

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Hoven, Noel Z.; Nyborg, Karine

## Working Paper The mixed effects of making contribution behaviors observable

Memorandum, No. 01/2025

**Provided in Cooperation with:** Department of Economics, University of Oslo

*Suggested Citation:* Hoven, Noel Z.; Nyborg, Karine (2025) : The mixed effects of making contribution behaviors observable, Memorandum, No. 01/2025, University of Oslo, Department of Economics, Oslo

This Version is available at: https://hdl.handle.net/10419/311630

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



# WWW.ECONSTOR.EU

# MEMORANDUM

No 01/2025 January 2025

The mixed effects of making contribution behaviors observable



This series is University of Department	published by the f Oslo of Economics	In co-operatio The Frisch Co Research	n with <b>entre for Economic</b>
P. O.Box 1095 Blindern		Gaustadalleén 21	
N-0317 OSLO Norway		N-0371 OSLO Norway	
Telephone: +47 22855127		Telephone:	+47 22 95 88 20
Fax: -	+ 47 22855035	Fax:	+47 22 95 88 25
Internet: h	ttp://www.sv.uio.no/econ	Internet:	http://www.frisch.uio.no
e-mail: e	condep@econ.uio.no	e-mail:	frisch@frisch.uio.no

## Last 10 Memoranda

No 03/2024	Effects of Compressed Work Schedules on Sickness Absence Elisabeth Fevang, Andreas Fidjeland, Karen Hauge, and Otto Lillebø		
No 02/2024	Tore Adam Reiremo The Anatomy of Mismatch in the College Market		
No 01/2024	Gunnar Bårdsen and Ragnar Nymoen U.S. wage-price dynamics, before, during and after COVID-19, through the lens of an empirical econometric model		
No 04/2023	Andreas Bjerre-Nielsen, Lykke Sterll Christensen, Mikkel Høst Gandil and Hans Henrik Sievertsen Playing the system: address manipulation and access to schools		
No 03/2023	Gunnar Bårdsen and Ragnar Nymoen Dynamic time series modelling and forecasting of COVID-19 in Norway		
No 02/2023	Hans A. Holter, Serhiy Stepanchuk and Yicheng Wang Taxation and Entrepreneurship in the United States		
No 01/2023	Pedro Brinca, João B. Duarte, Hans A. Holter and João G. Oliveira Technological Change and Earnings Inequality in the U.S.: Implications for Optimal Taxation		
No 05/2022	Bård Harstad A theory of pledge-and-review bargaining		
No 04/2022	Michael Graber, Magne Mogstad, Gaute Torsvik and Ola Vestad Behavioural responses to income taxation in Norway		
No 03/2022	Eskil Heinesen, Christian Hvid, Lars Kirkebøen, Edwin Leuven and Magne Mogstad Instrumental variables with unordered treatments: Theory and evidence from returns to fields of study		

Previous issues of the memo-series are available in a PDF® format at:http://www.sv.uio.no/econ/english/research/unpublished-works/working-papers/

## The mixed effects of making contribution behaviors observable

Noel Z. Hoven<sup>\*</sup>, Karine Nyborg<sup>†</sup>

2023

### Abstract

Voluntary contributions to public goods have been shown to increase when contribution behaviors become observable by peers. We examine the effect of social pressure on moral behaviour, using a framework distinguishing explicitly between observable and nonobservable forms of contribution. We show that even with moral motivation and social pressure, there is underprovision of the public good. Making some contribution behaviors observable does increase effort, but also causes misallocation of effort between observable and non-observable effort types. The latter negative effect can possibly outweigh the positive effect of increased effort on public good provision.

#### Keywords

Moral motivation  $\cdot$  Behavioural economics  $\cdot$  Warm glow  $\cdot$  Modeling visible altruism

#### Acknowledgements

Thanks to Snorre Kverndokk and Tone Ognedal for helpful suggestions. This work is partly based on Hoven (2022).

