FAERE
French Association
of Environmental and Resource Economists

Working papers

The importance of considering optimal government policy when social norms matter for the private provision of public goods

Guy Meunier – Ingmar Schumacher

WP 2017.17

Suggested citation:

G. Meunier, I. Schumacher *(2017)* The importance of considering optimal government policy when social norms matter for the private provision of public goods.
FAERE Working Paper, 2017.17.

ISSN number: **2274-5556**

www.faere.fr

The importance of considering optimal government policy when social norms matter for the private provision of public goods

Guy Meunier* Ingmar Schumacher†

September 7, 2017

*INRA–UR1303 ALISS guy.meunier@inra.fr and CREST

†IPAG Business School, 184 boulevard Saint-Germain, 75006 Paris, France. Tel.: +352 621264575. E-mail: ingmar.schumacher@ipag.fr.

For comments we would like to thank Geir Asheim, Kjell Arne Brekke, Snorre Kverndorkk, Florian Wagener, Cees Withagen, as well as participants at the EAERE 2017 conference in Athens, the ISEFI 2017 conference in Paris, the IRMBAM 2017 sub-conference in environmental economics in Nice, the Lameta EERN seminar in Montpellier and the Energy and Prosperity Chair seminar in Paris, as well as Francois Salanie and an anonymous referee from FAERE. Guy Meunier is grateful for the financial support of the Chair Energy and Prosperity.

Abstract

Social pressure can help overcome the free rider problem associated with public good provision. In the social norms literature concerned with the private provision of public goods there seems to be an implicit belief that it is best to have all agents adhere to the 'good' social norm. We challenge this view and study optimal government policy in a reference model (Rege, 2004) of public good provision and social approval in a dynamic setting. We discuss the problem with the standard crowding in and out argument and analyze the relationship with Pigouvian taxes. We show that even if complete adherence to the social norm maximizes social welfare it is by no means necessarily optimal to push society towards it. We stress the different roles of the social externality and the public good problem. We discuss the role of the cost of public funds and show how it can create path dependency, multiplicity of both optimal equilibria and optimal paths, and discuss the role of parameter instability. We argue that extreme care must be taken when formulating policies and subsequent results will fully depend on this formulation.

Keywords: government policy; social norms; public goods; optimal policy.
JEL: H23.

1 Introduction

In the social norms literature that is concerned with the private provision of public goods there seems to be an implicit belief that it is always best to have all agents adhere to the 'good' social norm of contributing to the public good.[1] Thus, this literature instigates an overall strong support for a government policy that induces society to move towards an equilibrium with full adherence to the norm. This is especially true for the case where society would otherwise be stuck with a 'bad' social norm of no one contributing. The question that we ask in this article is whether and when it is actually optimal from a society's point of view that a government nudges society to take up a certain social norm.

We start from a simple, yet widely accepted model of social norms (Rege, 2004), allow for optimal government policy, and then study how far this augments the tacit idea that a government should try to enforce an otherwise socially beneficial norm. We view the contribution of this article as emphasizing the need to move away from the simple crowding in and crowding out story and instead start to focus more on endogenous optimal government policies. Additionally, we send a note of caution by giving some first insights into how different setups of the government policy lead to strikingly different results regarding the optimal level of the social norm and private contributions.

Up to now, the literature on social norms and public goods has mostly concentrated on exogenous government policy. Researchers have emphasized whether or not a public policy would, via its effect on social norms, crowd in or crowd out private provisions (e.g. Nyborg, 2003; Nyborg and Rege, 2003; Bowles and Polania-Reyes, 2012). While this focus might help avoid the implementation of counterproductive policies, it lacks a normative content. Government policy is not exogenous but should be directed at achieving a specific target. In the tradition of welfare economics, this target should not be to crowd in or out

[1] A 'norm' is an informal rule of behavior (to be contrasted with an explicit legal rule), a 'social norm' is enforced through social sanctions whereas a 'moral norm' is enforced through self sanctions. The distinction between a moral and a social norm could be blurry, especially from an analytical perspective. The main feature we are concerned with is the presence of a social reinforcing effect: the more people adhere to the norm the larger the incentive to adhere to it.

private contributions, but to induce a socially optimal level of a public good and social norm. Once one accepts that government policy should be fully endogenized, then, as we show below, many of the commonly accepted results from the exogenous policy literature must be viewed in a new light.

The present work is at the crossroads of several strands of the literature, namely the analysis of voluntary contributions to public goods, social approval and the dynamic diffusion of pro-social norms. Several authors have proposed explanations for the observed fact that, even in large societies, people do contribute to public goods such as charities. Even though pure altruism can explain that people contribute to public goods, economic theory predicts that incentives to contribute vanish with increasing numbers of potential contributors (see Warr, 1983; Bergstrom et al., 1986). Sugden (1982) and Andreoni (1990) argue that there is a private benefit from contributing to public good, related to one's own contribution, that helps explain common patterns observed and soften the possible crowding out from public provision. Sugden (1984) proposes that individuals maximizes their utility subject to a moral constraint.

Holländer (1990) suggests the existence of a social effect and considers that people care about their relative contribution to a public good. Brekke et al. (2003) introduce the idea that people compare their contribution to an ideal one, which is depending on the norm. In a static framework, they emphasize the possible adverse effect of a public policy on contribution that would operate via a deterioration of the norm. Bruvoll and Nyborg (2004) analyze the influence of a change of the norm on behavior and welfare, and stress the psychological costs associated to strengthening the norm. Bénabou and Tirole (2006) model the signaling effect of contributing to a public good, altruism being socially rewarded and signaled through high contributions, where they emphasize that a subsidy to contributions reduces the altruism signaling incentive.

Bowles and Hwang (2008) also study optimal policy with public goods and social preferences. They assume that subsidies have a direct impact on social values, and analyze how this influences optimal policy design. In their setting, social preferences are given and do not depend on the interaction of individuals. We focus on the dynamic aspect of the diffusion of the norm and its social enforcement through interactions.

Some articles introduce the self-enforcing dynamic of a social norm: the more people adhere to a norm the higher the private benefit to adhere to it. Such mechanism can give rise to multiplicity of equilibria. This multiplicity of equilibrium translates into path dependency and historical lock-in in dynamic frameworks. Rege (2004) introduces social approval as a self-enforcing mechanism. When considering public policy she stresses that a subsidy can help unlock society from a zero-contributor situation and push it toward a full contribution equilibrium. Nyborg et al. (2006) provide a model of green consumption with social approval that shares similar features. The existence of multiple equilibria associated with social approval has also been studied by Lin and Yang (2006), who argue that only sizable subsidies may induce significant shifts in the equilibria.

The three above mentioned articles consider the ability of public policy to shift society to the full participation equilibrium without providing a welfare discussion.Furthermore, they advocate for temporary policies which would be enough as long as they help the norm to sufficiently penetrate through society to such an extent that it develops its own positive dynamics. This view implicitly tells us that there should be a cost attached to this policy, as why would one want to stop an otherwise beneficial policy. Thus we take the costly aspect of the policy intervention more serious and in so doing we show that the full contribution equilibrium, for which the policy was introduced in the first place, may not be optimal any longer. In fact, we show that in the case of a linear cost of public funds it is never optimal to induce the full participation equilibrium. The equilibrium share of contributors can be lower with the norm than without it because of the negative social externalities associated with approval and disapproval.

We subsequently relax the assumption of a linear cost of public funds and consider quadratic costs. In that case we show that it can be optimal to reach full participation, the subsidy being eventually zero. Thus, we argue that more general models of public funding could recover the result that the full contribution equilibrium is optimal and, at equilibrium, no policy intervention will be necessary. However, we also show that the full participation equilibrium need not be optimal. There may be a multiplicity of steady states and associated optimal trajectories, some of which may even be optimal at the same time. In other words, we observe the existence of Skiba points which implies that parameters and their stability may be equally important as initial conditions.

The article is set up as follows. In the next section 2 we start by introducing Rege (2004)'s original model. We then extend her model in section 3 by introducing endogenous government policy. In section 3.1 we look into further aspects such as a comparison to the Pigouvian tax and in section 4.2 we study different public funding cost structures. We furthermore discuss the problems of path dependency and multiplicity of equilibria for optimal policy in section 4.3. Section 5 concludes.

2 The background model

In this section we present the main ideas behind Rege (2004)'s model of a social norm that influences the incentives for the private provision of a public good, discuss the results on the exogenous government policy, and then study the implication of endogenizing this policy. We fully follow the notation in Rege (2004) for simplicity.

