991), Schwartz (1992) and Shweder et al. (1997).

This aspect of morality is particularly obvious with regard to virtues, both religious and secular. These virtues – such as those described in the Judeo-Christian Bible, the Mahabarata, Homer's Illiad or the Sunna of Muhammad – are exemplars that shape the perceptions, beliefs, emotions and habits required for leading a good life in particular social roles and particular historical and cultural contexts. These virtues cannot be reduced to a single principle, but are rather multiple directions that often hold one another in check (such as Aristotle's Golden Mean).[10]

Our analysis provides a rationale for the multiplicity of context-dependent moral values, consonant with virtue-based approaches to morality.[11] It helps explain why different virtues are prized across cultures, e.g. warriors, nomads, settled farmers and city dwellers focus on different sets of virtues.[12]

This moral psychological approach to morality stands in stark contrast to the Enlightenment presumption that moral principles are based on rational, impartial, universal principles about actions or about the consequences of actions – both generally specified independently of their social relationship contexts. In particular, deontology (such as the categorical imperative of Immanuel Kant) focuses on the context-independent intrinsic rightness or wrongness of actions. Consequentialism (such as the utilitarianism of Jeremy Betham or John Stuart Mill) judges actions only by their consequences for human wellbeing and these consequences are usually described in terms of context-independent utility).

Our analysis also parts company from the conventional deontological approach of viewing moral principles in "either-or" terms, such as in the assumption that the "ethic of justice" (suitably defined) is mutually exclusive of the "ethic of care." The either-or approach is also prevalent in behavioral economics, where moral principles are associated with other-regarding preferences. For example, in the inequity-averse preference function of Fehr and Schmidt (1999), individuals derive disutility from being worse off than others in their reference group ("envy") and disutility from being better off ("guilt"). In the menu-based reciprocity model of Rabin (1993), the preference function depends on individuals' kindness, determined by their choices in relation to the alternatives. In the personality-based preference function of Levine (1998), individuals value the payoffs of others. Under Rawlsian preferences (e.g. Charness and Rabin (2002)) an individual's utility depends on the welfare of the worst-off individual in the reference group.

In practice, however, most people who live under the influence of the Western Enlightenment feel drawn to both deontological and consequentialist principles. Some of these deontological principles may be in conflict with one another and their relative salience in people's minds often depends on their social contexts. Similarly, the relative salience of people's evaluation of consequences also depends on the contexts in which they find themselves.[13] Many well-known moral

---

[10] See also MacIntyre (1981).

[11] See, for example, MacIntyre (1981).

[12] For example, Nisbett and Cohen (1996).

[13] For example, Fox and Kahneman (1992).

dilemmas (such as the trolley problems)[14] arise from conflicts among these principles. The great debate in moral psychology between the "morality of justice"[15] and the "morality of care"[16] also arises from conflicting moral concerns, i.e., justice and fairness versus nurturance and avoidance of suffering.

In rationalizing the multiplicity of context-dependent moral motives, our analysis suggests that different moral principles should not be understood as mutually exclusive alternatives, but rather as components to be applied in combination with one another. In our analysis, multiple moral principles are required to internalize each of the externalities considered above. These moral principles need to be combined along well-defined lines in order for the appropriate social cooperation to take place.

Our analysis aims to capture the multiplicity of context-dependent moral motives by focusing on collective action problems involving positive and negative externalities. Moral motives are portrayed as generating moral incentives, which are psychic benefits and costs for internalizing these externalities. We examine how different economic and social contexts generate different externalities and how the internalization of different externalities calls for different moral motives, resting on different moral principles.

A *motive* is conceived in psychological terms as a force that gives direction and energy to one's behavior, influencing the objective of the behavior.[17] *Moral motives* are understood as motives associated with right or wrong actions or with good or bad outcomes, primarily within a social context. We conceive of moral motives as forces supporting our collective interests. *Moral incentives* are the psychic benefits and costs arising from the moral motives.

To illustrate how different contexts can give rise to different externalities, we focus on two types of context: (i) a *cooperative context* characterized by positive externalities from a public good and (ii) a *positionally competitive context* characterized by negative externalities from status seeking.

In practice, each of these contexts can take a variety of forms. Let us define a *first-order cooperative context* as one that generates a positive externality without strategic effects. Such an externality arises through individual contributions to a public good whose benefit to each individual depends on the sum of the individual contributions. We show that this externality can be internalized through *individual moral incentives*, i.e., incentives for individual decisions, independent of other people's decisions. The individual moral incentives can be interpreted as "self-imposed Pigouvian subsidies," the psychic equivalent to the monetary Pigouvian subsidies in the conventional analysis of externalities.

We define a *second-order cooperative context* as one that generates an externality through a strategic complementarity. For example, one individual's contribution to a public good may raise another individual's productivity in contributing to the public good. We show that this externality can be internalized through *interpersonally reinforcing moral incentives*, specifying a positive

[14] For example, Greene (2008) and Hauser (2006).
[15] For example, Kohlberg (1969) and Kohlberg, Levine and Hewer (1983).
[16] For example, Gilligan (1981).
[17] See Elliot and Covington (2001), following Atkinson (1964).

feedback between the individuals' contributions.

A positionally competitive context can be analyzed along analogous lines. A *first-order positionally competitive context* creates a negative externality without strategic effects. Such an externality can can arise through individual contributions to status competition, whereby an individual's gains depends on the difference between the individualistic payoffs. In this context, the externality can be internalized through *individual moral disincentives* that suppress competitiveness. These may be viewed as "self-imposed Pigouvian taxes."

A *second-order positionally competitive context* operates through strategic substitutabilities, whereby for example one individual's contribution to a collective good may reduce another individual's productivity in contributing to the good. In this case, the socially desirable moral incentives can be achieved through "*interpersonally compensating moral incentives*," specifying a negative feedback between the individuals' contributions.

In short, the contributions to the public good are generated by moral motives and associated incentives in the form of psychic benefits – some operating individually and others operating interpersonally. Similarly, contributions to the status contest generate moral motives linked to moral incentives in the form of psychic costs. The nature of these psychic benefits and costs depends on the nature of the externality generated by a particular context.

Our analysis shows how these moral motives may be interpreted in terms of the application of multiple moral values. To illustrate the multiplicity of moral approaches, we initially focus on three separate moral principles: the "principle of merit" (one's moral rewards related to one's merit), the "principle of need" (one's moral rewards related to one's satisfaction of others' needs) and the "principle of distributive justice" (one's moral rewards related to some principle of distributive equity). Later in our analysis, we consider two further illustrative principles: the principle of moral reciprocity (the Silver Rule) and the principle of moral redress (redressing the mistakes of others).

Our analysis suggests that different collective contexts give rise to different collective action problems which call for different sets of moral principles. The moral principles relevant in a cooperative context are different from those relevant in a positionally competitive context. Similarly, the principles required for an appropriate response to a first-order cooperative or competitive context are different from those required for an appropriate response to a second-order cooperative or competitive context.

From one perspective, our analysis is normative: it tells us what moral incentives should drive people's decision making in order for socially desirable outcomes to be achieved. From another perspective, it is also descriptive in an cultural evolutionary sense: Assuming that people tend to develop moral incentives that enable them to thrive, their decision making will come to reflect the moral incentives that our analysis derives, at least in a stable environment.

In a stable social, technological and physical environment, socially desirable collective practices tend to be adopted gradually through time in the process of cultural evolution and moral motives develop in tandem, enabling coordina-

tion for the sake of the common good.[18] In such an environment, people also have more opportunity to engage in deliberate moral cognition (using practical reasoning to address moral problems), rather than relying on automatically triggered moral emotions, and this also tends to promote coordination for the common good.[19] Under these circumstances, our analysis may be expected to offer descriptive insights into people's moral behavior.

However, in times of profound social, technological and physical change, it is of course possible that the moral motives underlying people's behavior patterns are no longer appropriate for new collective action problems they face. This is an example of evolutionary mismatch, with important implications for psychology and public policy.[20] Under these conditions, if there is broad agreement on social objectives (summarized in a social welfare function), our analysis offers prescriptive insights, indicating the moral motives that should be active to tackling the collective action problems. Our analysis indicates that the less the members of a social group can agree on their social objectives, the smaller the gains from their social cooperation and consequently the smaller the social group will tend to be.

In line with the evidence referenced below, moral motives can be primed and are highly context-specific. The role of politicians, business and civil society in priming motives, as well as their role in creating cooperative contexts that activate moral motives to tackle collective action problems, remains under-investigated in economic analysis.

Whereas moral psychology (investigating the motives underlying moral judgments) should not be confused with moral epistemology (concerned with the nature of moral knowledge), it is clear that moral psychology plays an important epistemological role. It is pointless to specify right and wrong actions that are psychologically infeasible to perform. Similarly, it is pointless to specify good and bad states that psychologically impossible to appreciate. Moral philosophies are useful only insofar as they can guide action, and thus no strong case can be made to take seriously those moral philosophies that are at odds with without psychological capacities to implement them.

The rest of the paper is organized as follows. Section 2 covers underlying ideas and related literature. Section 3 presents a general framework for the analysis that follows. Section 4 explores moral motives in a cooperative context and Section 5 does so in a competitive context. Section 6 applies our model to the moral implications of knowledge-biased technological progress. Finally Section 7 concludes.

---

[18] For example, Christakis (2019) and Henrich (2016, 2020).

[19] On the distinction between moral reasoning and moral emotions, see for example Greene (2013).

[20] For example, Li, van Vugt and Colarelli (2018).

2 Underlying Ideas

To place our analysis in context of mainstream economic and social theory, we examine the role of morality in neoclassical economics and in social relationship theory.

2.1 Morality in Neoclassical Economics

Neoclassical economics has divorced itself from morality from individual private-sector decision making. It does so through the following assumptions, usually made implicitly.

1. Neoclassical economics is strictly individualistic, based on the view that economic decisions are made by autonomous individuals. These individuals are assumed to have only individual objectives. The activities of social groups are conceived merely in terms of the sum of independent individual decisions. By contrast, moral principles frequently address communitarian concerns, representing individuals' collective objectives arising from individuals' participation in social groups. Moral principles are often concerned with precepts, rules, standards and codes of conduct that enable individuals to live fruitfully in larger groups.

2. Neoclassical economics is concerned primarily with how scarce resources are used to satisfy "wants," rather than to satisfy worthy goals or to promote wellbeing. Wants are conceived as being exogenously given, not subject to evaluation or influence. This focus is central to Lionel Robbin's dominant definition of economics as the science which "studies human behaviour as a relationship between ends and scarce means which have alternative uses" (Robbins (1932, p. 16)). Previously, economics had been conceived differently. Alfred Marshall, for example, considered economics to be the study of the "material requisites of well-being" (Marshall (1890, p.9)). "Well-being" or other worthy goals invite moral considerations; wants do not. The focus on wants implies that moral objectives are subsumed under "preferences" in economic analysis, alongside tastes. In this framework, the taste for ice cream plays an analogous role to the distaste for murder in guiding people's economic decisions.

3. The object of these wants are assumed to be commodities, not social relationships. Thereby neoclassical economics ignores that humans are social creatures with social aims. By contrast, morality often regulates social relationships, promoting the common goods within these relations.

4. Conventional neoclassical economics assumes that individuals are purely self-interested. Most moral issues arise in weighing self-interest against the interest of others. For this reason self-interest is considered to be problematic on moral grounds in other social sciences. Neoclassical economics does not provide the opportunity for exploring the moral desirability of

self-interest, since such self-interest is simply postulated as a fact of life or at least as a useful simplification for the study of economic behavior.