<sup>\*</sup>Noel.hoven@hotmail.com, ORCID: 0009-0003-9481-0097

<sup>&</sup>lt;sup>†</sup>Department of Economics, University of Oslo. Karine.nyborg@econ.uio.no, ORCID: 0000-0002-0359-548X

#### 1 Introduction

When one incurs inconvenience in the name of doing what is "right", is it to satisfy oneself or one's peers? As ever more greenhouse gasses are emitted (Olhoff and Christensen, 2021) and waste is generated (Kaza et al., 2018, p. 25), understanding people's motives for voluntary contributions to public goods become increasingly relevant.

Here, we propose a simple economic model capturing individuals' motivation to act morally while also being affected by social pressure. Our model predicts that a policy increasing the visibility of a specific contribution behavior does indeed increase that specific contribution behavior, but that there is a crowding-out effect as well: increased contributions of more visible behaviours come at the cost of decreased invisible effort, causing a less socially efficient combination of effort types in terms of public good production.

It is well documented that a non-negligible share of individuals are willing to contribute to public goods even under conditions of full anonymity (see, e.g., Chaudhuri, 2011). When public good provision cannot be observed by others, however, contributions must be motivated solely by internalized factors such as intrinsic altruism and moral motivation. A natural presumption, thus, is that voluntary contributions can be increased by making behaviors more easily observable.

Several empirical studies support the idea that once a given contribution behavior becomes more easily observable, participation in this same behavior increases. For example, Dannenberg, Johanson-Stenman, and Wetzel (2022) found that cinema-goers were significantly more likely to donate more to an art cinema if they were offered a "social status gift", such as a mug or a bag with text that broadcasted their generosity, than if they were offered the same gift but without the text marking them as a generous patron of the cinema. Similarly, Yoeli et al. (2013) found that participation in an electric utility's blackout prevention program increased substantially when one's neighbors could observe one's participation. However, as we argue below, if individuals substitute their efforts away from non-observable to observable efforts, making the composition of effort types less socially efficient as a result, increased visibility does not necessarily increase overall contributions. Models like Andreoni's impure altruism (1990) explain altruistic behaviour through the mechanism of a "warm glow", whereby the mere act of contributing voluntarily provides a private benefit to the individual. Further developing this, Brekke, Kverndokk and Nyborg (2003) propose a model of moral motivation in which the individual experiences increasing disutility, or cognitive dissonance, the further removed her actual behaviour is from her view of the morally ideal behavior. This ideal, in turn, is determined by asking oneself what would maximize social welfare if, hypothetically, everyone acted just like oneself, which can be viewed as based on the Kantian Categorical Imperative.

Our formal framework is based on the model of Brekke et al. (2003), adjusted to encompass the distinction between visible and non-visible efforts as well as a preference for social approval. In our model, the individual experiences increasing disutility as her actual behaviour strays from her own moral ideal, and also as her behaviour *in the eyes of others* strays from their ideal. Since others can only judge the individual based on what they observe, the inclusion of a social motivation for moral behaviour introduces an element of incomplete information.

Our analysis contributes to the literature on voluntary public good provision (Bergstrom et al., 1986; Andreoni (1988, 1990); Kahneman and Knetsch, 1992; Bénabou and Tirole, 2006; Brekke and Nyborg, 2008; Brekke and Nyborg, 2010; Chaudhuri, 2011; Ek, 2018). The inclusion of a preference for social approval also links our analysis to the extensive literatures within economics, psychology, philosophy and sociology on social norms (see, e.g., Farrow et al., 2017; Nyborg, 2018, 2020; Kahneman et al., 1986; Sugden, 2000; Cialdini et al., 1990, Bicchieri, 2006; Coleman, 1990). Young (2015, p. 360) describes social norms as "the unwritten codes and informal understandings that define what we expect of other people and what they expect of us." Social norms are generally taken to be enforced by the approval or disapproval of other people (Elster, 1989).

To our knowledge, however, none of the above-mentioned formal analyses consider specifically the potential substitution between observable and non-observable efforts. If the behaviour a norm concerns is not visible to others, it cannot be subjected to external enforcement: others are unable to distinguish between behaviour that is imperceptible to them and the absence of this behaviour. The analysis we are aware of that is closest to ours is an unpublished working paper by Fuhai Hong, Jean Tirole, and Chen Zhang (2023) (developed simultaneously with and independently of ours). Although their formal framework differs in several respects from ours, they find, like us, that observable effort crowds out non-observable efforts; moreover, they present experimental evidence supporting this prediction.