In Rege (2004) there exists a continuum of agents (on $[0, 1]$) who decide to either contribute ($g_i = 1$) or not contribute ($g_i = 0$) to a public good. x is the share of contributors, and $w(x)$ is the benefit of the public good. Agents have income I that they may use for consumption (at price 1) or for the public good at price $p > 0$, and they are also affected by a social approval $q_i(x)$. The utility function is assumed to be linear and of the form $U_i = I + w(x) - pg_i + q_i(x)$.

The social norm arises from the interaction of agents. A non-contributor feels disapproval, whereas a contributor feels approval if he is observed by a contributor. A person feels neither approval nor disapproval from non-contributors. The magnitude of the approval or disapproval feeling depends on the frequency of the behavior in society; it is proportional to the benefits for society of the social norm (by a factor $\lambda \equiv w(1) - p > 0$); and it depends on how many agents someone meets from one's own type, the share of which is $k \in (0, 1/2)$. This yields a social approval of $q_i(x) = \lambda(1-x)(k+(1-k)x)$ for contributors, while non-contributors obtain a disapproval of $q_i(x) = -\lambda x^2(1-k)$.[2]

In the static version, individuals play a coordination game in which they choose whether to be a contributor by maximizing their utility, the difference in utility being

$$\Delta U(x) = U^1(x) - U^2(x) = \lambda(k + (1-2k)x) - p. \tag{1}$$

Because of social approval this difference is increasing in the share of contributors. We define $\bar{x}$ to be the share of contributors that makes individuals indifferent between contributing and not contributing, and it is given by

$$\bar{x} \equiv \frac{p - \lambda k}{\lambda(1-2k)}.$$

If the share of contributors is larger than $\bar{x}$, then everyone prefers to be a contributor, while no one prefers to contribute otherwise. As stated in Proposition 1 in Rege (2004), if $\bar{x} \in (0, 1)$ there are three Nash equilibriums of the static game: $x = 0$, $x = 1$ and $x = \bar{x}$. This is the case if the following assumption is fulfilled:

Assumption 1 *We assume $p > \lambda k$ and $p < \lambda(1-k)$.*

Rege (2004) took this model a step further and, based on Weibull (1997), Börgers and Sarin (1997) and Taylor and Jonker (1978), allowed the social norm to evolve dynamically and endogenously. Clearly, if agents were fully aware of their preferences and could immediately adopt the social norm, then the social norm would instantly spread through society. This is unlikely to be realistic and it seems more appropriate to assume that agents are not well-informed about their preferences regarding the norm and can only slowly, through repeated interactions, change their behavior. This idea of dynamic evolution can be captured through the so-called replicator dynamics and is given by the following equation

$$\dot{x}_t = x_t(1-x_t)\Delta U(x_t). \tag{2}$$

[2]The components are for a contributor (resp. non contributor): $(1-x)$ (resp. $-x$) is the comparison between the observed behavior and the average; $(k+(1-k)x)$ (resp. (1-k)x) the number of contributors met. See Rege (2004) for futher details.

Based on these evolutionary dynamics Rege (2004) shows that the three potential equilibria identified above are still possible, but only two are stable ($x_t = 0$ and $x_t = 1$) while $x_t = \bar{x}$ is unstable. For $x_t < \bar{x}$ society converges over time to the non-contributor equilibrium, while for $x_t > \bar{x}$ society converges to the equilibrium where everyone contributes. In other words, if there are too few contributors in society then non-contributors have not enough incentives to become contributors. Similarly, contributors are not sufficiently approved if they meet too few other contributors, so that it may simply not be worthwhile for them to contribute any longer. As a result, society converges to a norm where nobody contributes. In contrast, if there are sufficient contributors in society then social approval and disapproval motivates non-contributors to become contributors.

Rege (2004) then investigates if the government, by introducing price subsidies s that are paid for via income taxes xs, can instigate a change in society that may induce convergence to the equilibrium with everyone contributing. We thus rewrite the utility of contributors as

$$U^1(x, s) = I - xs + w(x) - p + s + \lambda(k + (1-k)x)(1-x), \tag{3}$$

where $p - s$ is the original price p with the subsidy s, and xs is the income tax, while the utility of non-contributors becomes

$$U^2(x, s) = I - xs + w(x) - \lambda(1-k)x^2. \tag{4}$$

This augments the utility difference, which now becomes

$$\Delta U(x, s) = s + \lambda(k + (1-2k)x) - p. \tag{5}$$

Thus, if the subsidy is large enough ($s > p - \lambda k$) then this utility difference will be positive and society can converge to the high equilibrium. Furthermore, once the subsidy has been put in place for long enough such that the share of the contributors exceeds a certain threshold ($\bar{x}$), then even without this subsidy society would continue to converge to the equilibrium where everyone contributes to the public good. Thus, a subsidy will crowd-in voluntary contributions.

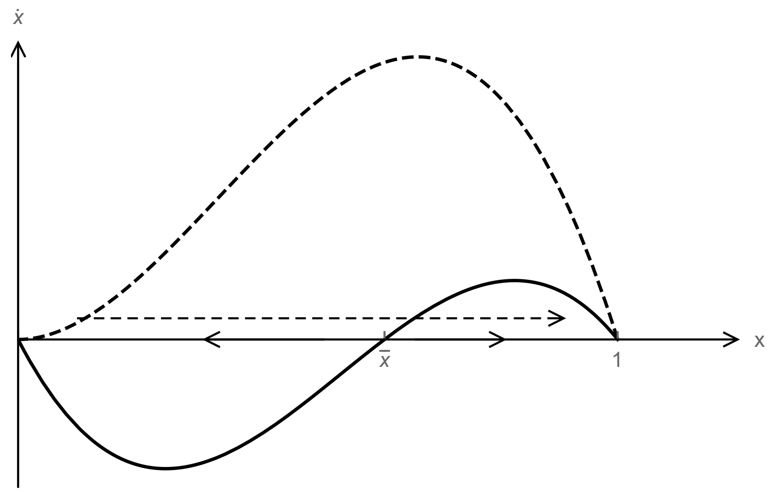


Figure 1: Influence of a fixed subsidy $s > p - \lambda k$ on the dynamic of the share of contributors. Source: Adapted from Rege (2004).

While it is good to know that government policy can induce changes in equilibria, it is also important to know whether and when this is actually optimal. We now turn to our contribution.

3 The implication of endogenous policy

We now go a step further and, based on the previous model, derive an intertemporal social welfare function that a policy maker would use in order to assess an optimal government policy. We take the simplest possible setting as this already yields the insights that we are after. We assume that agents continue to act myopically when adhering to the norm, and the evolution of the social norm follows the replicator dynamics. However, we assume that there exists a government who can introduce a policy. The policy maker is forward looking, has perfect foresight, and maximizes the infinite stream of the agents' utilities by appropriately setting incentives. The policy maker can give a subsidy s_t on contributions. We allow this subsidy to be negative, which gives the policy maker complete freedom over the direction in which he wants to push the production of the public good, and the social norm.

Assumption 2 *We assume the existence of exogenous bounds on* s_t *in the form of* $\underline{s} \leq s_t \leq \bar{s}$*, with* $\underline{s} < 0$*. We shall, for simplicity, assume that* $\underline{s} \leq p - \lambda(1-k)$ *and* $\bar{s} \geq p - \lambda k$*.*

These bounds are a natural restriction for public policy. We assume them to be large enough so that the policy maker is assured full flexibility over the influence of the policy.

Then based on the model introduced in the previous section we define the instantaneous gross welfare as the sum of agents' utilities: $V = xU^1(x,s) + (1-x)U^2(x,s)$ which only depends on x and after substitution this yields

$$V(x) = I + w(x) - px + \lambda x(1-x)k. \tag{6}$$

To gain some intuition on the influence of the norm in this welfare function, we take the derivative with respect to x, slightly rewrite it by making use of equation (5), and obtain

$$V'(x) = w'(x) + \Delta U(x,s) - s + \lambda \Big\{ x\Big((1-k)(1-x) - (k+(1-k)x)\Big) - (1-x)2(1-k)x \Big\}. \tag{7}$$

The term in curly brackets is the marginal external social impact of an increase in the share of contributors. It is the sum of the effect on contributors and on non-contributors, which derive from the increased probability of meeting a contributor and the influence on approval intensity. For non-contributors, the diffusion of the contributory behavior has a cost only: they meet more contributors and feel more disapproval. For contributors the effect is ambiguous as they benefit from meeting more contributors but their feeling of approval is reduced for each encounter. After simplification, the term in curly brackets reduces to $-\lambda x$, implying that, at the equilibrium when we expect $\Delta U(x,s) = 0$, social approval and disapproval represent an external cost.

