The tragedy of the commons and socialization: Theory and policy

Emeline Bezin, Grégory Ponthière

The Tragedy of the Commons and Socialization: Theory and Policy*

Emeline BEZIN† Gregory PONTHIERE‡

July 16, 2019

Abstract

We revisit the Tragedy of the Commons in a dynamic overlapping generations economy populated of shepherds who decide how many sheep they let graze on a common parcel of land, while relying on different forms of rationality (Nash players and Kantian players). We examine the dynamics of moral behaviors and land congestion when the prevalence of different types evolves over time following a vertical/oblique socialization process à la Bisin and Verdier (2001). We study the impact of a quota and of a tax on the congestion of land, and we show that introducing a quota may, in some cases, reduce the proportion of Kantians, and worsen the Tragedy of the Commons with respect to the laissez-faire. Ignoring the dynamics of moral traits may lead governments to implement policies that make the Tragedy worse than at the laissez-faire, even though such policies would work well for a fixed population composition.

Keywords: Tragedy of the Commons, Kantian rationality, cultural transmission, overlapping generations, environmental policy.

JEL classification codes: C62, D64, Q24, Z1.

*The authors would like to thank the Editor, Till Requate, and three anonymous referees for their valuable comments and suggestions. We also thank Ingela Alger, Stefan Amblec, Sylvain Chabé-Feret, Jocelyn Donze, Nicolas Gravel, Jean Hindriks, Bertrand Koebel, Mathieu Lefebvre, Fabien Moizeau, Yukihiro Nishimura, François Salanié, Motohiro Sato, Ingmar Schumacher and Anne Stenger, as well as participants of seminars at SURED 2016 (Banyuls sur Mer), Social Choice and Welfare 2016 (Lund), BETA (Strasbourg), TSE (Toulouse), CORE (Louvain-la-Neuve), CREM (Rennes), LAGV 2017 (Aix-en-Provence), Erich-Schneider-Seminar (Kiel University).

†Paris School of Economics and Centre National de la Recherche Scientifique (CNRS). [corresponding author] Email: emeline.bezin@gmail.com

‡University Paris East (ERUDITE), Paris School of Economics and Institut universitaire de France (IUF).

The Tragedy of the Commons and Socialization: Theory and Policy

July 22, 2019

Abstract

We revisit the Tragedy of the Commons in a dynamic overlapping generations economy populated of shepherds who decide how many sheep they let graze on a common parcel of land, while relying on different forms of rationality (Nash players and Kantian players). We examine the dynamics of moral behaviors and land congestion when the prevalence of different types evolves over time following a vertical/oblique socialization process à la Bisin and Verdier (2001). We study the impact of a quota and of a tax on the congestion of land, and we show that introducing a quota may, in some cases, reduce the proportion of Kantians, and worsen the Tragedy of the Commons with respect to the laissez-faire. Ignoring the dynamics of moral traits may lead governments to implement policies that make the Tragedy worse than at the laissez-faire, even though such policies would work well for a fixed population composition.

Keywords: Tragedy of the Commons, Kantian rationality, cultural transmission, overlapping generations, environmental policy.

JEL classification codes: C62, D64, Q24, Z1.

1 Introduction

Popularized by Harding (1968), the Tragedy of the Commons refers to a general class of situations where a common resource is overused by individuals who do not, when making their decisions, internalize the negative impact of their decisions on others' interests. The standard example consists of an economy of shepherds deciding how many sheep they let graze on the same common (unregulated) parcel of land. Each shepherd may not fully internalize the negative effect of raising the number of sheep on the output of other shepherds. As a consequence, the land is congested, the return per sheep is low, and there is an aggregate under-supply of output. This situation is not Pareto efficient.

The overuse of a common resource has been widely tested in experiments through common pool resource games.[1] The theoretical prediction of those games is that each player plays Nash, and extracts as much common resource as he can. However, although outcomes of common pool resource games vary depending on the particular rules of the game, experimental evidence shows that a significant proportion of players extract *less* than what Nash players maximizing their own payoffs would extract.[2]

A possible explanation for that paradoxical result may be that players play Nash, but exhibit inequality aversion, as in Fehr and Schmidt (1999). Another explanation is that the concept of Nash equilibrium is not adequate to describe how individuals behave. As an alternative, Rabin (1993) developed the concept of fairness equilibrium, i.e. a set of strategies and beliefs on other player's intentions such that (i) each player plays the best reply to the other players strategies given his beliefs on other players' intentions; (ii) those beliefs are actually verified. Alternatively, Laffont (1975) and, more recently, Roemer (2010), introduced the concept of Kantian rationality, which consists in choosing not the best strategy for oneself, but the best *generalizable* strategy, i.e. the best strategy if everyone played it as well.[3]

The coexistence of diverse forms of rationality when facing the Tragedy of the Commons raises several questions. First, is the presence of non-Nash rationality a purely temporary phenomenon? Will non-Nash players become extinct in the long-run? A second question concerns the ways in which governments can intervene to make the Tragedy less worse. How can public policies (e.g. quotas and taxes) reduce land congestion once various forms of rationality coexist?

In this paper, we revisit the Tragedy of the Commons in a dynamic economy composed of Nash and Kantian players. We propose to study the joint dynamics of land congestion and moral behaviors, to explore whether standard policies (quotas and taxes) can still achieve their goal of reducing land congestion in an economy whose partition into Nash and Kantian players varies over time, and is affected by those policies.

To this end, we develop a two-period overlapping generations model (OLG) composed of a population of shepherds, who decide how many sheep they let graze on a common parcel of land, while relying on different forms of rationality (Nash players and Kantian players). In line with the literature, the more sheep a shepherd has, the lower is the available land for each sheep, which reduces the output per sheep.

Moral behaviors are transmitted through a cultural transmission process à la Bisin and Verdier (2001) where the evolution of cultural traits is the result of family socialization actions and role modeling within the society. In that setting, parents can influence the probability to transmit their trait (here their rationality, Nash or Kantian) to their offspring, motivated by a form of imperfect altruism (i.e.,'imperfect empathy'). The dynamics of moral traits thus depends on parental socialization efforts which are functions

[1] See Walker and Gardner (1992), Ostrom, Gardner and Walker (1994), Keser and Gardner (1999), Ostrom et al (2002) and Apesteguia (2006).

[2] According to Ostrom et al (2002), players in a basic common pool resource game extract only about 50 % of the common resource, whereas the unique Nash equilibrium would consist in extracting 100 % of the common resource.

[3] Note that Kantian rationality does not assume limited information of the decision-maker; this consists of a different kind of rationality, which aims at selecting the best generalizable actions (under full information).

of the socio-economic conditions.

The key insight is that land congestion and the population composition in terms of moral behaviors are co-determined. In that framework, we first examine the existence and uniqueness of the temporary equilibrium: the equilibrium of the common pool resource game at a given period of time when the population includes both Kantian and Nash players . Then, when studying the dynamics of heterogeneity induced by our socialization framework, we examine long-run equilibria. We obtain three main results. First, we show that, at the laissez-faire, Kantians survive in the long-run for any initial distribution of cultural traits. This result is in line with the evolutionary game theory literature showing that cooperative behaviors can survive in the long-run (Alger and Weibull 2013, 2016a). In our framework, this result is due to the endogenous response of parents (i.e., socialization efforts) who care for their child.[4] Our second result and main contribution is to show that introducing a quota on the number of sheep per shepherd may, in some cases, worsen the Tragedy of the Commons, by reinforcing the congestion of land. Even though quotas reduce land congestion at period t, they may have a negative dynamic effect by lowering moral behaviors. In particular, weak quotas make Nash children better off which encourages Nash parents (resp. Kantian) to increase (resp. decrease) their socialization effort to transmit the Nash (resp. Kantian) trait. This unintended extensive margin effect sometimes outweighs the desired intensive margin effect which results in a rise in the aggregate number of animals in the long-run. We also examine the impact of a Pigouvian tax. Numerically, we find that unlike the quota, the monetary instrument fosters moral behaviors in the long-run. That is, monetary incentives crowd in while direct regulation crowds out moral behaviors. This qualitative change in the impact of the public policy can be rationalized as follows. The quota directly reduces the number of animals which, in the present framework, translates into a rise in the productivity per animal. This rise benefits more to agents who choose a higher number of animals, i.e., Nash players. This in turn strengthens socialization to Nash behaviors. By contrast, the tax indirectly regulates the total number of animals by affecting the cost of purchasing animals. As a consequence, there are two opposite effects affecting socialization to the Nash trait. As for the quota, there is a positive effect due to the rise in average productivity. However, there is an additional negative effect since the rise in the cost per animal negatively impacts individuals who choose more animals, that is Nash players. Overall, numerical simulations suggest that the second negative effect outweighs the first one so that the tax fosters Kantian morality.[5] Thus, in a world with strong political constraints, the tax performs better than the quota.

In the light of those results, this paper reexamines the Tragedy of the Commons by highlighting the interactions between land congestion, public policy and moral behaviors within a dynamic model of socialization and cultural transmission. As such, it casts some light on various real-world situations which share, with our stylized model, some salient characteristics. First, the cultural transmission of Kantian morality is in line with recent field experiments showing persistent differences in pro-social behaviors among different societies. Henrich et al (2001, 2005) realize a large cross-cultural study using ultimatum, public good, and dictator games. Subjects come from 15 small-scale societies (in Africa, Asia and South America) exhibiting a wide variety of economic and social conditions (hunt-gatherers, herders, farmers). They find substantial differences in cooperative behaviors across societies (in particular there is much more variability across groups than within groups). Second, our theory is consistent with empirical evidence suggesting that economic structures with different incentives produce different pro-social behaviors in the long-run. A study conducted by Thalem et al. (2014) reveals that current differences in individualistic values among Han Chinese have been shaped by

[4] Unlike what is found in Bisin and Verdier (2001), in the present framework cultural substitutability (i.e. parents whose type is largely prevailing in the society invest less in socializing their children, since they rely more on oblique socialization) is not necessary for the survival of moral behaviors because there are other substitution forces at work.

[5] Numerically, we also find that the higher the severity of land congestion at equilibrium (as measured by the elasticity of the output to the number of animals) the higher the crowding in effect of the tax.

distinct historical patterns of crop cultivation among Chinese provinces. More precisely, they find that in rice-growing southern provinces, which required intense cooperation (because of a need of elaborated irrigation networks and massive labor requirements), individuals are significantly less individualistic than in the wheat-growing north. Another piece of evidence comes from Henrich et al. (2005), who find that the higher the degree of market integration of one society, the higher the return to cooperation and the greater the expressed level of pro-sociality as measured by the generosity of offers and the refusal of unfair offers. Third, our paper, which examines how intrinsic motivations can conflict with policy intervention, can cast some light on case studies provided by Ostrom (2000). In long-surviving irrigation systems, Bardhan (1999) finds that the quality of maintenance of irrigation canals is lower in systems where irrigation rules have been enforced by a local government. In all of the villages where a government agency decides how water is to be allocated and distributed, farmers have a more negative attitude about the water allocation rules and they tend to contribute less to the local village fund. This is in line with our result showing that in some cases policy crowds out pro-social behaviors in the long-run.

The rest of the paper is organized as follows. Section 2 surveys the related literature. The model is presented in Section 3, which characterizes also the temporary equilibrium (under different possible distributions of moral traits). Section 4 studies the long-run dynamics of moral behaviors. The impact of introducing a quota on the dynamics of moral behaviors and on land congestion is examined in Section 5. Section 6 studies the influence of a Pigouvian tax on land congestion. Section 7 examines the robustness of our results to alternative frameworks. Section 8 concludes.