#### 2 A model of visible and invisible effort

#### 2.1 The setup

Our modelling framework follows Brekke et al. (2003) closely, except for two features: first, our distinction between visible effort, which can readily be observed by others, and invisible effort, observable only by the individual herself; second, our assumption that the individual's utility depends on social validation.

Consider a society of N identical individuals. The utility of individual i is given by

$$U_i = u\left(x_i, l_i, G, I_i, V_i\right) \tag{1}$$

where  $x_i$  is *i*'s consumption of private goods,  $l_i$  is her leisure time, *G* is the supply of a public good,  $I_i$  is the individual's self-image, and  $V_i$  is the her social image or external validation, encompassing the social rewards and punishments given to the individual by other individuals for her behaviour. The utility function is assumed to be strictly increasing and quasi-concave in all of  $(x_i, l_i, G, I_i, V_i)$ , and is linearly separable in  $I_i$  and  $V_i$ .

Like Brekke et al. (2003), we assume that the individuals' supply of labour, income, and private consumption are exogenous, allowing us to focus on the allocation of time between leisure and efforts to contribute to the public good. Each individual's time constraint is

$$l_i + hv_i + lv_i = T \tag{2}$$

where  $hv_i$  is *i*'s time spent on visible effort towards provision of the public good;  $lv_i$  is time spent on invisible effort; and *T* is the total time available to the individual after accounting for labour and sleep. Private contributions to the public good can come from either visible or invisible effort. Let  $\phi$  be a production function capturing the relationship between the individuals' effort, both visible and invisible, and the resulting provision of the public good:

$$g_i = \phi\left(hv_i, lv_i\right) \tag{3}$$

It is assumed that  $\phi(0, 0) = 0$ ,  $\frac{\partial \phi}{\partial hv} > 0$ ,  $\frac{\partial^2 \phi}{\partial hv^2} < 0$ ,  $\frac{\partial \phi}{\partial lv} > 0$ ,  $\frac{\partial^2 \phi}{\partial lv^2} < 0$ , and, for simplicity,  $\frac{\partial^2 \phi}{\partial hv \partial lv} = \frac{\partial^2 \phi}{\partial lv \partial hv} = 0$ . This is to say that  $\phi$  is additively separable and strictly increasing but with diminishing returns in both hv and lv. We also assume that one form of effort is not completely superior to the other, so that the most efficient investment of time required to reach a certain level of  $g_i$  will necessarily be some combination of both  $hv_i$ and  $lv_i$ .

Disregarding any exogenous provision by nature and/or the government, total public good supply is given by

$$G = \Sigma_i g_i \tag{4}$$

summing over all i = 1, ..., N.

As in Brekke et al. (2003), individual i's self-image depends on the distance between i's actual behaviour and i's view of the morally ideal behaviour, defined as the behaviour that would benefit society the most if everyone acted exactly like i. Here, however, we assume that self-image depends not on effort levels as such, but on the distance between i's actual provision of the public good and a morally ideal provision. Thus, i's self-image will depend not only on the amount of effort she invests in providing the public good, but also on how efficient that effort is. We assume that self-image is given by

$$I_{i} = f(g_{i}, g_{i}^{*}) = -a(g_{i} - g_{i}^{*})^{2}, \ a > 0$$
(5)

where  $g_i^*$  is i's endogenously determined morally ideal contribution to the public good.

As  $lv_i$  and  $hv_i$  have diminishing returns to scale, there must exist a pair of ideal effort levels,  $lv_i^*$  and  $hv_i^*$ , that correspond to

$$g_i^* = \phi\left(hv_i^*, lv_i^*\right) \tag{6}$$

such that  $g_i^*$  maximizes a welfare function, assumed to be utilitarian for simplicity,

$$W = \sum_{i=1}^{N} U_i \tag{7}$$

subject to equations (1)–(5) and (8) given that  $hv_i = hv_j$  and  $lv_i = lv_j$  for all  $j \neq i$ , j = 1, ..., N. We return to the determination of  $g_i^*$ ,  $hv_i^*$  and  $lv_i^*$  in the Results section below. In the following we assume that there is an interior solution to this maximization problem; i.e., that the public good is sufficiently socially important to justify at least some production costs in terms of lost leisure, but that it is never socially optimal that individuals spend all their available time contributing to the public good.