5. Neoclassical economics divorces morality from individual private-sector decision making through its implicit assignment of responsibilities in society: households are responsible for pursuing their self-interest, firms are responsible for maximizing their profits, and the government is responsible for setting the rules of the game so as to mitigate inefficiencies and inequities. Inefficiencies are viewed as a waste of resources and thus reducing inefficiencies is self-evidently right, requiring no serious moral analysis. Thus the arena for moral decisions is narrowed to issues of distributive justice in the hands of the government, to be addressed through democratic processes. Thereby the need for moral decisions by individuals in the private sector has been banished from neoclassical analysis.

Our analysis calls these assumptions into question, thereby creating space for morality in individual decision making. We takes into account individuals' collective goals (in contrast to Assumption 1). In this context, our analysis encompasses interpersonal relations as drivers of economic decision making (in contrast to Assumption 3). Consequently, it recognizes that people are not exclusively self-interested (in contrast to Assumption 4). Finally, it recognizes that individuals in the private sector are morally load-bearing and thus can participate in bearing moral responsibility for overcoming inefficiencies and inequities (in contrast to Assumption 5). (Assumption 2 lies outside the scope of our analysis, as we simply assume that the collective goals are worthy, rather than evaluating their worthiness.)

The rationale for the Invisible Hand through the first fundamental theorem of welfare economics – arguably the most important result of the economics discipline – arises from Walrasian general equilibrium analysis under the assumption that economic decisions are made under perfectly competitive conditions and in the absence of externalities. While neoclassical economics recognizes the existence of externalities arising from free-market activities, it places responsibility for addressing these problems in the hand of the government.

The second fundamental theorem of welfare economics addresses the problem of distributional inequities through the fiction of lump-sum transfers, which permit any distribution of wealth to be compatible with economic efficiency. In this context, there is no need to assess the desirability of alternative wealth distributions on moral grounds. When neoclassical economics recognizes that lump-sum transfers may be impracticable, it assumes once again that the correction of distributional inequities are the responsibility of government, to be fulfilled by choosing an appropriate point on an equity-efficiency tradeoff. This distributive policy choice is only avenue whereby morality enters neoclassical analysis.

Civil society is implicitly excluded from moral decisions regarding the correction of inefficiencies and inequities. In the presence of multitude of government failures (including preference-revelation problems, political self-interest, political myopia, regulatory capture, regulatory arbitrage, undue pressure group

8

influence, etc.) and tacit knowledge in the private sector to which government has no access, it does not appear reasonable to leave the job of correcting inefficiencies and inequities entirely to the government. Furthermore, individuals in the private sector are often acutely aware of, sensitive to and active in the internalization of externalities and the pursuit of equity.

On this account, it becomes important to rethink neoclassical economics' standard division of responsibilities among economic agents: households responsible for pursuing their self-interest; firms responsible for maximizing their profits; and the government responsible for correcting market failures and inequities. Our analysis, by contrast, examines the role of morality in guiding the behavior of private-sector individuals to address such collective action problems.

2.2 Social Relationships

One particularly important way whereby people coordinate their behavior to address externalities and inequities is to engage in cooperative social relationships. The public good in our analysis can be interpreted in terms of such relationships. After all, cooperative relationships involve positively valenced psychic exchanges among individuals that often do not receive monetary compensation. They are, in other words, a type of public good. These cooperative relationships may also mitigate the negative externalities arising from status competition in our analysis.

In neoclassical economics, externalities and inequities are defined with reference to behavior patterns that would arise from purely self-interested behavior. Social relationships affect people's preferences and when these relationships are cooperative, they enable people to internalize some externalities and overcome some inequities. Conversely, of course, adversarial relationships can create externalities and inequities.

Morality may be understood functionally as having two mutually reinforcing roles with regard to social relationships. First, moral motives promote cooperative social relationships, with different motives promoting different relationships. Second, these relationships, in turn, promote the salience of their associated moral motives.

This connection between moralities and social relationships helps explain the strong emotional force that moral principles often have. People's survival commonly depends on their ability to cooperate with one another; social relationships enable them to do so; and moral motives are the drivers and regulators of these relationships. The connection between moralities and social relationships also elucidates the reasons for moral conflict that may occur across cultural lines. Different cultures often face different collective action problems, calling for different moral motives, supported by different moralities and social relationships. An example is the contrast between the "culture of honor" in the American South and the liberal culture in the American North. According to Nisbet and Cohen (2004), the former was a response to the coordination problems of herders in the Southern states, while the latter was adapted to the coordination problems of growers in the Northern states. When these cultures

meet, they clash. Since their distinctive moral motives arose from a process of cultural evolution in response to collective action in different social contexts, Southerners and Northerners need not be aware of the functions their moralities serve. Instead, they view themselves as adhering to moral principles that are intrinsically worthwhile.

A clash of moralities can also arise when different cultures choose to solve the same collective action problem through different moral motives, which generate different social relationships. Fiske (1991) provides an illustrative example of a small town considering how to organize fire fighting. One option is to discuss the issue at meetings of the inhabitants, until a community consensus emerges. A second possibility is to let the leader of the community decide by fiat. A third option is to decide the issue by referendum. A final possibility is to let the market decide by letting people purchase fire insurance. Each of these coordination mechanisms is associated with distinctive moral motives.

To understand the connection between these coordination mechanisms, their associated social relationships and their underlying moral motives, it is useful to classify cooperative relationships in terms of the social-relational approach to moral psychology,[21] which posits four basic types of social relationship: communal sharing, authority ranking, equality matching and market pricing.

In "communal sharing," people perceive themselves to be undifferentiated components of a social entity, as may occur in families, friendship groups, military units, teams, nations, ethnicities and religions. In small-scale social relationships where group members have an intimate understanding of each other's wellbeing, communal sharing can be driven by the morality of care. In large-scale communities where this intimate understanding is lacking, people provide aid to one another in accordance with the morality of need. Both of these moralities serve to support the integrity of in-groups through a sense of collective responsibility.

"Authority ranking" is a hierarchical relationship, in which superiors have a sense of pastoral responsibility toward their subordinates, providing aid in accordance with their abilities, and the subordinates are obligated to respect their superiors. This coordination mechanism is appropriate when there are significant asymmetries of ability, information and power. Along the lines above, small-scale communities can motivate this social relationships through the morality of care, while large-scale communities can do so through the morality of merit.

"Equality matching" aims at achieving an even balance and in-kind reciprocity in social relationships. It is governed by the morality of equality-based forms of distributive justice.[22] Finally, under "market pricing," people are compensated in proportion to the benefits and costs they bestow on others. When this compensation is monetary, the social coordination is performed by the market, so that so that self-interest is sufficient,[23] though support for choosing this coordination mechanism may have to arise from the morality of merit.

---

[21] In particular, the relational models theory of Fiske (1991, 1992, 2011), Fiske and Haslam (2005), Haslam (1994, 2004) and Haslam and Fiske (1992).

[22] For example, Deutsch (1975) and Folger et al. (1995).

[23] In this case, the conventional neoclassical rationale for the Invisible Hand is applicable.

Each of these social relationships entails its own obligations and transgressions. The psychological function of morality in this context is to faciliate the creation and maintenance of these cooperative relationships. Different social relationships entail different moralities, which in turn support their distinctive social relationships. When one culture uses its own moral standards to judge the morality of another culture, the result may be bitter conflict. Recognizing the sources of divergent moral motives may help the parties understand the context-dependence of their moral disagreements and thereby come a step closer to mutual tolerance and respect.

In this context, it is clear how moral motives promote the internalization of positive externalities associated with relationship-driven public goods. The moralities of need and merit, as well as the morality of care, all support preferences that reward contributions to such public goods. They also serve to reduce the negative externalities associated with positional competition through psychic costs of anti-social behavior. Furthermore, the morality of distributive justice serves to bring individual preferences into line with social concerns regarding inequities. All these moralities limit the capacity for self-interest to subvert collective goals.

Our analysis below also illustrates how disagreements concerning distributive justice lead to reduced provision of the public good (which could represented cooperative social relationships). This implies that there is less benefit from cooperating within a social group. When this disagreement does not affect the cost of group formation and maintenance, lower benefits from cooperation tend to lead to smaller social groups.

3   Framework of Analysis

There are n individuals indexed $i = 1, 2, 3, \ldots . n$. Let the action of individual i, in general a vector, be denoted y_i. Let the actions of individuals other than i be denoted y_{-i}. Express the self-interested well-being of individual i, dependent on actions of all individuals, as $V_i(y_i, y_{-1})$, which is increasing and concave in y_i and increasing in y_{-1} for the case of positive externality of an individual's action on the wellbeing of others, but decreasing in y_{-1} for the case of negative externality.

When pursuing narrow self-interest, each individual i maximizes $V_i(y_i, y_{-1})$ with respect to y_i, with first-order conditions:

$$V_{iy_i}(y_i, y_{-i}) = 0, \ \ i = 1, 2, ..., n \tag{1}$$

Equations (1) solve for the Nash Equilibrium outcomes in actions.

Let group well-being, the social welfare function, be given by a weighted sum of individual well-beings:

$$S = \sum_j \alpha_j V_j(y_j, y_{-j}) \tag{2}$$

Maximizing group wellbeing with respect to y_i requires actions defined by the following first order conditions:

$$V_{iy_i}(y_i, y_{-i}) + \frac{1}{\alpha_i} \sum_{j \neq i} \alpha_j V_{jy_i}(y_i, y_{-i}) = 0, \ i = 1, 2, ..., n \tag{3}$$

The difference between the sets of first order conditions in (3) and (6) captures the misalignment of private incentives in the advancement of group wellbeing. Now consider the objective of a "moral individual" by adding a moral incentive to the individual's self-interested welfare: $V_i(y_i, y_{-1}) + r_i y_i$. The resulting first-order condtions are

$$V_{iy_i}(y_i, y_{-i}) + r_i = 0, \ \ i = 1, 2, ..., n \tag{4}$$

We will discuss below alternative interpretations and motivations of the moral incentives r_i, but formally the necessary conditions for the first order conditions in (4) to coincide with the first order conditions for group maximization in (3) are:

$$r_i^* = \frac{1}{\alpha_i} \sum_{j \neq i} \alpha_j V_{jy_i}(y_i, y_{-i}), \ \ i = 1, 2, ..., n \tag{5}$$

Under this condition, individual welfare maximization is consistent with maximization of the group well-being. We will explore the implications of (5) in detail in this paper.

When r_i is independent of actions y_i and y_{-i}, we can think of r_i as a self-imposed Pigouvian tax or subsidy. The moral incentive r_i is analogous to a "tax" (if negative) or a "subsidy" (if positive) on action y_i. It is not of course monetary cost or compensation for the action, from the state or from anyone else. It is "self-imposed" psychic cost or benefit, arising from the individual's cognitive or affective faculties.

This analytical framework is useful to set the scene for our analysis and we will return to it later. But we now proceed to various instructive special cases that allow for tractable solutions and more detailed investigation of moral incentives. Specifically, we are interested in how constellations of moral incentives are composed of constituent parts that appeal to different strands of moral reasoning.

4 Moral Motives in a Cooperative Context

4.1 Individual Moral Incentives in a First-Order Cooperative Context

Consider two individuals, indexed 1 and 2, who interact in a cooperative setting, characterized by their contributions to a public good. Individual i makes effort Y_i as contribution to a public good. The impact on the public good is

mediated by a productivity coefficient A_i so that the effective contribution of each individual to the public good is $A_i Y_i$. The overall public good is given by

$$G = A_1 Y_1 + A_2 Y_2. \tag{6}$$

The value of G to individual i is $B_i G$. The cost of effort Y_i for individual i is $\frac{1}{2} Y_i^2$. So the net benefit to individual i, the narrowly defined self-interest, is given by

$$V_i = B_i \left(A_1 Y_1 + A_2 Y_2 \right) - \frac{1}{2} Y_i^2. \tag{7}$$

If the two individuals maximize their self-interested wellbeing, their efforts are given by

$$Y_1^P = B_1 A_1 \tag{8}$$

$$Y_2^P = B_2 A_2 \tag{9}$$

where the superscript P indicates private optimization. Note that each individual's self-interested effort level depends only on their productivity and preference, with no reference to the other individuals.

