Moving toward Greener Societies: Moral Motivation and Green Behaviour

Lorenzo Cerda Planas*

Last revision: May 31st, 2015

Abstract

This paper intends to provide an alternative explanation for why societies exhibit varied environmental behaviours. To explore this idea, I use a Kantian moral approach at a microeconomic level. Under this premise, I show that two identical societies (according to income level and political system) might follow different paths with respect to their "green" behaviour. Additionally, I identify tipping points that could nudge a society from a polluting behaviour to a green one. I find that environmental perception, as well as how governments are elected, can be important factors in this shift.

JEL Classification: C62, D64, D72, H41, Q50.

Keywords: Green behaviour, Kantian morale, moral motivation, tipping points.

*Paris School of Economics, C.E.S. - Université de Paris 1 Panthéon-Sorbonne; 106-112, Boulevard de l'Hôpital, 75647 Paris Cedex 13, France. Email: lorenzo.cerda@gmail.com

Acknowledgments: I want to thank Bertrand Wigniolle, Eugenio Figueroa and Scott Barrett for their invaluable help and support. I would also like to thank Maria Kuecken for her thorough revision, as well as the various conference and seminar participants who have helped me formulate and articulate these ideas.

1 Introduction

This paper aims to provide an alternative explanation for why countries behave differently with respect to the environment and contributions to global pollution, although they might be quite similar from an economic development point of view. To explore this idea, I use a simple micro-founded model in which individuals derive utility from their own well-being as well as from a moral standpoint. The utility of the latter concept comes from the idea that individuals derive satisfaction from doing 'the right thing' (at least to some degree) – or, according to Immanuel Kant, from behaving according to the imperative principle. Being or acting green could fall into the category of such imperative principles. Using these concepts in addition to a simple political framework, I show that two equivalent societies (i.e., societies with the same income, political system, etc.) can reach two different environmental behaviour equilibria. I also locate the means of nudging a society from one equilibrium to another. Although I do not claim that this explanation is the only reason for why countries behave differently, this model provides a very simple rationale for why this could happen.

Different theories have been developed to explain, to some extent, the dissimilarity of green behaviour among similar countries. One set of theories revolves around a country's level of development. For example, the Environmental Kuznets Curve (EKC) relates a country's environmental behaviour to its income level, revealing an inverted U-shaped relationship between the two factors.(1) The drawback of such an approach is that it explains dissimilarities concerning local pollution, but it cannot explain the dissimilar behaviour when it comes to global pollutants (for example, CO_2 emissions per capita) among countries with similar income levels like the USA, Canada, Australia, and their European counterparts Sweden, Norway, and Finland.

(1)The EKC is similar to a traditional Kuznets curve. As a country gets richer and more developed, it begins to pollute more (as measured on a per-capita basis). After reaching a certain developmental level, society begins to designate more importance to the environment and therefore starts to pollute less as it becomes richer. Hence, we observe an inverted U-shape for the relationship between pollution and income per capita. Although the EKC was used to analyse local pollution, it could also be used to compare countries concerning global pollutants.

Another important strand of literature explains green behaviour based on moral motivation. This is due to the simple fact that green actions cannot be explained with a purely homo-oeconomicus theory. If individuals only abided by homo-oeconomicus principles, there would be no incentive for them to contribute to maintaining a good environment, since the gain for this action would be negligible. One starting point in the literature is the idea of 'warm-glow giving' (Andreoni (1990)), in which the sole act of giving provides utility to the agent. This notion has been further developed by Nyborg and Rege (2003) and Nyborg et al. (2006). The former study provides a comprehensive summary of different types of moral motivations (altruism models, social norm models, fairness models, models of commitment, and the cognitive evaluation theory), although it focuses on how these different types of moral motivations can crowd out private contributions. Nyborg et al. (2006) present a model in which societies become polarized due to the effects of peer pressure. In other words, a society can be completely green (when everyone protects the environment) or completely grey (when everyone chooses to pollute). While such an (extreme) outcome has not been observed in reality, it is clear that this social pressure exists.

I propose a different approach as follows: People are diverse and behave, up to some degree, according to something close to a Kantian imperative. That is, individuals try to do 'the right thing', even if it goes against their private (or 'selfish') homo-oeconomicus tendencies. This Kantian idea has been applied to economics in Brekke et al. (2003), Brekke and Nyborg (2008), Nyborg (2011), Wirl (2011), Roemer (2010), and Laffont (1975). The first three papers use the Kantian approach in order to define an ideal. In measuring his behaviour against this ideal, one gains utility from his self-image, which grows as his behaviour approaches the ideal. Similarly, Wirl (2011) talks about green players that find the social optimum following a Kantian approach. Roemer's paper (2010) uses the Kantian idea to create a rule in a Kantian manner. This is a game-theoretic approach which defines a Kantian equilibrium as one that no one wants to deviate from in the same *proportion*. In Laffont's paper (1975), agents take into account what would or should be their behaviour with the premise that everyone will do as they do. He then derives a market equilibrium and shows that if agents behave in a Kantian way, they internalize their externalities and the outcome is optimal.

In my model, I refer to Kantian people in a way similar to Laffont. Here Kantian means that I consider what I should do (the right thing), which might differ from what I would do if I were to act in a purely homo-oeconomicus way. In this sense, the agent has to know or find out what *the right thing* is. This is not exactly what Kant intended in his work (Kant et al. (2002)); he described duties (deontology), not responsibilities. He talked about a *maxim* (universal law) instead of *actions*. However, it may be argued that one succeeds or fails at meeting their responsibilities as a consequence of his or her actions. Now, if agents thought that their contribution to pollution (or its prevention) was negligible, they would not say (in a direct way) that they were responsible for climate change. However, it is widely accepted that we are all responsible. In that sense, to be Kantian (in this paper) means being a morally responsible person. It means making an assumption about the consequences of our actions (hence the word 'responsibility') using a Kantian way of thinking – "assuming everyone behaves as I do". We often hear others make statements like: *"I don't pollute (or I do recycle, etc.) because if everyone did so, the effects would be terrible, and I don't want to contribute to that".* This tells us that the person has assessed the situation, made a decision that the action is morally bad, and is taking the action of refraining from participating in it if at all possible. This practice of responsibility in moral systems may be explained by a variety of primary ethical principles, including the categorical imperative (Kant), utilitarianism, contractualism, cooperation, compassion, etc., as explained in Baumgärtner et al. (2014). In the present paper, I use the Kantian categorical imperative as a primary ethical principle. Supporting this idea, I can mention Bruvoll et al. (2000). They find that 88% of surveyed people in Norway, when asked about their motives for sorting waste, agreed or partly agreed with the following statement: *"I recycle partly because I think I should do what I want others to do"*. This is definitely a Kantian motivation.

In general, the choice to be green is a costly one; if it were not, of course, it would not make sense *not* to do so. Therefore, the agent is faced with a trade-off between doing the right thing and being negatively affected by the direct cost of this action. This cost is the one incurred when we compare the green action with the one that would be made in a purely homo-oeconomicus way. To be able to effectively assess the situation, the agent must know (or estimate) his moral 'obligation' and its cost. He then weighs the options and decides how to behave. As we can see, this is not a general rule (and therefore not precisely Kantian), since a specific person could be swayed toward taking part in green or grey behaviour depending on the cost and the *perceived* moral implication of his choice. For example, the implications of contaminating the environment when it is already quite polluted are quite different from those in a situation in which the environment is clean. It is clear, though, that different people weigh the cost of their green action and their moral responsibility in diverse ways. To account for this, I will assume that society is composed of a continuum of people, ranging from the purely homo-oeconomicus individual to the purely Kantian one.

As we can see, this is an evolving process. Individuals' awareness and moral motivations depend on the perception of environmental quality (or pollution). Changes to the environmental state and/or corrections to one's perceptions can cause awareness to vary. At the same time, public decisions are made by governments, which generally try to apply the best policies. The problem arises from the fact that global pollution is usually a prisoner's dilemma game (in the international arena), and in this case, it is 'optimal' for the country to pollute: It is the dominant strategy of the game. In this sense, governments are not likely to implement green policies (concerning global pollutants). However, a government may be willing to incur the cost of implementing a green policy if it has the support of its constituency, as Sweden did in the late 1990s.[(2)] Consequently, I introduce a simple political mechanism into the model. The government in power will put a green (or greener) policy in place if the people demand it. Therefore, I assume that if individuals

[(2)]Sweden was the first country to introduce a carbon tax in 1991, and they have been increasing it as time goes by. Currently, they are aiming to increase it again, in order to "point out how to achieve the 2050 vision of zero net GHG emissions" in their Climate Roadmap. (Energy Policies of IEA Countries, Sweden, 2013 review).

behave in a green way, their government is also likely to have a green policy in place. This is in line with what was observed in Germany by Comin and Rode (2013), who found that when people behaved in a greener way, green parties received more votes at elections. It is also in line with Leiserowitz (2006), who finds that *"egalitarian-value based"* individuals (similar to those described as embodying the Kantian morale) stand for green policies.

On the perception side, a poorer environmental state, which is expressed by Mother Nature through more frequent and severe climate events, triggers concern. This fact has been comprehensively studied by different surveys (as in Gallup and GlobeScan, among others: Gallup world poll [25], Globescan radar [26] and Extreme Weather and Climate Change in the American Mind April 2013 [31]) and verified using econometric techniques by Krosnick et al. (2006) and Zahran et al. (2006). Lee and Markowitz (2013) performed an overall analysis showing that individuals' environmental awareness and concern rises after major climate events, although typically only in the short term. As an illustrative example, after superstorm Sandy hit New York City President Obama said: "We must do more to combat climate change ... Now, it's true that no single event makes a trend. But the fact is, the 12 hottest years on record have all come in the last 15. Heat waves, droughts, wildfires and floods – all are now more frequent and more intense."(3) Although the president's making a statement does not necessarily guarantee that proper environmental behaviour will materialise immediately, it shows that awareness of climate issues has reached a significant segment of the population, and that these issues have been acknowledged by those in power.

Finally, I build a model that takes into consideration the concepts mentioned so far. By combining these elements – namely, different types of people with regard to their environmental behaviour (their Kantian 'structure'), the degree of environmentally induced awareness, and a simple political system – I find that two equivalent societies can reach two different environmental behaviour equilibria even if they share the same structural characteristics (with respect to income, political system, etc.). I also identify the possi-

(3)President Obama at the State of the Union Address. February 12th, 2013.

bility of switching from one equilibrium to another. In other words, there is a tipping point at which a society that is not behaving in an environmentally-friendly manner can be swayed toward green behaviour. The mechanics are related to the idea that having a more polluted environment makes people more aware, triggering green behaviour. It may be the case that this is not enough to switch the political praxis to a green one, and we could be left with a grey society; however, there could also be a situation in which the awareness levels and behaviour are such that a green(er) government gets elected, starting a process to tip the system.