As the individual judges herself, so too will others judge her contributions to the public good. Others, however, are limited to judging what they can observe, and so consider only the individual's visible effort:

$$V_{i} = \lambda \left( hv_{i}, hv_{i}^{*} \right) = -b \left( hv_{i} - hv_{i}^{*} \right)^{2}, \ b > 0$$
(8)

Each individual thinks others should also act so as to maximize social welfare and individuals therefore judge the observable effort of others based on their own ideal level of observable effort. <sup>1</sup>

Both eqs. (8) and (5) are concave, with vertices at  $hv_i = hv_i^*$  and  $g_i = g_i^*$ , respectively. The functional values are negative at all points except at their vertices (where they equal zero), increasing when  $hv_i < hv_i^*$  and  $g_i < g_i^*$ , respectively, and decreasing when  $hv_i > hv_i^*$  and  $g_i > g_i^*$ . Thus, the self-image and validation functions can be regarded as representing utility loss due to cognitive dissonance and social disapproval, respectively.

<sup>&</sup>lt;sup>1</sup>Note that with rational expectations and identical preferences, individuals may realize the solution to others' utility maximization problem and thus deduce their actual behaviour even for invisible efforts, taking this into account when judging them. In line with the idea that sanctioning is typically based on actual observation, not subjective beliefs, we disregard such sophistication here.

#### 2.2 Results

#### 2.2.1 The morally ideal contribution

To determine her ideal effort levels, the individual maximizes the welfare function (7) with respect to  $lv_i$  and  $hv_i$  subject to equations (1)–(4) and (6), given that  $lv_i = lv_j$ ,  $hv_i = hv_j$  for all  $i \neq j$ . Note that while eq. (7) implies that self-image and social image are considered welfare relevant (see Chan, 2024), this does not matter for our results: at the social optimum,  $g_i = g_i^*$ ,  $hv_i = hv_i^*$  and  $lv_i = lv_i^*$ , implying that  $I_i = V_i = 0$ .

The maximization yields a set of first-order conditions. The i subscript can be ignored for this analysis, as all individuals have been assumed identical. Subscripts will henceforth denote derivatives. The first-order conditions are

$$u_l = u_G N \phi_{lv} \tag{9}$$

and

$$u_l = u_G N \phi_{hv}. \tag{10}$$

The individual's morally ideal effort is defined by:

$$u_l = u_G N \phi_{lv} = u_G N \phi_{hv}. \tag{11}$$

This is to say that in her ideal world where all other individuals act like her, she would choose lv and hv such that her own marginal utility of leisure would equal the marginal utility to society of lv as well as hv.<sup>2</sup> Another way to regard this is that when the individual imagines that her behaviour guides that of all other individuals, the benefit to her of any contribution she makes to the public good will be magnified N times by the identical contributions of every other individual.

#### 2.2.2 Optimizing individual utility

Based on the ideal efforts defined by equation (11), the individual must now determine her actual behaviour. Since she is only making actual decisions about her own effort levels, she maximizes her own utility (equation (1)) with respect to lv and hv considering the behaviour of all other individuals as exogenous, subject to her time constraint and

<sup>&</sup>lt;sup>2</sup>That is,  $\phi_{lv^*} = \phi_{hv^*}$ 

equations (2)-(5) and (8). This yields the first order conditions

$$u_{l} = u_{G}\phi_{lv} + u_{I}\left(-2a\left(g - g^{*}\right)\right)\phi_{lv}$$
(12)

and

$$u_{l} = u_{G}\phi_{hv} + u_{I}\left(-2a\left(g - g^{*}\right)\right)\phi_{hv} + u_{V}\left(-2b\left(hv - hv^{*}\right)\right)$$
(13)

The individual's effort is determined by

$$u_{l} = u_{G}\phi_{lv} + u_{I}\left(-2a\left(g - g^{*}\right)\right)\phi_{lv}$$
  
=  $u_{G}\phi_{hv} + u_{I}\left(-2a\left(g - g^{*}\right)\right)\phi_{hv} + u_{V}\left(-2b\left(hv - hv^{*}\right)\right)$  (14)