2 Related literature

The present paper is related to several branches of the literature.

Firstly, our study is related to Curry and Roemer (2012), which focuses on the Tragedy of the Commons in a shepherds' economy, and studies the dynamics of heterogeneity (Nash *versus* Kantian players) by means of an evolutionary game approach. Curry and Roemer show that Kantian behavior can be evolutionary advantageous when players can observe the type of their opponents (and then play Nash when facing Nash players, and play Kantian when facing Kantian players), whereas the opposite holds when they cannot observe the type of their opponents. Other related papers are Sethi and Somanathan (1996) and Noailly et al (2007), who study the evolution of social norms concerning the exploitation of a common renewable pool of resource.[6] Our study differs from those papers: whereas those works adopt an evolutionary game approach based on the replicator dynamics, we study instead how a cultural transmission mechanism involving parental socialization efforts contributes to producing behavioral norms.

Secondly, our paper is also linked to the (more general) environmental economics literature on moral motivations.[7] Daube and Ulph (2016) consider a static economy with two goods (clean and dirty), where individuals, in the absence of morality, do not internalize the impact of their consumption decision on the environment. Daube and Ulph (2016) consider also the impact of morality, some individuals taking into account the moral value of deviating from selfish consumption choices through Kantian calculation.[8] Another related paper is Grafton et al (2017), who study games of climate change

[6]Note that those two papers include not only cooperators and defectors, but, also, enforcers. Noailly et al (2007) can be regarded as a variant of Sethi and Somanathan (1996), with the addition of a spatial structure limiting interactions. Another related paper is Noailly et al (2009), which differs from Noailly et al (2007) regarding the spatial structure of the game (torus versus circle), and also regarding the dynamics of heterogeneity, which is based not on the replicator's dynamics, but on pure imitation.

[7]A survey of environmental economics under other-regarding preferences and ethical attitudes is provided by Long (2016). That survey includes studies not only on Kantian morality, but, also, on inequality-aversion and altruism in environmental issues.

[8]The framework of Daube and Ulph (2016) was recently extended by Cerda Planas (2018), who introduced within that framework a simple political economy mechanism where the government's envi-

mitigation, with agents behaving either as Nash players or as Kantian players. Grafton et al (2017) characterize the Kantian-Nash (temporary) equilibrium, and study the dynamics of heterogeneity by using an evolutionary game approach. They show that, as the proportion of Kantians increases in the population, emissions decrease and the society is better off. Although our paper also studies the Kantian-Nash (temporary) equilibrium, it differs from those studies from a dynamic perspective, since we study the Tragedy of the Commons by relying on a model of socialization and cultural transmission (rather than on an evolutionary game approach). As such, our paper focuses on interactions between parental socialization decisions and the dynamics of land congestion.

Thirdly, our paper is also related to the literature on evolutionary dynamics focusing on the survival of cooperative behaviors (Weibull 1995, Sethi and Somanathan 2001, Cressman 2003, Lehmann et al 2015). Using a general model of evolutionary dynamics, Alger and Weibull (2013, 2016a) show that when there exists some assortativity in the process whereby individuals are matched (i.e., the probability to be matched with a given type is conditional on the individuals' type) some cooperative behaviors survive in the long run.[9] Our socialization framework amounts to consider endogenous degrees of assortativity (i.e., assortativity conditional on parental socialization efforts). In our framework, the optimal parental response to changes in the socioeconomic environment ensures the survival of moral behaviors. Moreover, the existence of endogenous socialization effort is also what lowers, in some cases, the efficacy of public policies.[10]

Fourthly, our paper complements the literature on the economics of culture and socialization (Bisin and Verdier 2001, 2011). In particular, recent studies, such as Bezin (2015, 2019) and Schumacher (2009, 2015), explore the joint dynamics of preferences and environmental variables within a socialization model. This literature applies to global environmental issues such as climate change, whereas we focus here on common-property resource problems. Moreover, while these studies emphasize multiplicative effects of environmental policies, we show that environmental regulation can, under some conditions, worsen the situation with respect to the laissez-faire.

Our paper is also related to the public economics literature showing that public policies (extrinsic incentives) generate unintended outcomes through their impact on intrinsic motivations (pro-social preferences). On the empirical side, there exist both laboratory and field experiments (surveyed in Bowles, 2008) showing that explicit incentives generally crowd out intrinsic motives (being sometimes counterproductive) but, in other cases crowd in pro-social preferences. Bowles (2008) emphasizes two causes of incentive effects on pro-social preferences. First, there is a short-run effect, which he relates to state-dependent preferences: incentives act as a signal affecting the behavioral salience of an individual's social preferences.[11] Second, there is a long-run effect going through endogenous preferences, which captures situations in which incentives affect the process of preference-updating by which individuals acquire new tastes or social norms that will persist over long periods. On the theoretical side, both types of mechanisms have been examined. Brekke et al. (2003) propose a theory of crowding out where preferences are state-dependent. They consider a model of public good provision where individuals want to conform to a morally ideal effort. They show that public policies sometimes crowd out private contributions by reducing that moral ideal. Bowles and Hwang (2008) provide a general model of state-dependent preferences and adopt a normative standpoint by studying the design of optimal incentives for the provision of a public good. They show that, in the presence of crowding out effects, a sophisticated social planner (one who is

ronmental policy reacts to the demand of the population. But that paper takes the distribution of traits as given, unlike our socialization framework.

[9] In particular, they show that the *homo-moralis*, a weighted sum of Nash trait and Kantian trait with weights reflecting the degree of assortativity, is evolutionary stable.

[10] Alger and Weibull (2016b) discuss some policy implications of their model, though they do not consider the response of moral behaviors evolution to those policies (which is responsible for our counterintuitive result).

[11] Bowles distinguishes between three mechanisms: incentives convey information, incentives trigger moral (dis)engagement, and incentives affect self-determination.

aware of incentives effects on motivation) may make either greater or lesser use of standard incentives than the naive planner depending on the nature of the problem. Other theoretical papers examine how incentives affect long-term changes in motivations by altering how agents learn preferences (Bar-Gill and Ferstman, 2005, Rege, 2004). Bar-Gill and Ferstman examine the effect of a subsidy to private contributions to a public good in a model where some individuals care about social status and these preferences are endogenously determined by an evolutionary selection mechanism. They show that the endogeneity of preferences weakens the effectiveness of policies by crowding out pro-social behaviors. In a similar framework where agents want to conform to a social norm (instead of caring for status), Rege (2004) shows that subsidies crowd in voluntary contributions. As the above mentioned papers, our work follows the logic of the "Lucas Critique" in the context of endogenous preferences. We differ from that literature by considering here another mechanism through which policy affects the dynamics of pro-social behaviors: that mechanism arises through the impact of policy on parental socialization decisions.[12] In this setting, we find that regulation through a quota may crowd out, while monetary incentives tend to crowd in intrinsic motivation. Finally, from a policy perspective, our paper is also linked to recent works on Kantian equilibrium and regulation (De Donder and Pestieau 2015; De Donder and Roemer 2016). We complement those papers by introducing socialization, and by examining how the effectiveness of policies is affected by the endogeneity of moral behaviors.

3 The model

We consider a 2-period OLG economy with a fixed piece of land of finite surface. Each cohort is a continuum of agents of size 1. Individuals live two periods. Period 1 is childhood, during which no decision is made. Period 2 is adulthood, during which individuals produce some output (e.g. meat or wool) by means of land and some input (e.g. sheep).

The population is composed of two types of agents $i \in \{K, N\}$, who differ regarding the decision rule they follow when choosing their actions:

- Kantian individuals (denoted type-K): those individuals, when making their decisions, choose the best generalizable strategy (i.e. the strategy that is the best for themselves in the hypothetical case where everyone would choose the same strategy);
- Nash individuals (denoted type-N): those individuals, when making their decisions, choose the best strategy for themselves, conditionally on the strategies of others.

Let us denote by q_t the proportion of Kantians in the mature generation at time t.

3.1 Production

Any individual j (with type $i \in \{K, N\}$) produces an output y_t^j by means of some input (animals) e_t^j and land s_t^j. The total surface of land is equal to $S > 0$, and is of homogeneous quality. For the sake of analytical tractability, it is assumed throughout the paper that the surface of land is constant over time, and is also of constant quality.[13] Land is publicly owned: no producer has any property right on it. The amount of land used by each producer/shepherd is supposed to be strictly proportional to the relative number of animals used by the producer, in proportion to the total number of animals used in the population.

[12] Our paper is also related to Ponthiere (2013), who studied how the public provision of long term care can, by affecting parental incentives to socialize children, affect the prevalence of altruism.

[13] This amounts to suppose that this natural resource is fully regenerated after each period of time. See Section 7.2 on the consequences of relaxing that assumption.

The production function for each producer j is:

$$y_t^j = F\left(e_t^j, s_t^j\right) \tag{1}$$

where $s_t^j = \frac{e_t^j S}{E_t}$ is the amount of land used by a producer j, while $E_t \equiv \int_{j=0}^{1} e_t^j dj$ is the aggregate number of animals. Since there is no property right, the amount of land used by one producer is strictly proportional to the share of the number of animals he uses within the total number of animals. Substituting for s_{jt}, we have:

$$y_t^j = F\left(e_t^j, \frac{e_t^j}{E_t} S\right)$$

Let us normalize S to 1 and assume that $F(\cdot)$ is homogeneous of degree 1. The output of producer j can be re-written as

$$y_t^j = \frac{e_t^j}{E_t} G(E_t), \tag{2}$$

where $G(E_t) \equiv E_t F\left(1, \frac{1}{E_t}\right)$ which is the functional form used by Curry and Roemer (2012). Note that we have $G'(E_t) > 0$ and $G''(E_t) < 0$.

3.2 Preferences

Individual preferences have two components.

On the one hand, individuals derive utility from consumption which depends on the number of animals they choose. This consists of the self-oriented component of individual preferences.

On the other hand, individuals are altruistic toward their offspring and derive some utility from their children's welfare. This is the children-oriented component of preferences.[14] Parents can influence the welfare of their children by exerting a socialization effort.

The two maximization problems, i.e., the choice of the number of animals and the socialization choice can be treated separately in, respectively, Sections 3.3-3.4 and 3.5. This is because we assume that socialization costs enter separately into preferences. It is thus possible to study the choice of the optimal number of animals - and the associated equilibrium - for a given partition of the population, and then to study the choice of socialization efforts for a given equilibrium in terms of number of animals.

The self-oriented component of individual utility is given by:

$$U(e_t^j) \equiv \frac{e_t^j}{E_t} G(E_t) - c e_t^j - v\left(e_t^j\right) \tag{3}$$

where the first term is the benefit per animal (which is proportional to the average productivity), the second term is the cost of buying animals and the third term is the cost of raising animals. We assume, $v^{'} > 0$, $v^{''} > 0$, $v(0) = 0$, $v^{'}(0) = 0$.[15]

The children-oriented component of individual utility is studied in Section 3.5.

[14] We prefer to use the expression children-oriented component of preferences instead of altruistic component of preferences as altruism toward children is not the only form of other-regarding preferences in our framework.

[15] The cost of raising animals stands for a cost of labor (it could alternatively be introduced by assuming that each shepherd disposes of one unit of time which can be used either to rise production or to enjoy leisure time). The introduction of this cost allows us to avoid total dissipation of the rent by Nash players, which occurs with open access and free entry (i.e., the existence of a continuum of agents), see Sethi and Somanthan (1996). This assumption makes our results more comparable to a model with a finite set of agents. Note, however, that without that cost one could show that our main results would hold, i.e., survival of Kantian agents, crowding out of moral motivations by regulation instruments and crowding in by economic incentives (tax).