The logical question that arises is: How can a society be swayed from grey to green? To answer this, I analyse the influence of two factors: individuals' perception of pollution and the existing political system. A shock to the perception of pollution (such that individuals become more aware of or concerned with environmental issues) is an effective mechanism to induce tipping. Moreover, a political framework in which coalitions are more likely to exist eases the shift from a grey to green society and vice-versa. This is mainly due to the fact that a more 'continuous' political spectrum allows a society to shift toward a relatively greener government and, from there, to greener and greener governments in a cascading process. Applied at a government level, this cascading idea is similar to the one developed by Kuran (1991).(4) Continuing with the political angle, the literature has also addressed the determinants of green behaviour by comparing political systems.(5) Unfortunately, these results present neither a clear nor consistent view of how political systems might influence a society's green behaviour. Since the previous idea is

(4)Kuran talks about the collapse of Eastern Europe's communist regimes. He divides the society into 10 types of people, ranging from those who are more in favour of a communist government, to those completely opposed to it. He shows that if some sort of threshold is crossed, protests can begin, which can encourage those initially less likely to go against the incumbent regime to join in protesting. This process can lead to a cascading effect, which can in turn trigger the collapse of the whole regime.

(5)Persson et al. (2000) used a theoretical model to show that presidential regimes should produce an under-provision of public goods (thus leading to a dirtier environment). On the other hand, Bernauer and Koubi (2004) found the opposite result. They use an econometric study to find evidence that presidential democracies provide more public goods than do parliamentarian democracies. More recently, Saha (2007) tested the previous hypotheses empirically. She finds that the electoral system has no effect on any of the environmental public good supply indicators and that the nature of the political regime has no significant impact either.

closely related to this literature, it might explain, at least partially, the results found in this literature.

To finalise the model, and acknowledging the existence of another psychological ingredient, I add to the model *peer effects* or *social approval*. They can act as secondary motivators, as in the case of Nyborg et al. (2006). Incorporating this concept into the model reveals that, in fact, an 'ideological' peer effect[(6)] makes the transition from a grey to green society a more difficult task to achieve. This comes from the notion that if the society is primarily grey, the agent will have to bear his economic cost ***plus*** the new peer pressure cost in order to behave in a green way, thus making the shift harder to accomplish.

The paper is structured as follows: First, Section 2 presents the model and its main features. Section 3 shows possible tipping points and demonstrates how the system can be nudged. Section 4 introduces the concept of social approval as a psychological driver of behaviour. Section 5 concludes and presents a brief discussion of the model.

2 The Model

The people

I assume that people care about their own utility and about social utility, as mentioned in the Introduction. The fact that individuals can have different ***attitudes*** regarding social well-being causes heterogeneity among agents. Therefore, we can write each agent's utility as:

$$U(\cdot) = (1 - \alpha)\, u_p + \alpha\, u_s \tag{2.1}$$

where u_p and u_s are the private and social utility respectively. The parameter α ($0 \leq \alpha \leq 1$) represents how homo-oeconomicus ($\alpha \to 0$) or Kantian ($\alpha \to 1$) the person is, hence his attitude. An attitude can be defined as an inherent trait formed from a combination

(6)This means that green people prefer others to behave in a green way as well, and grey people prefer others to behave in a grey way. Of course, it might be the case that being grey is always considered 'bad', even for grey people, as in the case of smoking.

of cultural background and education. It relates to and influences an agent's moral responsibility. For the purposes of this study, $\alpha = 0$ means that the agent does not care at all for the rest of the society, whereas $\alpha = 1$ means that an agent is the most morally responsible (Kantian). Moreover, the society is composed of a continuum of people, finite in number, each one matching a value of α in a biunivocal correspondence where α has some distribution f_α.[7]

Both u_p and u_s are constructed in the same way. They have a consumption part $u(\cdot)$ and a damage part $d(p_t(\cdot))$:

$$u_i = u(\cdot) - d(p_t(\cdot)) \tag{2.2}$$

The first part, $u(\cdot)$, is the classic consumption utility with $u' > 0$ and $u'' < 0$. The damage term $d(p_t(\cdot))$ also has its classic properties of $d' \geq 0$, $d'' > 0$ and $d'(0) = 0$, where $p_t(\cdot)$ denotes the pollution level at time t. I return to these functions in the following pages.

Goods and Pollution

In this simple framework, each person can either buy green products (x) or grey products (y). From a consumption point of view, the products are perfect substitutes. However, there are two differences. The first is that the green one does not pollute, whereas the grey one does. The other difference is that the green product is more expensive than its grey counterpart.[8] I assign a normalized price of 1 to the grey good and a price of $(1+\rho)$ to the green one. Therefore, the value of ρ represents the extra amount (with respect to the whole original price) to be paid for a green product. Since the grey product pollutes, I also denote with γ the impact on the environment of the consumption (or more accurately, production) of this type of product. For simplicity, the agent will only choose

[7] As will be made clear in the following pages, the distribution of α will not change the main results, but assuming a uniform distribution will certainly ease the subsequent calculations and simulations. The analysis will be performed using a uniform distribution. On the other hand, different distributions will simply change the place of the tipping point and the conditions needed to tip, as in Kuran (1991).

[8] If there were a green good that was cheaper than its grey counterpart, agents would automatically choose that good instead of the grey one for purely economic reasons. If that were the case, we could recalculate the pollution produced by a new representative grey good and return to the set-up presented here.

one or the other, not a mix.[9] The agent's income is also normalized to 1.

Hence, the agent can be a 'grey' consumer, $(x, y) = (0, 1)$, or a 'green' consumer, $(x, y) = (\frac{1}{1+\rho}, 0)$. The pollution equation will be the standard one:

$$p_t = (1 - \delta)p_{t-1} + \gamma \cdot y_t^s \tag{2.3}$$

where p_t is the pollution level at time t, δ is the natural decay of pollution level (due to natural absorption), and y_t^s is the *society's* mean grey consumption at time t, which is just the average of grey consumption by all agents. I will come back to this term in the following pages.

Public Concern

I will call 'public concern' the part of the utility function that pertains to social welfare. Each agent is weighted by the parameter α, which captures the degree to which he or she is concerned about public well-being or the relative weight they assign to being morally responsible. In order to model the public concern and the impacts of this concern on the agent's behaviour, I use the Kantian morale previously described. More precisely, the agent will consider u_s as if everyone else were behaving as he or she is. This does not mean that the agent is actually expecting that everyone will behave exactly as he or she does. However, it allows us to mimic the decision process by modelling the utility of doing the right thing, which is in line with the Kantian idea. In other words, this part of the utility function is modelled as though the agent assumes that everyone is behaving as he or she does in order to make a decision about how to behave. Strictly speaking, this is not what Kant meant with his categorical imperative, as already discussed in the Introduction. His was not an heteronomous ethic. In this sense, the present formulation

[9]This actually does not make any difference, except for the extra complexity. If the agent can choose a mix of green and grey products, all results *hold*. The behaviour functions and the resulting dynamics turn out to be the same. A more extended explanation and mathematical development can be obtained upon request.

is not categorical, autonomous or independent of external influence; on the contrary, it depends on the environment's quality. Rather, my formulation is in line with the one used by Laffont (1975), which borrows the idea of choosing the good (ethical) rule when assuming that everybody behaves as oneself does hence the term 'Kantian morale'.

Regarding pollution, the level considered by the agent is the (estimated) pollution level p_t^e. This, in turn, depends on two factors: the perceived (past) pollution level, p_{t-1}^p, and the assumed emissions. At this point I assume that the agent has perfect information about the past pollution level, $p_{t-1}^p = p_{t-1}$. Recalling the pollution equation (2.3), we get the following relationship:

$$p_t^e(y^s) = (1-\delta)p_{t-1}^p + \gamma \cdot y^s \tag{2.4}$$

Now I return to the private and social utility functions. For u_p, the agent understands that he is atomistic with respect to the society, and hence he knows that his contamination is negligible with respect to the total emissions. This translates to a damage term $d(p_t(y))$ that does not vary with his individual decision y, but depends only on the society's behaviour y^s. Since I will compare the agent's two options, and because the previous term does not vary with the agent's decision, I drop it from the following equations (it cancels out).

On the other hand, u_s is the social utility taken into account by the agent when using a Kantian view. In other words, the agent considers that everyone behaves as he or she does, implying that society's emissions y^s will follow their choice y, as well as x^s with x. Putting all of the pieces together and rewriting the previous equations, we get:

$$u_p = u(x,y) \qquad \text{and} \qquad u_s = u(x,y) - d(p_t^e(y^s = y)) \tag{2.5}$$

Finally, we can plug these results into the agent's utility function 2.1, arriving at:[10]

$$U(y) = (1-\alpha)\cdot[u(x,y)] \;+\; \alpha\cdot[u(x,y) - d(p_t^e(y^s\!=\!y))]$$

$$U(y) = u(x,y) - \alpha\cdot d(p_t^e(y)) \tag{2.6}$$

$$U(y) = u(x,y) - \alpha\cdot d[(1-\delta)p_{t-1} + \gamma\cdot y] \tag{2.7}$$

This formulation is in harmony with the standard representation of green behaviour. Since a purely homo-oeconomicus approach cannot explain this type of behaviour, we must consider a moral motivation, as stated in the Introduction. People behave in a green way because they think it is the right thing to do, not because it is in their best economic interest to do so. But what is the right thing to do in a framework like this one? To tackle this question, I have used the ideas of Immanuel Kant. In his exploration of what was 'good' and 'bad', he devised the idea that a good action was one that could be tested as a maxim rule, one that everyone would follow, known as the categorical imperative (Kant et al. (2002)). If this rule makes the society better off, then it is a good rule to follow, meaning that following it makes us good people. In order to use this idea, for the present formulation, this dictum can be translated into: *Which general rule of action should I follow to maximize social welfare, as I perceive it, given that everyone acts according to the same general rule?*[11][12] We can see now that this approach fits the model quite well. The agent is considering the social well-being u_s in his personal utility function. If he or she thinks that everyone behaves as they do, when estimating the implication of other's actions with respect to social well-being, they are operating within a Kantian vision.

Naturally, this representation could be considered naive in face of reality. Why should each individual expect others to behave as he or she does? Most individuals do not, in

[10] Recall that the pair (x,y) can be either $(0,1)$ or $(\frac{1}{1+\rho},0)$.

[11] The original categorical imperative (or one of the original versions) was: "So act as if the maxim of your action were to become through your will a universal law of nature." Kant et al. (2002).