Now let group wellbeing be defined by a weighted sum of the individual net benefits:

$$\Omega = V_1 + \alpha V_2$$

where α is the weight of individual 2's wellbeing relative to that of individual 1.

We begin with the assumption there is agreement on this relative weight. (The issue of disagreement and its implications is taken up later in this section.) Then the wellbeing of the social group is maximized, with respecct to the efforts of both individuals when

$$Y_1^* = A_1 \left(B_1 + \alpha B_2 \right) \tag{10}$$

$$Y_2^* = A_2 \left(B_2 + \frac{1}{\alpha} B_1 \right) \tag{11}$$

We refer to these as socially optimal contributions to the public good. Each individual's contribution depends on both individuals' preferences and both individual's weights in the group. It should be clear that, comparing (8) with (10), and (9) with (11), group optimal contributions are greater than self-interested contributions, because they take into account the positive externality of one individual's contribution on the wellbeing of the other through the public good.

4.2 Moral Incentives

Finally, suppose that the individuals are motivated both by self-interest and moral values, which induce them to take the public interest into account. Let

us represent the moral incentives of individuals 1 and 2 as R_1Y_1 and R_2Y_2, respectively, so that the individualize maximize:

$$M_1 = V_1 + R_1Y_1 \tag{12}$$
$$M_2 = V_2 + R_2Y_2 \tag{13}$$

The moral incentives R_1 and R_2 can be interpreted as the perceived reward from contributing effort towards the public good, or "doing the right thing". In the context of a model containing only linear relations – e.g. the benefit from public goods linearly related to the contributions, social welfare linearly related to individual utilities, and so on – R_1 and R_2 can be interpreted of course as a "self-imposed Pigouvian subsidies" on effort towards producing the public good. These give rise to the morality-adjusted solutions:

$$Y_1^M = B_1A_1 + R_1 \tag{14}$$
$$Y_2^M = B_2A_2 + R_2 \tag{15}$$

Compared to (12) and (13), each individual's effort is adjusted by the "moral incentives" R_1 and R_2.

Starting from zero, increases in the moral weights R_1 and R_2 given to public-good contributions will naturally increase individual effort from the pure self-interested level. In order for these effort levels to coincide with the group optimal levels (10) and (11), the moral incentives will have to be set at:

$$R_1^* = \alpha B_2 A_1 \tag{16}$$
$$R_2^* = \frac{1}{\alpha}B_1 A_2 \tag{17}$$

In effect, if each individual applies a self-imposed Pigouvian subsidy on effort at the rates given above, the effort levels will be as in (10) and (11), the socially optimal outcome.

We proceed to interpret these moral incentives in terms of two alternative approaches: justice and care. The justice approach is commonly articulated in deontologically, in terms of principles of justice. They are the outcome of rational deliberation, applied to "thin" social relations, appropriate for social groups in which people have limited personal experience of one another and thus limited knowledge of the wellbeing of other group members. Under these circumstances, they may achieve their common purpose through an appropriate application of moral principles that have normative force independently of the people to whom they are applied. The associated moral incentives can be conceived as arising through a process of cultural evolution of a population with stable traits in a stable environment.

By contrast, the caring approach is relevant for social groups in which members have "thick" social relations – ranging from wide-ranging personal relation-

ships[24] (inducing them to engage in effective perspective-taking and profound empathic concern) to cultural affinities (enabling more modest perspective-taking and concern) – a different normative approach may be more effective, namely, one that rests on one's direct participation in the wellbeing of others. These thick social relations are associated with a "thick," particularist morality, which is based on different evaluative lines from the "thin," universalist morality that we apply to strangers in other cultures.[25]

In short, the justice approach is particularly appropriate to civil societies, whereas the caring approach pertains primarily to communities.

4.3   The Justice Approach

Interpret the moral incentives through principles of justice, we will focus on three principles, which we define in restrictive terms: the principles of merit, need and distributive justice.

According to the "*principle of merit*," everyone should be rewarded in accordance with one's merit.[26] Though this principle is usually articulated in the external domain, where merit is rewarded through compensation and social distinctions, our analysis applies it to the internal domain, where merit receives psychic benefits.[27] In cooperative settings, the psychic benefits can arise from the "warm glow" of contributing to the common good in accordance with one's ability and from social prestige.[28] Alternatively, there may be psychic costs of shame and guilt[29] from failing to contribute in accordance with one's merit. In competitive settings, the psychic benefits may rise from feelings of "noblesse oblige" or pride that one has not exploited one's advantage over others even though it was in one's power to do so.

Psychologically, the principle of merit is associated with the "achievement motive." People thus motivated show preferences for challenges or competitive situations[30] and tend to engage in persistent, ambitious, or dominant actions[31]. When achievement-motivated individuals receive positive feedback on their productivity, they tend to solve more tasks in the same time as individuals who lack this motive[32].

The "*principle of need*" prescribes that everyone should be rewarded in accordance with their need.[33] It provides the moral basis for means-tested un-

---

[24] For example, Selman (2008).

[25] These notions of thick and thin morality are taken from Walzer (1994).

[26] See, for example, Sandel (2020).

[27] External rewards arising for moral reasons – employers compensating employees of high ability on moral grounds rather than profitability, universities offering places to the most able students, welfare policy measures designed to promote equality of opportunity – play potentially important roles in the functioning of economies. An analysis of these rewards lies beyond the scope of this paper.

[28] See, for example, Henrich and Gil-White (2002).

[29] For example, Tangney and Dearling (2002).

[30] Winter (1973) and Schultheiss & Brunstein (2010).

[31] Winter (1973); (Pang, 2010); for a review see Hall, Stanton & Schultheiss (2010).

[32] Lowell (1952); French (1956)

[33] For example, Gilligan (1982).

employment and pension benefits, whereas contribution-based unemployment and pension benefits rely on the principle of merit. Our analysis focuses on the morally motivated psychic benefits from contributing to the common good or suppressing competitiveness in accordance with the needs of others.

Psychologically, the principle of need is associated with the compassion motive.[34] People primed with the compassion motive with the name of a secure attachment figure are more willing to help people in distress.[35]

The "*principle of moral reciprocity*" requires that we repay others in accordance with what they have done to us. Under pure reciprocity, the repayment is equivalent; under "moral reciprocity," by contrast, the degree of repayment depends on morally relevant characteristics of the individuals involved. (For example, needier people may be required to repay a benefit in accordance with their means.) The principle of moral reciprocity has been called the Silver Rule,[36] to distinguish it from the Golden Rule (which prescribes benevolence, rather than reciprocity). An example of this rule is generous tit-for-tat.[37] In many religious traditions, the requisite degree of generosity is related to one's merit.

Psychologically, the principle of moral reciprocity – or fairness – is associated with several motives, particularly the affiliation motive to reward fairness, the anger (threat approach) motive to punish fairness violations and the fear (threat avoidance) motive to avoid such punishment. Individuals driven by the affiliation motive experience the need to belong, to be accepted, and the urge for relatedness.[38] When these needs are not met, they experience self-reported separation distress.[39] The affiliation motive induces individuals to initiate social interactions and maintain the cohesion of social groups.[40] It is also responsible for anxiety-driven efforts to meet affiliative needs.[41]

Fairness violations are commonly perceived as social threats by third parties, who commonly respond with anger.[42] The anger motive involves aggressive behavior and increased proclivity towards risky decisions[43] that are often destructive, intimidating or antisocial.[44] The fear motive induces individuals to flee, to freeze, or to appease.[45] Emotionally, fear-motivated people experience

---

[34] This is to be distinguished from the affiliation motive, dealing with the need to belong to a social group (Weinberger, Cotler, & Fishman, 2010) and Reiss (2004).

[35] Mikulincer, Shaver, Gillath & Nitzberg (2005).

[36] For example, Hirshleifer (1985).

[37] For example, Wedekind and Milinski (1996).

[38] Atkinson et al. (1954); Boyatzis (1973); Baumeister & Leary (1995).

[39] Eisenberger, Lieberman & Williams (2003).

[40] Atkinson et al. (1954); Lansing & Heyns (1959); Exline (1962); Walker & Heyns (1962); McAdams & Constantian (1983); McClelland (1985); Weinberger, Cotler, & Fishman (2010).

[41] Weinberger, Cotler, & Fishman (2010).

[42] Fessler (2010).

[43] Kornadt, (1984); Fessler, Pillsworth, & Flamson (2004); Leith & Baumeister (1996); for a review see Carver & Harmon-Jones (2009).

[44] Tedeschi, Smith & Brown (1974); Small & Lerner (2008); Eimontaite, Nicolle, Schindler, & Goel (2013).

[45] For reviews see Adolphs (2013); LeDoux, (1998).

anxiety or panic.[46]

The "*principle of distributive justice*" creates the obligation to treat people's needs in accordance with some principles of distributive equity.[47] Our analysis focuses on a particular aspect of distributive equity: interpersonal comparisons of utility with regard to the moral drivers of economic decisions. For example, under strict egalitarianism, each person's material wellbeing must be treated equally; under Benthamite utilitarianism, the material wellbeing of the poor should receive more weight than the material wellbeing of the rich (in line with the declining marginal utility of wealth); under Rawls' difference principle, inequalities in material wellbeing should be designed to benefit the least advantaged members of society; and so on. In this sense, the principle of justice may be considered separately from the principles of merit and need, since the former is concerned with interpersonal comparisons of utility, whereas the latter is concerned with rewards for particular activities (contributions to public goods and suppression of status seeking).

Psychologically, the principle of distributive justice is associated is the same motives as the principle of moral reciprocity, but these motives are applied toward different ends (distributive justice rather than moral reciprocity). The combination of the anger and fear motives helps promote the principle of distributive justice by eliciting punishments (through the anger motive) and simultaneously generating the sensitivity to these punishments (through the fear motive).

In this context, our analysis makes two significant claims: (1) To coordinate economic decisions in the public interest, the various principles of justice need to be combined in well-defined ways, not viewed as mutually exclusive alternatives. (2) The socially desirable combination of justice principles is context-specific, i.e. it depends on the nature of the collective action problem to be addressed.

The justice principles underlying the moral incentives above may be summarized as follows:

- The greater is an individual's productivity (A_i), the greater is the individual's moral incentive (R_i^*). Thus people of greater ability should receive a greater moral reward for their contribution to the public good and, correspondingly, feel a greater obligation to contribute to the common good. This is the "morality of merit."

- The greater is an individual's need (B_i), the greater is the other individual's moral incentive $(R_j^*, \text{ for } i \neq j)$. Thus a person should recognize a greater sense of social responsibility, the greater are the needs of others in his or her social group. The greater the needs of others, the greater should be the moral reward for satisfying these needs by contributing to the common good. This is the "morality of need."

- The greater is the relative weight of an individual's wellbeing in social wellbeing (such as α for individual 2), the lower is that individual's moral

---

[46] Avram et al., (2010); Schaefer et al. (2010); Stemmler et al. (2001).
[47] Summarized, for example, in Fleurbaey (2004).

incentive (R_2^*) and the greater is the other individual's moral incentive (R_1^*). Thus the importance of an individual's wellbeing in society justifies lower contributions to the common good. This is the "morality of distributive justice." The weights of individuals' wellbeing in social wellbeing can be underpinned by a variety of moral values. For example, Benthamite utilitarianism under declining marginal utility of wealth implies that poorer individuals should be given greater weight in the social utility function.