3.3 The shepherd's problem

3.3.1 Nash behavior

Nash players solve the standard maximization of utility problem, subject to market prices, taking all others' actions as given. The producer j of type N solves

$$\max_{e_t^j} \frac{e_t^j}{E_t} G(E_t) - c e_t^j - v\left(e_t^j\right) \tag{4}$$
$$\text{where } E_t = \int_{j=0}^{1} e_t^j dj$$

Note that since there is a continuum of agents, the marginal cost from the rise in the aggregate number of animals is negligible for a Nash player. This is in line with the atomism assumption discussed in Daube and Ulph (2016).[16]

Denote by $e_t^{j,N}$, the optimal number of animals of the producer j of type N. Since we shall focus on symmetric Nash equilibria, we can write $e_t^{j,N} = e_t^N$. This optimal number of animals is implicitly defined as:

$$\frac{G(E_t)}{E_t} - c - v'(e_t^N) = 0, \tag{5}$$

or is equal to zero if $\frac{G(E_t)}{E_t} - c \leq 0$.

Note that due to the concavity of G, at an interior solution we have:

$$\frac{de_t^N}{dE_t} < 0,$$

that is the optimal number of animals e_t^N is decreasing in the expected number of animals of other producers. This result is due to land congestion: the higher the total number of animals, the lower the productivity per animal and the lower the number of animals chosen by a producer of type N. The optimal number of animals e_t^N also depends on q_t, the fraction of Kantian agents in the population, since it affects the aggregate number of animals.

3.3.2 Kantian behavior

Kantian rationality is about choosing the best generalizable actions, i.e., select the number of animals such that, if chosen by all individuals, this would lead to the largest level of well-being.

The producer j of type K chooses the number of animals that maximizes its own well-being provided all other individuals enjoy the same number of animals. The problem can be written as:

$$\max_{e_t^j} \frac{e_t^j}{E_t} G(E_t) - c e_t^j - v\left(e_t^j\right) \tag{6}$$
$$\text{s.t. } E_t = e_t^j$$

The optimal number of animals $e_t^{j,K}$ is the same for all type-K agents, i.e., $e_t^{j,K} = e_t^K$. This optimal number of animals is implicitly given by

$$G'(e_t^K) - c - v'(e_t^K) = 0. \tag{7}$$

Clearly, the optimal number of animals of a Kantian does not vary with q_t, or with a change in Nash players' optimal numbers of animals.

[16] We discuss this assumption in Section 7.4.

3.4 Temporary equilibria

We now characterize the temporary equilibrium, i.e., the equilibrium of the common pool resource game at a given period of time, the population composition in terms of moral traits being taken as given (socialization choices and the dynamics of moral traits will be studied in Sections 3.5 and 4).

We will first consider the existence of an equilibrium when the population is composed exclusively of Kantian individuals or Nash players. We will then consider the existence of an equilibrium in mixed societies.

3.4.1 Kantian equilibrium

Let us first consider an economy that is composed only of Kantian individuals.

Definition 1 *A (symmetric) Kantian equilibrium ($q = 1$) is a uniform distribution of animal numbers with $e_t^i = e_t^{K*}$ such that:*

$$G'(e_t^{K*}) - c - v'(e_t^{K*}) = 0.$$

Proposition 1 *In an economy with $q = 1$, there exists a unique Kantian equilibrium.*

The proof of Proposition 1 is trivial as Kantian agents choose their optimal number of animals independently from other agents. Note that since Kantian individuals internalize the effects of their decisions on land congestion, they choose the socially optimal number of animals. We thus have that the Kantian equilibrium is Pareto efficient. Kantian behavior solves the Tragedy of the Commons.[17]

3.4.2 Nash equilibrium

Let us now consider the case of an economy composed exclusively of Nash players.

Definition 2 *A symmetric Nash equilibrium ($q = 0$) is a uniform distribution of animal numbers with $e_t^{j,N} = e_t^{N*}$ $\forall j$, such that:*

$$\frac{G\left(e_t^{N*}\right)}{e_t^{N*}} - c - v'(e_t^{N*}) = 0,$$

Proposition 2 *In an economy with $q = 0$, there exists a unique symmetric Nash equilibrium where the optimal effort of each producer is given by $e_t^{N*} \in]0, \widetilde{e_t^N}[$, where $\widetilde{e_t^N}$ is such that $\frac{G\left(\widetilde{e_t^N}\right)}{\widetilde{e_t^N}} - c = 0$.*

Proof. There exists a unique symmetric Nash equilibrium if and only if the equation

$$\frac{G\left(e_t^N\right)}{e_t^N} - c - v'(e_t^N) = 0,$$

admits a unique solution. Define the function $\Phi : \mathbb{R}^+ \to \mathbb{R}$ given by

$$\Phi(e_t^N) \equiv \frac{G\left(e_t^N\right)}{e_t^N} - c - v'(e_t^N).$$

We have

$$\Phi'(e_t^N) = \frac{\left(G'\left(e_t^N\right)e_t^N - G\left(e_t^N\right)\right)}{\left(e_t^N\right)^2} - v''(e_t^N) < 0.$$

[17] That result has been highlighted in the literature, e.g., Curry and Roemer (2012), Roemer (2010, 2015).

Furthermore, we have $\Phi(0) > 0$. At $e_t^N = \tilde{e_t^N}$ where $\tilde{e_t^N}$ is such that $\frac{G(\tilde{e_t^N})}{\tilde{e_t^N}} - c = 0$, we have $\Phi(\tilde{e_t^N}) < 0$. We deduce that there exists a unique $e_t^{N*} \in]0, \tilde{e_t^N}[$ such that $\Phi(e_t^{N*}) = 0$. ■

Due to land congestion, the game exhibits strategic substitutability which in turn ensures the existence and uniqueness of the symmetric Nash equilibrium.

Since individuals do not internalize the impact of their decision on the congestion of land, the Nash equilibrium is not Pareto efficient, and players are trapped in a prisoner's dilemma. This is the problem of the Tragedy of the Commons, as studied in Sethi and Somanathan (1996) and Curry and Roemer (2012).

3.4.3 q equilibrium with Kantians

Let us now characterize symmetric equilibria when some individuals behave like Nash players, whereas other individuals are Kantian.

Definition 3 *A symmetric q equilibrium with Kantian individuals is a distribution of animal numbers where a proportion q_t of producers has $e_t^j = e_t^{K*}$ and a proportion $1 - q_t$ of producers has $e_t^j = e_t^{N*}$ where e_t^{K*} is given as in equation (7) and e_t^{N*} is given by:*

$$\frac{G(E_t)}{E_t} - c - v^{'}(e_t^{N*}) = 0, \quad \textit{with} \quad E_t = q_t e_t^{K*} + (1 - q_t) e_t^{N*}.$$

Proposition 3 examines the existence and uniqueness of a symmetric q equilibrium with Kantians.

Proposition 3 *(i) In an economy with $0 < q < 1$, there exists a unique symmetric q equilibrium with Kantians. The optimal number of animals of Nash players e_t^{N*} belongs to the interval $]0, \tilde{e^N}(q_t)[$, where the function $e^N : [0,1] \to \mathbb{R}^+$, is implicitly given by*

$$\frac{G\left((1 - q_t)e^N(q_t) + q_t e_t^{K*}\right)}{(1 - q_t)e^N(q_t) + q e_t^{K*}} - c = 0.$$

(ii) The number of animals chosen by Nash agents at the q equilibrium is strictly larger than the number of animals chosen by Kantian agents.

Since optimal numbers of animals chosen by Nash players are strategic substitutes and since Kantian agents have no reaction function, the game with a proportion q of Kantian agents exhibits strategic substitutability and thus admits a unique equilibrium. Moreover, Kantian agents act as a social planner by internalizing all externalities and thus choose a strictly lower number of animals.

Corollary 1 *(i) The number of animals chosen by Nash players at the q equilibrium is an increasing function of q_t, the fraction of Kantian agents.*

(ii) The aggregate number of animals is a decreasing function of q_t the fraction of Kantian agents in the population.

Since Kantian agents choose a lower number of animals than Nash players, when the fraction of Kantian agents rises, then the marginal productivity of land is higher so that, at equilibrium, all Nash agents increase their optimal number of animals. Item (ii) reveals that the aggregate number of animals is a decreasing function of q_t. Note that the rise in q_t has two opposite effects on the aggregate number of animals (and thus on land congestion). There is an extensive margin effect, i.e., a higher fraction of Kantian agents reduces the aggregate number of animals since Kantian agents purchase fewer animals. There is an intensive margin effect, as Nash players react to an increase in the fraction of Kantian by increasing their optimal number of animals. The intensive margin effect is always lower than the extensive margin effect. This comparative static results is similar to the one found in Grafton et al. (2017) who consider strategic interactions between Kantian and Nash players in a game of climate change mitigation.

3.5 The parent's problem: socialization

Up to now, we considered only the behavior of shepherds concerning the purchase of animals and the production of output. However, in our economy, shepherds have another activity, which consists in socializing their children. We assume that socialization takes place according to a socialization process *à la* Bisin and Verdier (2001) where socialization inside the family (i.e., vertical socialization) and socialization within the society (i.e., oblique socialization) interact. Children are born without type. Each parent has one child. Parents can invest in a socialization effort $\tau_t^i \in [0,1]$ which incurs a disutility cost given by:

$$\frac{\kappa\left(\tau_t^i\right)^2}{2} \tag{8}$$

The socialization takes place in two stages. First, the child is subject to socialization by his parent. With a probability τ_t^i, a child with a parent of type i will take directly the trait of his parent. With a probability $1-\tau_t^i$, vertical socialization fails, and the child picks up the trait of a role model drawn randomly in the cohort of the parent.

Let us denote by $P_t^{ii'}$ the probability that a child with a parent of type i adopts type i', $i, i' \in \{K, N\}$. The transition probabilities are given by

$$P_t^{KK} = \tau_t^K + \left(1-\tau_t^K\right) q_t = 1 - P_t^{KN} \tag{9}$$

$$P_t^{NK} = \left(1-\tau_t^N\right) q_t = 1 - P_t^{NN}. \tag{10}$$

An important feature of cultural transmission models is that the probability that the parent of type i directly transmits trait i to his/her child is endogenous. Motivating by increasing the welfare of their child, parents can influence that probability. Following Bisin and Verdier (2001), we introduce a friction in parental altruism which is called 'imperfect empathy'. We assume that parents are aware of the different traits that their child can adopt and they correctly anticipate the number of animals that a child with trait $i \in \{K, N\}$ will choose when adult. However, they evaluate this choice through the filter of their own 'subjective moral rule'. Let $V^{ii'}$ be the gain for a parent of type i to have a child of type i', $i, i' \in \{K, N\}$. We have:

$$V^{KK} = G\left(e_t^{K*}\right) - ce_t^{K*} - v\left(e_t^{K*}\right), \tag{11}$$

$$V^{KN} = G\left(e_t^{N*}\right) - ce_t^{N*} - v\left(e_t^{N*}\right), \tag{12}$$

$$V^{NK} = \frac{e_t^{K*}}{E_t^*} G\left(E_t^*\right) - ce_t^{K*} - v\left(e_t^{K*}\right), \tag{13}$$

$$V^{NN} = \frac{e_t^{N*}}{E_t^*} G\left(E_t^*\right) - ce_t^{N*} - v\left(e_t^{N*}\right). \tag{14}$$

where $E_t^* \equiv q_t e_t^{K*} + (1-q_t)e_t^{N*}$. Look at equations (11), and (12) which stand for the gain for a Kantian parent when having a Kantian (respectively Nash) child. Depending on the type of his/her child, the parent correctly anticipates that the child chooses either e_t^{K*} or e_t^{N*}. However, whatever the type of the child, the parent assesses the value of this choice through the lens of its own moral rule, that is as if every other agent in the society were also making this choice (no matter what agents are actually choosing).[18]

[18] Note that Kantian parents do not care about what actually happens at equilibrium would it be for them or for their child (i.e., they choose the number of animals that would maximize their welfare if everybody were choosing the same number of animals). One could have considered more sophisticated

Therefore, when the parent has a child of type K then the child obtains the maximal utility from the point of view of the parent (i.e., e_t^{K*} solves problem (6)). However, when the Kantian parent has a child of type N, he assesses this choice supposing that every other agent were choosing e_t^{N*}. This is clearly sub-optimal from the point of view of the Kantian parent. Kantians then have an incentive to transmit trait K. Imperfect empathy implies 'cultural intolerance', i.e., a desire for transmitting one's own moral trait.