[12] For a well-written essay on the relationship between the Kantian imperative and climate change, see Rentmeester (2010).

fact, believe this. As noted in Brekke and Nyborg (2008), "...the categorical imperative defines one's moral responsibility vis-a-vis society without referring to others' *actual* behaviour, there is no presumption ... that he thinks others will *in fact* follow his example". Now, it is reasonable to say that everyone is different with respect to this (Kantian) moral responsibility. Some people are indeed more responsible than others. The weighting parameter α accounts for this fact. Larger values of α mean that the agent is being more responsible (Kantian) than homo-oeconomicus. Hence, those who care nothing for public well-being (or at least, do not behave as if they care) have $\alpha \approx 0$. Those who care (are responsible) and behave the best have $\alpha \approx 1$.[13]

Agent's Behaviour

Depending on his attitude toward the environment, concern and behaviour will emerge with different strengths. For example, someone with a stronger green attitude will have higher levels of concern for the environment and hence a greater response toward environmental conservation (or green behaviour). Following the model, the agent can choose to behave in a way that is green $(y = 0)$ or grey $(y = 1)$. In this case:

$$\text{Green (y=0):} \qquad u\left(\tfrac{1}{1+\rho}\right) - \alpha \cdot d[(1-\delta)p_{t-1} + 0]$$

$$\text{Grey (y=1):} \qquad u(1) - \alpha \cdot d[(1-\delta)p_{t-1} + \gamma]$$

The agent will behave in a more green manner if the first term is greater than or equal to the second one. Applying this inequality to the previous equations and rearranging

(13) There is also another way of tackling this diversity: We might think that each person cares about social well-being with the same intensity, but that the parameter α instead reflects how 'naive' (or optimistic) each person is. Although this is *not* the same idea stated here, a development and a proof of its equivalence is given in the Appendix A.

the terms, we get:

$$\alpha \underbrace{\left[d((1-\delta)p_{t-1}+\gamma)-d((1-\delta)p_{t-1})\right]}_{\substack{\Delta d\text{ : social cost of}\\ \text{behaving grey}}} \geq \underbrace{u(1)-u(\tfrac{1}{1+\rho})}_{\substack{\Delta u\text{ : cost of}\\ \text{behaving green}}} \tag{2.8}$$

$$\alpha\ \Delta d \geq \Delta u$$

At this point, two points are worth mentioning:

- $\Delta d(\gamma, \delta, p_{t-1}) = d[(1-\delta)p_{t-1}+\gamma] - d[(1-\delta)p_{t-1}]$ is increasing in p_{t-1} (since $d'' > 0$).
- $\Delta u(\rho) = u(1) - u(\frac{1}{1+\rho})$ is increasing in ρ.

The first observation implies that having higher perceived pollution levels, p^p_{t-1} (which is equal to p_{t-1}), will yield more people adopting green behaviour. This is simply because the condition in 2.8 is met for lower values of α when $\Delta d(\gamma, \delta, p_{t-1})$ increases. Therefore, a higher proportion of society will choose to behave in a green way. The second point corresponds to the obvious fact that the more expensive the green product is (higher values of ρ), the higher the cost (in terms of consumption) that will be borne by agents exhibiting green behaviour.(14)

Therefore we can see that for a given price of the green product and a perceived pollution level, there is a value of α^* that *divides* the society in two: those behaving in a green way ($\alpha \geq \alpha^*$) and those behaving in a grey one ($\alpha < \alpha^*$). Hence we can define a function $\theta(p_{t-1}, \rho)$ that tells us the proportion of people exhibiting grey behaviour for the values

(14) An interesting feature to note is that if we consider consumption levels as proportional to some income level w (i.e. comparing $u(w)$ and $u(\frac{w}{1+\rho})$), the cost of behaving in a green way, Δ_u, could be increasing, decreasing or independent of the income level w, depending on the functional form of $u(\cdot)$. Since in this formulation I explicitly leave aside the income effect, I use for simulations the case where $u(w) = \ln(w)$, which gives us a Δ_u that is independent of w. For details see Appendix B.

of p_{t-1} and ρ, as:[15]

$$\theta(p_{t-1},\rho) = \min\left(\frac{u(1)-u(\frac{1}{1+\rho})}{d[(1-\delta)p_{t-1}+\gamma]-d[(1-\delta)p_{t-1}]}\ ,\ 1\right) = \min\left(\frac{\Delta u(\cdot)}{\Delta d(\cdot)}\ ,\ 1\right) \tag{2.9}$$

I note two important things about this new function $\theta(\cdot)$:

- There is a level of p_{t-1} ($p_{t-1_{min}}$) below which everyone's behaviour is grey. In other words, the environment is clean enough that no one 'cares' about it:

$$\text{Setting } \alpha = 1 \quad \rightarrow \quad d[(1-\delta)p_{t-1_{min}}+\gamma]-d[(1-\delta)p_{t-1_{min}}] = \Delta u$$

- There will be always some people exhibiting grey behaviour:

  We can always find $\alpha < \epsilon$, such that $\alpha\ \Delta d < \Delta u$, for any given $p_{t-1} > 0$ and $\rho > 0$. It is easily verified when using $\alpha = 0$: the agent has no incentive at all to behave in a green way.

We can graph this function with respect to p_{t-1}, for a given value of ρ, as in Figure 1.

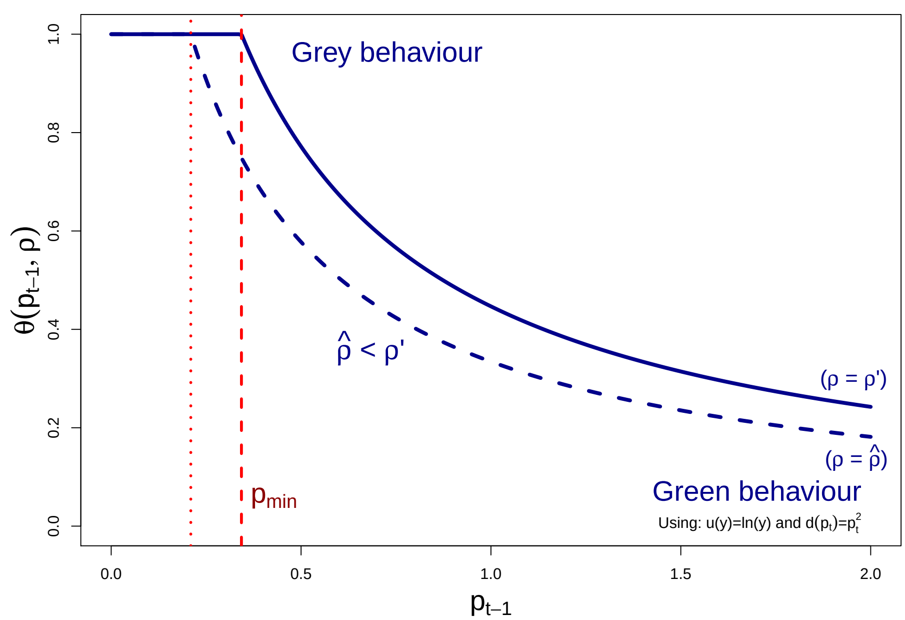


Figure 1: Share of grey behaving people w/r to perceived pollution level.

[15]For the present and following definitions, I use a uniform distribution of α. If this were not the case, we would have a different function $\theta(\cdot)$, but it would still retain the subsequent properties and results.

As we can observe via the dashed line in the illustration, a lower value of ρ (having a cheaper green product) will mean that more people adopt green behaviour. We can also notice that below a threshold level of pollution p_{min}, everyone behaves in a grey way: $\theta(p_{t-1}, \rho) = 1$. The intuition for this is quite straightforward: The environment is too clean to 'care' about, and therefore no one does. Alternatively, since the environment is so clean, it is simply too expensive to behave in a green way. We can also note that as the environment grows worse, the society becomes greener. This last observation comes from the fact that $\Delta d(\gamma, \delta, p_{t-1})$ is an increasing function of p_{t-1}. However, it is also intuitive. As the environment worsens, more people (those with less of a green attitude) become more aware and begin to prefer green behaviour.

It is important to note that the green behaviour is not coming from the agent thinking that his contribution to the environment will make things better (he assumes his contributions are negligible). The previous results comes from the idea that, as the (perceived) environment worsens, the agent's moral motivation increases with them (Δ_d) and therefore more people exhibit green behaviour. In the same fashion, if behaving green gets cheaper (lower values of ρ and hence lower Δ_u), and having the same moral motivation, more people will behave in a green way.

Now we can return to the pollution evolution (Equation 2.3). Rearranging the terms and defining the *change in pollution* as $\Delta p_t = p_t - p_{t-1}$, we have:

$$\Delta p_t = \underbrace{\gamma\ \theta(p_{t-1}, \rho)}_{\substack{\text{Actual} \\ \text{emissions}}} - \underbrace{\delta p_{t-1}}_{\substack{\text{Natural} \\ \text{absorption}}} \tag{2.10}$$

The first term corresponds to present emissions: It is the impact of consumption/production on the environment (γ) multiplied by the share of people behaving grey ($\theta(\cdot)$) multiplied by their grey consumption, which is 1. The second term is the natural absorption of the pollutant.

We graph these two terms in Figure 2.[16] As we can see from the figure, starting from

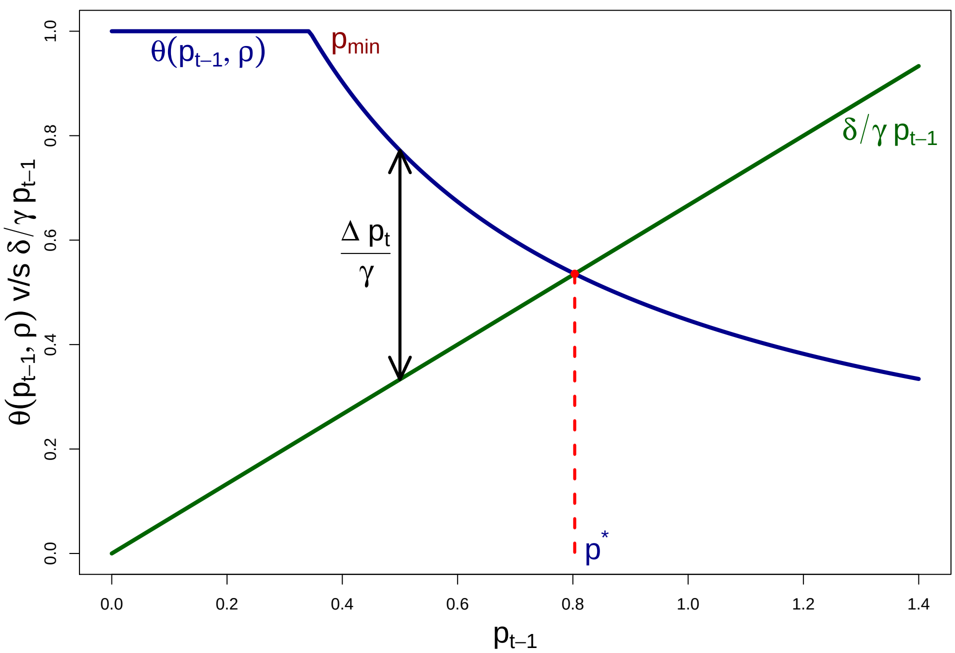


Figure 2: Evolution of pollution.

a low level of pollution, $\Delta p_t > 0$, meaning that we will be polluting faster than what Mother Nature can absorb in the same period of time. At the beginning, for low levels of p_{t-1}, the society will behave in a way that is completely grey, $\theta(\cdot) = 1$, and from a point, p_{min}, we will see more and more people behaving in a way that is green (decreasing section of curve $\theta(\cdot)$). This curve will cross the straight line $\delta/\gamma\ p_{t-1}$ which represents the amount of pollution captured in a natural form. At this point, $\Delta p_t = 0$ means that the system stops evolving. It is easy to see that this equilibrium point is stable, since going further to the right will make $\Delta p_t < 0$.