In sum:

Proposition 1 *In order to achieve socially optimal contributions to the provision of a standard public good (for which one's benefit depends on the sum of the individuals' contributions), an individual's moral rewards must be proportionate to (1) the individual's merit, (2) the needs of others and (3) the social importance of others.*

These results are striking, since they stand in contrast to much conventional moral reasoning. For instance, it is common to depict the morality of merit (reward people in accordance with their abilities) as opposed to the principle of need (reward people in accordance with their needs). Moreover, the principle of distributive justice is commonly viewed as quite separate from the principles of need or merit. The results above, however, suggest that these various moral principles should be combined in the appropriate proportions in order for socially optimal outcomes to be achieved.

The implications are profound and far-ranging. For example, in *The Tyranny of Merit,* Sandel argues that rewarding people in accordance with their abilities leads to a neglect of the public good, since it induces the winners to regard their winnings as a just reward for their merit and the losers to blame themselves for their lack of success. Our analysis implies that, with regard to public good provision, Sandel's argument is well-founded, on two counts. First, our analysis does not at all yield the principle of giving people greater rewards merely for greater abilities. Instead, our analysis implies that people should receive greater psychic rewards for greater ability only if this leads to greater contributions to the public good. Second, receiving psychic rewards solely on the basis of ability (i.e., $R_i^* = A_i$) does indeed not yield the social optimum. But this does not mean that merit should be ignored in pursuing social purpose; rather, rewards to merit should be supplemented by rewards for the satisfaction of needs and for deference to social importance.

4.4 Moral Disagreement and the Limits of Social Identity

Now suppose that people disagree with regard to their relative importance in social welfare. One common form of such disagreement arises from the influence of self-interest.[48] Insofar as each individual accords herself greater importance

---

[48] Another form of common moral disagreement - not analyzed in this paper - arises from adherence to conflicting moral principles.

in her subjective social welfare function than other individuals do so in their subjective social welfare functions, this moral disagreement has important implications for the size of a social group contributing to a common public good.

Let α_1 and α_2 represent the beliefs of individuals 1 and 2, respectively, concerning the weight of individual 2 relative to individual 1 in social welfare, and let both individuals be driven by moral concerns, as depicted by their moral incentives. Since the behaviors of individuals 1 and 2 are driven by their beliefs α_1 and α_2, respectively, their moral incentives become:

$$
\begin{align}
R_1^\beta &= \alpha_1 B_2 A_1 \tag{18}\\
R_2^\beta &= \frac{1}{\alpha_2} B_1 A_2 \tag{19}
\end{align}
$$

where the superscript β denotes that the moral incnentives are belief-driven.

The associated contributions to the public good are:

$$
\begin{align}
Y_1^\beta &= A_1 \left(B_1 + \alpha_1 B_2 \right) \tag{20}\\
Y_2^\beta &= A_2 \left(B_2 + \frac{1}{\alpha_2} B_1 \right) \tag{21}
\end{align}
$$

The resulting provision of the public good is

$$
G^\beta = A_1^2 \left(B_1 + \alpha_1 B_2 \right) + A_2^2 \left(B_2 + \frac{1}{\alpha_2} B_1 \right) \tag{22}
$$

The associated subjective level of wellbeing for individuals i $(i = 1, 2)$:

$$
M_i^\beta = B_i G^\beta - \frac{1}{2} \left(Y_i^\beta \right)^2 + R_i^\beta Y_i^\beta \tag{23}
$$

When both individuals are self-interested, $\alpha_1 < 1$ and $\alpha_2 > 1$. Furthermore, as the individuals become more self-interested, α_1 falls and α_2 rises. The fall in the subjective weight α_1 leads to a fall in individual 1's contribution to the public good : $\frac{\partial Y_1^\beta}{\partial \alpha_1} = A_1 B_2 > 0$. The associated fall in individual 1's maximized subjective wellbeing is:

$$
\begin{align}
\frac{\partial M_1^{\beta*}}{\partial \alpha_1} &= \frac{\partial G^\beta}{\partial \alpha_1} - Y_1^\beta \frac{\partial Y_1^\beta}{\partial \alpha_1} + Y_1^\beta \frac{\partial R_1^\beta}{\partial \alpha_1} + R_1^\beta \frac{\partial Y_1^\beta}{\partial \alpha_1} \\
&= A_1^2 B_2 - A_1^2 \left(B_1 + \alpha_1 B_2 \right) B_2 + A_1^2 \left(B_1 + \alpha_1 B_2 \right) B_2 + \alpha_1 B_2^2 A_1^2 \\
&= A_1^2 B_2 + \alpha_1 B_2^2 A_1^2 = A_1^2 B_2 \left(1 + \alpha_1 B_2 \right) > 0
\end{align}
$$

Analogously, a rise in α_2 leads to a fall in individual 2's public good provision: $\frac{\partial Y_1^\beta}{\partial \alpha_1} = -\frac{1}{\alpha_2^2} A_2 B_1 > 0$. The associated fall in individual 2's maximized subjective wellbeing is:

$$\begin{aligned}
\frac{\partial M_1^{\beta*}}{\partial \alpha_2} &= \frac{\partial G^\beta}{\partial \alpha_2} - Y_2^\beta \frac{\partial Y_2^\beta}{\partial \alpha_2} + Y_2^\beta \frac{\partial R_2^\beta}{\partial \alpha_2} + R_2^\beta \frac{\partial Y_2^\beta}{\partial \alpha_2} \\
&= -\frac{A_2^2 B_1}{\alpha_2^2} + A_2 \left(B_2 + \frac{1}{\alpha_2} B_1 \right) \frac{A_2 B_1}{\alpha_2^2} - A_2 \left(B_2 + \frac{1}{\alpha_2} B_1 \right) \frac{B_1 A_2}{\alpha_2^2} - \frac{B_1^2 A_2^2}{\alpha_2^2} \\
&= -\frac{A_2^2 B_1}{\alpha_2^2} - \frac{B_1^2 A_2^2}{\alpha_2^2} = -\frac{A_2^2 B_1}{\alpha_2^2} \left(1 + B_1 \right) < 0
\end{aligned}$$

In short, when both individuals become more self-interested, they both reduce their contributions to the public good and their levels of subjective wellbeing consequently fall.

Let the subjective value added from group identity for individual i be specified as the difference between the individual's subjective wellbeing from morally motivated behavior and the individual's subjective wellbeing from purely self-interested behavior: $\varpi_i^{\beta*} = M_i^{\beta*} - V_i^{\beta*}$. Then, along the same lines above, it can be shown that as both individuals become more self-interested, each individual's value added from group identity falls:

$$\frac{\partial \varpi_i^{\beta*}}{\partial \alpha_1} > 0, \ \frac{\partial \varpi_i^{\beta*}}{\partial \alpha_2} < 0$$

so that a fall in the weight α_1 and a rise in the weight α_2 both lead to a fall in the individuals' subjective value added.

Furthermore, suppose that the cost of creating and sustaining group identity for each individual is a constant C. Then the individuals will form the social group when their combined subjective values added from group identity is greater than or equal to this cost:

$$\sigma = \varpi_1^{\beta*} + \varpi_2^{\beta*} \geq C.$$

Next, we extend our analysis to social groups of arbitrary size. To fix ideas, let individual 1 be the individual around whom the social group is formed. Let us rank individuals in decreasing order of combined subjective values added:

$$\sigma = \sigma(N), \ \ \sigma' < 0,$$

where N is the number of people in the social group.

Then the size of the equilibrium social group is

$$\sigma^* = \sigma(N^*) = C.$$

Then a rise in the degree of self-interest leads to a downward shift of the $\sigma = \sigma(N^*)$ curve and thus to a fall in the group size N^*.

In sum,

Proposition 2 *The greater the degree to which individuals' evaluations of social welfare is motivated by self-interest, the lower are their contributions to the public good. Consequently, the lower is their combined value added from group identity and the smaller is the equilibrium size of the social group that contributes to the public good.*

4.5 The Caring Approach

As noted, the caring approach is appropriate to communities with "thick" social ties, enabling their members to participate in each other's wellbeing. Such participation is articulated in the "*principle of care*," which prescribes altruism[49] and is articulated in the Golden Rule (treating others as one wants to be treated oneself), espoused by most of the world's religions. This morality is supported psychologically by the care motive, involving emotions of warmth, friendship, affection, or love.[50] Care-motivated individuals show positive concern with the well-being of others.[51] Such individuals tend to have a reduced cognitive focus on their own needs in favor of others' needs.[52] They tend to engage in non-instrumental interpersonal sharing and to make choices that benefit another individual or group.[53]

Among thick social groups, pursuing the morality of care has important advantages over deontological moralities, particularly in times of change. In such times, the danger of evolutionary mismatch in the pursuit of abstract deontological principles is relatively large, whereas direct participation in the wellbeing of other people whom one knows well permits flexible adjustment to such contextual change.

In the analysis below, we portray individual i's participation in the wellbeing of individual j by including j's self-interested utility function, as perceived by i, in i's utility function. In general, j's self-interested utility is weighted relative to i's self-interested utility: equally weighted utilities stand for perfect altruism; when i weights her own self-interested utility more than j's self-interested utility, the care motive is muted. This approach can be articulated through a consequentialist approach to the morality of care.

The utility functions of individuals 1 and 2 may be expressed as

$$U_1 = (1 - \kappa_1) \left(B_1 \left(A_1 Y_1 + A_2 Y_2 \right) - \frac{1}{2} Y_1^2 \right) + \kappa_1 \alpha \widetilde{V}_{21} \tag{24}$$

$$U_2 = (1 - \kappa_2) \alpha \left(B_2 \left(A_1 Y_1 + A_2 Y_2 \right) - \frac{1}{2} Y_1^2 \right) + \kappa_2 \widetilde{V}_{12} \tag{25}$$

where the parameter κ $(0 \leq \kappa \leq \frac{1}{2})$ denotes the degree of care, $\widetilde{V}_{21}$ is the utility attributed to individual 2 by 1, and $\widetilde{V}_{12}$ is the utility attributed to individual 1 by 2. In particular,

$$\widetilde{V}_{21} = \left(\widetilde{B}_{21} \left(A_1 Y_1 + \widetilde{A}_{21} Y_2 \right) - \frac{1}{2} Y_2^2 \right)$$

$$\widetilde{V}_{12} = \left(\widetilde{B}_{12} \left(\widetilde{A}_{12} Y_1 + A_2 Y_2 \right) - \frac{1}{2} Y_2^2 \right)$$

---

[49] For an excellent overview, see Ricard (2015).
[50] McAdams & Powers (1981); McAdams, (1989); Weinberger, Cotler, & Fishman (2010).
[51] McAdams, Healy, & Krause (1984).
[52] Batson et al. (1987).
[53] French, (1956, 1958); McAdams (1989); Klimecki et al. (2012).

where $\widetilde{B}_{21}$ and $\widetilde{A}_{21}$ are attributions to 2 by 1, and $\widetilde{B}_{12}$ and $\widetilde{A}_{12}$ are attributions to 1 by 2.

Regarding the degree of care, at one extreme, $\kappa = 0$ represents no care (pure self-interest), while at the other extreme, $\kappa = \frac{1}{2}$ represents perfect care (equivalent to the social welfare function).

The corresponding contribution levels are

$$
\begin{aligned}
Y_1^C &= B_1 A_1 + \frac{\kappa_1}{1 - \kappa_1} \alpha \left(\widetilde{B}_{21} A_1 \right) \\
Y_2^C &= B_2 A_2 + \frac{\kappa_2}{1 - \kappa_2} \frac{1}{\alpha} \left(\widetilde{B}_{12} A_2 \right)
\end{aligned}
$$

Obviously, the greater is the degree of care and the more accurate are the attributions $\widetilde{B}_{21}$ and $\widetilde{B}_{12}$, the closer these effort levels come to their social optimal ones. When $\kappa_1 = \kappa_2 = \frac{1}{2}$, $\widetilde{B}_{21} = B_2$ and $\widetilde{B}_{12} = B_1$, then these effort levels are identical to their socially optimal counterparts: $Y_1^C = Y_1^*$ and $Y_2^C = Y_2^*$.