Now, look at equations (14), and (13) which capture the gain for a Nash parent when having a Nash (respectively Kantian) child. Again, depending on the type of his/her child, the parent is able to correctly foresee that the child chooses either e_t^{N*} or e_t^{K*} depending on his/her trait. But whatever the type of the child, the parent assesses this choice taking all other members' actions as given. When the child is Nash (14), the child derives the optimal utility level from the point of view of the parent. However, when he/she is Kantian, the parent does not apply the Kantian moral rule but the Nash one. This behavior leads to (13): the Kantian choice is clearly sub-optimal from the point of view of a Nash parent. Just as Kantian parents, Nash parents then have an incentive to transmit the Nash trait.[19]

A parent of type i, determines his/her optimal socialization effort τ_t^i by solving:

$$\max_{\tau_t^i} \qquad -\frac{\kappa\left(\tau_t^i\right)^2}{2} + P^{ii}V^{ii} + P^{ii'}V^{ii'},$$

where P^{ii} and $P^{ii'}$ are given by equations (9) and (10) while V^{ii} and $V^{ii'}$ are given by equations (11), (12), (13) and (14).

The optimal socialization effort of type-K (resp N) parents are respectively given by:

$$\tau_t^{K*} = \frac{(1-q_t)}{\kappa}(V^{KK} - V^{KN}), \tag{15}$$

$$\tau_t^{N*} = \frac{q_t}{\kappa}(V^{NN} - V^{NK}). \tag{16}$$

First note that the socialization effort of a type-i parent at time t is a decreasing function of the proportion of type-i agents in the population at time t. This is because when the proportion of type-i agents increases, parents of type-i can more heavily rely on oblique transmission to have children of type i and thus have lower incentives to exert a costly socialization effort. This feature, which is called the 'cultural substitution property', is standard in the cultural transmission literature. A new feature in the present framework, is that the socialization efforts of parents depend on the extend of land congestion through parents' anticipation of their child's welfare.[20]

The dynamics of the fraction of Kantian agents is described by the following equation

$$q_{t+1} = q_t + q_t(1-q_t)\left(\tau_t^{K*} - \tau_t^{N*}\right), \tag{17}$$

where τ_t^{K*} and τ_t^{N*} are given by equations (15) and (16).

4 Long run dynamics

The previous section (subsection 3.4) aimed at studying temporary equilibria, that is equilibria under a given partition of the population into groups of Nash players and

agents who choose the number of animals maximizing their welfare if- not everybody- but only Kantians were choosing the same number of animals (which is what actually happens at equilibrium). We provide more details about that case in sub-section 7.1.

[19] For the sake of simplicity, we also assume that parents, when evaluating the gains for having children of their type, form myopic anticipations regarding the number of animals chosen by their children. This allows us to avoid self-fulfilling prophecies (see Bisin and Verdier, 2000, for a formal treatment of self-fulfilling expectations in a socialization framework).

[20] Note that moral rules (i.e., adopting a Nash or Kantian behavior) do not affect the socialization decisions of agents. Actually, the socialization effort of one parent is not affected by socialization efforts of other parents because the probability to have a child with the same trait depends on q_t and not on q_{t+1}.

Kantian players. In this section, we adopt a dynamic perspective. At each period of time t, given the distribution of traits q_t, there is an equilibrium for animal numbers which in turn affects the future composition of the population q_{t+1}. There is a two-way interaction between the equilibrium choices of animals and the population composition in terms of moral traits.

Let $\Omega : [0,1] \to [0,1]$ be the function such that $q_{t+1} = \Omega(q_t)$. The dynamics of Kantian morality is described by the equation $q_{t+1} = \Omega(q_t)$. Let us now examine the long-run behavior of this economy. Our results are summed up in Proposition 4 below.

Proposition 4 *There exist two unstable equilibria $q = 0$ and $q = 1$ and at least one interior locally stable equilibrium. The interior equilibrium is unique if*

$$\underbrace{\left[G\left(e^{K*}\right) - e^{K*}\frac{G\left(E^*\right)}{E^*} + e^{N*}\frac{G\left(E^*\right)}{E^*} - G\left(e^{N*}\right)\right]}_{-}$$
$$\underbrace{+(1-q)\left[-\frac{de^{N*}}{dq}\left[G'\left(e^{N*}\right) - c - v'\left(e^{N*}\right)\right]\right]}_{+}$$
$$\underbrace{-q\left[\frac{dE^*}{dq}\left[e^{N*} - e^{K*}\right]\frac{G'\left(E^*\right) - \frac{G(E^*)}{E^*}}{E^*}\right]}_{-} < 0, \quad \forall q \in [0,1].$$

When the interior equilibria is unique it is globally attracting.

A first noticeable result which appears in Proposition 4 is that there will be some heterogeneity in the long-run (i.e., stable interior equilibrium) so that moral behaviors survive in the long-run. This result is due to the existence of cultural substitution. Whenever the fraction of type-i agents is very small, transmission of trait i (resp. i') by role models (i.e., oblique socialization) is very low (resp. high). Hence, type-i parents cannot anymore rely on socialization by the society to have children of type i, while those of type i' can fully rely on oblique socialization to have children of their type. As a consequence, type-i (resp. i') parents have incentives to exert a high (resp. low) socialization effort which prevents the disappearance of trait i. Hence, cultural substitution favors the survival of Kantian morality in the long-run.

When the condition stated in Proposition 4 holds, there exists a unique globally attracting steady state. In that case, a rise in q_t has a negative impact on the future fraction of Kantian agents: there exists *substitutability* in the dynamics of Kantian morality. Looking at this condition, one can see that several counteracting forces shape the dynamics of moral behaviors.

There are two forces which favor dynamic substitutability and thus the existence of a unique interior equilibria. The first force, which is captured by the former term of the sum, is the cultural substitution effect. As stated before, through this effect, any rise in q_t decreases (resp. increases) the socialization effort of Kantian (resp. Nash) agents and thus gives rise to dynamic substitutability.

The second force favoring dynamic substitutability impacts socialization effort through changes in the gain associated to the type of the child. This force is captured by the third term of the sum, which is negative. In this set-up, when q_t rises, the socialization effort of Nash parents increases, which, in turn, decreases the future fraction of Kantian agents. The intuition goes as follows. As stated in Corollary 1, a rise in q_t reduces the aggregate number of animals, which increases the marginal product per animal. From the point of view of Nash parents, this reduction benefits to both Nash children and Kantian ones (see equations (13) and (14)). However, since Nash children choose a higher number of animals, they gain more from the rise in the marginal product per animal.

Finally, the last force, captured by the second term of the sum (which is positive) is a force to dynamic complementarity. It measures the impact of a rise in q on the socialization effort of Kantian parents. When q rises, Nash children increases their optimal number of animals and so move away from the optimum for Kantian parents which is e^{K*}. Hence, the rise in q decreases the gain, for Kantian parents, to have Nash children compared to that of having Kantian children which encourages those parents to increase their socialization effort.

Remark: As explained earlier, the cultural substitution property is sufficient for the preservation of Kantian morality in the long-run. Importantly, though, in our framework it is not necessary. Without cultural substitution, the dynamics of cultural traits would be given by equation (17) where socialization efforts are given by[21]

$$\begin{aligned}\tau_t^{K*} &= \frac{1}{\kappa}(V^{KK} - V^{KN}), \\ \tau_t^{N*} &= \frac{1}{\kappa}(V^{NN} - V^{NK}).\end{aligned}$$

In such a case, one easily shows that $q = 0$ and $q = 1$ are still steady states of the economy. Using a similar reasoning as for Proposition 4, the steady state $q = 1$ is stable if

$$\begin{aligned}&\frac{1}{\kappa}\left[-\frac{de^{N*}}{dq}|_{q_t=1}\left[G'\left(e^{N*}\right) - c - v'\left(e^{N*}\right)\right]\right] \\ &-\frac{1}{\kappa}\left[\frac{dE^*}{dq_t}|_{q_t=1}\left[e^{N*} - e^{K*}\right]\frac{G'\left(e^{K*}\right) - \frac{G\left(e^{K*}\right)}{e^{K*}}}{e^{K*}}\right] > 0.\end{aligned}$$

As shown in Proposition 4, the first term of this sum is positive and the second term is negative so that the sum can be positive. For instance, suppose that $G(E_t) = AE^{1/2}$, $v(e_t) = (2\Psi/3)(e_t)^{3/2}$, $c = 0$. In this case we have

$$\begin{aligned}&\frac{de^{N*}}{dq} = \frac{(e^{N*} - e^{K*})(q_t e^{K*} + (1 - q_t)e^{N*})^{-2}}{1 + (1 - q_t)(q_t e^{K*} + (1 - q_t)e^{N*})^{-2}}, \\ &e^{K*} = \frac{A}{2\Psi} \text{ and } e^{N*} = 2\frac{A}{\Psi} \text{ at } q = 1.\end{aligned}$$

We can show that the above condition holds whenever

$$\Psi > 0.75A^2.$$

When Ψ is high, the reaction of Nash players to a rise in q is strong. In that case, the number of animals chosen by Nash children is so far from their optimum that when q increases, parental preferences for having Kantian children become stronger. This in turn implies a strong enough reaction from the part of Kantian parents to allow the survival of Kantian morality.

At this stage, it is useful to contrast our results with the existing literature on the dynamics of moral behaviors under the Tragedy of the Commons. Our result of existence of heterogeneity in the long-run differs from Curry and Roemer (2012), who obtain, on the basis of an evolutionary game approach, that Nash players become extinct in the long-run when the Kantians can observe the type of their opponent in the contest, whereas Kantian players become extinct in the long-run when they cannot observe the type of their opponent. In a socialization framework, no matter the structure of information, both types of moral behaviors may survive due to parental socialization efforts.

[21] The cultural substitution effect could be neutralized by assuming positive cost externalities of socialization across individuals of the same cultural trait as in Olivier et al (2008).

5 Policy (1): quotas

Up to now, we considered an economy at the laissez-faire, without public intervention. However, the laissez-faire situation gives rise, in the present context, to a socially suboptimal result: the Tragedy of the Commons. In our context, the laissez-faire is characterized by a paradox: shepherds are buying, on average, too many animals, and the total product, in terms of wool, is too low, due to the congestion on land which reduces output per animal.

Let us now consider the impact of public policy on social outcomes. A first candidate for policy in the context of the Tragedy of the Commons is the imposition of a quota on the number of animals. We now suppose that the government fixes a quota equal to $\bar{e}$, which satisfies the following constraints:

$$e^{K*} \leq \bar{e} < e_t^{N*} \quad \forall q_t. \tag{18}$$

The lower bound of the quota, e^{K*}, coincides with the social optimum. Because of economic and political constraints, in reality, quotas exceed that optimal level. The second part of the inequality requires that Nash players are strictly constrained by the quota for any distribution of moral behaviors. In what follows, let us denote q^Q/q^{LF} (resp. E^Q/E^{LF}) the long-run distribution of moral behaviors (resp. aggregate number of animals) when a quota is implemented/at the *laissez-faire*.