[16]The figure has been rescaled by a factor of $1/\gamma$ in order to use the same previous Figure 1 and for simplicity in coming sections.

Let us now rotate the graph counter-clockwise, which will make it easier to analyse for further discussion. In doing so, we obtain Figure 3, which notes

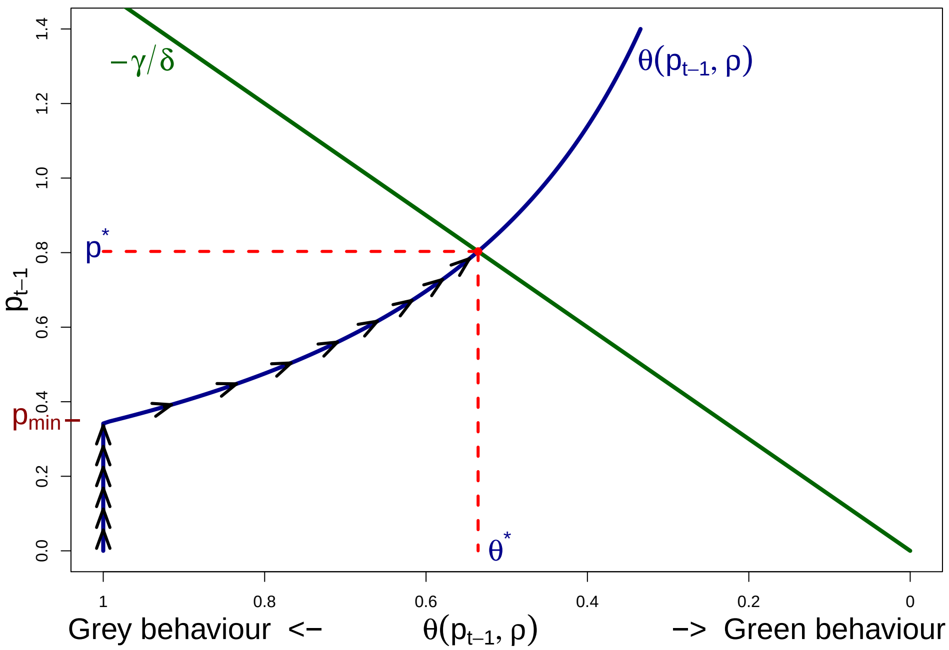


Figure 3: History of pollution.

the pollution level on the y-axis and how green the society is on the x-axis (the greenest being on the right side). The straight line (denoted by $-\gamma/\delta$) is again the natural absorption of pollution. The negative sign and change in terms comes from the rotational process. Since we know that when it crosses the curve $\theta(p_{t-1}, \rho)$ the system is at equilibrium, it is handy to leave it in the graph. In the same fashion, $\theta(p_{t-1}, \rho)$ in Figure 3 is the rotated version of $\theta(p_{t-1}, \rho)$ in Figure 2. Using an 'historical' approach, we begin from the lower-left corner and we then follow the arrows. At this starting point, there is no pollution and therefore everyone exhibits grey behaviour, $\theta(\cdot) = 1$ (recall that the x-axis is inverted). This leads to increasing levels of pollution up to the point where people begin to care about the issue, p_{min} (the kink to the right). As society continues to pollute, it becomes more aware of the issue and becomes greener. Society then reaches a point where emissions are equal to natural absorption, at (θ^*, p^*). At this point, a proportion $(1 - \theta^*)$ of the society is aware enough to exhibit green behaviour, which translates into an emission level equal to what Mother Nature can absorb at that pollution level p^*.

Endogenizing ρ: Introducing a political framework

I now introduce a simple political framework with two parties: the green party and the grey party. They only differ in that each party has a different environmental policy. For simplicity, I assume that the grey party does nothing about the environment, whereas the green party will implement a green-oriented policy. In this set-up, a green policy will simply be some tax/subsidy scheme to reduce the price gap between the green and grey products. This policy could be accomplished by taxing the grey products, subsidizing the green ones, or using both instruments at the same time. In other words, the green government will lower ρ, while the grey government will alter nothing. Other ways of making the green behaviour easier (cheaper) can be introduced by a green government. Recycle points could be closer to people's homes (which is not the case in most countries in the world), causing those who did not recycle previously (because it was too costly) to begin to to so. The idea is to reflect that having a green government will make green behaviour easier. In this simple model, it means lowering the cost of the green product by lowering parameter ρ.

It is implicitly assumed in this simple framework that as $\theta(p_{t-1}, \rho)$ decreases (and the society becomes greener), the government will implement a policy where ρ gets smaller. I base this assumption on an environmentally-ideological basis in the sense that green people will prefer green policies of this type. This idea also is supported by the results found by Comin and Rode in their paper "From green users to green voters" (2013). They show that as more German families started using photovoltaic panels, the Green Party received better results in the elections. But it is also a quite comprehensible assumption. Green people behave in a green way because they think it is the right thing to do (given the current environment) and they will support a government that eases this behaviour.

Following the previous reasoning, we know that the elections will be won by the choice of the median voter. In other words, when the share of green people is bigger than some threshold $(1 - \overline{\theta})$ (and, hence, $\theta \leq \overline{\theta}$), a green government will be elected. In a simple case (for graphical illustration) where α is uniformly distributed, we have that $\overline{\theta} = 1/2$.

On the other hand, we have seen that if the value of ρ decreases, the function $\theta(p_{t-1}, \rho)$ will change, as in Figure 1. Now, since this policy is active only when $\theta(\cdot) \leq \overline{\theta}$ (on the right side of the graph), we have a $\theta(p_{t-1}, \rho)$ function with a discrete jump at $\overline{\theta}$, as in Figure 4.

We notice that we have two equilibria: θ_1^* a 'grey' equilibrium and θ_2^* a 'green' one. In the example depicted, $\theta_1^* < \overline{\theta}$. This gives us the possibility of two equilibria.

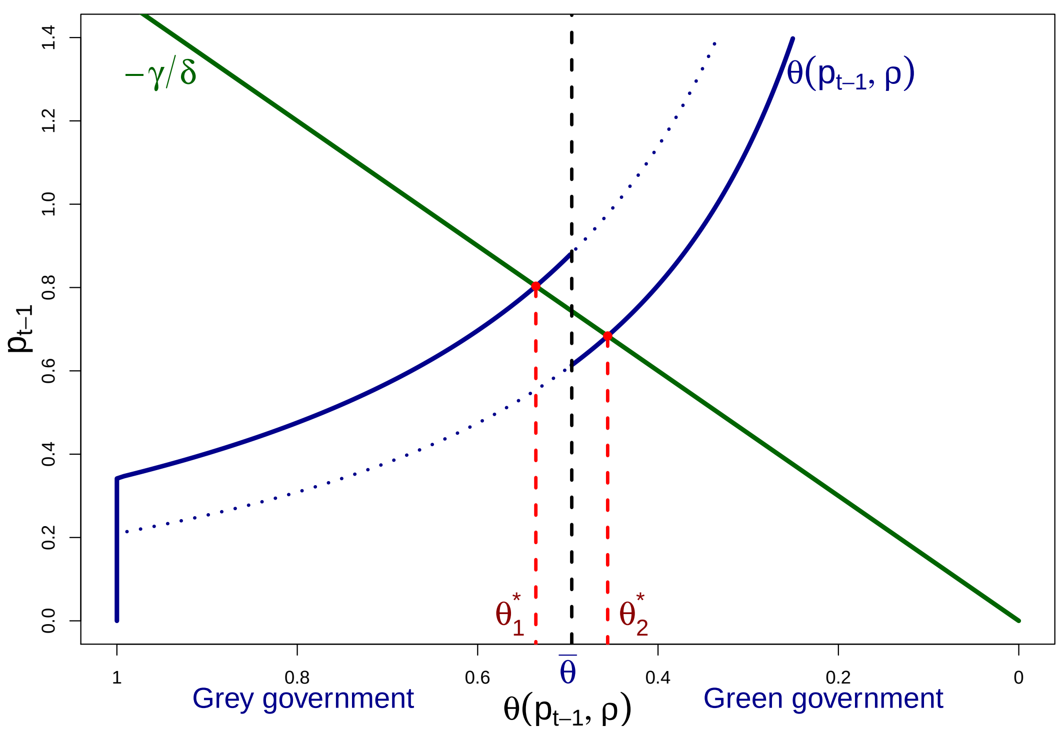


Figure 4: A green government.

It is interesting to note that the multiple equilibria occur *despite* the decrease in pollution levels. Having a better environment gives agents less moral motivation to exhibit green behaviour. But since there is a green government in place, behaving in a green way is also cheaper (the jump if $\theta(p_{t-1}, \rho)$ function), an effect that overrides the decrease in moral motivation. It could also be the case that some grey people (those close to $\overline{\theta}$) would like to have a green government, even if their personal behaviour is grey. This could be due to the fact that these people feel the moral responsibility, but are unable or unwilling to bear the extra financial cost of behaving in a green way. Although this possibility might be true, I focus on the political model because it is simpler and, moreover, it is easy to realize that introducing this into the model would not change the dynamics of the society (we would simply see that everything shifts).