Nevertheless, it is important to emphasize that the psycho-social mechanisms underlying the morality of care are quite distinct from those underlying the deontological moral principles above. Whereas the deontological principles result from the rational application of moral precepts (e.g. reward merit, reward need, defer to accept standards of social standing), the morality of care involves empathic concern and the imaginative attribution of mental states to other people. These abilities can rest on distinct capacities, such as to simulate other people's mental states, i.e. putting ourselves in other people's shoes (along the lines of "simulation theory"),[54] or to infer others' mental states from token behavior patterns (along the lines of "theory theory")[55], possibly along Bayesian lines.[56]

4.6 Interpersonal Moral Incentives in a Second-Order Co-operative Context

Thus far moral principles have been interpreted as self-imposed Pigouvian taxes and subsidies, levied on contributions to the public good. But this interpretation holds only for models in which there is no interaction among individuals in public good provision. We now consider the case in which the two individuals' contributions to the public good are strategic complements in production, so that the productivity of one individual's contribution to the public good depends positively on the contribution of the other individual. In particular, let the production function for the public good be

$$
G = Y_1 + Y_2 + E Y_1 Y_2 \tag{26}
$$

---

[54] See, for example, Batson (1991), de Vingnemont and Singer (2006), Galese, Keysers and Rizzolatti (2004) and Preston and de Wall (2002).

[55] See, for example, Bonawitz , van Schijndel, and Schulz (2012).

[56] For example, Gopnik (2011) and Ullman and Tenembaum (2020).

where E is a positive parameter representing the magnitude of the strategic complementarity. In what follows, it will be useful to interpret this public good in terms of a mutually beneficial social relationship between the two individuals, such as in a relationship of care.

Assuming again that individual i's benefit from G is $B_i G$ and that the effort cost of effort Y_i for individual i is $\frac{1}{2} Y_i^2$, then individual i's self-interested utility is

$$V_i = B_i \left(Y_1 + Y_2 + E Y_1 Y_2 \right) - \frac{1}{2} Y_i^2. \tag{27}$$

Let social welfare be

$$\Omega = V_1 + V_2 = \left(B_1 + B_2 \right) \left(Y_1 + Y_2 + E Y_1 Y_2 \right) - \frac{1}{2} Y_1^2 - \frac{1}{2} Y_2^2$$

The first-order conditions for social optimality yield the individuals' optimal contributions to the public good:

$$
\begin{aligned}
Y_1^* &= \left(B_1 + B_2 \right) \left(1 + E Y_2 \right) & (28) \\
Y_2^* &= \left(B_1 + B_2 \right) \left(1 + E Y_1 \right) & (29)
\end{aligned}
$$

We now represent the objective of each morally-driven individual in terms of psychic rewards from contributing to the public good, both individually (R_i per unit of contribution to Y_i) and in interaction with the other individual ($S_i Y_j$ per unit of contribution to Y_i, $i \neq j$), where R_i and S_i are individual and interpersonal moral incentives, respectively:

$$
\begin{aligned}
M_1 &= B_1 \left(Y_1 + Y_2 + E Y_1 Y_2 \right) + R_1 Y_1 + S_1 Y_1 Y_2 - \frac{1}{2} Y_1^2 & (30) \\
M_2 &= B_2 \left(Y_1 + Y_2 + E Y_1 Y_2 \right) + R_2 Y_2 + S_2 Y_1 Y_2 - \frac{1}{2} Y_2^2 & (31)
\end{aligned}
$$

The interpersonal moral incentives explicitly call for coordination of the individuals' contributions. This requirement reflects a common feature of moral judgments, namely, prescribing behavior not just for oneself individually, but also for one's interactions with others. The interpersonal moral incentives may be characterized as "moral reciprocity," whereby the degree of reciprocity depends on particular morally relevant characteristics (such as need or merit).

The first-order conditions for moral behavior imply that

$$
\begin{aligned}
Y_1^M &= B_1 \left(1 + E Y_2 \right) + R_1 + S_1 Y_2 & (32) \\
Y_2^M &= B_2 \left(1 + E Y_1 \right) + R_2 + S_2 Y_1 & (33)
\end{aligned}
$$

Interpreting the public good as a mutually beneficial social relationship, these conditions include the principle of moral reciprocity (the Silver Rule): each individual has a moral incentive to reciprocate the other individual's contribution to the relationship, weighted in accordance with morally relevant factors.

For individuals 1 and 2, the following values of the moral incentives ensure socially optimal behavior (so that the individuals' first-order conditions for moral behavior (equations (32) and (33)) are equivalent to the first-order conditions for social optimality (28) and (29)):

$$R_1 = B_2, \quad S_1 = EB_2 \tag{34}$$

$$R_2 = B_1, \quad S_2 = EB_1 \tag{35}$$

This means that each individual i's contribution to the public good should be determined by

- the other individual's need ($R_i = B_j$, $i \neq j$) and

- a principle of moral reciprocity, in which the degree of reciprocation depends on two morally relevant characteristics: the other individual's need (B_j) and the complementarity (EY_j).

In sum:

Proposition 3 *Strategic interdependence in the social context calls for strategic interdependence of moral incentives. In particular, a strategic complementarity in the provision of the public good calls for the inclusion of the principle of moral reciprocity into individuals' moral incentives. The degree of reciprocation depends positively on another's need and on the magnitude of the strategic complementarity.*

It is important to recognize the significance of this principle. Given the behavior of individual j, individual i has a moral incentive to respond reciprocally. This holds regardless of whether individual j is behaving in the socially optimal way.[57]

Naturally, under the assumption that both individuals are behaving optimally, it is possible to specify each individual's moral incentives in purely individualistic terms. In particular, let the moral welfare function be $V_i = B_i (Y_1 + Y_2 + EY_1Y_2) - \frac{1}{2}Y_i^2 + R_iY_i$. Then it can be shown that the socially optimal moral incentives for the two individuals are:

$$R_1 = B_2 \left(1 + E \frac{(B_1 + B_2)(1 + E(B_1 + B_2))}{1 - E^2(B_1 + B_2)} \right)$$

$$R_2 = B_1 \left(1 + E \frac{(B_1 + B_2)(1 + E(B_1 + B_2))}{1 - E^2(B_1 + B_2)} \right)$$

But this specification of moral incentives is not just impracticably complicated; it also misses an essential point: The socially desirable moral incentive calls for moral reciprocity regardless of whether the other individual is behaving optimally. In the real world, optimal behavior of others cannot be taken for granted.

---

[57] Furthermore, note that the principle of need plays a dual role here, directly and via the degree of interdependence. Since we assumed that the individuals have identical productivity and receive the same weight in the social welfare function, the principles of merit and distributive justice have no role to play.

5 Moral Behavior in a Positionally Competitive Setting

5.1 Individual Moral Incentives in a First-Order Competitive Context

Now consider two individuals interacting in a competitive setting characterized by status seeking, whereby one individual's gain generates a loss for the other individual. Ostentatious consumption is a prime example of behavior in such a setting. Individual i's effort in this status competition is y_i and the individual's ability (i.e. productivity in such competition) is a_i. For simplicity, suppose that status is measured by the difference between the two individuals' contributions. The cost of effort y_i for individual i is $\frac{1}{2}y_i^2$. Then i's utility from positional competition is

$$v_i = \pi \max\left(a_i y_i - a_j y_j, 0\right) - \varepsilon \max\left(a_j y_j - a_i y_i, 0\right) - \frac{1}{2}y_i^2, \quad i \neq j. \tag{36}$$

where π is the pride parameter $(\pi > 0)$ and ε is the envy parameter $(\varepsilon > 0)$. Boyce et al. provide evidence that $\varepsilon > \pi$. To fix ideas, we assume that individual 1 has greater ability than individual 2.

When both individuals maximize their self-interested utility, their effort levels are

$$y_1^P = \pi a_1 \tag{37}$$

$$y_2^P = \varepsilon a_2 \tag{38}$$

As above, group wellbeing is the weighted sum of the individual self-interested utilities:

$$\omega = v_1 + \alpha v_2$$

where we assume agreement on α, the weight of individual 2's wellbeing relative to that of individual 1.

Maximizing group wellbeing with respect to individual efforts,

$$y_1^* = a_1 \left(\pi - \alpha\varepsilon\right) \tag{39}$$

$$y_2^* = a_2 \left(\varepsilon - \frac{1}{\alpha}\pi\right) \tag{40}$$

Comparing (37) with (50), and (38) with (51), we see that in the competitive setting, self-interested effort on the part of each individual exceeds socially optimal effort. The reason is that status seeking behavior is asssociated with a negative externality: more status gained by one individual means less status for the other.

25

Next consider the effort decisions by moral individuals, whose behavior is driven by both self-interest and the moral incentives r_1 and r_2, for individuals 1 and 2 respectively. The individuals' utility functions are

$$m_1 = v_1 + r_1 y_1 \tag{41}$$
$$m_2 = v_2 + r_2 y_2 \tag{42}$$

Maximizing these objectives, the moral effort levels become

$$y_1^m = \pi a_1 + r_1 \tag{43}$$
$$y_2^m = \varepsilon a_2 + r_2 \tag{44}$$

These moral incentives are set at the socially optimal levels when

$$r_1^* = -\alpha \varepsilon a_1 \tag{45}$$
$$r_2^* = -\frac{1}{\alpha} \pi a_2 \tag{46}$$

Since there is no interaction among individuals on the production and consumption sides, the moral incentives can be interpreted as self-imposed Pigouvian taxes. These incentives compensate for the negative externality from status seeking.

To mitigate the damage from this negative externality, the moral incentive now represses competition among individuals. It has been argued that repression of competition is one of the major forces in the evolution of human cooperation, alongside kin selection and reciprocal altruism.[58] Repression of competition aligns aligns individual interests with group interests, because when competition has been repressed, an individual can make herself better off primarily by promoting the interests of her group. Thereby repression of within-group competition promotes group selection. Alexander (1979, 1987) argued that repression of within-group competition shaped social structures over the past millenia, favoring successful groups and limiting the opportunities for reproductive dominance. Maynard Smith (1988) provided a general formulation of the principle that group-level selection requires suppression of within-group competition.[59]

Moral principles underlying the repression of competition are well-known. Adam Smith's "impartial spectator" represents one prominent approach, namely, if a person "If he would act so as that the impartial spectator may enter into the

---

[58] See, for example, Frank (2003).

[59] "One can recognize in the evolution of life several revolutions in the way in which genetic information is organized. In each of these revolutions, there has been a conflict between selection at several levels. The achievement of individuality at the higher level has required that the disruptive effects of selection at the lower level be suppressed. (Maynard Smith 1988, pp. 229–230).

principles of his conduct, which is what of all things he has the greatest desire to do, he must, upon this, as upon all other occasions, humble the arrogance of his self-love, and bring it down to something which other men can go along with." (Smith 1996, p. 120). Another major approach is the "veil of ignorance" of Rawls' (1971), whereby a just society follows rules that individuals regard as fair behind a veil of ignorance concerning their position in society. This reasoning also suppresses selfish competition among individuals.

In this context, it is noteworthy that the composition of the socially optimal moral incentives is the same as for the cooperative context, except that now the moral incentive promotes the suppression of competition – the control of greed – rather than the expansion of cooperation in public good provision:

- The greater is an individual's ability (a_i), the greater should be the moral reward for suppressing greed. In other words, merit obligates: greater ability calls for more forebearance in pursuit of self-interest.

- The greater is an individual's need (b_i), the more the other individual is obligated to suppress greed. The morality of benevolence promotes voluntary restraint in exploiting one's position of status.