Let us now consider the impact of imposing the quota $\bar{e}$ on the long-run aggregate number of animals E^Q. The derivative of E^Q with respect to the quota is given by

$$\frac{dE}{d\bar{e}} = (1 - q_t) - \frac{dq^Q}{d\bar{e}}(\bar{e} - e^{K^*}).$$

There are two distinct effects. The first positive effect is the intensive margin effect. When the quota increases, since Nash players choose the number of animals allowed by the quota, the number of animals chosen by those agents rises which increases the aggregate number of animals. The second effect which has ambiguous sign is the extensive margin effect. In our model, when the quota increases, the long-run fraction of Kantian agents is modified which in turn affects the aggregate number of animals.

Let us now examine the sign of $dq^Q/d\bar{e}$. The long-run equilibrium q^Q is implicitly given by $\Gamma^Q(q^Q) = 0$ where

$$\Gamma^Q(q^Q) = (1 - q^Q)\left(V^{KK} - V^{KN}\right) - q^Q\left(V^{NN} - V^{NK}\right),$$

with

$$\begin{aligned}
V^{KK} &= G\left(e^{K*}\right) - ce^{K*} - v\left(e^{K*}\right), \\
V^{KN} &= G\left(\bar{e}\right) - c\bar{e} - v\left(\bar{e}\right), \\
V^{NK} &= \frac{e^{K*}}{E^{Q*}} G\left(E^{Q*}\right) - ce^{K*} - v\left(e^{K*}\right), \\
V^{NN} &= \frac{\bar{e}}{E^{Q*}} G\left(E^{Q*}\right) - c\bar{e} - v\left(\bar{e}\right), \\
E^{Q*} &= q^Q e^{K*} + (1 - q^Q)\bar{e}.
\end{aligned}$$

Using Proposition 4 and since $d\bar{e}/dq_t = 0$, one can deduce the following.

Corollary 2 *There exists a unique globally stable stationary equilibrium* $q^Q \in]0, 1[$.

Using the implicit function theorem around the stationary equilibrium value of the fraction of Kantian agents, we find

$$\frac{dq^Q}{d\bar{e}} = -\frac{\partial \Gamma^Q / \partial \bar{e}}{\partial \Gamma^Q / \partial q^Q}.$$

Given Corollary 2, $\partial\Gamma^Q/\partial q^Q < 0$ so that $dq^Q/d\bar{e}$ is the sign of $\partial\Gamma^Q/\partial\bar{e}$. Using the FOC for both types of agents, we find

$$\frac{\partial\Gamma^Q}{\partial\bar{e}} = \underbrace{-(1-q^Q)\left(G^{'}(\bar{e}) - c - v^{'}(\bar{e})\right)}_{+}$$
$$\underbrace{-q^Q(1-q^Q)\frac{1}{E^{Q*}}\left(\frac{G^{'}(E^{Q*})}{E^{Q*}} - G(E^{Q*})\right)(\bar{e} - e^{K*})}_{+}$$
$$\underbrace{+q^Q\left(\frac{G\left(E^{Q*}\right)}{E^{Q*}} - c - v^{'}(\bar{e})\right)}_{-}.$$

The quota affects the long-run distribution of moral traits through several forces. First, a more stringent quota (i.e. a decrease in $\bar{e}$) negatively impacts the long-run fraction of Kantian agents by encouraging Kantian parents to reduce their socialization effort. The reason is that the quota decreases the loss when having a child who plays Nash as the latter, by choosing a lower number of animals, comes closer to the optimum for Kantian parents. This force is captured by the first term. The two subsequent terms capture the impact of the policy on socialization efforts of Nash parents. A rise in the quota's stringency has an ambiguous impact on Nash parents' socialization effort. On the one hand, the quota increases the gain to have Nash children by decreasing the aggregate number of animals. A decrease in the total number of animals benefits to each type of children by increasing the gain per animal purchased. However, since Nash children have more animals, they benefit more from this rise. On the other hand, the quota decreases the gain to have Nash children for Nash parents as the number of animals purchased by those children is reduced. Now, we can show the following.

Proposition 5 *In an economy with a quota* $\bar{e} \in \left[e^{K*}, e^{N*}\right[$*, there exists some threshold* $\tilde{e}$ *such that for any* $\bar{e} \in [\max\{\tilde{e}, e^{K*}\}, e^{N*}[$,

$$q^Q < q^{LF}.$$

When the quota is not restrictive enough, the negative effect on the socialization effort of Nash parents decreases (because the choice made by Nash children is closer to the optimum e^{N*}). The quota decreases the socialization effort of Kantian parents and increases the socialization effort of Nash parents, so that it negatively affects the fraction of Kantian agents.

Finally, let us examine whether the imposition of the quota $\bar{e}$ on the number of sheep per shepherd contributes to reduce the total number of animals, and, hence, allows the society to reach a Pareto improvement.

Proposition 6 *In an economy with a quota* $\bar{e} \in \left[e^{K*}, e^{N*}\right[$*, there exists* $\tilde{e}^{'}$ *such that for any* $\bar{e} \in [\max\{\tilde{e}^{'}, e^{K*}\}, e^{N*}[$,

$$E^Q > E^{LF}.$$

Let us illustrate this proposition by a simple numerical example. For that purpose, Figure 1 shows the total number of animals (y axis) as a function of the prevailing quota $\bar{e}$ (x axis). We see that the relationship between the total number of animals and the quota is non-monotonic. The extreme point on the right of the figure is the total number of animals when the quota is not constraining for Nash players. Then, when we move progressively to the left (and consider thus lower values for the quota $\bar{e}$), the quota becomes constraining for Nash players. The total number of animals associated with the quota is increasing with the quota $\bar{e}$, up to $\bar{e} = 3.3$. When we consider even more restrictive quotas, then the total number of animals falls.

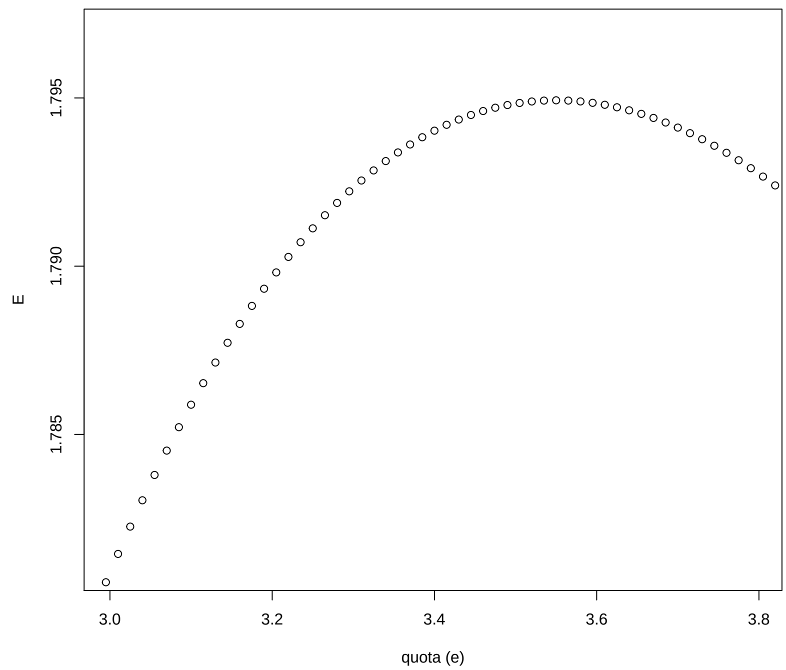

Figure 1: Total number of animals as a function of the quota ($G(E) = E^{1/2}$, $v(e) = (2/3)e^{3/2}$, $c = 1/2$, $\kappa = 200$).

In that numerical example, any quota that restricts the number of animals purchased by Nash players by less than 13 % contributes not to reduce, but to *raise* the total number of animals with respect to the laissez-faire. In other words, imposing quotas that are not strong enough worsens the Tragedy of the Commons.

This theoretical result echoes empirical evidence from both lab experiments and real-world situations. Cardenas et al. (2000) study the effects of external control (i.e. quotas) on environmental quality in rural settings of the developing world. They run experiments in three rural villages of Colombia where subjects were asked to decide how much time they would spend collecting firewood from a surrounding forest, while realizing that this activity has an adverse effect on local water quality because of soil erosion. First, they find that individuals' withdrawal systematically deviates from the individually selfish level. Second, they confronted subjects with a modest (i.e., lower than the optimum) government-imposed quota on the amount of time that can be spent collecting firewood. Individuals initially responded by restricting their withdrawals to close to the group optimum. However, after several periods, regulatory control caused subjects to make choices that were closer to their pure Nash strategies. This crowding out result echoes the result of Proposition 6 even though the underlying mechanism is certainly different that the one considered in that paper.

Our result is also consistent with findings from empirical field research described in Ostrom (2000). In a study of 48 long-lived irrigation systems in India, Bardhan (1999) finds that irrigation canals are significantly less successfully managed where farmers perceive the rules to be made by a government agency. In villages where rules have been enforced by an external agency, he finds that farmers have more negative attitudes toward water allocation rules and contribute less to the local village fund.

Finally, we rely on numerical simulations to examine how the efficacy of a quota policy is affected by socialization. We compare the effectiveness of quota policies in the short-run (i.e., the effectiveness of the policies if moral traits were fixed) to the effectiveness of such policies in the long-run (i.e., the effectiveness of the policies once we include endogenous moral traits). To do so, we compute the percentage reduction in the aggregate number of animals when a quota is implemented (with respect to the laissez-faire) in the short-run and in the long-run. Figure 2 displays these two quantities as a function of the quota $\bar{e}$. One can see that whatever the value of the quota $\bar{e}$, the effectiveness of the policy is lower in the long-run. Even when the policy has the intended effect to reduce the aggregate number of animals in the long-run, once we account for socialization, it will never performed as well as in the short-run. In other words, whatever the stringency of the policy, socialization negatively affects the efficacy of quota policy. This is because, in this numerical example, for any values of the quota, the policy reduces the long-run fraction of Kantian agents through its impact on socialization. Even when the net effect of the policy (i.e. sum of the intensive and extensive margin effect) is positive, the negative extensive margin effect reduces the efficacy of the policy in the long-run.

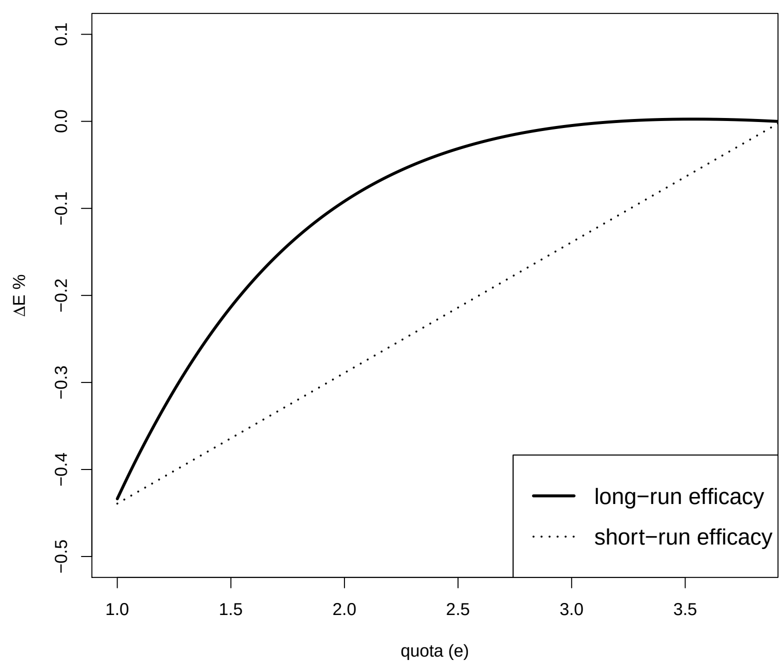

Figure 2: Long-run efficacy of a quota ($G(E) = E^{1/2}$, $v(e) = (2/3)e^{3/2}$, $c = 1/2$, $\kappa = 200$).