Now let us suppose that we are in a political framework with more parties involved (in the environmental spectrum). This assumption comes from the fact that, in the political arena, other topics are also involved in voters' (and politicians') decisions, such as income distribution, educational policies, etc. If this is the case, we can expect to have more 'jumps' in the dynamics, on the left and right side of $\overline{\theta}$. This could be due to:

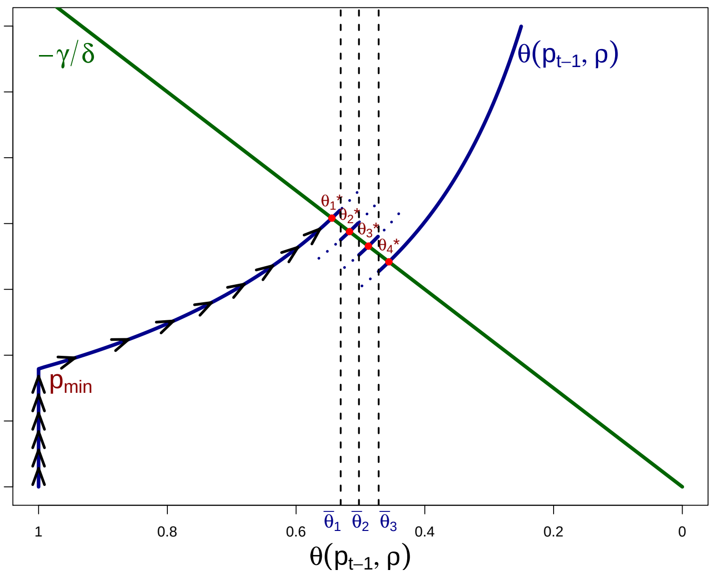
- The fact that coalitions might form in order to attract this 'multi-dimensional' median voter.
- A party or coalition might be willing to implement some green policy (a ρ level between the no policy and the full policy) in order to gain green voters. They are trying to attract voters who also share other political dimension(s) with this party or coalition.

If this is the case, we could have the following set-up:

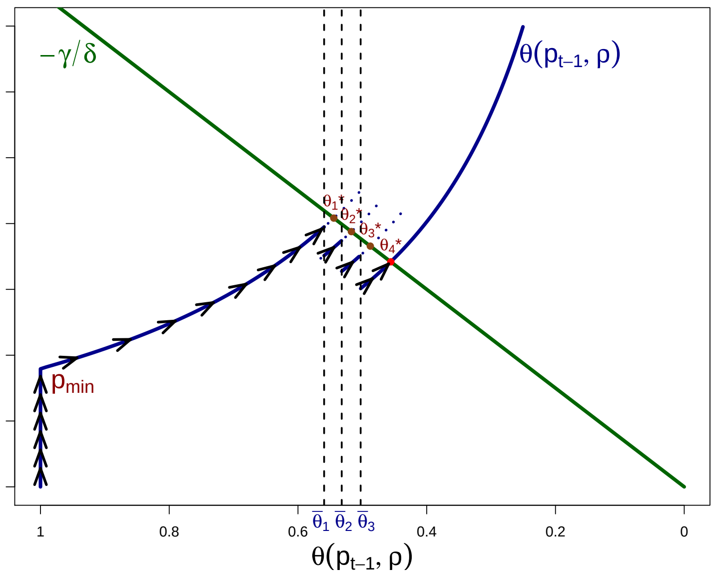
- $\theta > \overline{\theta}_1 \quad \to \rho_1$ (100% grey government)
- $\overline{\theta}_1 \geq \theta > \overline{\theta}_2 \quad \to \rho_2$ (partially grey government)
- $\overline{\theta}_2 \geq \theta > \overline{\theta}_3 \quad \to \rho_3$ (partially green government)
- $\overline{\theta}_3 \geq \theta \quad \to \rho_4$ (100% green government)

with $\overline{\theta}_1 > \overline{\theta}_2 > \overline{\theta}_3$ and $\rho_1 > \rho_2 > \rho_3 > \rho_4$. In this case, we can have different outcomes with $\overline{\theta}_1 \lesseqgtr \theta_1^*$, $\overline{\theta}_2 \lesseqgtr \theta_2^*$ and $\overline{\theta}_3 \lesseqgtr \theta_3^*$.

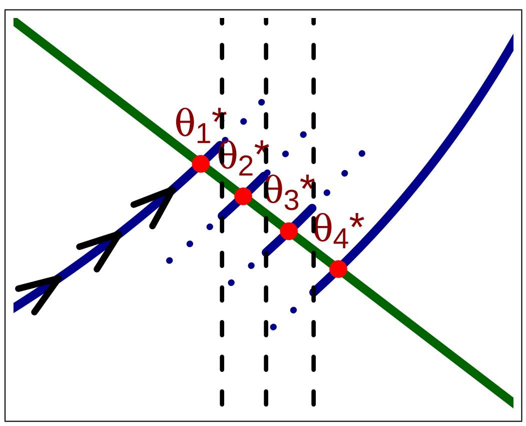
Two cases are depicted in Figures 5a and 5b:



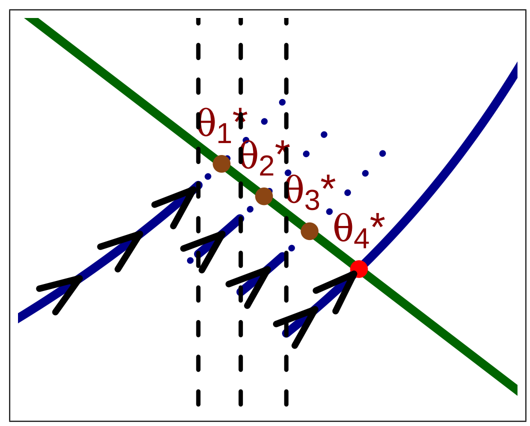
(a) $\theta_1 > \theta_1^*$, $\theta_2 > \theta_2^*$ and $\theta_3 > \theta_3^*$



(b) $\theta_1 < \theta_1^*$, $\theta_2 < \theta_2^*$ and $\theta_3 < \theta_3^*$



(c) Zoom of Figure 5a.



(d) Zoom of Figure 5b.

Figure 5: Different possibilities for multiple equilibria.

Observing the arrows drawn in Figure 5a, we can see how a natural evolution of the environment and society's behaviour could progress. As before, the system starts from the bottom left corner, where there is no pollution and everyone behaves in a grey way. Again, at some point (p_{min}), some people start to exhibit green behaviour and the system begins to shift to the right (always increasing). In this first case (Figures 5a and 5c), the system will arrive at equilibrium θ_1^*. However, if the conditions for Figure 5b (and 5d) are met ($\overline{\theta}_1 < \theta_1^*$, $\overline{\theta}_2 < \theta_2^*$ and $\overline{\theta}_3 < \theta_3^*$), the story becomes a different one. When moving toward θ_1^*, the system crosses the $\overline{\theta}_1$ threshold, jumping to a lower level of ρ (ρ_2) and not reaching the equilibrium θ_1^*. But it now moves toward θ_2^*. This process is repeated with $\overline{\theta}_2$ and $\overline{\theta}_3$ to finally arrive at equilibrium point θ_4^*. In other words, as the society becomes greener in its behaviour, greener governments get elected. If the conditions are 'right', this might produce a cascading process that concludes with the (fully) green government in power.

However, the political framework might not be the only source of change in ρ. Another way of endogenizing ρ is from an evolution in the production process of the green products. As society becomes greener and more green products are bought, it would be logical to expect that: a) economies of scale begin to appear and; b) technological innovations occur in green product production methods (in contrast to a more mature grey production technology). Again, this last point could be 'artificially' induced by a political agreement in response to the society's demands. In the end, either method of endogenizing ρ leads to the same result: The system has a tipping point from which it moves toward a greener equilibrium.

3 A Green Nudge

As we saw in Figures 5a and 5b, all we need to nudge the system is to shift the threshold levels $\overline{\theta}_i$ or to move the curve $\theta(\cdot)$, even temporally. In this previous example, the two societies were not precisely the same, but I choose this framework in order to exemplify how a 'jump' from a grey equilibrium θ_1^* into a greener one θ_4^* might occur. To do so, we should recall that function $\theta(\cdot)$ depends on the price (differential) of the green product ρ as well as on the perceived pollution. This last term, as its name indicates, has to do with how people perceive the pollution levels, which I initially assumed to be correctly perceived ($p_{t-1}^p = p_{t-1}$ used in equation 2.4). Now we can assume that the agent has a perception bias Ω, giving us the following relationship: $p_{t-1}^p = \Omega\, p_{t-1}$. Changing the value of Ω will shift the curve $\theta(p_{t-1}, \rho)$ left or right. Hence, if we are able to change this bias, intentionally or by chance, then we could push the system from the grey equilibrium (or path) into the green one. As we expect, this variation does not need to be permanent. When the system has crossed the last threshold ($\bar{\theta}_3$), then there is no longer a need for such a nudge to sustain the new equilibrium.[17]

The change in Ω could be due to different causes. The perception of pollution levels depends on information. A higher level of exposure to information about climate change, which identifies the true environmental conditions, might alter this parameter. Extreme natural events, such as hurricanes hitting more frequently and severely (as with superstorm Sandy), could have the same impact (Lee and Markowitz (2013)). We have also seen how public attitudes dramatically changed after the Deepwater Horizon oil spill and Fukushima Daiichi nuclear accident. These are not exactly the same phenomena modelled here, but the effect of external information on people's awareness and behaviour is the same. In the wake of the Fukushima Daiichi nuclear accident, the fear of nuclear disasters occurring elsewhere grew considerably, to the point where Germany permanently shut down eight of its reactors and pledged to close the rest by 2022. A similar sentiment arose in Italy, where more than 94% of voters opposed the government's plans to resume nuclear power generation in a June 2011 referendum. However, other major global players like United States did not change their policies in reaction to this event – not, at least, purely for that reason.

However, increasing information or changing perceptions are not the only ways of inducing change. Other avenues exist for nudging the system. One such channel could be political 'noise'. Specifically, a country could be close to its tipping point when a greener government is elected due to non-environmental reasons (left vs. right, social reforms, etc.). When this government implements green policy measures, this action could also trigger a cascading path into the green equilibrium. Independent studies by NGOs or the media might also yield the same type of nudge by increasing awareness about the environment and, hopefully, causing constituencies to push for changes to public policy.

[17] We can see from this reasoning that it is not necessary to know exactly where the thresholds are, but only to be aware of the ability to nudge the system toward a greener equilibrium given these thresholds.

In the same vein, it is worth noting that a multi-threshold situation meaning one in which there is more than one $\overline{\theta}_i$, eases the switching process. If we have a similar situation with more thresholds levels, centred in $1/2$, for example (meaning half on the $\theta(\cdot) < 1/2$ side and the other half on the $\theta(\cdot) > 1/2$ side), it follows that the cascading process begins at a lower level of greenness of the society (starting at higher levels of $\theta(\cdot)$). Also, it is easy to realise that the jump needed to switch regimes is bigger in the binary case compared to the more continuous one. In this sense, having a political arena where coalitions are more likely to form, could facilitate this switching process, increasing its likelihood of happening.