- The greater is the relative weight of an individual's wellbeing in social wellbeing (such as α for individual 2), the more the other individual needs to suppress greed and the less the individual in question needs to do so. For example, if poorer individuals receive relatively more weight in social wellbeing, then the poor individuals have less obligation to suppress their greed and the rich individuals have more obligation to do so.

It is worth noting that our model of the competitive setting (represented by equation (36)) focuses exclusively on the negative externality from status seeking: one individual's gain in status is another individual's loss. In practice, however, status competition often also has socially beneficial effects, as in the case of competition for a useful invention.

5.2   Positional Competition in Public Good Provision

Thus far, positional competition has been associated solely with negative externalities. In practice, however, it can also generate positive externalities, as when people compete in the production of public goods. Competition in the discovery and production of new vaccines and other medical break-throughs are good examples. These phenomena can be conceived as composite cases, in which the negative externalities from status-seeking are combined with the positive externalities from public good provision. Specifically, we now assume that the individual who wins the status tournament (in terms of achieving higher $a_i y_i$) produces a public good that benefits both individuals. Let a_i be individual i's expected productivity in the competition for status and let A_i and A_j be that individual's probability of winning the status tournament times the availability

of the public good for individuals i and j, respectively. Then individual i's self-interested utility be expressed as

$$v_i = b_i \left(a_i y_i - a_j y_j \right) + B_i \left(A_i y_i + A_j y_j \right) - \frac{1}{2} y_i^2, \quad i \neq j. \tag{47}$$

The self-interested effort levels are

$$y_1^P = b_1 a_1 + B_1 A_1 \tag{48}$$

$$y_2^P = b_2 a_2 + B_2 A_2 \tag{49}$$

Maximizing group wellbeing $\omega = v_1 + \alpha v_2$, the socially optimal effort levels are

$$y_1^* = (b_1 a_1 + B_1 A_1) + \alpha (A_1 B_2 - a_1 b_2) \tag{50}$$

$$y_2^* = (b_2 a_2 + B_2 A_2) + \alpha (A_2 B_1 - a_2 b_1) \tag{51}$$

When moral individuals maximize their objectives – $m_1 = v_1 + r_1 Y_1$ and $m_2 = v_2 + r_2 Y_2$ – their effort levels become

$$y_1^m = (b_1 a_1 + B_1 A_1) + r_1 \tag{52}$$

$$y_2^m = (b_2 a_2 + B_2 A_2) + r_2 \tag{53}$$

Thus the socially optimal moral incentives are

$$r_1^* = \alpha (A_1 B_2 - a_1 b_2) \tag{54}$$

$$r_2^* = \alpha (A_2 B_1 - a_2 b_1) \tag{55}$$

Note that the sign of the moral incentive for each individual depends on whether the expected benefit of the public good is greater or less than the expected cost of the quest for status.

5.3 Interpersonal Moral Incentives in a Second-Order Competitive Context

In the analysis above, positionally competitive contexts were interpreted as ones that call for status seeking. We now consider an alternative interpretation of positional competition, namely, contexts in which individuals' public good contributions are strategic substitutes. In particular, let the production function for the public good be

$$G = Y_1 + Y_2 - A Y_1 Y_2. \tag{56}$$

Let social welfare be

$$\begin{aligned} \Omega &= V_1 + V_2 \\ &= (B_1 + B_2)(Y_1 + Y_2 - AY_1Y_2) - \frac{1}{2}Y_1^2 - \frac{1}{2}Y_2^2 \end{aligned}$$

As above, let the objectives of morally driven individuals be represented by equations (12) and (13). Then the socially optimal moral incentives for individuals 1 and 2 are:

$$R_1 = B_2, \quad S_1 = -AB_2$$

$$R_2 = B_1, \quad S_2 = -AB_1.$$

Note that whereas the principle of moral reciprocity applies in the presence of a strategic complementarity, the opposite principle applies when people's contributions to the public good are strategic substitutes. This may be called the "*principle of moral redress*:" the less one individual contributes to the public good, the more another individual is obligated to contribute. This principle is not surprising, since one individual's contribution reduces the other individual's productivity in contributing to the public good. Under these circumstances, one individual's failure to contribute adequately to the public good should be "redressed" by the other individual. The degree of redress should depend on the magnitude of the strategic substitutability (AY_j) and the other's need (B_j).

In sum:

Proposition 4 *When people's contributions to the public good are strategic substitutes, individuals are enjoined to pursue the principle of moral redress, i.e. compensating for others' miserliness and taking advantage of others' generosity. The degree of redress should depends positively on others' need and the magnitude of strategic substitutability.*

As the externalities in the first- and second-order interactive contexts do not interact in our model, the optimal individual moral incentives (arising from the first-order interactive contexts) are separable from the optimal interpersonal moral incentives (arising from the second-order interactive contexts). On this account, the optimal moral incentives associated with other combinations of cooperative and competitive contexts can be straightforwardly inferred from the analysis above.

6   Application: Moral Implications of Technological Market Bias

On the basis of the analysis above, we now inquire how morlaity should evolve in response to the phenomenon of technological market bias, whereby the productivity growth generated by technological advances falls more on market activities than on prosocial relationships.[60]

---

[60] For example, Snower and Bosworth (2016).

The rationale for this phenomenon is analogous to the "Baumol effect:" The amount of time required for prosocial relationships (such as relations with family and friends) has changed much less than the amount of time to produce many commodities. For example, while the productivity in producing computing power has increased more than a million-fold over the past half century, the time required for successfully raising one's children has not fallen to anything approaching one millionth.

Just as technological change is often more closely associated with market activities than with prosocial relationships, so is positional competition. The reason is that positional competition generally requires a measuring rod whereby status may be assessed. Market activities provide such a measuring rod: prices. Prosocial activities are not only harder to measure; measurement is not relevant to their performance. Consequently, people tend to compete in terms of income and conspicuous consumption, not in terms of the affection they receive from family and friends.

Consider an economy that produces two outputs: a public good (which may be interpreted as a prosocial relationships) and a private good. For simplicity, let the private good be devoted entirely to positional competition.[61] Each individual has a fixed endowment (normalized to unity), which is divided between her contribution to the public good and her expenditure on the private good. For individual i, let the consumption of the private good be y_i and the contribution to the public good be $(1 - y_i)$.

As above, we can restrict ourselves to consideration of two individuals. The magnitude of the public good is

$$G = A_1 (1 - y_1) + A_2 (1 - y_2)$$

where A_i is the productivity of individual i.

Individual i's utility from the public good is $B_i G$. Let her direct utility from the consumption good is $\ln y_i$.

Furthermore, let individual 1 have a greater endowment than individual 2: $y_1 > y_2$. We consider two costs of inequality, one falling on the poorer individual (ε) and the other on the richer individual (γ). We may think of these costs as "envy" and "guilt" in the context of inequality aversion context of Fehr and Schmidt (1999), psychological and social costs of social fragmentation,[62] and so on.[63] Individual 1 experiences "guilt," with utility $\gamma (C_1 y_1 - C_2 y_2)$, and Individual 2 experiences "envy," with utility $\varepsilon (C_1 y_1 - C_2 y_2)$, where $C_i > 0$ is the productivity of agent i with respect to the private good and γ, ε are positive parameters.

---

[61] Provided that positional competition does not decline with technological advance, the inclusion of non-positional goods does not change the qualitative conclusions of our analysis. (See Snower and Bosworth (2016).)

[62] Examples include rising crime, falling trust and political gridlock. See, for instance, Atkinson (2015).

[63] Alternatively, we may think of the parameter γ as representing "pride," a psychic benefit gained by the winner of the positional competition. This parameter then has a negative value in our analysis.

Then the utilities of individuals are

$$
\begin{aligned}
V_1 &= B_1 \left(A_1 \left(1 - y_1 \right) + A_2 \left(1 - y_2 \right) \right) + \ln \left(C_1 y_1 \right) - \gamma \left(C_1 y_1 - C_2 y_2 \right) \\
V_2 &= B_2 \left(A_1 \left(1 - y_1 \right) + A_2 \left(1 - y_2 \right) \right) + \ln \left(C_2 y_2 \right) - \varepsilon \left(C_1 y_1 - C_2 y_2 \right)
\end{aligned}
$$

Under self-interested behavior, each individual maximizes her utility with respect to her activities, so that

$$
\begin{aligned}
\widehat{y}_1 &= \frac{1}{B_1 A_1 + C_1 \gamma} \\
\widehat{y}_2 &= \frac{1}{B_2 A_2 + C_2 \varepsilon}
\end{aligned}
$$

It can be shown that the social optimum is achieved by maximizing social welfare $(S = \alpha_1 V_1 + \alpha_2 V_2)$ with respect to y_1 and y_2:

$$
y_1^* = \frac{1}{A_1 \left(B_1 + \frac{\alpha_2}{\alpha_1} B_2 \right) + C_1 \left(\gamma + \frac{\alpha_2}{\alpha_1} \varepsilon \right)}
$$

$$
y_2^* = \frac{1}{A_2 \left(\frac{\alpha_1}{\alpha_2} B_1 + B_2 \right) - \left(\frac{\alpha_1}{\alpha_2} C_2 \gamma + C_2 \varepsilon \right)}
$$

Moreover, moral individuals have the following objectives:

$$
\begin{aligned}
V_1 &= B_1 G + \ln \left(C_1 y_1 \right) - \gamma \left(C_1 y_1 - C_2 y_2 \right) + r_1 \left(1 - y_1 \right) \\
V_2 &= B_2 G + \ln \left(C_2 y_2 \right) - \varepsilon \left(C_1 y_1 - C_2 y_2 \right) + r_2 \left(1 - y_2 \right)
\end{aligned}
$$

where r_1 and r_2 are the moral incentives of individuals 1 and 2, respectively. These individuals choose the following activity levels:

$$
\begin{aligned}
y_1^m &= \frac{1}{B_1 A_1 + C_1 \gamma + r_1} \\
y_2^m &= \frac{1}{B_2 A_2 - C_2 \varepsilon + r_2}
\end{aligned}
$$

Then it can be shown that the socially optimal moral incentives are:

$$
r_1 = \frac{\alpha_2}{\alpha_1} \left(A_1 B_2 + C_1 \varepsilon \right)
$$

$$
r_2 = \frac{\alpha_1}{\alpha_2} \left(A_2 B_1 - C_2 \gamma \right)
$$

Interpreting the parameters γ and ε in terms of the social costs of inequality, we find that:

- the greater the social cost of inequality falling on the advantaged agent (i.e. the greater is γ), the more the disadvantaged agent is relieved of moral responsibility to contribute to the public good, and

- the greater the social cost of inequality falling on the disadvantaged agent (i.e. the greater is ε), the more the greater is the moral responsibility of the advantaged agent to contribute to the public good.

In accordance with the technological market bias and the positional competition bias hypotheses, technological advance can be captured primarily by a rise in the parameter C_i, representing the productivity in producing the positional good.[64]

This implies a rise in the moral incentive r_1 and a fall in the moral incentive r_2. Insofar was the technological advance falls primarily on the better-endowed individual, the rise in r_1 will exceed the fall in r_2 in absolute magnitude. In words, for social optimality to be achieved, the richer individual's moral responsibility for contributing to the prosocial relationship rises continually, whereas the poorer individual's moral responsibility for such contributions falls continually.

Insofar as technological advance falls on the prosocial relationships as well, this development is captured by a rise in the parameter A_i, putting upward pressure on the moral responsibility of both agents, in proportion to their productivity in generating these prosocial relationships.