In each period the policy maker balances the budget but whenever he raises taxes to pay for the price subsidy then he incurs a deadweight loss. We represent this cost of public funds by $\gamma > 0$, i.e. raising \$1 of public money costs society \$$(1+\gamma)$ because of distortionary taxation. Conversely, if the government taxes the public good ($s < 0$) then this reduces the taxpayers' burden by $sx(\gamma+1)$.[3] The underlying assumption is that taxing a public good leads to less distortionary taxation than e.g. a tax on labor. This, in a sense, gives this model a bit of a partial equilibrium character, but it significantly simplifies the analysis and it is one of the most commonly used approaches that set up government policy.[4]

Hence we can now write the policy maker's objective function, which is given by

$$W = \int_{t=0}^{+\infty} e^{-\phi t}\big[V(x_t) - \gamma x_t s_t\big]\,dt, \tag{8}$$

where $\phi > 0$ is the discount rate. Based on this setup there is only one state equation

$$\dot{x} = x(1-x)\Delta U(x,s). \tag{9}$$

The policy maker then maximizes equation (8) with respect to s_t, subject to the constraint (9) and with bounds on subsidies in the form of $\underline{s} \leq s_t \leq \bar{s}$.

Using equation (5) and subsituting s into the expression (8) of the objective function yields

$$W = \int_{t=0}^{\infty} e^{-\phi t}\left\{V(x) - \gamma x_t\Big[p - \lambda(k+(1-2k))x_t\Big] - \gamma\frac{\dot{x}}{1-x}\right\}dt \tag{10}$$

Due to the linearity of the control s_t we know that the solution to this is a *Most Rapid Approach Path*. Maximizing the objective function we obtain the Euler equation, and denoting the optimal solution by x^*, we then get

$$V'(x^*) - \gamma\Big[p - \lambda k - 2\lambda(1-2k)x^*\Big] - \frac{\gamma\phi}{1-x^*} = 0. \tag{11}$$

[3]It is also possible to introduce a tax on non-contributors, if such a tax t is available, the difference in utility becomes $\Delta U(x, s+t)$ and the total budget$xs - (1-x)t = x(s+t) - t$, it is *as if* every agent was taxed t and contributors receive a subsidy $s+t$ which becomes the relevant variable to be set by the regulator.

[4] In a later section we consider a non-linear deadweight loss which provides a more general equilibrium character.

This optimal level of contributors is depending on three terms. The first is the effect on per period welfare $V(x)$ of an increase in the number of contributors. This term can be further decomposed as $V'(x) = w'(x) - p + \lambda(1-2x)k$, the sum of the marginal benefits from the public good, the cost of contributing and the marginal social approval effect. Note that this last social effect can be either positive or negative depending on how many people are already contributing to the public good. The second term is the effect on the per period cost of the total subsidy xs. The subsidy necessary to sustain a given share of contributors is decreasing in the share of contributors, an effect which is due to the social approval. The third term is related to the impatience of the social planner, it is the discounted cost of the subsidy. The more impatient the social planner the lower the benefits from increasing the number of contributors. This marginal cost grows to infinity as x becomes close to 1, for then the norm spreads slowly in the society (there are few non-contributors to be converted) and it becomes very costly to further incentivize agents to take up the norm. This time effect crucially hinges on the existence of the deadweight loss that the policy maker bears each instance.

Assumption 3 *We assume that* $w'(0) > \gamma\phi + (1+\gamma)(p - \lambda k)$ *and* $w'''(x) = 0$.

This insures that there is a unique interior solution to x^*. While this assumption is not necessary and only sufficient, it simplifies the subsequent analysis and allows us to focus more clearly on the essential results. In section 4.3 we discuss the issue of multiplicity more deeply.

Proposition 1 *The optimal solution to the maximization problem* (10) *is a* Most Rapid Approach Path. *Given Assumptions 1, 2 and 3, the optimal policy consists in :*

- *If* $x_t < x^*$ *then* $s_t = \bar{s}$, *and* $\dot{x} > 0$;
- *If* $x_t > x^*$ *then* $s_t = \underline{s}$, *and* $\dot{x} < 0$;
- *And once* $x_t = x^*$ *the steady state solution is*

$$s^* = p - \lambda(k + (1-2k)x^*) \tag{12}$$

in which x^* solves (11).

Proof 1 Due to the linearity of the control s_t it is clear that the solution to the optimal control problem is a Most Rapid Approach Path. The Euler equation then is given by (11). The properties of this Euler equation are as follows. Define $SL(x) \equiv w'(x) - (1+\gamma)(p-\lambda k) - 2\lambda x(k-\gamma(1-2k))$, and $SR(x) \equiv \frac{\gamma\phi}{1-x}$. Then we obtain $SR(0) = \gamma\phi > 0$, $SR(1) = \infty$, $SR'(x) = \frac{\gamma\phi}{(1-x)^2} > 0$ and $SR''(x) = 2\frac{\gamma\phi}{(1-x)^3} > 0$. Furthermore, $SL(0) = w'(0) - (1+\gamma)(p-\lambda k)$, $SL(1) = w'(1) - (1+\gamma)(p-\lambda k) - 2\lambda(k-\gamma(1-2k))$, $SL'(x) = w''(x) - 2\lambda(k-\gamma(1-2k))$ and $SL''(x) = w'''(x)$. Thus it is clear that Assumption 3 insures a unique interior equilibrium. Then it is straight forward to see that for $x_t < x^*$ we have $s_t = \bar{s}$, while for $x_t > x^*$ we obtain $s_t = \underline{s}$. For $x_t = x^*$ we have the steady state solution $s^* = p - \lambda(k + (1-2k)x^*)$, where x^* solves (11). ■

At the steady state the corresponding subsidy can be positive or negative. If it is positive, the cost of further increasing the subsidy should be equalized with the discounted value of the benefits from increasing the share of contributors. If it is negative, the public good is taxed and the benefits from this tax should be compared with the losses from decreasing the pool of contributors. We can directly see that $s^* < 0$ if and only if $x^* > \bar{x}$. The policy maker applies a tax if $x_t > x^*$, or a subsidy if $x_t < x^*$. This stands in stark contrast to the general belief that government policy should enforce the full contribution equilibrium and where thus a tax (a negative subsidy) was never even considered. However, in our case the government understands that a tax on the price would help reduce the deadweight loss from other policies elsewhere and this thus provides incentives for the government to not induce the $x = 1$ equilibrium.[5] The assumption that the tax receipts can be used to reduce the deadweight loss elsewhere in society is a standard approach in the public economics literature and indeed an empirically verified regularity (Laffont, 2005). There are obviously further and different ways in which the costs of the government policy can be introduced in a model, and in section 4.2 we look more closely

[5]It must be emphasized that this result arises from a reduction in the deadweight loss that is outside of this model, but without loss to this main result we could easily extend the model to encompass this in a more general equilibrium setting.

into this. We shall show that this result may persist but then depends on parameter configurations.

Two results related to the dynamics are worth stressing: First, the policy enforces the stability of the targeted x^*, and any deviation would trigger an adjustment of the subsidy to ensure that society comes back to x^*. The original social norm tipping point $\bar{x}$ is no longer tipping the system since now the policy maker directs the social norm. Indeed, x^* would be the tipping point associated to s^*, if s^* were fixed once and for all. At the optimal policy the system is no longer tipping.

First, if the subsidy s^* is fixed once and for all, the associated interior share x^* is unstable and society likely converges toward one of the corner 0 or 1. However, since a deviation from this share prompts an adjustment of the subsidy, the steady state is stabilized. When the policy can be adjusted the instability of the interior equilibrium disappears.

Second, it is never optimal to push society toward $x = 1$ because of the cost of public fund and the diffusion dynamics. If x is close to 1 then there are few non-contributors left, and it is too expensive to subsidize contribution to convert them into contributors. Conversely, to tax contributors has a small effect on the dynamic of the norm and is then justified by the cost of public funds.

We now show that, despite the fact that the full contribution equilibrium maximizes social welfare, the presence of the costly public policy may (this depends on the shape of the deadweight loss as we show later) imply that this is finally not the policy maker's preferred equilibrium. For this we define the first best share of contributors as the share x^{FB} that maximizes $V(x)$ defined by equation (6).