4 Social Approval and Social Pressure

So far I have used an *absolute* moral gain, meaning that the moral or green motivation of the agents comes only from an inner motivation. The agent acts in a given manner because they believe it is the right thing to do (the Kantian idea referred to above). I introduce now what I will call a *'relative'* moral gain. In this case, the agent also derives utility from being accepted by his peers and society in general. This idea has already been discussed in Hollander (1990), Nyborg et al. (2006) and Rege (2004). The concept is quite simple: I get positive feedback from people behaving as I am behaving.[18] Therefore I call $u_a(\cdot)$ the satisfaction from *social approval* that agents get, which will in turn depend on his behaviour and on society's behaviour $\theta(\cdot)$. A corresponding weighting parameter β is introduced, producing the following version of the agent's utility function:[19]

$$U(\cdot) = u_p + \alpha\, u_s + \beta\, u_a \tag{4.1}$$

[18] It is interesting to note that this peer effect is also influenced by how 'public' our green actions are. In other words, we can declare that we behave in a quite green way, when maybe in reality we do not. A clear example of this can be found in Byrnes et al. (1999), which uses a real-life example from an electric utility green pricing program. Despite this fact, it is undoubtable that social pressure exists, especially considering the important and ever-present role of social networks in society and self-representation.

[19] It could be assumed that the agent interacts more with people behaving as he is doing, and therefore only has some probability of meeting with a random person, as in Bisin and Verdier (2000). It turns out that the resulting dynamics are the same, with a minor redefinition of the parameter β.

Concerning the form of u_a, I will assume that the agent gets a social reward from people behaving as he does.[(20)] In the interest of clearer notation, I will denote $\theta(\cdot)$ with θ_t, which is the share of grey people in the society at time t. The function $v(\cdot)$ will transform the share of people behaving as the agent behaves into social pressure, which is a strictly increasing function with $v(0) = 0$ and $v(1) = 1$ for normalization. In other words, this function translates the social behaviour θ_t into social pressure. Hence we have a piece-wise function:

$$u_a(y, \theta_t) = \begin{cases} v(1-\theta_t) & \text{if } y = 0 \qquad \text{(behaving green)} \\ v(\theta_t) & \text{if } y = 1 \qquad \text{(behaving grey)} \end{cases} \tag{4.2}$$

We can immediately notice that if the society is mainly exhibiting grey behaviour ($\theta_t > 1/2$), the agent will *tend* to act in a grey way, and vice-versa. Recalling condition 2.8 for green behaviour and updating it to this new set-up, we have:

$$\alpha\ \Delta d(\gamma, \delta, p_{t-1}) \ + \ \underbrace{\beta\,[v(1-\theta_t) - v(\theta_t)]}_{\substack{\Delta a\,:\,\text{difference} \\ \text{in social approval}}} \ \geq \ \Delta_u(\rho) \tag{4.3}$$

In other words, when $\Delta_a > 0$, which occurs when $\theta_t < 1/2$, a bigger share of the society will behave in a green manner, and vice-versa.[(21)]

[(20)]I therefore disregard the case of negative social pressure from people *not* behaving as the agent does. It is easy to see, though, that including this second effect would not change the results.

[(21)]Since $v' > 0$, the point when $\Delta_a = 0$ (neutral social approval: $v(1-\theta_t) = v(\theta_t)$) has to be when $\theta_t = 1/2$. Hence, if $\theta_t < 1/2 \rightarrow \Delta_a > 0$ and vice-versa. In addition, this is a recursive way of defining the new function $\theta(\cdot)$. Another way of modelling this feature, would be to say that the agent reacts to θ_{t-1}, the share of grey people in the previous period. This would introduce a difference equation into the system, where lags would play a role too. In order to keep the model simple, I use the case where the agents instantly responds to society's behaviour.

We now observe two things now:

- The minimum pollution where people start behaving green will be higher than the case without social pressure: $p'_{min} > p_{min}$

  If everyone is exhibiting grey behaviour and is affected by this peer effect (or *force*), the agent will have less incentive to behave in a green manner. In other words, the pollution will have to be higher in order for someone to care about it *and* endure this (new) social pressure. To find p'_{min}, we again set $\alpha = 1$ and solve: $\Delta_d(\gamma, \delta, p'_{min}) - \beta = \Delta_u$

- If the peer effect is too strong (β bigger than some threshold $\overline{\beta}$), the society will either behave in a completely green or grey way, switching at some pollution level $\overline{p}$ when $\beta = \overline{\beta}$.

  This follows the same intuition as the result obtained in Nyborg et al. (2006). They get a society with two extreme equilibria, completely green or completely grey, with a third unstable equilibria in between. The intuition is the following: Starting with the case in which everyone behaves in a grey way, when peer pressure is too strong, the pollution will increase and still no one will exhibit green behaviour. But at some (much) higher pollution level $\overline{p}$, some people, even bearing the high peer pressure, will start behaving in a green manner. In doing so, the peer pressure will be reduced and therefore more people will also behave in a green way. There will be a domino effect, with more people becoming green and reversing the peer effect towards being green, and therefore reaching a full green society. We notice that the new function $\theta(\cdot)$ is structurally different than its previous version, meaning that it cannot be worked out from the original set-up. For the proof of the last statement, see Appendix C.

In order to make this idea clearer, I graph the changes obtained when peer pressure is present. Figure 6 is the new version of Figure 1 found on page 15.

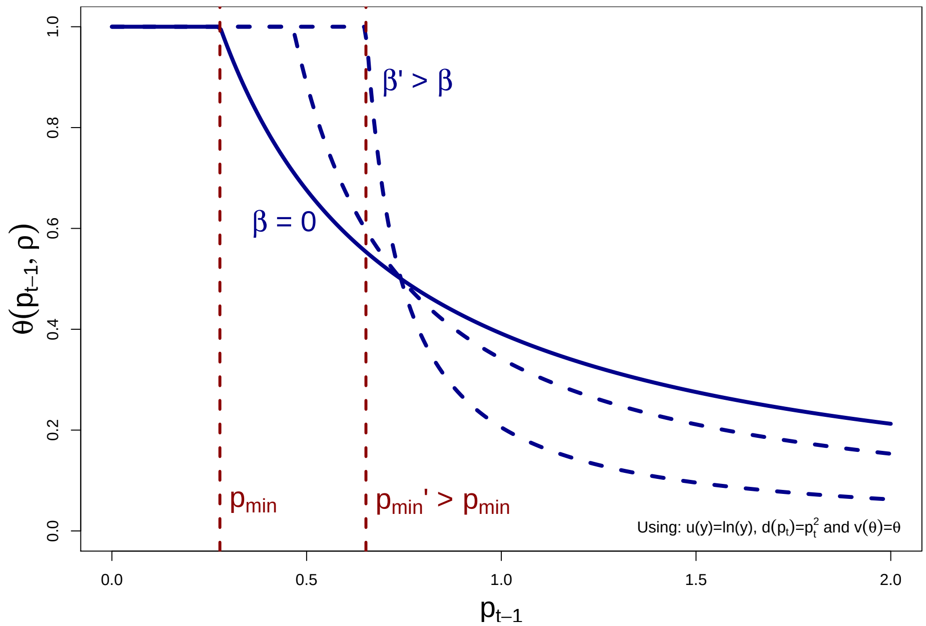


Figure 6: Share of grey behaving people with social approval.

First, we can notice that the three examples cross at a specific point on the graph. This feature is simply the fact that when $\theta_t = 1/2$, the peer effect disappears (it becomes neutral). Therefore, for any $\beta \leq \overline{\beta}$, the different curves should intersect at this point. We can use this fact to find the value of $\overline{p}$ mentioned before: It is just the pollution when $\theta(\overline{p}) = 1/2$, with the 'original' version of $\theta(\cdot)$. When $\beta > \overline{\beta}$, we will have a case with hysteresis. In this circumstance, the pollution level will have to be bigger than $\overline{p}$ in order to switch from grey to green. But once the society has switched to green, in order to switch back to grey the pollution level will have to be smaller than $\overline{p}$, hence the hysteresis effect. It is easy to note at this point, that the original version of $\theta(\cdot)$ and the actual one are structurally different.

Below I rotate the figure again in order to analyse how a society could evolve in the presence of peer pressure. As we can notice in Figure 7 (the updated version of Figure 4, found in page 20), the basic idea is the same. Now, however, the equilibrium points have moved apart from each other.

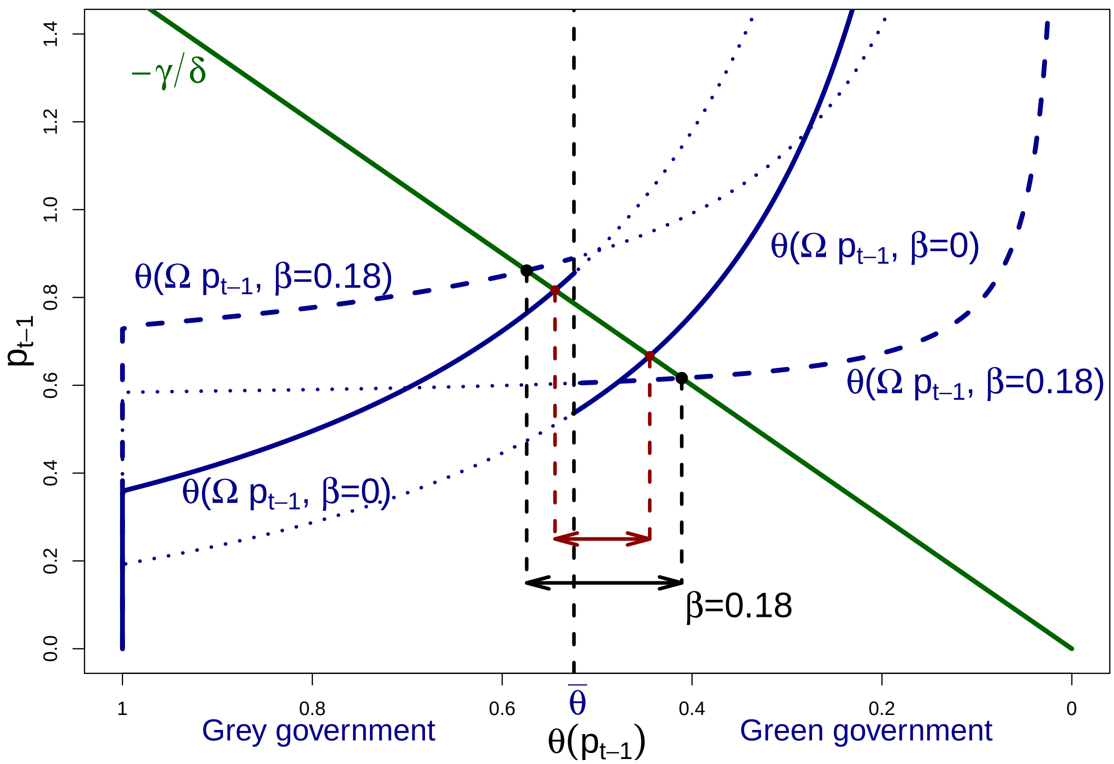


Figure 7: Two equilibria with social approval.