6 Policy (2): Pigouvian taxes

When facing the Tragedy of the Commons, another standard policy tool consists in introducing a Pigouvian tax on the number of animals. In our economy, Nash players do not, when choosing the number of animals, internalize the impact of the number of animals on the society. In order to make producers internalize the social consequences of their decisions, a government can introduce a tax on animals. If one denotes such a tax by $\top$, the price of each animals then increases from c to $c(1+\top)$. Just like for the quota case, the introduction of taxes has two distinct effects on the number of animals: one short-run intensive margin effect (as the cost of input increases producers have incentives to reduce the number of animals purchased), one long-run extensive margin effect since, by affecting incentives to socialize children, the tax affects the future population composition

Regarding the intensive margin effect, one expects that, by increasing the price of animals, the tax is likely, by a substitution effect, to reduce the number of animals purchased by each shepherd, and, hence, to reduce the total number of animals, to make it closer to the social optimum level. Let us examine how each type of agents reacts to a rise in the tax rate in the short-run.

$$\begin{aligned}
\frac{de_t^{K*}}{d\top} &= -\frac{-c}{G''(e_t^{K*}) - v''(e_t^{K*})} < 0, \\
\frac{de_t^{N*}}{d\top} &= -\frac{-c + q_t \frac{1}{E_t}\left(G'(E_t) - G(E_t)/E_t\right) de_t^{K*}/d\top}{(1-q_t)\frac{1}{E_t}\left(G'(E_t) - G(E_t)/E_t\right) - v''(e_t^{N*})}.
\end{aligned}$$

Note that, whereas the sign of $de_t^{K*}/d\top$ is clearly negative, the sign of $de_t^{N*}/d\top$ is ambiguous due to the fact that, as Kantians reduce their number of animals, Nash players react by *increasing* their number of animals. Since our focus is on socialization, we will, in the rest of this section, assume that this indirect effect is of smaller magnitude than the direct effect of the tax, so that the tax reduces Nash players' number of animals, i.e., $de_t^{N*}/d\top < 0$.[22]

Even though the intensive margin effect is negative, we do not know the impact of a tax on the aggregate number of animals, because of the extensive margin effect. Let us now examine the direction of this second effect, that is the impact of a rise in $\top$ on the long run fraction of Kantian $q^\top$. The long-run equilibrium is unique if the condition stated in Proposition 4 is satisfied. It is implicitly given by

$$\Gamma^\top(q^\top) = 0,$$

where

$$\Gamma^\top(q^\top) = \frac{(1-q^\top)}{\kappa}(V^{KK} - V^{KN}) - \frac{q^\top}{\kappa}(V^{NN} - V^{NK}),$$

with

$$\begin{aligned}
V^{KK} &= G(e^{K*}) - c(1+\top)e^{K*} - v(e^{K*}), \\
V^{KN} &= G(e^{N*}) - c(1+\top)e^{N*} - v(e^{N*}), \\
V^{NN} &= e^{N*}G(E^*)/E^* - c(1+\top)e^{N*} - v(e^{N*}), \\
V^{NK} &= (e^{K*})G(E^*)/E^* - c(1+\top)e^{K*} - v(e^{K*}).
\end{aligned}$$

Using the implicit function theorem, we obtain:

$$\frac{dq^\top}{d\top} = -\frac{\partial\Gamma^\top/\partial\top}{\partial\Gamma^\top/\partial q^\top}.$$

[22]For instance, suppose that $G(E_t) = E_t^\alpha$. Our assumption holds if and only if

$$q_t(E_t)^{\alpha-2}(1-\alpha) < \alpha(1-\alpha)(e_t^{K*})^{\alpha-1} + v''(e_t^{K*}).$$

This inequality holds at $q_t = 0$ and at $q_t = 1$ since then $E_t = e_t^{K*}$. In addition, the derivative of the LHS with respect to q_t is equal to

$$(E_t)^{\alpha-2}(1-\alpha) + q_t(2-\alpha)(e_t^{K*} - e_t^{N*})(E_t)^{\alpha-2}(1-\alpha),$$

which is positive. Thus our assumption is satisfied.

We have $\partial\Gamma^{\top}/\partial q^{\top} < 0$. Hence, the sign of $dq^{\top}/d\top$ is the sign of $\partial\Gamma^{\top}/\partial\top$. We obtain the following:

$$\frac{\partial\Gamma^{\top}}{\partial\top} = \underbrace{c(e^{N*} - e^{K*})}_{+} \underbrace{-(1-q^{\top})\left(G'(e^{N*}) - c(1+\top) - v'(e^{N*})\right)\frac{de^{N*}}{d\top}}_{-}$$
$$\underbrace{+\, q^{\top}\left(G(E^*)/E^* - (c+\top) - v'(e^{K*})\right)\frac{de^{K*}}{d\top}}_{-}$$
$$\underbrace{-q^{\top}\frac{1}{E^{*2}}\left(G'(E^*)E^* - G(E^*)\right)(e^{N*} - e^{K*})\frac{dE^{N*}}{d\top}}_{-}.$$

The first term captures the direct effect of the tax on utilities of both types of children. Whatever the parent's type, the tax increases the utility of Kantian children with respect to that of Nash children. The tax increases the cost per animal purchased. Since, the number of animals purchased by Nash players is higher than that of Kantian agents (whatever the parent's type), Nash children suffer more from the rise in the cost per animal. This first direct effect increases (resp. reduces) incentives to transmit the Kantian (resp. Nash) trait and positively affects the future fraction of Kantian agents. The three negative terms capture the indirect effect of a rise in the tax through changes in optimal choices of both types of agents. When the tax rises, the number of animals chosen by Nash players decreases, which makes their choice of animals closer to the optimal choice for Kantians. This effect reduces the loss of Kantian parents when they have Nash children, which negatively affects $q^{\top}$. Second, a rise in the tax rate $\top$ also decreases the number of animals chosen by Kantian agents, which makes the choice of Kantian children more distant from that of Nash children and increases cultural intolerance (i.e. preferences for Nash children) of Nash parents. Finally, all else equal, the tax decreases the total number of animals, which increases the return per animal purchased. Since Nash players choose a higher number of animals, they benefit more from this rise.

To study the overall impact of the tax on the long-run proportion of Kantian agents, we rely on numerical simulations. In Figure 3, we draw the long-run fraction of Kantian agents as a function of the tax rate $\top$ where the latter varies between one and one hundred per cent of the cost of animals. In that numerical example, one can observe that the long-run fraction of Kantian agents is monotonically increasing with the tax. Hence, the extensive margin effect is positive meaning that socialization *amplifies* (rather than counteracts) the effect of the policy. Let us now examine how this result is impacted when some parameters vary. In particular, we examine the impact of variations in α, which measures the severity of land congestion at equilibrium. To do so, we perform the elasticity of the long-run fraction of Kantians to the tax rate $\top$ for each value of $\top$ (which is denoted $\epsilon(q^{\top}, \top)$). The different curves on Figure 4 correspond to different values of α (this was set to 1/2 on Figure 3). First, one can remark that for each value of α, the tax has a positive impact on the long-run proportion of Kantian agents. Moreover, the lower the value of α, i.e., the stronger the severity of land congestion at equilibrium, the higher the positive impact of the tax on the long-run proportion of Kantians (i.e., the elasticity of the long-run proportion of Kantians to the tax increases). Hence, in that framework, the more severe the Tragedy of the Commons at equilibrium, the stronger the crowding in of moral behaviors by the tax and the more effective is the tax instrument to tackle over-exploitation.

As for the quota case, we now compare the short-run efficacy to the long-run efficacy of the policy. To do so, we compute the percentage reduction in the aggregate number of animals in the short-run to the same percentage reduction in the long-run. Figure 5 displays these two quantities as a function of the tax rate. For the tax case, numerical simulations generate quite different results than for the quota. While the long-run efficacy of quotas was lowered compared to the short-run, in those simulations, the long-run

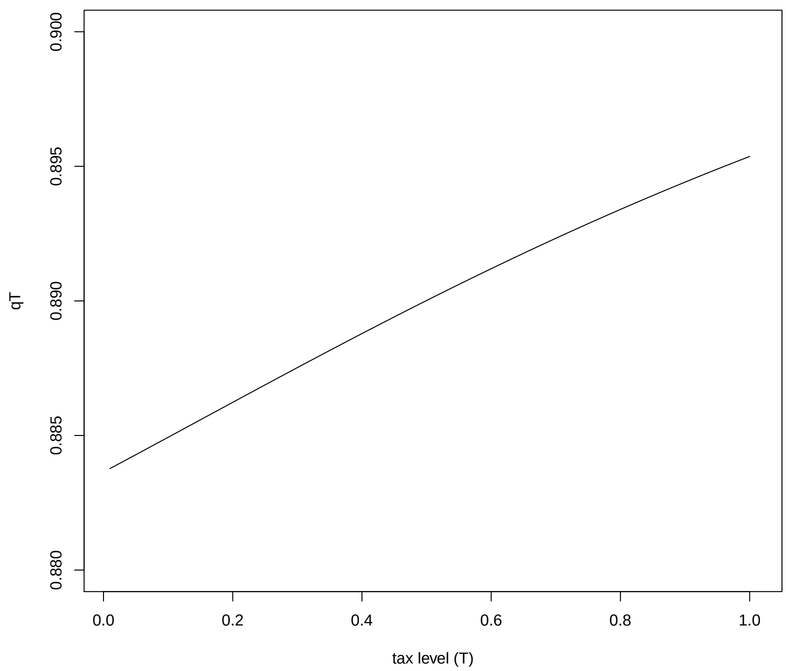

Figure 3: Long-run fraction of Kantian agents as a function of the tax rate $\top$ ($G(E) = E^{1/2}$, $v(e) = (2/3)e^{3/2}$, $c = 1/2$, $\kappa = 200$).

efficacy of taxes is always higher than in the short-run. Hence, while quotas crowd out moral motivations, taxes crowd in intrinsic motives. This qualitative change in the impact of the public policy can be rationalized as follows. We showed in Section 5 that quotas, which directly reduce the number of animals at time t, sometimes rise the aggregate number of animals in the long-run because they trigger an increase in the short-run productivity per animal. This rise benefits more to Nash players since they choose a higher number of animals. The productivity change in turn strengthens socialization to Nash behaviors and negatively affects Kantian morality in the long-run. By contrast, the tax indirectly regulates the total number of animals by affecting the cost of purchasing animals. As a consequence, there are two opposite effects affecting socialization to the Nash trait.[23] As for the quota, there is a positive effect due to the rise in average productivity in the short-run. However, there is an additional negative effect since the rise in the cost per animal negatively impacts individuals who choose more animals, that is Nash players. Overall, numerical simulations suggest that the second negative effect outweighs the first one so that the tax fosters Kantian morality. Hence, in our set-up, while formal regulation and intrinsic incentives are likely to be substitutes, monetary incentives and intrinsic motivations are likely to be complements.