Assumption 4 *$w'(1) - p > \lambda k$.*

Hence we assume that we are in a world where full adherence to the norm seems socially desirable, and, by itself, the social norm can potentially move society toward this situation if $x(0) > \bar{x}$.[6] In such a case, it is true that the full contribution equilibrium made possible by the social norm is desirable.

However, this is not necessarily true once we consider costly regulatory intervention. With costly policy, the optimal share of contributors becomes contingent on the regulatory tools used. More specifically, even if Assumption 4 is true then despite this our result above shows that the optimal equilibrium is below the full contribution one, $x^* < x^{FB} = 1$, and thus our regulator never pushes society toward the $x = 1$ equilibrium.[7]

We thus conclude that, despite the assumption that the full contribution equilibrium is socially optimal without government intervention, the existence of this policy intervention changes this result. The full contribution equilibrium, for which authors in the literature on the private contribution of public goods suggest government intervention (Rege (2004); Nyborg et al. (2006); Lin and Yang (2006)), may not be optimal any longer once one seriously considers (costly) government intervention.

3.1 Relation with Pigouvian tax

In general, research that deals with public intervention and social norms discusses whether or not the public policy has a crowding in or crowding out effect. One of the reasons for the focus on crowding in and out is that it is acknowledged that actions undertaken due to social norms have an intrinsically superior value than the same actions induced by public interventions. Furthermore, the argument stands that under crowding out of the social norm a public policy may only have very limited impact in general.

We believe this focus is too limited because it ignores the nature of public interventions: namely to correct market failures and to induce socially optimal outcomes. Based on this line of thought we here make the bridge from the crowding in and out literature to

[6] Note that with $\lambda = w(1) - p$, as assumed by Rege (2004), and $k < 1/2$, the above assumption is satisfied if $w(x)$ is linear with respect to x (which is not the case in Rege (2004)).

[7] If one assumes the absence of costly public funds, the regulator sets a high subsidy to move society toward $x = 1$ as fast as possible whether there is a social norm or not, convergence is accelerated by the social norm which is welfare improving.

the Pigouvian tax literature. In fact, the textbook way to correct for externalities is a Pigouvian tax. In the model that we study a policy maker faces two externalities, the externality of the public good and the one of the social norm.[8] The positive externality associated with contribution to the public good justifies a subsidy. At the same time the social norm may help to overcome the free-riding problem, but it also introduces a social externalities related to social approval and disapproval. Thus it needs to be optimally managed as well.

Some readers now may argue that we endow the policy maker with only one instrument to deal with both externalities. However, we want to emphasize that this is precisely where Rege (2004)'s model comes in handy. Both the externality of the social norm and the public good problem are defined by one variable only, namely x. In other words, both externalities only depend on the single variable x. Hence one policy tool is sufficient in this case.

We now relate our previous results to Pigouvian taxes. We can only do this comparison at equilibrium where $x_t = x^*$, and thus constant, simply because we know that during transition the MRAP implies that s_t will be at either its lower or upper bound. In order to clarify the relationship with a Pigouvian tax we combine equation (11) and (12), assuming there exists an interior solution to equation (11), to obtain

$$s^* = \frac{w'(x^*) - \lambda x^*}{1+\gamma} - \frac{\gamma}{1+\gamma}\left(\frac{\phi}{1-x^*} - \lambda(1-2k)x^*\right). \tag{13}$$

The formula we obtain here is very much akin to the Ramsey formula of optimal taxation.[9] The term $(w'(x) - \lambda x)/(1+\gamma)$ is the marginal benefit from the public good in public monetary units, its difference with the optimal subsidy being the implicit tax

[8]In fact, while our policy maker wants to deal with those two externalities his public policy introduces a third market failure, namely a deadweight loss or gain due to the public funds. We will study the implication of not having this externality, but for a smoother working of the model we require costly public funds, otherwise there would never be an interior solution to the model. Plus, of course, it is a realistic feature of the model.

[9]For $\phi = 0$, Equation (13) can be rewritten in the Lerner-Ramsey form, defining $\epsilon = (p - s^*)/(\lambda(1 - 2k)x^*)$ (the elasticity of the steady state share of contributors with respect to the cost of contributing),

on contributory behavior. Thus, s^* can then be related to the standard Pigouvian tax (or subsidy) at equilibrium. There are three components that play a role: the marginal benefit of the public good, the costs (or benefits) of public funds, and the social norm. We now study the role of the different components.

First, we set λ and γ equal to zero, and denote the optimal solution in this case as $x^*_{\gamma\lambda}$. This case corresponds to one where public funds do not come at a cost and the social norm does not evolve. Equation (13) is then simply $s^*_{\gamma\lambda} = w'(x^*_{\gamma\lambda})$, and the optimal subsidy is the Pigouvian subsidy which is equal to the marginal external benefit, and at equilibrium obviously also equal to the price p.[10] This extreme case corresponds to a world without a social norm and without costly government funds and it prescribes an optimal Pigouvian tax along the lines of the standard public good literature Samuelson (1954). Note that, under Assumption 4, thus if $w'(1) > p$, then it is clear that the optimal solution would be to have $s_t = \bar{s}$, $\forall t$. In this case the corner $x = 1$ will be approached over time and the policy maker makes sure that society stays there.

Let us align this point more closely with the public goods literature that considers costly government funds. In this case $\gamma > 0$, but still $\lambda = 0$, and the policy maker would want to achieve an optimal level of the public good x_λ corresponding to

$$w'(x_\lambda) = \frac{\gamma\phi}{1 - x_\lambda} + (1 + \gamma)p.$$

Thus, in this case $w'(x_\lambda) > p$ and hence the solution for x_λ will be lower compared to $x_{\gamma\lambda}$ as now the costs of the public funds make policy intervention at equilibrium more costly. Thus the costs of public funds essentially create a wedge between the price of the public

then

$$\frac{1}{p - s^*}\Big[\frac{w'(x^*) - \lambda x^*}{1 + \gamma} - s^*\Big] = -\frac{\gamma}{1 + \gamma}\frac{1}{\epsilon}.$$

[10] It is clear that this result is not fully mathematically correct but corresponds more to a limiting case. To be precise, if $\gamma = \lambda = 0$, then welfare W will be maximized at $w'(x) = p$. As there is now no cost to policy, then it is clear that the optimal policy will be a bang-bang solution, with $s_t = \bar{s}$ for $x_t < (w')^{-1}(p)$, $s_t = s^*_{\gamma\lambda}$ for $x_t = (w')^{-1}(p)$, and $s_t = \underline{s}$ otherwise. This result obviously obtains since the government knows that equation 5 still applies, meaning that agents decide according to the utility differences.

good p and the marginal benefit to each agent. In addition, and more importantly, these costs interact with the dynamic diffusion of the norm which explains the presence of the discount rate in the formula above. As is also clear, it is not optimal to push society towards the full contribution equilibrium, which implies that the result $x^* < 1$ is not directly related to the presence of the social norm but the combination of the replicator dynamics and the linear cost of public funds.

Assume now that the social norm plays a role, $\lambda > 0$, but for clarity that the cost of public funding is zero, $\gamma = 0$, such that equation (13) becomes $s^*_\gamma = w'(x^*_\gamma) - \lambda x^*_\gamma$. From equation (7) we know that $-\lambda x$ is the social external cost at the intertemporal equilibrium. In this case the optimal subsidy encompasses two Pigouvian terms, the external benefit of the public good, as well as the social external cost. However, social benefits and costs also play another role due to the utility difference. Even though the optimal subsidy is lower with the social norm than without it the optimal share of contributors might well be larger with the social norm.[11]

Whether the presence of the social norm justifies a higher or lower optimal share of contributors depends on several factors, the negative social externality but higher internalized incentive to contribute, and, in addition, the lower regulatory costs when $\gamma > 0$. The optimal share of contributors can still be larger with the social norm than without it. It is illustrated in Figure (2), in which x^* and x_λ are depicted as a function of the cost of public fund. In the Figure, Assumption 4 is satisfied ($x^{FB} = 1$). For a small cost of public funds, the optimal share of contributors is larger without the social norm than with it. The benefits associated with the lower subsidy allowed by the social norm are not sufficient to compensate for the social external costs.[12] The comparison is reversed for large cost of public funds.

[11]The total social marginal benefits of an increase of x is $\lambda(1-2x)k$, of which $\lambda[(1-2x)k + x]$ are internalized when $\Delta U = 0$, and $-\lambda x$ are external costs.

[12]Note that for $x^* > \bar{x}$ the equilibrium subsidy is positive.