We can understand this effect as coming from some kind of *'attractors'* situated in each extreme ($\theta = 0$ and $\theta = 1$). Since in this particular case each equilibrium point is situated in either 'half' of the portrait ($\theta > 1/2$ and $\theta < 1/2$), they shift to the left and right side, respectively. Therefore, the resulting equilibria are more separated than before: The societies now behave dissimilarly. We can observe this in the two arrows drawn in the figure.

Recalling the cascading effect seen in the preceding pages, we notice now that this peer effect has detrimental consequences: The threshold needed to switch regimes is bigger ($\bar{\theta}$ has to be bigger), and at the grey equilibrium, the pollution is higher. For this reason it would be desirable (if our aim is to switch to the green side) to eliminate this grey peer effect. This idea may sound a bit too optimistic, but fortunately, history may be on our side. Thirty years ago, smoking was considered trendy. We were in the presence of the same effect: Being a smoker was fashionable, and others were encouraged to behave as smokers did. Today, this pro-smoking peer effect has almost completely disappeared or even become negative. Being a smoker is not considered good by anyone, even smokers! It is easy to see that a negative grey peer effect will have a positive outcome regarding the equilibrium and the chances of nudging the system.

Hence, if we can induce this negative grey social pressure (for example, with appropriate advertising), we could again induce this cascading process. Looking at the green side of the system, we might also note that the green peer effect is beneficial for society, at least in pollution terms, and it would drive society toward lower levels of contamination. A second effect also shows up: Being in presence of a green policy (lower levels of ρ) and having a peer effect ($\beta \gg 0$) pushes the system into a switching type. In other words, having a strong green social pressure and having cheap green products encourages a great share of people to behave in a green fashion. We can observe this in the shape of the curve $\theta(p_{t-1}, \Omega, \beta = 0.18)$, which resembles a step function. We can again recall the smoking example: Smoking is now socially frowned upon.

5 Conclusions

This model attempts to explain why similar developed countries take different actions with respect to the environment. To do so, I model society as a composition of different types of people, from those who possess stronger attitudes toward the environment up to those who do not care at all about the environment. Green behaviour derives from being morally responsible. People with a stronger green attitude follow, up to some extent, a (green) Kantian imperative which makes them more prone to exhibit green behaviour. With this set-up, people can either behave in a green fashion (i.e., contribute to the environment) or in a grey way. The model reveals that the same society can arrive at two different equilibria, which might explain differences in environmental policies. This framework might also explain why countries with low income levels sometimes care about the environment more than their developed counterparts, as shown in Lee and Markowitz (2013).

Using the result of multiple equilibria, I have also shown that it is possible to switch from a grey trajectory to a green one. Providing information to people, for example, can raise their awareness about the environment and increase the chances of social change. This outcome reinforces the findings of Corbett and Durfee (2004) and Dunwoody (2007), who show that mass media has a large influence over social concerns, which can result in changes to a country's environmental behaviour and, eventually, its legal framework.

Is the nudge idea a concept that could be applied to explain different behaviours between Europe and the United States with respect to environmental issues such as carbon dioxide emissions? It might be the case that Europe's history of acid rain changed European perceptions and awareness about transnational pollution. When European countries faced acid rain caused by their neighbours' emissions, they became quite aware of the issue and established international environmental agreements, although only regarding these countries. This was not the case in North America, since acid rain was primarily resolved via nationwide actions of the US government. But this could be a bifurcation from an European grey path and could actually explain, up to some extent, the difference in behaviour between North America and Europe.

The present model does not try to provide a full explanation of this phenomenon, but rather to give some alternative insights of what could be occurring. The whole model assumes that each person's green attitude is fixed. In other words, it is as if people were born with this trait and there is no chance of changing it over time. Obviously, this is an extreme assumption. However, the idea here is to create a model that can explain the faster changes in green behaviour observed over the last several years.

However, we could introduce a second slower evolutionary process of green attitude α (rather than using changes in behaviour directly). This process could be understood by considering how education, for example, changes attitudes over time. Such a change would be of the type described and modelled in Bisin and Verdier (2000). New generations approach environmental issues in a different way than their older cohorts, especially when they are taught about the environment (and the complications from environmental damage) from childhood. A possible empirical result of this effect is studied in Hersch and Viscusi (2005). The authors show that younger generations are greener than older generations. Adding this extension to the model would not change the previous results (especially the one concerning the nudge), and it would moreover reinforce the idea of how specific countries treat the environmental issue differently from a 'cultural' point of view. This last point could be a track for possible extensions of the present work.

It could also be interesting to explore the relationship between people's concern and their actions. In some cases it seems that the public is "concerned but unmoved," as stated by Oppenheimer and Todorov (2006). It seems as though this connection could be linked to people's values (Leiserowitz (2006)), which could be developed using a model with an evolution of attitudes.

Another extension might concern the relationship between the change in social awareness and/or behaviour and a more complex political framework model. A further line of work could be to verify this model with some empirical information. Unfortunately, some variables used in this set-up (for example, the concern of people and green attitude) are hard to properly measure. Even so, it would be a worthy venture to pursue.

References

James Andreoni. Impure altruism and donations to public goods: A theory of warm-glow giving? *Economic Journal*, 100(401):464–77, June 1990.

S. Baumgärtner, T. Petersen, and J. Schiller. Bringing norms into action – the concept of responsibility. Working paper, University of Lünenburg, 2014.

Thomas Bernauer and Vally Koubi. On the political determinants of environmental quality. *annual meeting of the American Political Science Association, Hilton Chicago and the Palmer House Hilton, Chicago, Illinois*, 2004.

Alberto Bisin and Thierry Verdier. A model of cultural transmission, voting and political ideology. *European Journal of Political Economy*, 16(1):5–29, March 2000.

Kjell Arne Brekke and Karine Nyborg. Attracting responsible employees: Green production as labor market screening. *Resource and Energy Economics*, 30(4):509–526, December 2008.

Kjell Arne Brekke, Snorre Kverndokk, and Karine Nyborg. An economic model of moral motivation. *Journal of Public Economics*, 87(9-10):1967–1983, September 2003.

Annegrete Bruvoll, Bente Halvorsen, and Karine Nyborg. Household sorting of waste at source. *Economic Survey*, 4(2000):26–35, 2000.

Brian Byrnes, Clive Jones†, and Sandra Goodman‡. Contingent valuation and real economic commitments: Evidence from electric utility green pricing programmes. *Journal of Environmental Planning and Management*, 42(2):149–166, 1999.

Diego Comin and Johannes Rode. From green users to green voters. Working Paper 19219, National Bureau of Economic Research, July 2013.

Julia B Corbett and Jessica L Durfee. Testing public (un) certainty of science media representations of global warming. *Science Communication*, 26(2):129–151, 2004.

Sharon Dunwoody. The challenge of trying to make a difference using media messages. *Creating a climate for change*, pages 89–104, 2007.

Joni Hersch and W. Kip Viscusi. The generational divide in support for environmental policies: European evidence. NBER Working Papers 11859, National Bureau of Economic Research, Inc, December 2005.

Heinz Hollander. A social exchange approach to voluntary cooperation. *American Economic Review*, 80(5):1157–67, December 1990.

I. Kant, A.W. Wood, and J.J.B. Schneewind. *Groundwork for the Metaphysics of Morals*. Rethinking the Western Tradition. Yale University Press, 2002. ISBN 9780300094879.

Jon A Krosnick, Allyson L Holbrook, Laura Lowe, and Penny S Visser. The origins and consequences of democratic citizens' policy agendas: A study of popular concern about global warming. *Climatic change*, 77(1-2):7–43, 2006.

Timur Kuran. Now out of never: The element of surprise in the east european revolution of 1989. *World Politics*, 44:7–48, 10 1991. ISSN 1086-3338.

Jean-Jacques Laffont. Macroeconomic constraints, economic efficiency and ethics: An introduction to kantian economics. *Economica*, 42(168):430–37, November 1975.

Tien Ming Lee and Ezra Markowitz. Disparity in the predictors of public climate change awareness and risk perception worldwide. 2013.

Anthony Leiserowitz. Climate change risk perception and policy preferences: the role of affect, imagery, and values. *Climatic change*, 77(1-2):45–72, 2006.

Karine Nyborg. I don't want to hear about it: Rational ignorance among duty-oriented consumers. *Journal of Economic Behavior & Organization*, 79(3):263–274, August 2011.

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

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

M. Oppenheimer and A. Todorov. Global warming: The psychology of long term risk. *Climatic Change*, 77(1-2):1–6, 2006. ISSN 0165-0009.

Torsten Persson, Gérard Roland, and Guido Tabellini. Comparative politics and public finance. *Journal of Political Economy*, 108(6):pp. 1121–1161, 2000. ISSN 00223808.

Gallup World Poll. `http://www.gallup.com/strategicconsulting/en-us/worldpoll.aspx`.

GlobeScan Radar. `http://www.globescan.com/expertise/trends/globescan-radar.html`.

Mari Rege. Social norms and private provision of public goods. *Journal of Public Economic Theory*, 6(1):65–77, 2004. ISSN 1467-9779.

Casey Rentmeester. A kantian look at climate change. *Essays in Philosophy*, 11(1):76–86, 2010.

John E Roemer. Kantian equilibrium. *The Scandinavian Journal of Economics*, 112(1):1–24, 2010.

Sarani Saha. Democratic institutions and provision of public good. University of California at Santa Barbara, Economics Working Paper Series qt55f3c17g, Department of Economics, UC Santa Barbara, April 2007.

Extreme Weather and Climate Change in the American Mind April 2013. `http://environment.yale.edu/climate-communication/article/extreme-weather-public-opinion-April-2013`.

Franz Wirl. Global warming with green and brown consumers. *Scandinavian Journal of Economics*, 113(4):866–884, December 2011.

Sammy Zahran, Samuel D Brody, Himanshu Grover, and Arnold Vedlitz. Climate change vulnerability and policy support. *Society and Natural Resources*, 19(9):771–789, 2006.

A 'Kantian' index and 'Naiveté' index equivalence.