7   Concluding Thoughts

To gain perspective, it is useful to place our contribution in the broad sweep of moral thinking. Morality appears to have embedded in human thinking and decision making since the beginnings of humanity. Around 400,000 years ago Homo Heidelbergensis began collaborative hunting and foraging.[65] At this time, collaboration must have become obligate, since it was vital for survival. Socially cooperative individuals had better chances of finding a mate and thereby cooperation emerged through social selection. This social cooperation, build on joint intentionality, was the source of morality. Around 150,000 years ago, growing tribal groups began splitting into smaller units. Individuals must have felt strong affiliation with their sub-group and somewhat weaker affiliation with their super-group, which needed to be distinguished from other tribes with which one may be in competition. These multilevel affiliations were the beginnings of "culture," from which elaborate moral systems grew. With the invention of agriculture, people formed larger, sedentary groups that had sophisticated divisions of labor and responsibility, reinforced by "objective moralities" whereby people shared

---

[64] Insofar as technological advance falls on the prosocial relationships as well, it is captured by a rise in the parameter A_i.

[65] Stiner and Kuhn (2006).

common conceptions of being both judgers and judged.[66] These moral systems could be understood in terms of highly context-dependent behavioral rules.

During the Axial Age (from around 800 to 200 BCE), many societies developed systems of virtue ethics,[67] which encompassed not just rules of behavior, but also worldviews that shaped perceptions, intuitions and emotions in terms of the relevant physical and social contexts. Within these contexts, people aspired to multiple virtues, to be achieved through ongoing practice and the development of the relevant skills.[68] Virtue-based ethical systems remained dominant throughout the Middle Ages and are enjoying a revival in philosophy nowadays.[69]

In the 18th century, philosophers of the European Enlightenment began constructing ethical systems from secular first principles. The two leading approaches that emerged were deontology (judging actions in terms of whether they are right or wrong) and consequentialism (judging actions in terms of their consequences). Unlike virtue ethics, the underlying principles here were abstract and universal, based on logical reasoning or calculation. Morality became focused on "What should I do?", rather than "Who should I become?"

Moral psychology was narrowed along analogous lines, focusing primarily on quandries concerning deontological versus consequentialist choices (e.g. whether or not to throw a switch that would divert a trolley to kill one person in order to save five others)[70] and justice versus care. For example, Kohlberg (1969) claimed that moral development in children was about the development of reasoning concerning justice, whereas Gilligan (1981) claimed that such development included the "ethic of care."

Meanwhile anthropologists and cultural psychologists focused on the distinction between individualism and collectivism in the formation of identity.[71] With regard to social affiliations, it is useful to distinguish between what Tönnies (2001/1887) called "Gemeinschaft" (community) and "Gesellschaft" (civil society). The former involves thick ties in communities of place and belief, whereas the latter refers to the thin ties among people who live in close proximity but make their own choices, provided that they don't harm others. Moral psychology has been mainly about the obligations that create Gesellschaft, such as respecting the rights of others, not harming others and helping people in need.

More recently, moral and evolutionary psychologists have focused on the function of morality, namely, to promote cooperation among people (particularly in in-groups) and suppress destructive self-interest.[72] This is the approach pursued in this paper. For this purpose, we began our analysis by considering a public good that is intrinsically worthwhile (without specifying the rationale for

---

[66] Tomasello (2018).
[67] Aristotle (1941).
[68] MacIntyre (1981), Churchland (1998), McDowell (1979).
[69] For example, Chappell (1996) and Crisp (1996).
[70] For example, Greene (2008), Hauser (2006).
[71] For example, Shweder and Bourne (1984) and Triandis (1995).
[72] For example, Haidt (2007).

this assessment) and a social welfare function that specifies how the individual objectives of different people are to be weighted in assessing social wellbeing. On this basis, we investigated how people need to cooperate for the optimal provision of the public good. Finally, we derived the moral incentives that lead to this socially desirable outcome.

We then conduct an analogous line of inquiry with regard to positional competition. The underlying insight is that social cooperation is often undermined through positional competition. In fact, many of our moral successes and understood in terms of our capacity to promote social cooperation while suppressing our drive to seek status. With regard to our social welfare function, we derived the moral incentives that lead to the socially optimal suppression of positional competitiveness.

It is commonly observed that economic and social decisions in our daily lives are riddled with moral dilemmas and it is common to feel the tug of conflicting moral demands. For example, should people be supported in accordance with their need or rewarded in accordance with their merit? How should we weigh the rights of individuals and communities in decisions on public infrastructure? To what degree should income be redistributed when redistribution hurts economic efficiency? Is theft condonable if the needs of the perpetrators are substantially greater than those of the victims? What are the moral limits of markets? Should we be able to sell our organs, pay mercenaries to fight our wars, transact pollution rights, sell citizenship to immigrants, pay for basic health care, and so on?

A reason why we perceive these and other issues to be moral dilemmas is that we generally assume moral principles as being mutually exclusive when they pull in contrary directions. In other words, we adopt an "all or nothing" approach to moral principles, e.g. we follow either the ethic of care or the ethic of justice. Another reason why we are so vulnerable to moral dilemmas is our Enlightenment-driven assumption that moral principles are to be abstract and universal, not context-dependent.

Our analysis calls both these assumptions into question. The socially desirable moral motives that we derive can be interpreted in terms of context-dependent combinations of deontological moral principles. Different social contexts generate generate different collective action problems, which call for different combinations of deontological moral principles. These principles are particularly applicable in the context of Gesellschaft, wherein the members of civil society have little intimate knowledge of each other's sources of wellbeing.

In the context of Gemeinschaft, where people can engage in the requisite perspective-taking concerning the wellbeing of other group members, the optimal provision of the public good and the optimal suppression of positional competitiveness can be achieved through the requisite degree of perspective-taking combined with the morality of care (conceived in consequentialist terms).

In short, our analysis builds bridges between the major approaches to morality – deontology, consequentialism and virtue ethics – and between various deontological principles:

- Deontological principles need to be combined in context-dependent ways to achieve socially desirable outcomes for collective action problems.

- The deontological and consequentialist approaches to morality are alternative routes to achieving desirable social outcomes, but under different informational conditions.

- The context-dependence of socially desirable moral principles harkens back to the essential features of virtue ethics: One is meant to pursue multiple virtues and the relevant virtues depend on our physical and cultural setting as well as our social roles within our social hierarchies. In practice, the appropriate combination of moral principles is something that can be learnt only through practice, following in the footsteps of moral exemplars.

The main contribution of our analysis is to analyze the link between social contexts, socially desirable moral motives and decision making. The exploration of a wider range of social contexts and their moral implications is a subject for future research.

Our analysis suggests that including moral motives in the analysis of economic decisions serves several important purposes: it gives useful insights for understanding how people behave in addressing their collective action problems, it provides guidelines for how people ought to apply moral principles in response to new collective action problems, and it provides a starting for rethinking the division of responsibility between government and civil society.

8   References

References

[1] Adolphs, R. (2013). The biology of fear. Current Biology, 23(2), R79–93. doi:10.1016/j.cub.2012.11.055

[2] Aristotle (1941), Nichomachean Ethics (trans. By W.D. Ross), New York: Random House.

[3] Atkinson, A. (2015), Inequality: What Can Be Done? Cambridge, MA: Harvard University Press.

[4] Atkinson, J. W. (1964). An Introduction to Motivation. Princeton: Van Nostrand.

[5] Atkinson, J. W., Heyns, R. W., & Veroff, J. (1954). The effect of experimental arousal of the affiliation motive on thematic apperception. The Journal of Abnormal and Social Psychology, 49(3), 405.

[6] Batson, C. D. (1991). The altruism question: Toward a social-psychological answer. Lawrence Erlbaum Associates, Inc.

[7] Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and Empathy: Two Qualitatively Distinct Vicarious Emotions with Different Motivational Consequences. Journal of Personality, 55(1), 19–39. doi:10.1111/j.1467-6494.1987.tb00426.x

[8] Batson, C. D., O'Quin, K., Fultz J., Vanderplas M., & Isen A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. Journal of Personality and Social Psychology, 45(3), 706–718. doi:10.1037/0022-3514.45.3.706

[9] Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. Psychological bulletin, 117(3), 497.

[10] Bonawitz EB, van Schijndel TJ, Friel D, Schulz L (June 2012). "Children balance theories and evidence in exploration, explanation, and learning" (PDF). Cognitive Psychology. 64 (4): 215–234.

[11] Boyatzis, R. E. (1973). Affiliation motivation. Human motivation: A book of readings, 252-276.

[12] Boyce, C. J., Brown, G. D. A., & Moore, S. C. (2010). "Money and happiness: Rank of income, not income, affects life satisfaction." Psychological Science, 21(4), 471-475.

[13] Chappell, T.D.J. (1996), Values and Virtues: Aristotelianism in Contemporary Ethics, Oxford: Oxford University Press.

[14] Christakis, N. (2019); Blueprint: The Evolutionary Origins of a Good Society, N.Y, N.Y.: Little, Brown Spark.

[15] Churchland, P.M. (1998), "Toward a Cognitive Neurobiology of the Moral Virtues," Topoi, 17, 83-96.

[16] Crisp, R. (ed) (1996), How Should One Live? Essays on the Virtues, Oxford: Oxford University Press.

[17] de Vignemont, F. & Singer, T. (2006). "The empathic brain: How, when and why?" (PDF). Trends in Cognitive Sciences. 10 (10): 435–441.

[18] Deutsch, M. (1975), „Equity, Equality and Need: What Determines Which Value Will Be Used as the Basis of Distributive Justice?" Jourjal of Social Issues, 31, 137-149.

[19] Eimontaite, I., Nicolle, A., Schindler, I., & Goel, V. (2013). The effect of partner-directed emotion in social exchange decision-making. Frontiers in Psychology, 4, 469. doi:10.3389/fpsyg.2013.00469

[20] Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. Science, 302(5643), 290-292.

[21] Elliot, A. J., & Covington, M. V. (2001). Approach and avoidance motivation. Educational Psychology Review, 13(2), 73-92.

[22] Exline, R. V. (1962). Effects of need for affiliation, sex, and the sight of others upon initial communications in problem solving groups. Journal of Personality, 30(4), 541-556.

[23] Fehr, E. and K.M. Schmidt (1999), "A Theory of Fairness, Competition and Cooperation," Quarterly Journal of Economics, 114(3), 817-868.

[24] Fessler, D. M. T., Pillsworth, E. G., & Flamson, T. J. (2004). Angry men and disgusted women: An evolutionary approach to the influence of emotions on risk taking. Organizational Behavior and Human Decision Processes, 95(1), 107–123. doi:10.1016/j.obhdp.2004.06.006

[25] Fiske, A.P. (1991), Structure of Social Life, New York: Free Press.

[26] Fiske, A.P. (1992), "The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations," Psychological Review, 99, 689-723.

[27] Fiske, A.P., and Haslam, N. (1997), "The Stuructre of Social Substitutions: A Test of Relational Models Theory," European Journal of Social Psychology, 27, 725-729.

[28] Fleurbaey, M. (2004), "Normative economics and theories of distributive justice," In John Bryan Davis & Alain Marciano (eds.), The Elgar Companion to Economics and Philosophy. Edward Elgar. pp. 132

[29] Folger, R., B.H. Sheppard, and R.T. Buttram (1995), "Equity, Equality and Need: Three Faces of Social Justice," in B. Bunker and J.Z. Zucker (eds), Conflict, Cooperation and Justice, San Francisco, CA: Jossey-Bass, 261-289.

[30] Fox, C.R., and D. Kahneman (1992), "Correlations, causes and heuristics in surveys of life satisfaction," Social Indicators Research, 27, 221–234, https://doi.org/10.1007/BF00300462.

[31] Frank, Steven A. (2003), "Repression of Competition and the Evolution of Cooperation," Evolution, 57(4), 693-705.

[32] French, E. G. (1956). Motivation as a variable in work-partner selection. The Journal of Abnormal and Social Psychology, 53(1), 96.