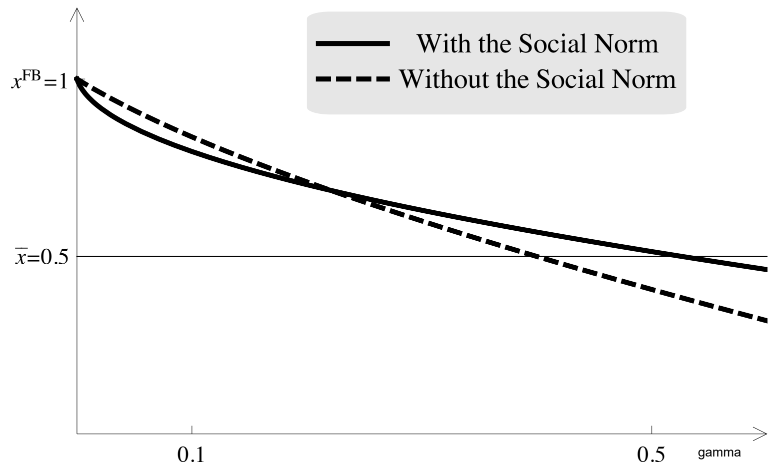


Figure 2: Optimal share of contributors as a function of the costs of public funds with the social norm (thick line) and without it (dashed line). The figure is obtained for $w(x) = (2 - x/2)x$, $p = 0.5$, $k = 0.4$, $\lambda = w(1) - p = 1$, $\phi = 1$

4 Further considerations

Our main results above clearly indicate that it is necessary to study the role of *optimal* government policy on the incentives for private contribution to public goods when social norms matter. When we talk about public policy, then there are some further issues that come to mind that we have not addressed above. In particular, we shall look at the implication of public debt, the role of the deadweight loss, as well as path dependency and multiple equilibria.

4.1 Public debt

It is clear that public debt should play a role for several reasons. Firstly, public debt will help to alleviate the exogenous bounds on the policy in case our Assumption 2 would not hold. Our bounds are exogenously assumed to be $\underline{s} \leq p - \lambda(1-k)$ and $\bar{s} \geq p - \lambda k$. Then let us take the case where, for example, $\bar{s} < p - \lambda k$. In this situation the subsidy is not large enough to induce a positive utility difference even for low levels of the social norm. As a result, the policy maker cannot induce a change in the social norm and public policy would not have any effect. However, imagine now that the policy maker can raise public debt and thereby endogenize these bounds. If this debt comes at a sufficiently low price (interest rate), then the policy maker would clearly be inclined to raise debt in order to be able to influence the norm.

Another reason for which public debt may play an important role is that the Ricardian equivalence is not going to hold in this setting. Agents are not forward looking enough to know that this debt needs to be financed, and subsequently will not be able to take the impact of the debt on their optimal decisions into account. Hence, debt will be an effective tool to overcome the potential bounds on public policy if Assumption 2 would not hold.

4.2 The cost of public funds

In the preceding sections, we made the common assumption of a linear costs of public funds. This linearity led to the result that it is never optimal to push society toward the $x = 1$ equilibrium. This assumption is justified in a partial equilibrium setting when the cost mainly comes from taxation deadweight losses (e.g. Laffont, 2005). In contrast, in a more general equilibrium setting, one can consider that whether the policy maker introduces a tax or a subsidy, there are administrative costs, related to enforcement and tax collection. This approach is similar to Bowles and Hwang (2008).

Thus, our objective here is threefold. First, since we want to motivate readers to further investigate the role of optimal policy in the social norm literature, we want to show that assuming linear or non-linear costs of public interventions can lead to substantial differences in the results. For example, we shall show that moving to a non-linear modeling of the cost can make the $x = 1$ social norm level, in contrast to the linear case, an optimal equilibrium. Second, our motivation is to move away from this somewhat partial equilibrium argument that founded our linear deadweight loss model, towards a general equilibrium model. While this may be an argument of semantics mostly, it may be a more appealing setting to the macroeconomic readership. Third, this non-linear case allows us to derive additional results regarding the choice to move between equilibria and the importance of initial conditions.

Let us assume that there is a cost to collect subsidies that is a quadratic function of the subsidy per individual so that the total cost is now given by $x\gamma s^2/2$.[13] The objective of the social planner is then to maximize

$$W_t = \int_{t=0}^{\infty} \Big(V(x_t) - \gamma x_t s_t^2/2 \Big) e^{-\phi t}. \tag{14}$$

The policy maker then maximizes equation (14) subject to s_t and the constraint (9). We delegate the derivations to the Appendix and only present the main results here.

After maximization we can derive a system of differential equations in $\{x, s\}$, which is given by

$$\begin{aligned} \dot{x} &= x(1-x)\Big(s - p + \lambda(k + (1-2k)x) \Big), & (15) \\ \dot{s} &= -xs\big(s - p + \lambda(k + (1-2k)x)\big) \\ &\quad + s\big(\phi - (1-2x)\big(s - p + \lambda k\big) - (2-3x)\lambda(1-2k)x\big) & (16) \\ &\quad - \frac{1-x}{\gamma}\Big(w'(x) - p - \lambda(1-2x)k - \gamma s^2/2 \Big). \end{aligned}$$

This system completely describes the dynamics of x_t and s_t. It gives rise to three potential candidates for steady states.

The first candidate is the $x = 0$ equilibrium. While we know that $x_t = 0$ is one of the potential steady state solutions for $\dot{x}_t = 0$, we also know that s_t at $x_t = 0$ is a variable that the policy maker can choose freely as it does not entail a social cost. We need to figure out whether the dynamic system would make us approach this steady state. In the convergence to this steady state the necessary conditions must be fulfilled. Whether

[13] One suggestion that we came across is to use the form $x(\gamma_0 s + \gamma_1 s^2)$ in order to link the linear and the quadratic case. However, the problem here is that this functional form only makes sense for positive s.

or not convergence to this steady state is optimal will depend on the shape of the phase curves and thus the associated dynamics.

Whether it can be optimal locally, for small initial level of the norm, to go toward $x = 0$ depends notably on $w'(0)$. If the marginal benefit from the public good is large at $x = 0$ it is then never optimal, whatever x_0, to go toward $x = 0$, the policy maker should then sufficiently subsidize the behavior to help overcome the social trap. Conversely, if $w'(0)$ is small, then the policy maker should not prevent the norm from disappearing, but should still set a subsidy in order to slow the process (see Appendix for further analysis). It is illustrated in Figure 4b in which for small initial x_0 the green trajectory should be followed.

The second candidate is the equilibrium of $x = 1$. This is the steady state where the policy maker would push for the highest level of the social norm in society. Substituting the $x_t = 1$ solution into the dynamic system yields the logical optimal solution $s_t = 0$. We know that from the threshold level $\bar{x}$ onwards the social norm is self-enforcing, then it is clear that positive subsidies after x_t crossed this threshold are only necessary in order to push x_t faster towards its steady state. Intuitively, the reason for a positive level of s_t for $x_t > \bar{x}$ is then only that the deadweight loss is small compared to the higher social norm (which would anyway have occurred).

The Jacobian around the $\{x_2, s_2\} = \{1, 0\}$ steady state is given by

$$\mathcal{J}\Big|_{(x_2,s_2)} = \begin{bmatrix} p - (1-k)\lambda & 0 \\ \frac{w'(1)-p-k\lambda}{\gamma} & \phi \end{bmatrix}.$$

As this is a lower triangular matrix we have that the eigenvalues are given by $EV_1 = p - (1-k)\lambda$ and $EV_2 = \phi$. This steady state is saddle path stable if $p < (1-k)\lambda$, which applies given Assumption 1. Conclusively, it is now possible that this steady state is optimal, which stands in stark contrast to the linear deadweight loss case.

The intuition is restored with the quadratic cost specification: it can be optimal (at least locally) to introduce a subsidy to stimulate the norm and then progressively phase out the policy as the norm spreads through society. The explanation for the difference between the quadratic and linear cost is related to the marginal cost to push the norm as the x gets close to 1. As x converges toward 1, the subsidy is progressively reduces to zero and so does the marginal cost per contributor γs, consequently the annualized cost to push the norm $\gamma s/(1-x)$ stays bounded in contrast to the linear cost case.

Hence, we find that not only is it important to acknowledge that there is a wider need to study the implication of optimal policy in the social norms literature, but we also observe that the way we model this public intervention can yield vastly different results given the optimal strategy that a policy maker may want to pursue.