The complementarity between economic incentives and intrinsic motivations is consistent with some empirical findings. For instance D'Haultfeuille et al. (2011) examine the impact of the French bonus-malus, a feebate that provides a financial reward for low-CO2-emitting vehicles, on consumers valuation of those vehicles. They find that the feebate had a crowding-in effect in addition to its price effect. The economic policy triggered a change in preferences which accounts for 40 % of the overall decrease in average CO2 emissions of new cars in the period considered.

[23]There are additional effects but they are of a second order magnitude due to the concavity of the utility functions.

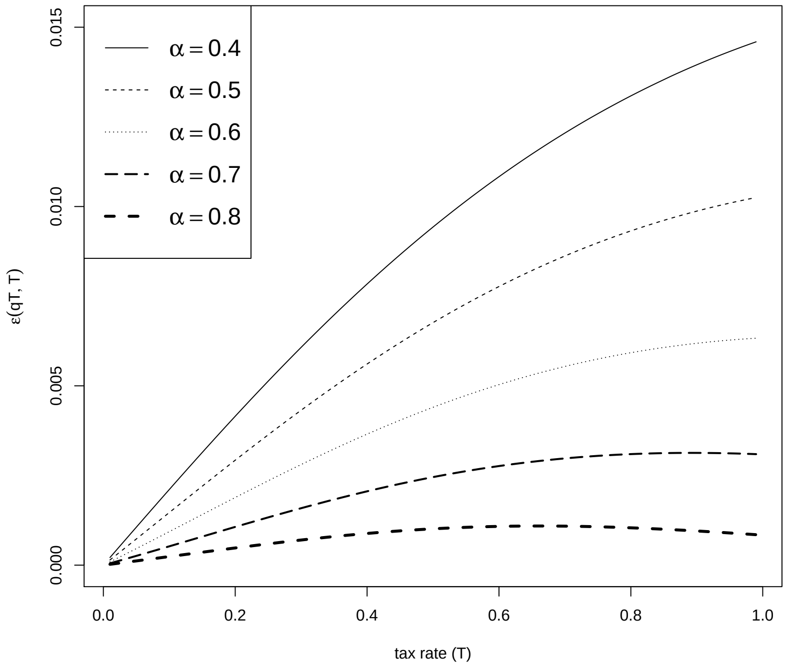

Figure 4: Elasticity of the long-run fraction of Kantian agents to the tax rate $\top$ as a function of the tax rate ($G(E) = E^{\alpha}$, $v(e) = (2/3)e^{3/2}$, $c = 1/2$, $\kappa = 200$).

7 Alternative modeling assumptions

7.1 Kantian behavior

We defined Kantian rationality as selecting the best generalizable action without taking into account that a fraction $1 - q_t$ of individuals behave in a different way. One may wonder to what extent our results are robust to an alternative definition of Kantian rationality. For instance, one may consider an 'Impure Kantian', which would choose the best action but generalizable *only* to her group (i.e. Kantian agents), while taking the Nash players' actions as given.

At the temporary equilibrium, the main difference is that Impure Kantians now react to the behavior of Nash players: the chosen number of animals will now depend on the distribution of moral traits in the population. An implication of this feature for socialization is that now the socialization effort of Kantian parents depends on the partition of the population not only through the cultural substitution effect but also through the gains associated to the type of the child.[24]

At the policy level, replacing Pure Kantians by Impure Kantians does not affect our results qualitatively. In particular, it is still the case that a too weak quota can worsen the Tragedy of the Commons, notably because Nash parents react to weak quotas by increasing their socialization effort (this is due to the rise in the average productivity just as in the baseline model).

[24] See Bezin and Ponthière (2016) for the associated conditions of existence, uniqueness and stability of the stationary equilibrium.

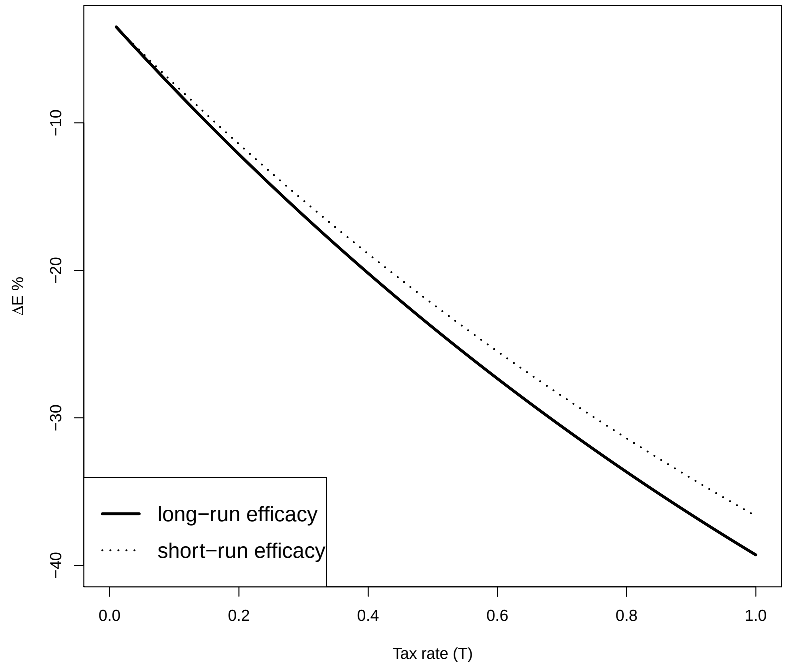

Figure 5: Long-run efficacy of the tax rate ($G(E) = E^{1/2}$, $v(e) = (2/3)\psi e^{3/2}$, $c = 1/2$, $\kappa = 200$).

7.2 Endogenous soil quality

Throughout this paper, we implicitly assumed that land regenerates perfectly from one period to the next independently from the extent of land congestion. This assumption allowed us to keep the dynamic system one-dimensional. This modeling assumption, also adopted by Curry and Roemer (2012), is a plausible proxy for animal grazing. However, it would be interesting to consider a more general setting where the resource is not perfectly renewable, as in Sethi and Somanathan (1996).

In addition to forces at work in our model, introducing an endogenous quality of land would add three extra forces in the dynamic system (which would then be two-dimensional). The first additional force would consist in the impact of land quality on parental socialization decisions. In a nutshell, a higher quality of land is likely to be more beneficial to Nash farmers, because they have more animals which would favor Nash players in the socialization process. The second force is the feedback from the distribution of moral behaviors to the quality of land. A rise in the proportion of Nash individuals contributes to deteriorate the quality of land. The third additional force is related to the intrinsic dynamics of land quality. Under a low speed of land regeneration, a high congestion of land may lead to a persistently low land quality.

When considering the long-run dynamics, the first two extra forces tend to create dynamic substitutability, which makes the emergence of a unique equilibrium more likely. On the contrary, the third extra force favors the emergence of a multiplicity of equilibria. This creates the possibility of low-quality traps, which would become particularly problematic in a society with a low initial proportion of Kantian agents.

7.3 Political economy

What would happen if individuals were voting on the tax per animal? Let us consider first the behavior of Kantians. Focusing on the self-oriented component of preferences, it can be shown that Kantians would vote for zero tax, since they already choose the optimal number of animals, which makes the tax redundant. This illustrates well that assuming Kantian rationality is distinct from assuming altruism or pro-social behaviors in general. This result echoes the findings of Alger and Weibull (2013): Kantian morality does not necessarily lead to the best social outcome from a consequentialist perspective. Consider the voting behavior of Nash players. Given that he anticipates that the tax will reduce land congestion and increase the return per animal, a Nash player votes for a positive tax.

From a static perspective, the political equilibrium depends on the partition of the population into Nash and Kantian traits. A positive tax emerges only if there is a majority of Nash individuals. Regarding long-run dynamics, one could observe an interesting substitution between morality and policy. A majority of Nash individuals will lead to a positive tax that tends to reduce the proportion of Nash players at the next period (through the parental socialization decisions), which progressively reduces the appeal of the tax. This mechanism goes on until Kantians have a political majority, which leads to a zero tax. In turn, the disappearance of the tax favors the diffusion of the Nash trait. There seems to be substitution forces between morality and policy.

7.4 A finite number of agents

We consider a continuum of agents for technical reasons: this assumption makes our stock variable q_t continuous, which allows us to rely on classical mathematical tools when studying the dynamics. If, instead, one considered a finite set of agents n, then Nash players would internalize the impact of their own actions on average productivity. Kantian agents' utility would be unchanged (except that it would depend on the parameter n). There would be strategic substitutability in shepherds' choices of animals, so that one could show the existence of a unique static equilibrium.

In addition, the same forces shape the dynamics, so that our previous results would hold. Take, for instance, the condition of Proposition 4. The first negative term, which captures the cultural substitution effect would still exist (since parents of type $i \in \{K, N\}$ still enjoy having children of type i). The second term, which captures the change in the gain to have Kantian children for Kantian parents resulting from a rise in q_t, would be unchanged, since Kantian agents' utility is unchanged in that alternative set-up. The third effect captures the impact of a change in q_t on the gain to have Nash children for Nash parents. It can be decomposed into two distinct effects. This first effect is due to the change in E_t resulting from an increase in q_t. This effect would remain positive, since an increase in the productivity per animal benefits more to Nash children (who choose more animals). The second effect comes from the change in the optimal number of animals chosen by each type of agent following the rise in q_t. This effect would still be nil, since (i) Kantian agents are insensitive to a change in q_t, (ii) the number of animals chosen by Nash players maximizes their utility, so that any change in that optimal number due to a change in q_t does not affect Nash players' utility. Hence, the same forces as for the case with a continuum of agents would shape the dynamics and allow for the existence of a stable interior stationary equilibrium.

Looking at the condition stating that the quota has a negative impact on the long-run fraction of Kantians, one finds that the two first (positive) effects would still exist, while the third effect would still be measured by the FOC for Nash assessed at $e^{N*} = \bar{e}$. Even though the FOC for Nash players is different from the one in our benchmark model (since Nash players now internalize the impact of their own action on average productivity), this effect goes to zero when the quota is low, i.e., when $\bar{e}$ tends to e^{N*}. Then, for some values of the quota, the two first forces would prevail implying that the quota crowds out moral behaviors.

8 Conclusion

Experiments of common pool resource games show that individuals, when facing the possibility to extract a common resource at the expense of others, do not extract as much as what rational self-oriented Nash players would extract in theory. One possible rationalization of those facts consists in assuming that not all agents have the same rationality.

The goal of this paper was to reexamine the Tragedy of the Commons in a context where individuals have different rationalities, some being Nash players, whereas others are Kantian players, and where the partition of the population into those different types of rationality follows a vertical/oblique socialization process à la Bisin and Verdier (2001). Our main result is that, quite paradoxically, introducing a quota on the number of sheep per shepherd can, in some conditions, lead to a rise in the extent of congestion of land, that is, can reinforce the Tragedy of the Commons. The reason is that such a quota can, in some cases, reduce the proportion of Kantians in the population. Such an undesirable composition effect (extensive margin) - which is absent in a static world with a fixed partition of the population - can overcome the social benefits from forcing Nash players to buy fewer sheep, and lead to an even more congested common land. This paradoxical result suggests that policy makers, when facing a Tragedy of the Commons, should be extremely cautious before imposing weak quotas, because in some cases the cure could make the disease even more severe. Although weak quotas can, because of an extensive margin effect, worsen the Tragedy of the Commons, we showed numerically that even a low tax on animals can reduce the congestion of land. Hence, in a world where governments face difficulties in imposing strong quotas/taxes, the tax dominates the quota, by allowing for a reduction of land congestion.