[33] Gallese, V.; Keysers, C. & Rizzolatti, G. (2004). "A unifying view of the basis of social cognition". Trends in Cognitive Sciences. 8 (9): 396–403.

[34] Gilligan, C. (1981), In a Different Voice: Psychological Theory and Women's Development, Cambridge, MA: Harvard University Press.

[35] Gilligan, C. (1982), In a Different Voice: Psychological Theory and Women's Development, Cambridge: Harvard University Press.

[36] Gintis, H. (2009), Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences, Princeton, NJ: Princeton University Press.

[37] Gopnik, A. (2011), The Theory Theory 2.0: Probabilistic Models and Cognitive Development. Child Development Perspectives, 5: 161-163.

[38] Greene (2013), Moral Tribes, London: Atlantic Books.

[39] Greene, J. (2013), Moral Tribes: Emotion, Reason, and the Gap Between Us and Them, London: Atlantic Books.

[40] Haidt, J. (2007), "The New Synthesis in Moral Psychology," Science, 316, 998-1002.

[41] Hall, J. L., Stanton, S. J., & Schultheiss, O. C. (2010). Biopsychological and neural processes of implicit motivation. In O. Schultheiss & J. Brunstein (Eds.), Implicit motives (pp. 279-307). Oxford: Oxford University Press.

[42] Hall, Stanton & Schultheiss, 2010

[43] Haslam, N. (1994), "Categories of Social Relationship," Cognition, 53, 59-90.

[44] Haslam, N. (2004), Relational Models Theory: A Contemporary Overview, Mahwah, NJ: Erlbaum.

[45] Haslam, N., and A.P. Fiske (1992), "Implicit Relational Prototypes: Investigating Five Theories of Cognitive Organization of Social Relationships, Journal of Experimental Social Psychology, 28, 441-474.

[46] Hauser (2006), Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong, New York: Harper Collins.

[47] Henrich, J. (2016), The Secret of Our Success, Princeton, N.J.: Princeton University Press.

[48] Henrich, J. (2020), The Weirdest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous, London: Penguin.

[49] Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. Evolution and Human Behavior, 22(3), 165-196. doi:10.1016/S1090-5138(00)00071-4

[50] Hofstede, G. 2001. Culture's consequences: Comparing values, behaviors, institutions and organizations across nations. Thousand Oaks, CA: Sage.

[51] Jack Hirshleifer (1985), "The Expanding Domain of Economics," American Economic Review, 75 (6), 53-69.

[52] Klimecki, O. M., Leiberg, S., Lamm, C., & Singer, T. (2013). Functional Neural Plasticity and Associated Changes in Positive Affect After Compassion Training. Cerebral Cortex, 23(7), 1552–1561. doi:10.1093/cercor/bhs142

[53] Klimecki, O. M., Leiberg, S., Ricard, M., & Singer, T. (2014). Differential pattern of functional brain plasticity after compassion and empathy training. Social Cognitive and Affective Neuroscience, 9(6), 873–879. doi:10.1093/scan/nst060

[54] Kohlberg (1969), "Stage and Sequence: The Cognitive-Developmental Approach to Socialization," in D.A. Goslin (ed), Handbook of Socialization Theory and Research, Chicago: Rand McNally.

[55] Kornadt, H. J. (1984). Motivation theory of aggression and its relation to social psychological approaches. In A. Mummendey (Ed.), Social psychology of aggression (pp. 21-31). Berlin/Heidelberg: Springer.

[56] Lansing, J. B., & Heyns, R. W. (1959). Need affiliation and frequency of four types of communication. The Journal of Abnormal and Social Psychology, 58(3), 365

[57] LeDoux, J. (1998). The emotional brain: The mysterious underpinnings of emotional life. Simon and Schuster.

[58] Leith, K. P., & Baumeister, R. F. (1996). Why Do Bad Moods Increase Self-Defeating Behavior? Emotion, risk taking, and self-regulation, Journal of Personality and Social Psychology, 71(6), 1250–1267.

[59] Li, N.P., M. van Vugt, and S.M. Colarelli (2018), "The Evolutionary Mismatch Hypothesis: Implications for Psychological Science," Current Directions in Psychological Science, 27(1), 38-44. https://journals.sagepub.com/doi/abs/10.1177/0963721417731378?journalCode=cdpa#:~:text=Evolutio

[60] Lowell, 1952;

[61] Lowell, E. L. (1952). The effect of need for achievement on learning and speed of performance. The Journal of Psychology, 33(1), 31-40.

[62] MacIntyre, A. (1981), After Virtue, Notre Dame: University of Notre Dame Press.

[63] MacIntyre, A. (1981), After Virtue, Notre Dame: University of Notre Dame Press.

[64] Maynard Smith, J. (1988), "Evolutionary progress and levels of selection." Pp. 219–230 in M. H. Nitecki, ed. Evolutionary progress. Univ. of Chicago Press, Chicago, IL.

[65] McAdams, D. P. (1989). Intimacy: The need to be close. Doubleday & Co.

[66] McAdams, D. P., & Constantian, C. A. (1983). Intimacy and affiliation motives in daily living: An experience sampling analysis. Journal of personality and social psychology, 45(4), 851.

[67] McAdams, D. P., & Powers, J. (1981). Themes of intimacy in behavior and thought. Journal of Personality and Social Psychology, 40(3), 573.

[68] McAdams, D. P., Healy, S., & Krause, S. (1984). Social motives and patterns of friendship. Journal of Personality and Social Psychology, 47(4), 828.

[69] McAdams, Healy, & Krause, 1984

[70] McClelland, D. C. (1985). How motives, skills, and values determine what people do. American Psychologist, 40(7), 812.

[71] McDowell, J. (1979), "Virtue and Reason," Monist, 62, 331-350.

[72] McGuire, L., Rizzo, M. T., Killen, M., & Rutland, A. (2018), "The development of intergroup resource allocation: The role of cooperative and competitive in-group norms," Developmental Psychology, 54(8), 1499–1506. https://doi.org/10.1037/dev0000535

[73] Mikulincer, M., Shaver, P. R., Gillath, O., & Nitzberg, R. A. (2005). Attachment, caregiving, and altruism: boosting attachment security increases compassion and helping. Journal of personality and social psychology, 89(5), 817.

[74] Mikulincer, M., Shaver, P. R., Gillath, O., & Nitzberg, R. A. (2005). Attachment, caregiving, and altruism: boosting attachment security increases compassion and helping. Journal of personality and social psychology, 89(5), 817.

[75] Nisbett, R.E., and D. Cohen (2004), Culture of Honor: The Psychology of Violence in the South, Boulder, CO: Westview Press.

[76] Pang, J. S. (2010). The achievement motive: a review of theory and assessment of n Achievement, hope of success, and fear of failure. In O. Schultheiss & J. Brunstein (Eds.), Implicit motives (pp. 30-71). Oxford: Oxford University Press

[77] Parks CD, and AD Vu (1994), "Social Dilemma Behavior of Individuals from Highly Individualist and Collectivist Cultures," Journal of Conflict Resolution. 1994;38(4):708-718. doi:10.1177/0022002794038004006

[78] Preston, S.D. & de Wall, F.B.M. (2002). "Empathy: its ultimate and proximate bases" (PDF). Behavioral and Brain Sciences. 25 (1): 1–20, discussion 20–71.

[79] Rawls, J. 1971. A Theory of Justice. Harvard Univ. Press, Cambridge, MA.

[80] Reiss (2004)

[81] Reiss, S. (2004). Multifaceted nature of intrinsic motivation: The theory of 16 basic desires. Review of General Psychology, 8(3), 179.

[82] Ricard, M. (2015), Altruism (trans. By C. Mandell and S. Gordon), New York: Little, Brown and Company.

[83] Robbins, L. C. (1932). An Essay on the Nature and Significance of Economic Science. London: Macmillan.

[84] Sandel, M. (2020), The Tyranny of Merit, London, Allen Lane.

[85] Schultheiss, O. C., & Brunstein, J. C. (2001). Assessment of implicit motives with a research version of the TAT: Picture profiles, gender differences, and relations to other personality measures. Journal of personality assessment, 77(1), 71-86.

[86] Selman (2008), "Through Thick and Thin," Human Development, 51, 318-325. https://www.jstor.org/stable/26764873

[87] Shweder, R.A., and E. Bourne (1984), "Does the Concept of the Person Vary Cross-Culturally?" in R. Shweder and R. LeVine (eds), Culture Theory, Cambridge: Cambridge University Press, 158-199.

[88] Singer, T. et al. (2006). Empathic neural responses are modulated by the perceived fairness of others. Nature, 439, 466–469.

[89] Snower, D.J., and S. Bosworth (2016), "Identity-Driven Cooperation versus Competition", American Economic Review, Papers and Proceedings, 2016. http://hdl.handle.net/10419/130402

[90] Sober, E., and D.S. Wilson (1998), Unto Others: The Evolution and Psychology of Unselfish Behavior, Cambridge, MA: Harvard University Press.

[91] Small, D. A., & Lerner, J. S. (2008). Emotional Policy: Personal Sadness and Anger Shape Judgments abou t a Welfare Case. Political Psychology, 29(2), 149–168. doi:10.1111/j.1467-9221.2008.00621.

[92] Smith, Adam (1759/1996), The Theory of Moral Sentiments. Sentry Press, New York. Reprint of 1853 ed., published by Henry G. Bohn, London, first published in 1759.

[93] Stemmler, G., Heldmann, M., Pauls, C. A., & Scherer, T. (2001). Constraints for emotion specificity in fear and anger: The context counts. Psychophysiology, 38(2), 275–291. doi:10.1111/1469-8986.3820275

[94] Stiner, M.C., and S. Kuhn (2006), "Changes in the 'Connectedness' and Resilience of Paleolithic Societies in Mediterranean Ecosystems," Human Ecology, 34(5), 693-712.

[95] Tangney, J.P, and R.L. Dearling (2002), Shame and Guilt, London: Guilford Press.

[96] Tedeschi, J., Smith, R., & Brown, R (1974). A reinterpretation of research on aggression. Psychological Bulletin, 81(9), 540–562.

[97] Tomasello, M. (2016), A Natural History of Human Morality, Cambridge, MA: Harvard University Press.

[98] Tönnies, F. (2001/1887), Community and Civil Society (trans. By J. Harris and M. Hollis), Cambridge: Cambridge University Press.

[99] Triandis, H. (2002), "Individualism-Collectivism and Personality," Journal of Personality, 69(6), 907-924.

[100] Triandis, H.C. (1995), Individualism and Collectivism, Boulder, CO: Westview.

[101] Ullman, Timer D; Tenembaum, Joshua B (2020). "Bayesian Models of Conceptual Development: Learning as Building Models of the World". Annual Review of Developmental Psychology. 2 (1): 553–558.

[102] Walker, E. L., & Heyns, R. W. (1962). An anatomy for conformity. Englewood Cliffs.

[103] Walzer, M. (1994), Thick and Thin: Moral Argument at Home and Abroad, Notre Dame, IN: University of Notre Dame Press.

[104] Wedekind C, Milinski M. Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus Generous Tit-for-Tat. Proc Natl Acad Sci U S A. 1996;93(7):2686-2689. doi:10.1073/pnas.93.7.2686. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC39691/

[105] Weinberger, J. Cotler, T. & Fishman, D. (2010). The duality of affiliative motivation. In O. Schultheiss & J. Brunstein (Eds.), Implicit motives. Oxford: Oxford University Press.

[106] Weinberger, J. Cotler, T. & Fishman, D. (2010). The duality of affiliative motivation. In O. Schultheiss & J. Brunstein (Eds.), Implicit motives. Oxford: Oxford University Press.

[107] Wilson, D.S. (2015), Does Altruism Exist? New Haven: Templeton Press.

[108] Winter, D. G. (1973). The power motive. New York: Free Press.