The third candidate is the interior equilibrium characterized by $\Delta U(x,s)=0$, that is, $s=p-\lambda(k+(1-2k)x)$. Substitute this into equation (17) evaluated at steady state, using $V'(x)=w'(x)-p+\lambda(1-2x)k)$ and defining $s(x)=p-\lambda(k+(1-2k)x)$, gives us

$$V'(x)-\gamma\left[\frac{1}{2}s(x)^2+xs(x)s'(x)\right]=\frac{\gamma s(x)\phi}{1-x}. \tag{17}$$

The parallel with equation (11) is then clear, the bracketed term is the derivative of the public cost $\gamma xs(x)^2/2$. The right hand side is the annualized marginal cost of an increase of x,and contrary to the linear cost case, it is proportional to the subsidy s.

This interior steady state equation is rather complicated and allows for a multitude of combinations of interior steady states with a wide variety of dynamics. The interior steady state can be unique and stable or unstable, or there can exist interior multiple steady states which are stable or unstable with or without complex dynamics, or it is also possible that no interior steady state exists at all. Finally, there is the possibility of Skiba points, such that there exists an initial condition $x(0)$ for which it is optimal to converge to either of the various equilibria. We discuss this in the next section.

4.3 Path dependency and parameter stability

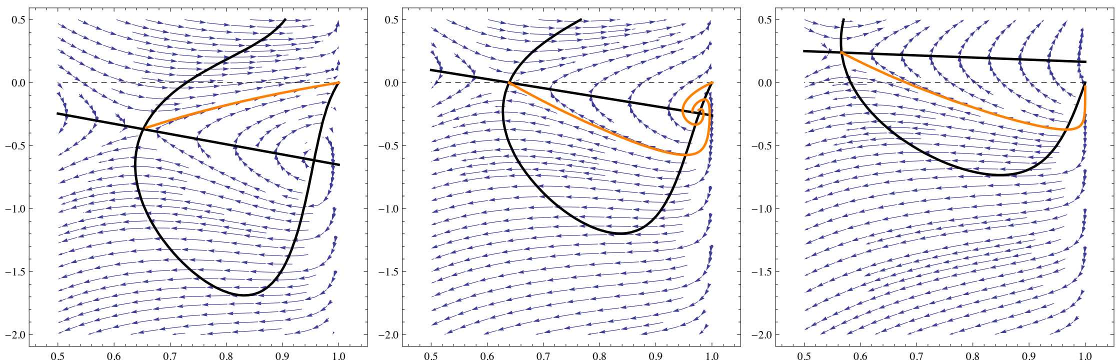
One issue that we have so far avoided is path dependency and parameter stability. It is a well-known result that initial conditions matter already without government intervention. For example, as Rege (2004) has shown, if the initial distribution of the social norm in society is favorable (meaning $x_0 > \bar{x}$), then society will converge to the full contribution equilibrium. Thus, whatever path led society to this initial condition, its subsequent evolution is fully depending on that level.

At the same time, it is clear that society first needs to develop a certain social norm, and these developments need to be done from scratch. In other words, society would be expected to start around the $x = 0$ equilibrium. Thus, for many social norms that have similar qualitative features such as ours, one can expect that, without some further incentives to initially adopt the norm, then no one in society would ever adhere to it. This, obviously holds especially true for the type of norm as developed in Rege (2004), where for a low initial distribution of the norm ($x_0 < \bar{x}$) society would never adopt it without something that provides further impetus of some sort.

In our extension above we have argued that public policy may want to provide such an incentive, and this incentive should be introduced in a socially-optimal way. As we argued, depending on parameters and functional forms, a policy maker would find it optimal to make society adhere fully to the social norm, to have no one adhere to the social norm, or any conceivable intermediate result. While it may be possible for a well-informed policy maker to know what would be the optimal policy, we shall now present a further complication which makes it very difficult to judge as to what is the correct policy. We shall do this with the social welfare function, equation (14), with the squared costs in mind.

As we have argued above, in the case of the squared costs we can easily identify a variety of potential steady states,[14] some of which have properties that give difficulties to policy choices. To be specific, let us look at Figures 3a to 3c. Note that x is only depicted between 0.5 and 1. As we can see, there are at maximum three potential equilibria, one being the corner equilibrium ($x = 1$, $s = 0$), the other being a saddle-path stable interior equilibrium, and between this interior equilibrium and the corner equilibrium there may exist an unstable spiral equilibrium. In the case of Figure 3b, around this unstable spiral

[14] Multiple steady states are also possible with the linear cost case, but we have yet to be able to show the existence of bifurcations and other qualitative changes to the dynamics in that case.



(a) Figure 3a (b) Figure 3b (c) Figure 3c

Figure 3: Phase diagrams for the dynamic system derived from (14) with constraint (9). On the x-axis we have x_t, on the y-axis s_t. The full thick lines denote the $\dot{x}_t = 0$ and $\dot{s}_t = 0$ phase curves, the orange lines show the stable manifolds. We assume $w(x) = ax^b$. In Figure 3a we use the parameters $a = 2.8$, $b = 0.3$, $p = 0.7695$, $k = 0.3$, $\gamma = 1$, $\phi = 0.08$. The unstable manifold of the interior steady state connects with the stable manifold of the corner steady state. In Figure 3b we use the parameters $a = 2.8$, $b = 0.3$, $p = 1$, $k = 0.3$, $\gamma = 1$, $\phi = 0.08$. There is a Skiba point around the unstable, spiraling interior steady state. In Figure 3c we use the parameters $a = 2.8$, $b = 0.3$, $p = 1.1$, $k = 0.45$, $\gamma = 1$, $\phi = 0.08$. The unstable manifold of the corner steady state connects with the stable manifold of the interior steady state.

equilibrium there is a Skiba point (see e.g. Wagener, 2003). A Skiba point is an initial share of contributors x_0, such that it is optimal to converge to either of both surrounding stable equilibria (the saddle-path stable interior equilibrium or the corner). However, for a small change in the initial condition to the left or right, only one of the equilibria is optimal. This is an example of a path dependency which shows that it is vital for a policy maker to precisely know the depth of the social norm throughout society.

Thinking this a step further, it also means that there is a historical lock-in, or social trap, even from the perspective of a policy maker searching for an optimal policy. Thus, for a low initial distribution of the social norm the policy maker may not find it optimal to induce a high equilibrium in society as it is simply too costly or time intensive. Instead, in this case the policy maker may view a low distribution of the social norm as socially optimal. Hence, social traps may be the result of this setting.

Furthermore, for a small change in parameters the system can undergo a significant qualitative change (a bifurcation), implying that a previously targeted equilibrium may no longer be optimal. Hence, it is vital to have a very precise estimate of the parameters, as any smaller variation in a parameter could mean that the policy maker has to completely reverse his previous policy. For example, moving from Figure 3a to Figure 3b we only changed k by 0.15, but this changed the optimal equilibrium from the corner one to having both the corner and the saddle-path stable interior equilibrium being optimal. Another smaller change in k (by 0.05) implies that the interior state becomes optimal. Conclusively, both a deep knowledge on the structural parameters as well as a sufficient parameter stability seem essential for policy intervention to be successful.

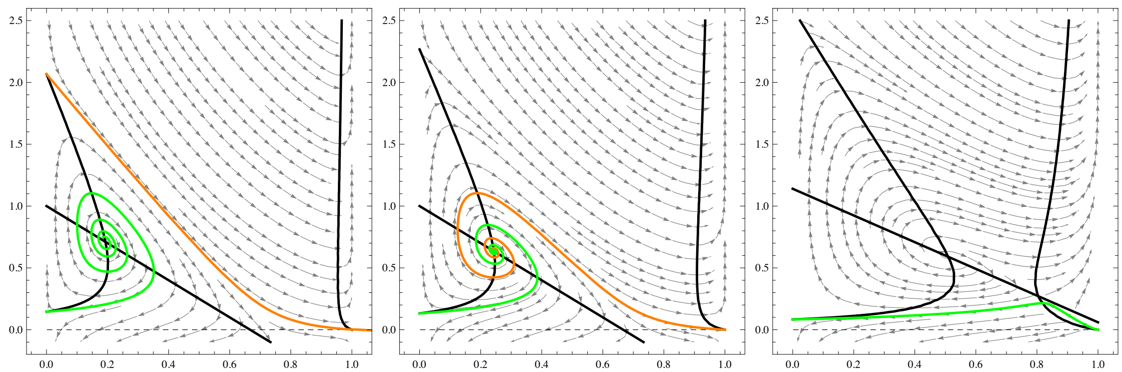
Few comments may help interpret the Figure 3. First, with the specification used ($w'(0) = +\infty$) it is never optimal to converge towards $x = 0$. For small initial values of the norm the policy maker implements a subsidy that is progressively reduced until x reaches the first (small x) equilibrium. Concerning the $x = 1$ equilibrium, $w'(1)$ is smaller than p which means that without costly public fund it would not be optimal to implement $x = 1$.[15] It is why the $\dot{s} = 0$ curve reaches point $(x = 1, s = 0)$ from below which corresponds to a tax.[16] The policy maker actually slows the diffusion of a too strong social norm if the norm is already well developed, he implements a small tax that is progressively reduces (higher branch of the spiral). For intermediate initial values of the norm, he reduces it by setting a large tax (lower branch of the spiral), which is also progressively reduces. For an initially low distribution of the norm it is optimal to subsidize it and progressively reduce the subsidy.