Equation (14) describes a Nash equilibrium where the individual seeks to allocate her time such that her marginal utilities of leisure, invisible effort and visible effort are equal. The individual's marginal utility of leisure equals her marginal benefits of invisible effort, consisting of her own public good benefit plus the marginal self-image gain. Also, the marginal utility of leisure must equal the marginal benefits of visible effort, which in turn consists of the marginal public good benefit to herself, the marginal self-image gain, and the marginal social validation gain. We assume that the individual's maximization problem has an internal solution; i.e., that it will be optimal for her to exert some effort of both types but not to contribute all her available time.

As a simplifying assumption, let us now assume that the individual considers her own effort to have a negligible effect on the total provision of the public good, G. Note that this assumption is made only when deciding what to actually do herself, not when determining the morally ideal efforts, since if everyone followed the individual's example the effects would hardly be negligible. This means that  $u_G \phi_{lv}$ ,  $u_G \phi_{hv} \approx 0$ . The simplified first-order condition describing the individual's actual allocation of her time is then

$$u_{l} = u_{I} \left( -2a \left( g - g^{*} \right) \right) \phi_{lv} = u_{I} \left( -2a \left( g - g^{*} \right) \right) \phi_{hv} + u_{V} \left( -2b \left( hv - hv^{*} \right) \right).$$
(15)

The marginal utility of leisure must equal the marginal utility of invisible effort through improved self-image which must, in turn, equal the marginal utility of visible effort through improved self-image and social validation.

#### 2.2.3 Underprovision of the public good

From equation (15) it can be shown that the individual, when maximizing her own utility, will underprovide the public good. That is, the individual's provision g of the public good will never reach the morally ideal provision  $g^*$ : If  $g \ge g^*$ , then the individual's marginal benefit of invisible effort,  $u_I (-2a (g - g^*)) \phi_{lv}$ , would be weakly negative. This cannot be the case in an interior utility maximum: Equation (15) states that the individual will choose hv and lv such that the marginal benefit of invisible effort equals the marginal utility of leisure, which is, by assumption, strictly positive. Thus, we replicate the finding of Brekke et al. (2003) that  $g < g^*$  in the Nash equilibrium. Since  $g^*$  is the individual provision that would be socially optimal if everyone did their part, and individuals are identical, this also means that the total supply of the public good, as given by eq. (4), is suboptimal.

#### 2.2.4 Allocation to visible and invisible forms of effort

Given the above result, we are able to show that the individual will exert less than their ideal dictates of invisible as well as visible effort. Moreover, the preference for social validation leads to misallocation of effort between visible and invisible activities.

Rearranging equation (15) gives

$$u_I \left(-2a \left(g - g^*\right)\right) \left(\phi_{lv} - \phi_{hv}\right) = u_V \left(-2b \left(hv - hv^*\right)\right) \tag{16}$$

Consider first the possibility that  $hv = hv^*$ . In equation (16) this would imply that  $(hv - hv^*) = 0$  and consequently the right-hand side of the equation equals zero. This means that on the left-hand side either  $u_I$ ,  $(g - g^*)$  or  $(\phi_{lv} - \phi_{hv})$  would need to equal zero<sup>3</sup>.  $u_I$  is strictly positive, and  $(g - g^*)$  is negative (since we have already demonstrated that  $g < g^*$ ). If  $(\phi_{lv} - \phi_{hv}) = (\phi_{lv} - \phi_{hv^*}) = 0$ , then:

$$\phi_{hv} = \phi_{hv^*} = \phi_{lv} = \phi_{lv^*}$$

since in the social optimum, the marginal productivity of the two effort types are equal

 $<sup>^{3}-2</sup>a$  is constant and negative, as a > 0

(eq. (11)). However,  $hv = hv^*$  and  $lv = lv^*$  would mean that  $g = g^*$ , which we have shown cannot occur in the Nash equilibrium. Thus it follows that  $hv \neq hv^*$ .