As stated in the main section, there is also another way of tackling the diversity among agents. We might suggest that each person cares about social well-being with the same intensity, but that the parameter α instead reflects a person's naiveté or optimism. From this idea, we can use α as the share of individuals within society that each person thinks (or hopes) will behave as he does. I show here that this approach is equivalent to the approach outlined in the main section of the paper. In the main text, the original utility function is:

$$U(y) = u(y) - \alpha \cdot d[(1-\delta)p_{t-1} + \gamma \cdot y] \tag{A.1}$$

Now we might think of α as a 'naiveté' measure instead of a 'Kantian attitude' measure. This means that α will now reflect which proportion of the society the agent is expecting (or hoping) to behave as he does. In this new case, we get the following formulation:

$$U_2(y) = u(y) - d[(1-\delta)p_{t-1} + \alpha \cdot \gamma \cdot y] \tag{A.2}$$

Following again the same reasoning of the main section, I posit that the agent will behave in a green manner if $U_2(0) \geq U_2(1)$. We proceed again in the same fashion by rearranging terms and getting, for both cases (original and new), the following conditions for green behaviour:

$$\underbrace{\alpha\left[d((1-\delta)p_{t-1} + \gamma) - d((1-\delta)p_{t-1})\right]}_{\substack{\Delta d(\alpha,\gamma,\delta,p_{t-1}) \text{ : social cost of} \\ \text{behaving grey}}} \geq \underbrace{u(1) - u(\tfrac{1}{1+\rho})}_{\substack{\Delta u(\rho) \text{ : cost of} \\ \text{behaving green}}} \tag{A.3}$$

$$\underbrace{\left[d((1-\delta)p_{t-1} + \alpha \cdot \gamma) - d((1-\delta)p_{t-1})\right]}_{\substack{\Delta_2 d(\alpha,\gamma,\delta,p_{t-1}) \text{ : new version of} \\ \text{social cost}}} \geq \underbrace{u(1) - u(\tfrac{1}{1+\rho})}_{\substack{\Delta u(\rho) \text{ : same as} \\ \text{before}}} \tag{A.4}$$

We can again define an α^* that divides the society into those whose behaviour is green and grey. The only thing to do now is to check if this new function $\theta_2(\cdot)$ has the same properties as the original one $\theta(\cdot)$. Since the right hand sides of the inequalities are the same, I will only focus on the left hand sides. In the original version, we had the following properties:

$$\frac{\partial \Delta d(\alpha,\gamma,\delta,p_{t-1})}{\partial p_{t-1}} > 0 \qquad \frac{\partial \Delta d(\alpha,\gamma,\delta,p_{t-1})}{\partial \alpha} > 0 \qquad \frac{\partial^2 \Delta d(\alpha,\gamma,\delta,p_{t-1})}{\partial \alpha \;\partial p_{t-1}} > 0 \tag{A.5}$$

It is easy to verify that the same properties will hold for the case of $\Delta_2 d(\alpha,\gamma,\delta,p_{t-1})$. Hence we arrive at a new $\theta_2(\cdot)$ with the same properties of $\theta(\cdot)$ (although **not** the same function).

We can finally verify the two remarks made about $\theta(\cdot)$ for $\theta_2(\cdot)$:

- There is a level of $\Omega\, p_{t-1}$ $(\Omega(p_{t-1})_{min})$ below which everyone exhibits grey behaviour (the environment is clean enough such that no one 'cares' about it):

  $$\text{Setting } \alpha = 1 \quad \rightarrow \quad d(\Omega(p_{t-1})_{min}) + \gamma) - d(\Omega(p_{t-1})_{min})) = \Delta u$$

  This will actually gives us the **same** level of $\Omega(p_{t-1})_{min}$ as before.

- There will be always some people who exhibit grey behaviour:

  We can again find $\alpha < \epsilon$, such that $\Delta_2 d(\alpha, \Omega\, p_{t-1}) < \Delta u$, for any given $\Omega p_{t-1} > 0$ and $\rho > 0$. In the same manner, we can see that since $\Delta_2 d$ is continuous in α and that $\Delta_2 d(\alpha = 0, \Omega\, p_{t-1}) = 0$, then there exists an ϵ such that $\Delta_2 d(\epsilon, \Omega\, p_{t-1}) < \Delta u$ for given $\Omega p_{t-1} > 0$ and $\rho > 0$.

We can finally observe a graph showing both versions of the function $\theta(\cdot)$:

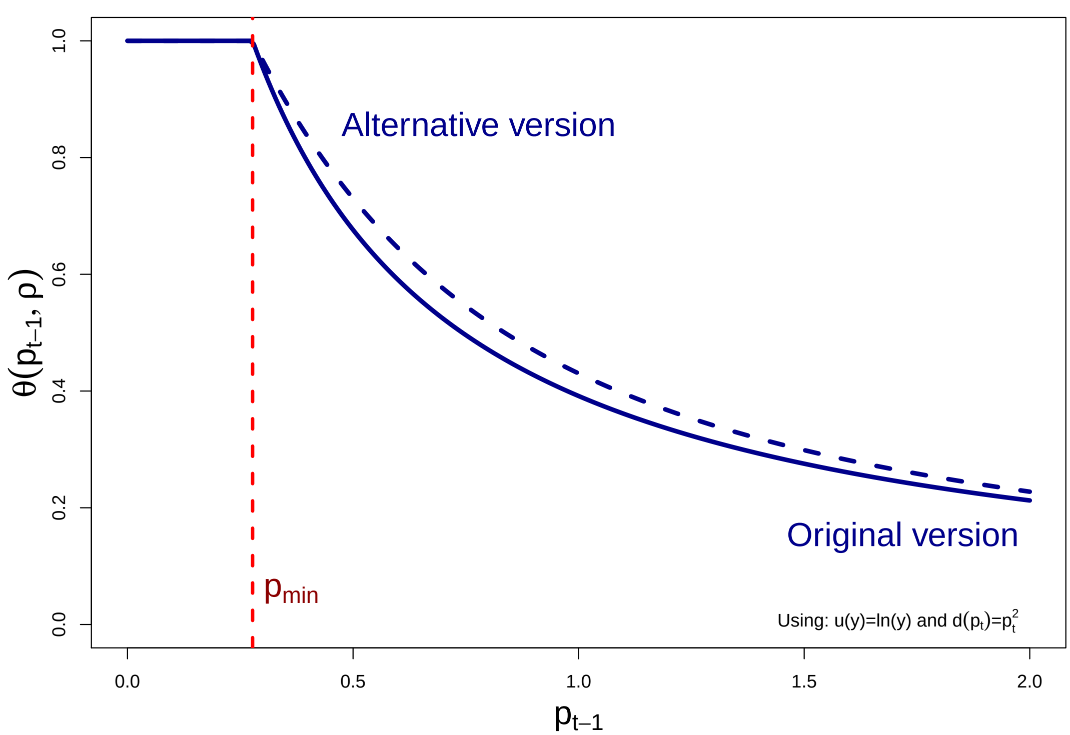


Figure 8: Alternative case where α is the agent's 'naiveness'.

A third and final option might be to suggest that both mechanisms (how Kantian each person is and how naive they are) are in place. For simplicity, I skip this alternative. However, if an individual's Kantian tendency is positively correlated to their optimism, the present results should hold.

B Cost of behaving green under other consumption utility functions.

Depending on the functional form of the consumption utility function, the cost of behaving in a green fashion Δ_u can be an increasing, decreasing, or constant function of the income level w. To observe this feature, I provide four examples using four different consumption utility functions.

$$u_1(c) = \ln(c) \tag{B.1}$$

$$u_2(c) = \sqrt{c} \tag{B.2}$$

$$u_3(c) = \frac{c-1}{c} \tag{B.3}$$

$$u_4(c) = \ln(c+K) \tag{B.4}$$

with K a positive constant.

Setting the grey and green consumption levels to w and $w/(1+\rho)$ respectively, we get:

$$\begin{aligned} \Delta_u^1 &= \ln(w) - \ln(w/(1+\rho)) \\ &= \ln\left(\frac{w}{w/(1+\rho)}\right) \\ \Delta_u^1 &= \ln(1+\rho) \end{aligned}$$

$$\begin{aligned} \Delta_u^3 &= \left(\frac{w-1}{w}\right) - \left(\frac{w/(1+\rho)-1}{w/(1+\rho)}\right) \\ \Delta_u^3 &= \frac{\rho}{w} \end{aligned}$$

$$\begin{aligned} \Delta_u^2 &= \sqrt{w} - \sqrt{w/(1+\rho)} \\ &= \sqrt{w}\left(1 - (\sqrt{1+\rho})^{-1}\right) \\ \Delta_u^2 &= \sqrt{w}\left(\frac{\sqrt{1+\rho}-1}{\sqrt{1+\rho}}\right) \end{aligned}$$

$$\begin{aligned} \Delta_u^4 &= \ln(w+K) - \ln\left(\frac{w}{1+\rho} + K\right) \\ \Delta_u^4 &= \ln\left(1 + \frac{\rho}{1 + \frac{K(1+\rho)}{w}}\right) \end{aligned}$$

We can observe that Δ_u^1 is constant, meaning it is independent of the value of w. In the second case, Δ_u^2 is increasing proportional to $\sqrt{w}$ (the term in parentheses is positive). In the third case, it is clear that Δ_u^3 is decreasing with w. The last case is a mixture between the first and second cases: Δ_u^4 is increasing with w but, as w grows, Δ_u^4 tends to a fixed value, $\ln(1+\rho)$, the same obtained in the first case.

C An (indeed) different dynamics with social approval.

As stated in section 4, adding a social approval gain to the agent's utility function does actually change (structurally) the function $\theta(\cdot)$, and therefore the dynamics. Adding social approval to the utility function will lead to a different functional form in the sense that the new $\theta(\cdot)$ function cannot be found by modifying the original parameters and/or by changing the (shape of the) consumption part $u(\cdot)$.

In order to prove this, we recall that the original function $\theta(\cdot)$ comes from Inequality 2.8 (page 14). There, I proved that for a given value of ρ and a pollution level p_{t-1}, there will always be some people exhibiting grey behaviour. I will prove now that this is not the case with social pressure, since it can be the case where for given values of ρ and p_{t-1} (and β), the society can become completely green. Recall the new condition for green behaviour, as in Inequality 4.3 (page 26):

$$\alpha\ \Delta_d(\gamma, \delta, p_{t-1}) \ +\ \underbrace{\beta\,[v(1-\theta_t) - v(\theta_t)]}_{\substack{\Delta a\,:\ \text{difference} \\ \text{in social approval}}} \ \geq\ \Delta_u(\rho) \tag{C.1}$$

Let us verify whether this hypothesis of everyone behaving in a green manner ($\theta_t = 0$) is true. In this case, the social approval will be equal to: $\Delta_a(\theta_t = 0) = v(1) - v(0) = 1$. Therefore, the previous condition becomes:

$$\alpha\ \Delta_d(\gamma, \delta, p_{t-1}) \ +\ \beta \ \geq\ \Delta_u(\rho)$$

Now, to verify that everyone is behaving in a green way, we can simply verify this condition for the least green person, the one purely motivated by economics, who has $\alpha = 0$. Therefore we get:

$$\beta \ \geq\ \Delta_u(\rho) \tag{C.2}$$

which is just the minimum weight of the social approval parameter in the agents' utility function. In other words, with this level of influence of peer pressure in agents' utility, even the pure homo-oeconomicus person bears the cost of green behaviour, only because of a social pressure source, and not because of a Kantian incentive, since he does not have one.