Technically to get a spiral and a Skiba point around the $x = 1$ equilibrium one needs that $w'(1) < (p - \lambda k)$ (so that the $\dot{s} = 0$ curve cross the $\dot{x} = 0$ curve from above) and the policy associated is a tax near the corner. However, it is also possible to get a spiral and

[15] With the specification used for the Figure 3: $w'(1) = b \times a = 0.3 \times 2.8 = 0.84$ which is smaller than p.

[16] Locally, at $x = 1$, the derivative with respect to x of the curve $\dot{s} = 0$ is $-1/(\phi\gamma)(w'(1) - p - \lambda k)$.

complex dynamics with a subsidy along the optimal trajectory as illustrated in Figure 4, for the case of a linear $w(x)$.



(a) Figure 4a (b) Figure 4b (c) Figure 4c

Figure 4: Phase diagrams for the dynamic system derived from (14) with constraint (9). On the x-axis we have x_t, on the y-axis s_t. The full thick lines denote the $\dot{x}_t = 0$ and $\dot{s}_t = 0$ phase curves, the orange and green lines show the stable manifolds. We assume $w(x) = ax$. In Figure 3a we use the parameters $a = 4$, $p = 1.5$, $k = 0.2$, $\gamma = 20$, $\phi = 0.106$. In Figure 3b we use the parameters $a = 4$, $p = 1.5$, $k = 0.2$, $\gamma = 20$, $\phi = 0.2$. In Figure 3c we use the parameters $a = 3.3$, $p = 1.5$, $k = 0.2$, $\gamma = 20$, $\phi = 0.2$. .

5 Conclusion

The literature on the private provision of public goods has mostly settled on social norms as a reason for which private agents would provide public goods. In this literature it has also been emphasized that there is room for public policy to induce the 'good' social norm of everyone contributing. As a result, the literature has to a large extend focused on whether or not public policy crowds in or out private provisions. In this article we have argued that it is not enough to focus on crowding in or out of private provisions as an argument for or against public policy. Instead, we argued that public policy needs to be assessed on the grounds of whether or not it is actually optimal from a society's perspective.

In order to study this we extended the model developed in Rege (2004) and introduced endogenous public policy. We showed that in very simple settings where the public policy is subject to a linear deadweight loss then it is not optimal to induce everyone to adhere to the social norm. Furthermore, we have shown that a Most Rapid Approach Path is the optimal solution and thus convergence and speed of convergence depends on the bounds of the public policy. In other words, results depend on in how far the policy maker can subsidize or tax private contributors. Optimality of the corner or interior solutions in the social norm then naturally depend on a variety of parameters.

Extending this simple model of a social norm to the case of non-linear deadweight losses turns out to have surprisingly complicated dynamics once a policy maker wants to take optimal policy into account. Here, we find a variety of potential outcomes, from a case with no interior equilibrium being optimal, to one with only a unique interior equilibrium, Skiba points and traps in social norms. Furthermore, while the equilibrium where everyone fully adheres to the social norm is always a potential equilibrium and it is, in fact, always locally stable, it does not need to be the optimal equilibrium.

In practical terms this result suggests that it is important to investigate the social optimality of government interventions in social norms. This has significant implications for example for the fashionable nudging (Thaler Richard and Sunstein Cass, 2008), for the analysis of social norms and public policy, and for in how far the government should intervene when it comes to the private provision of public goods. Furthermore, the analytical results on multiple steady states and various dynamics already show that the practical difficulties of judging the optimal policies could be very large. This certainly points to a stringent research agenda.

There are many applications and extensions that come to mind. Practically one would, for example, expect that a policy maker could be ignorant of a social norm that evolves through society. This case would correspond to one with asymmetric or limited information on the policy maker's side. The main issue would then be that a policy maker, oblivious to the fact that there is a social norm, would nevertheless set a certain policy, but only understand over time that his criterion was false. This could have important repercussions for the evolution of the norm, which could potentially not only crowd out the social norm but additionally result in sub-optimal policy decisions. The question would be whether one could design optimal policy rules despite having limited information on a social norm.

References

Andreoni, James. 1990. "Impure altruism and donations to public goods: A theory of warm-glow giving." *The economic journal*, 100(401): 464–477.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and prosocial behavior." *The American economic review*, 96(5): 1652–1678.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian. 1986. "On the private provision of public goods." *Journal of public economics*, 29(1): 25–49.

Börgers, Tilman, and Rajiv Sarin. 1997. "Learning through reinforcement and replicator dynamics." *Journal of Economic Theory*, 77(1): 1–14.

Bowles, Samuel, and Sung-Ha Hwang. 2008. "Social preferences and public economics: Mechanism design when social preferences depend on incentives." *Journal of Public Economics*, 92(8): 1811–1820.

Bowles, Samuel, and Sandra Polania-Reyes. 2012. "Economic incentives and social preferences: substitutes or complements?" *Journal of Economic Literature*, 50(2): 368–425.

Brekke, Kjell Arne, Snorre Kverndokk, and Karine Nyborg. 2003. "An economic model of moral motivation." *Journal of public economics*, 87(9): 1967–1983.

Bruvoll, Annegrete, and Karine Nyborg. 2004. "The cold shiver of not giving enough: on the social cost of recycling campaigns." *Land Economics*, 80(4): 539–549.

Holländer, Heinz. 1990. "A social exchange approach to voluntary cooperation." *The American Economic Review* 1157–1167.

Laffont, Jean-Jacques. 2005. *Regulation and development.*: Cambridge University Press.

Lin, Chung-cheng, and CC Yang. 2006. "Fine enough or don't fine at all." *Journal of Economic Behavior & Organization*, 59(2): 195–213.

Nyborg, Karine. 2003. "The impact of public policy on social and moral norms: some examples." *Journal of Consumer Policy*, 26(3): 259–277.

Nyborg, Karine, Richard B Howarth, and Kjell Arne Brekke. 2006. "Green consumers and public policy: On socially contingent moral motivation." *Resource and energy economics*, 28(4): 351–366.

Nyborg, Karine, and Mari Rege. 2003. "Does public policy crowd out private contributions to public goods." *Public Choice*, 115(3-4): 397–418.

Rege, Mari. 2004. "Social norms and private provision of public goods." *Journal of Public Economic Theory*, 6(1): 65–77.

Samuelson, Paul A. 1954. "The pure theory of public expenditure." *The review of economics and statistics*, 36(4): 387–389.

Sugden, Robert. 1982. "On the economics of philanthropy." *The Economic Journal*, 92(366): 341–350.

Sugden, Robert. 1984. "Reciprocity: the supply of public goods through voluntary contributions." *The Economic Journal*, 94(376): 772–787.

Taylor, Peter D, and Leo B Jonker. 1978. "Evolutionary stable strategies and game dynamics." *Mathematical Biosciences*, 40(1-2): 145–156.

Thaler Richard, H, and R Sunstein Cass. 2008. "Nudge: Improving decisions about health, wealth, and happiness."

Wagener, Florian OO. 2003. "Skiba points and heteroclinic bifurcations, with applications to the shallow lake system." *Journal of economic dynamics and control*, 27(9): 1533–1561.

Warr, Peter G. 1983. "The private provision of a public good is independent of the distribution of income." *Economics letters*, 13(2-3): 207–211.

Weibull, Jörgen W. 1997. *Evolutionary Game Theory.*: MIT press.

A Linear cost

In this section we derive the Euler equation (11) using standard Hamiltonian techniques in order to provide intuition for the result.