The individual will not exert strictly more visible effort than her moral ideal prescribes, either. If we imagine that  $hv > hv^*$ , it must necessarily be the case that  $lv < lv^*$ such that  $g < g^*$ , since we have shown that there will be an overall underprovision of the public good. Assuming  $hv > hv^*$  and  $lv < lv^*$  in turn implies that

$$\phi_{hv} < \phi_{hv^*} = \phi_{lv^*} < \phi_{lv}$$

 $\phi_{hv} < \phi_{lv}$ 

If  $hv > hv^*$ , the right-hand side of equation (16) would be negative. Since  $\phi_{hv} < \phi_{lv}$ , however,  $(\phi_{lv} - \phi_{hv}) > 0$ , making the left-hand side of equation (16) positive. Thus, the equality cannot hold, so  $hv \neq hv^*$  in the Nash equilibrium, even if lv is low enough to keep  $g < g^*$ .

Knowing that  $hv < hv^*$ , the right side of equation (16) is positive for the individual's chosen values of hv and lv. For the left-hand side of the equation to be positive as well it must be the case that  $\phi_{lv} > \phi_{hv}$  such that  $(\phi_{lv} - \phi_{hv}) > 0$ . The individual's higher marginal productivity on invisible effort means that she could provide the same amount of the public good for less total effort if she invested less in visible effort and more in invisible effort. Thus, the individual will always choose to provide an inefficient (with regards to provision of the public good) combination of visible and invisible effort.

If both activities contributed equally to the welfare of society, and the individual had to choose whether to spend her next hour on a socially visible moral activity or an invisible one, she would always choose to spend her time on the visible activity. Intuitively, this is because her contribution will have the same effect on her self-image regardless of what she chooses, but if she chooses the visible activity she additionally improves her social image.

#### 2.2.5 Effect of social validation on total effort contributed

Even if effort is underprovided and misallocated compared to the first-best social optimum, social validation does provide an incentive to exert effort towards public good provision. Hence, it may seem reasonable to expect that observability does indeed increase the individual's total public good provision g. As it turns out, this is not necessarily the case.

We can analyze this by imagining that the visible activity becomes invisible, yielding no social validation. This is behaviourally equivalent to assuming that b = 0 (contrary to the assumption made earlier that b > 0), i.e. that the individual does not care about social validation at all. Would provision g then be reduced, compared to the case where b > 0?

Let us continue to call the two activities hv and lv, although for the moment none of them will be observable to others. Recall that eq. (15), characterizing the individuals' utility maximizing allocation of time, is given by

$$u_{l} = u_{I} (-2a (g - g^{*})) \phi_{lv}$$
  
=  $u_{I} (-2a (g - g^{*})) \phi_{hv} + u_{V} (-2b (hv - hv^{*}))$ 

The right hand side of this equation represents the individual's marginal benefits of visible effort. If b = 0, the positive term  $+u_V (-2b (hv - hv^*))$  disappears. For the first and third part of eq. (15) to still be equal, the individual will change her allocation of time in order to decrease her marginal utility of leisure, which can only be done by increasing her leisure time - corresponding to less time for effort. For the last two terms to be equal, we must now have  $\phi_{hv} = \phi_{lv}$ ; that is, there will no longer be misallocation of effort between the two activities, so a larger share of the effort will be in terms of lv.

It is not clear, however, whether this decrease in efforts due to less observability would also cause less overall provision of the public good. The reason is that although effort has decreased overall, the effort is now better balanced between the two activities. Which effect dominates depends on the specific form of the production function  $\phi(hv, lv)$ .

Although the above analysis was based on the removal of observability (or the pref-

erence for social validation), conclusions will of course simply be reversed when it comes to making an activity observable that was previously not: Observability of one of the alternative contribution activities, when combined with a preference for social validation, does increase total time spent on contribution activities. However, while possible, it is not necessarily the case that observability increases the overall provision of the public good, since it also causes misallocation of the alternative effort types.

#### 2.3 Variable visibility

As a simplifying assumption, we have been treating visibility as a binary state: an individual's effort is either visible or invisible to others. It is relatively straightforward, however, to generalize the model to include a larger number of activities with a varying degree of observability. Some efforts, such as sorting household waste for recycling, may be visible only to one's family or neighbors; others are visible to larger groups of individuals, such as owning and/or driving an electric vehicle, which might be visible also to others on the road, colleagues, friends, and so on.