The objective is to maximize welfare given by (8) with respect to s_t subject to $\dot{x}_t = x_t(1-x_t)\Delta U(x_t, s_t)$ and with bounds on subsidies in the form of $\underline{s} \le s_t \le \bar{s}$. The Hamiltonian is

$$\mathcal{H}(x, s, \mu) = V(x) - \gamma x s + \mu x(1-x)\Delta U(x, s).$$

the derivative with respect to s is $\mathcal{H}_s = -\gamma x + \mu x(1-x)$, so that

$$s = \begin{cases} \bar{s} & \text{if } \mu(1-x) > \gamma \\ \underline{s} & \text{if } \mu(1-x) < \gamma \\ \textit{indeterminate} & \text{if } \mu(1-x) = \gamma \end{cases} \tag{18}$$

the evolution of μ is given by

$$\begin{aligned} \dot{\mu} = \phi\mu - \mathcal{H}_x &= \mu\phi - \mu\big[(1-2x)\Delta U + x(1-x)\Delta U_x\big] - \big(V'(x) - \gamma s\big) \\ &= \mu\phi - \mu\big[(1-2x)\Delta U + x(1-x)\Delta U_x\big] - V'(x) + \gamma\big[\Delta U - x\Delta U_x + (p - \lambda k)\big] \\ &= \mu\phi + [\gamma - \mu(1-2x)]\Delta U - [\gamma + \mu(1-x)]x\Delta U_x - V'(x) + \gamma(p - \lambda k) \end{aligned}$$

if $s \in (0,1)$ then $\mu(1-x) = \gamma$ and $\dot{\mu} = \gamma \dot{x}/(1-x)^2 = x\Delta U/(1-x)$ injecting this in the above equation gives

$$\frac{x}{(1-x)}\Delta U = \frac{\gamma\phi}{1-x} + \frac{\gamma}{1-x}\Big[x\Delta U - 2(1-x)\Delta U_x\Big] - V'(x) + \gamma(p - \lambda k)$$

that is $V'(x) - \gamma\big[p - \lambda k - 2\lambda(1-2k)x\big] = \frac{\gamma\phi}{1-x}$.

B Square costs

We write the Hamiltonian as

$$\mathcal{H} = V(x) - \gamma x s^2/2 + \mu x(1-x)\Delta U(x,s).$$

First order conditions yield

$$\begin{aligned} s &= \frac{\mu(1-x)}{\gamma}, & (19)\\ \dot{\mu} &= \mu\Big[\phi - (1-2x)\big(s - p + \lambda k\big) - (2-3x)\lambda(1-2k)x\Big] - [V'(x) - \gamma s^2/2] & (20) \end{aligned}$$

Differentiating equation (19) wrt time and solving for $\dot{\mu}$ yields

$$\dot{\mu} = \frac{\gamma}{1-x}\dot{s} + \frac{\gamma s}{(1-x)^2}\dot{x}.$$

We now obtain a system of differential equations in $\{x, s\}$, which is given by

$$\dot{x} = x(1-x)\Delta U \tag{21}$$

$$\dot{s} = s\phi - s(1-x)[\Delta U + x\Delta U_x] - \frac{1-x}{\gamma}\Big[V'(x) - \frac{\gamma}{2}s^2\Big] \tag{22}$$

The elements of the Jacobian matrix are

$$J_{xx} = (1-2x)\Delta U + x(1-x)\Delta U_x = (1-x)[\Delta U + x\Delta U_x] - x\Delta U \tag{23}$$

$$J_{xs} = x(1-x) \tag{24}$$

$$\begin{aligned} J_{ss} &= \phi - (1-x)[\Delta U + x\Delta U_x] - s(1-x) - (1-x)\frac{W_{xs}}{\gamma} \\ &= \phi - (1-x)[\Delta U + x\Delta U_x] \end{aligned} \tag{25}$$

$$\begin{aligned} J_{sx} &= s[\Delta U + x\Delta U_x] - s(1-x)[2\Delta U_x + x\Delta U_{xx}] + \frac{1}{\gamma}[(V' - \gamma s^2/2) - (1-x)V\prime\prime] \\ &= s[\Delta U + (3x-2)\Delta U_x] + \frac{1}{\gamma}[(V' - \gamma s^2/2) - (1-x)V\prime\prime] \end{aligned} \tag{26}$$

There are three candidates for steady states:

candidate 1: $x = 0$

If there is an optimal trajectory that converges toward $x = 0$, the associated s is bounded and converges toward the solution of the quadratic equation

$$0 = s\big(\phi - (s - p + \lambda k)\big) + \frac{1}{\gamma}\Big(\gamma s^2/2 - \lambda k - w'(0) + p\Big),$$

- There is a solution to this equation iff
$$w'(0) < \frac{\gamma}{2}\big(p - \lambda k + \phi\big)^2 + p - \lambda k.$$

- If $w'(0) < \frac{\gamma}{2}\big(p - \lambda k + \phi\big)^2 + p - \lambda k$, the two solutions are
$$s_{0\pm} = \phi + p - \lambda k \pm \sqrt{(\phi + p - \lambda k)^2 - \frac{2}{\gamma}\left[w'(0) - (p - \lambda k)\right]}$$

- For a solution s, the associated Jacobian matrix is lower triangular ($J_{xs} = 0$), the two eigenvalues are $\Delta U(0,s) = s - (p - \lambda k)$ and $\phi - \Delta U(0,s) = \phi + (p - \lambda k) - s$. There is a saddle path if $\Delta U(0,s) < 0$ or $\Delta U(0,s) > \phi$.

- s_{0-} is the only candidate, it is a saddle if $\Delta U(0, s_{0-}) < 0$ that is $s_{0-} < p - \lambda k$ or

$$\phi < \sqrt{(\phi + p - \lambda k)^2 - \frac{2}{\gamma}\left(w'(0) - (p - \lambda k)\right)}$$

or

$$\frac{2}{\gamma}\Big[w'(0) - (p - \lambda k)\Big] < (p - \lambda k)\Big[2\phi + (p - \lambda k)\Big]$$

- s_{0+} cannot be the limit of an optimal trajectory: s_{0+} is strictly larger than $p - \lambda k$, so that $\Delta U(x, s_{0+}) > 0$ in a neighborhood of 0. Along an optimal trajectory converging toward $(s_{0+}, x = 0)$, x_t is eventually decreasing (since $x_t > 0$) so $\Delta U(x_t, s_t)$ is eventually negative, a contradiction.
- Then $\Delta U(0, s_{0-}) < \phi$ and $(x = 0, s = s_{0-})$ is a saddle if and only if $\Delta U(0, s_{0-}) < 0$.

candidate 2: $x = 1$

- Substituting the $x_t = 1$ solution into the dynamic system yields the logical optimal solution $s_t = 0$.
- The Jacobian around the $\{x_2, s_2\} = \{1, 0\}$ steady state is given by

$$\mathcal{J}\Big|_{(x_2,s_2)} = \begin{bmatrix} p - (1-k)\lambda & 0 \\ \frac{w'(1) - p - k\lambda}{\gamma} & \phi \end{bmatrix}.$$

- As this is a lower triangular matrix we have that the eigenvalues are given by $EV_1 = p - (1-k)\lambda$ and $EV_2 = \phi$.
- This steady state is saddle path stable if $p < (1-k)\lambda$, which applies given Assumption 1.
- If $p < (1-k)\lambda$, in a neighborhood of $x = 1$ it is optimal to go toward $x = 1$.

candidate 3: $x = \frac{p - s - \lambda k}{(1-2k)\lambda}$, then x is solution of equation (**??**). There can be multiple solutions of this equation. The Jacobian at a solution is given by equations (23) to (25). The trace is equal to ϕ, the eigenvalues are:

$$EV_{\pm} = \frac{1}{2}\left[tr \pm \sqrt{tr^2 - 4D}\right]$$

in which D is the determinant of the Jacobian.

- If $D < 0$ there is a saddle equilibrium which is then possibly locally associated to an optimal trajectory.
- If $D > 0$ the equilibrium is unstable, with spirals if $D > \phi^2/4$.

The determinant is

$$D = x(1-x)\left\{\Delta U_x[\phi - (1-x)x\Delta U_x] - s(3x-2)\Delta U_x + [V' - \gamma s^2/2 - (1-x)V'']/\gamma\right\} \tag{27}$$

with $\Delta U_x = \lambda(1-2k)$ is positive, and decreasing with k. Note that $x(1-x)$ is maximized at $x = 1/2$ for a value of $1/4$; $3x-2$ is negative for $x < 2/3$.