Instead of one visible and one invisible type of effort, thus, consider an individual who engages in a number of different contribution activities. Let her self-image benefit be given by equation (5) as before, while the individual's contribution is a function  $\Phi$  of all her efforts similar to equation (4):

$$g = \Phi\left(e_1, e_2, \dots, e_n\right) \tag{17}$$

where subscripts now denote effort type,  $\Phi(0, 0, ..., 0) = 0$ ,  $\frac{\partial \Phi}{\partial e_i} > 0$ ,  $\frac{\partial^2 \Phi}{\partial e_i^2} < 0$ , and  $\frac{\partial^2 \Phi}{\partial e_i \partial e_j} = \frac{\partial^2 \Phi}{\partial e_j \partial e_i} = 0$ , where i, j = 1, ..., n and  $i \neq j$ . The function  $\Phi$  is additively separable and strictly increasing with diminishing returns in all  $e_i$ . Social verification could then be modelled as

$$V = \lambda(e_1, \dots, e_n, e_1^*, \dots, e_n^*)$$
  
=  $-b \sum_i \omega_i (e_i - e_i^*)^2, \ 0 \le \omega_i \le 1, \ 0 < b$  (18)

where  $0 \le \omega_i \le 1$  are weights to capture  $e_i$ 's degree of visibility. The case with only two types of effort where one is unobservable and the other is fully observable, would then

correspond to the model introduced above.

#### 3 Conclusion

Voluntary provision of public goods can be characterized as a prisoners' dilemma game: everyone would benefit from coordinating to provide the socially optimal public good supply, but because they are unable to do so, their individual motivations trap them in an inefficient Nash equilibrium. Previous research (Brekke et al., 2003) has established that this holds, although to a less severe extent than with narrow self-interest, even if people are morally motivated, feeling discomfort when contributing less than they consider morally ideal.

Social acceptance is crucial to human beings; indeed, neurological research indicates that social exclusion triggers similar processes in the brain as pain caused by physical injury (Lieberman, 2013). It thus seems natural to expect that by making at least some contribution behaviors observable, allowing for the additional incentive of social pressure, the prisoners' dilemma may be overcome.

However, given the formal framework presented above, this expectation is only partly justified. First, even when individuals are morally motivated and some contribution activities are observable by peers, provision of the public good is suboptimal. Second, if a previously unobservable contribution activity becomes observable, the total time spent contributing to the public good does increase. Third, there is nevertheless also a counteracting effect: The additional incentive of social pressure causes individuals to substitute between activities, engaging relatively less in the unobservable type of effort. This can cause a misallocation of effort between the different activities; and potentially, this negative effect could dominate the positive impact of increased efforts.

#### References

- Andreoni, J. (1988). Privately provided public goods in a large economy: The limits of altruism. Journal of Public Economics, 35(1), 57–73. https://doi.org/10.1016/ 0047-2727(88)90061-8
- Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warmglow giving [Publisher: [Royal Economic Society, Wiley]]. The Economic Journal, 100(401), 464–477. https://doi.org/10.2307/2234133
- Bénabou, R., & Tirole, J. (2006). Incentives and Prosocial Behavior. American Economic Review, 96(5), 1652–1678. https://doi.org/10.1257/aer.96.5.1652
- Bergstrom, T., Blume, L., & Varian, H. (1986). On the private provision of public goods. Journal of Public Economics, 29(1), 25–49. https://doi.org/10.1016/0047-2727(86)90024-1
- Bicchieri, C. (2006). The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press. https://doi.org/10.1017/CBO9780511616037
- Brekke, K. A., Kverndokk, S., & Nyborg, K. (2003). An economic model of moral motivation. Journal of Public Economics, 87(9), 1967–1983. https://doi.org/10.1016/ S0047-2727(01)00222-5
- Brekke, K. A., & Nyborg, K. (2008). Attracting responsible employees: Green production as labor market screening. *Resource and Energy Economics*, 30(4), 509–526. https://doi.org/10.1016/j.reseneeco.2008.05.001
- Brekke, K. A., & Nyborg, K. (2010). Selfish bakers, caring nurses? A model of work motivation. Journal of Economic Behavior & Organization, 75(3), 377–394. https: //doi.org/10.1016/j.jebo.2010.05.008
- Chan, N. W. (2024). Pigouvian Policies under Behavioral Motives [Publisher: The University of Chicago Press]. Journal of the Association of Environmental and Resource Economists, 11(1), 97–135. https://doi.org/10.1086/725445
- Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, 14(1), 47–83. https: //doi.org/10.1007/s10683-010-9257-1
- Cialdini, R., Reno, R., & Kallgren, C. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. Journal of Personality and Social Psychology, 58, 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015
- Coleman, J. S. (1990). Foundations of Social Theory [Google-Books-ID: a4Dl8tiX4b8C]. Harvard University Press.

- Dannenberg, A., Johansson-Stenman, O., & Wetzel, H. (2022). Status for the good guys: An experiment on charitable giving [\_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecin.1 *Economic Inquiry*, 60(2), 721–740. https://doi.org/10.1111/ecin.13039
- Ek, C. (2018). Prosocial behavior and policy spillovers: A multi-activity approach. Journal of Economic Behavior & Organization, 149, 356–371. https://doi.org/10.1016/j. jebo.2018.01.006
- Elster, J. (1989). Social norms and economic theory. *Journal of Economic Perspectives*, 3(4), 99–117. https://doi.org/10.1257/jep.3.4.99
- Farrow, K., Grolleau, G., & Ibanez, L. (2017). Social norms and pro-environmental behavior: A review of the evidence. *Ecological Economics*, 140, 1–13. https://doi. org/10.1016/j.ecolecon.2017.04.017
- Hong, F., Tirole, J., & Zhang, C. (2023). Prosocial Behavior in Public and Private Spheres: Theory and Evidence. Unpublished note, Toulouse School of Economics.
- Hoven, N. Z. (2022). Modelling Visible Altruism [Master's thesis] [Accepted: 2022-08-25T22:00:22Z]. Retrieved March 8, 2024, from https://www.duo.uio.no/handle/ 10852/95726
- Kahneman, D., & Knetsch, J. L. (1992). Valuing public goods: The purchase of moral satisfaction. Journal of Environmental Economics and Management, 22(1), 57– 70. https://doi.org/10.1016/0095-0696(92)90019-S
- Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a Constraint on Profit Seeking: Entitlements in the Market [Publisher: American Economic Association]. *The American Economic Review*, 76(4), 728–741. Retrieved April 21, 2024, from https://www.jstor.org/stable/1806070
- Kaza, S., Yao, L. C., Bhada-Tata, P., & Van Woerden, F. (2018, September 20). What a waste 2.0: A global snapshot of solid waste management to 2050 (Accepted: 2018-08-31T16:25:22Z). World Bank. Washington, DC. https://doi.org/10.1596/978-1-4648-1329-0
- Lieberman, M. D. (2013). Social: Why our brains are wired to connect [Pages: x, 374]. Crown Publishers/Random House.
- Nyborg, K. (2018). Social norms and the environment. Annual Review of Resource Economics, 10(1), 405–423. https://doi.org/10.1146/annurev-resource-100517-023232
- Nyborg, K. (2020). No man is an island: Social coordination and the environment. Environmental and Resource Economics, 76(1), 177–193. https://doi.org/10.1007/s10640-020-00415-2

- Olhoff, A., & Christensen, J. (2021, October 25). Emissions gap report 2021, chapter 1. Retrieved April 6, 2022, from http://www.unep.org/resources/emissions-gapreport-2021
- Sugden, R. (2000). The motivating power of expectations. In J. Nida-Rümelin & W. Spohn (Eds.), *Rationality, rules, and structure* (pp. 103–129). Springer Netherlands. https://doi.org/10.1007/978-94-015-9616-9\_7
- Yoeli, E., Hoffman, M., Rand, D. G., & Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment [Publisher: Proceedings of the National Academy of Sciences]. Proceedings of the National Academy of Sciences, 110 (supplement\_2), 10424–10429. https://doi.org/10.1073/pnas.1301210110
- Young, H. P. (2015). The evolution of social norms. Annual Review of Economics, 7(1), 359–387. https://doi.org/10.1146/annurev-economics-080614-115322