Whereas the above results are positive, one may be curious to know more about the design of optimal policies in our framework. If one adopts a static perspective (the partition being taken as given) and focuses only on the self-oriented component of preferences, then it easy to show that the first-best policy consists of a quota equal to the optimal number of animals for Kantians.[25] If one now considers the dynamic framework, two difficulties arise. First, parents do not have the same preferences concerning the type of children, so that it is not straightforward to aggregate individual utility functions. A second, more fundamental issue lies in the fact that a purely consequentialist social welfare criterion may be questioned in the present context, where individuals differ about rationality and care about the decision process (which they want to transmit to their children) and not only about economic outcomes.

All in all, the main contribution of this paper is to show that, in order to deal with the Tragedy of the Commons, governments should, when designing a policy intervention, study the consequences while taking into account its effect on the dynamics of moral behaviors. Ignoring that dynamics may lead governments to implement policies that make the Tragedy worse than at the laissez-faire, even though such policies would have worked well for given moral traits.

9 References

Alger, I., Weibull, J. (2013). Homo Moralis—Preference Evolution Under Incomplete Information and Assortative Matching, *Econometrica*, 81(6), 2269-2302.

Alger, I., Weibull, J. (2016a). Evolution and Kantian morality. *Games and Economic Behavior*, 98: 56-67.

Alger, I., Weibull, J. (2016b). Morality: evolutionary foundations and policy implications, mimeo.

[25] In a static setting, the optimal tax could be derived from a simple social planning problem where the government would act as a Stackelberg leader and would internalize the reaction of the two types of agent to the tax. Given that the tax, unlike the quota, cannot target Nash players only, the optimal tax would be clearly dominated by the optimal quota.

Apesteguia, J. (2006). Does information matter in the Commons? Experimental Evidence. *Journal of Economic Behavior and Organization*, 60: 55-69.

Bar-Gill, O., Ferstman, C. (2005). Public policy with endogenous preferences. *Journal of Public Economic Theory*, 7 (5), 841-857.

Bardhan, P. (1999). Water community: An empirical analysis of cooperation on irrigation in South India. *Berkeley: University of California, Department of Economics, Working paper.*

Bell, A.V, Richerson, P.J., McElreath, R. (2009). Culture rather than genes provides greater scope for the evolution of large-scale human prosociality, *Proceedings of the National Academy of Sciences*, 106(42), 17671-17674.

Bezin, E. (2015). A cultural model of private provision and the environment. *Journal of Environmental Economics and Management*, 71: 109-124.

Bezin, E. (2019). The economics of green consumption, cultural transmission and sustainable technological change. *Journal of Economic Theory*, 181: 497-546.

Bezin and Ponthière (2016). The tragedy of the Commons and Socialization. Theory and Policy. PSE Discussion Paper 2016-28.

Brekke, K. A., Kverndokk, S., Nyborg, K. (2003). An economic model of moral motivation. *Journal of public economics*, 1967-1983.

Bisin, A., Verdier, T. (2000). A model of cultural transmission, voting and political ideology, *European Journal of Political Economy*, 16(1), 5-29.

Bisin, A., Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences, *Journal of Economic Theory*, 97: 298-319.

Bisin, A., Verdier, T. (2011). The Economics of cultural transmission and socialization. In J. Benhabib, A. Bisin and M. Jackson (eds), *Handbook of Social Economics*, North Holland.

Bjornerstedt, J., Weibull, J. (1994). Nash equilibrium and evolution by imitation. Discussion paper, Department of Economics, Stockholm University.

Bowles, S. (2008). Policies designed for self-interested citizens may undermine the moral sentiments: Evidence from economic experiments. *Science*, 320, 5883.

Bowles, S., Hwang, S-H. (2008). Social preferences and public economics: mechanism design when social preferences depend on incentives. *Journal of Public Economics*, 92, 1811-1820.

Brekke, K., Kverndokk, S., Nyborg, K. (2003). An economic model of moral motivation. *Journal of Public Economics*, 87 (9-10): 1967-1983.

Cardenas, J. C., Stranlund, J., Willis, C. (2000). Local environmental control and institutional crowding-out. *World development*, 28(10): 1719-1733.

Cerda Planas, L. (2018). Moving Toward Greener Societies: Moral Motivation and Green Behaviour. *Environmental and Resource Economics*, 70, 835-860.

Cressman, R. (2003). *Evolutionary dynamics and extensive form games.* Cambridge: MIT Press.

Curry, P., Roemer, J. (2012). Evolutionary stability of Kantian optimization. *Hacienda Publica Espanola / Review of Public Economics*, 200 (1), 131-148.

Daube, M., Ulph, D. (2016). Moral behavior, altruism and environmental policy. *Environmental and Resource Economics*, 63 (2), 505-522.

De Donder, P., Pestieau, P. (2015). Lobbying, Family Concerns, and the Lack of Political Support for Estate Taxation. *Economics & Politics*, 27(3): 389-403.

De Donder, P., Roemer, J. (2016). An allegory of the political influence of the top 1%, *Business and Politics*,18, 1, 85–96.

D'Haultfeuille, X. Durrmeyer, I., Fevrier, P. (2011) Le coût du bonus-malus écologique: que pouvait-on prévoir? *Revue Economique*, 62, 491-499.

Fehr, E., Schmidt, K. (1999). A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics*, 114 (3): 817-868.

Gale, J., Binmore, K., Samuelson, L. (1995). Learning to be imperfect: the ultimatum game. *Games and Economic Behavior*, 8 (1):

56-90.

Grafton, R.Q., Kompas, T., Long, N.V. (2017). A brave new world? Kantian-Nashian interaction and the dynamics of global climate change mitigation. *European Economic Review*, 99, 31-42.

Harding R. (1968). The tragedy of the commons. *Science*, 162: 1241–1248.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *The American Economic Review*, 91(2): 73-78.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., Henrich, N. S. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and brain sciences*, 28(6): 795-815.

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation, *Journal of Economic Behavior & Organization*, 53(1), 3-35.

Keser, C., Gardner, R. (1999). Strategic behavior of experienced subjects in a common pool resource game, *International Journal of Game Theory*, 28: 241-252.

Laffont, J.J. (1975). Macroeconomic constraints, economic efficiency and ethics: an introduction to Kantian economics. *Economica*, 168: 430-437.

Lehmann, L., Alger, I., Weibull, J. (2015). Does evolution lead to maximizing behavior? *Evolution*, 69 (7): 1858-1873.

Long, N. V. (2016). The impacts of other-regarding preferences and ethical choice on environmental outcomes: A review of the literature. *Strategic Behavior and the Environment*, 6(1-2), 1-35.

McCay, B., Acheson, J. (1987). *The Question of the Commons. The culture and ecology of communal resources.* University of Arizona Press, Tucson.

Noailly, J., Van den Bergh, J., Withagen, C. (2007). Spatial evolution of social normsin a common-pool resource game. *Environmental and Resource Economics*, 36, 113-141.

Noailly, J., Van den Bergh, J., Withagen, C. (2009). Local and global interactions in an evolutionary resource game. *Computational Economics*, 33, 155-173.

Olivier, J., Thoenig, M., Verdier, T. (2008). Globalization and the dynamics of cultural identity. *Journal of International Economics*, 76(2), 356-370.

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of economic perspectives*, 14(3), 137-158.

Ostrom, E., Gardner, R., Walker, J. (1994). *Rules, games, and common-pool resources.* The University of Michigan Press, Ann Arbor.

Ostrom, E., Dietz, T., Dolsak, N., Stern, P., Stonich, S. and Weber, E. (2002) *The Drama of the Commons*. Washington: National Research Council.

Ponthiere, G. (2013). Long term care, altruism and socialization. *B.E. Journal of Economic Analysis and Policy - Advances*, 14 (2): 429-471.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83 (5): 1281-1302.

Rege, M. (2004). Social norms and private provision of public goods. *Journal of Public Economic Theory*, 6(1), 65-77.

Roemer, J. (2010). Kantian equilibrium. *Scandinavian Journal of Economics*, 112: 1-24.

Roemer, J. (2015). Kantian optimization: A microfoundation for cooperation. *Journal of Public Economics*, 127, 45-57.

Schumacher, I. (2009). The dynamics of environmentalism and the environment. *Ecological Economics*, 68(11), 2842-2849.

Schumacher, I. (2015). The endogenous formation of an environmental culture. *European Economic Review*, 76: 200-221.

Sethi, R., Somanathan, E. (1996). The evolution of social norms in common property resource use. *American Economic Review*. 86 (4): 766-788.

Sethi, R., Somanathan, E. (2001). Preference evolution and reciprocity. *Journal of Economic Theory*. 97: 273-297.

Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., Kitayama, S. (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science*. 344(6184): 603-608.

Walker, J., Gardner, R. (1992). Probabilistic destruction of common-pool resources: Experimental evidence. *Economic Journal*, 102: 1149-1161.

Weibull, J. (1995). *Evolutionary game theory.* Cambridge: MIT Press.

A Appendix

A.1 Proof of Proposition 3

The proof of the first item of Proposition 3 follows from the same arguments as those used in the proof of Proposition 2.

We show the second item by contradiction. Suppose that

$$\begin{aligned} &e_t^K > e_t^N, \\ \Leftrightarrow &v(e_t^K) > v(e_t^N), \\ \Leftrightarrow &G'(e_t^K) > \frac{G(E_t)}{E_t} \end{aligned}$$

Note that the function $\frac{G(e)}{e}$ is decreasing in e. Since $e_t^K > E_t > e_t^N$ by assumption, we have

$$\frac{G(e_t^K)}{e_t^K} < \frac{G(E_t)}{E_t}$$

Moreover, due to the concavity of G we have $G'(e) < \frac{G(e)}{e}$, implying

$$G'(e_t^K) < \frac{G(E_t)}{E_t},$$

which contradicts the above so that we deduce that

$$e_t^K < e_t^N.$$

A.2 Proof of Corollary 1

Item (i). The implicit function theorem in a neighborhood of e_t^{N*} gives

$$\frac{de_t^{N*}}{dq} = -\frac{(e_t^{K*} - e_t^{N*})\frac{G'(E_t)E_t - G(E_t)}{E_t^2}}{(1-q_t)\frac{G'(E_t)E_t - G(E_t)}{E_t^2} - v''(e_t^{N*})}.$$

Because of the concavity of the function G we have

$$G'(E_t)E_t - G(E_t) < 0.$$

We deduce that the denominator is negative. The numerator is positive since we show in the second item of Proposition 3 that $e_t^{K*} - e_t^{N*} < 0$. We deduce that $\frac{de_t^{N*}}{dq} > 0$.

Item (ii).

$$\frac{dE_t^*}{dq} = -(e_t^{N*} - e_t^{K*}) + (1-q_t)\frac{de_t^{N*}}{dq}.$$

Remark that

$$\frac{de_t^{N*}}{dq} = \frac{(e_t^{N*} - e_t^{K*})}{1 - q_t + v''(e_t^{N*})/(\frac{G(E_t) - G'(E_t)E_t}{E_t^2})} < \frac{(e_t^{N*} - e_t^{K*})}{1 - q_t},$$

since $v''(e_t^{N*})/(\frac{G(E_t) - G'(E_t)E_t}{E_t^2}) > 0$ so that $\frac{dE_t^*}{dq} < 0$.

A.3 Proof of Proposition 4

The dynamics of q_t is given by:

$$q_{t+1} = q_t + q_t(1-q_t)\left[\frac{(1-q_t)}{\kappa}(V^{KK} - V^{KN}) - \frac{q_t}{\kappa}(V^{NN} - V^{NK})\right]$$

Obviously, there exist two stationary equilibria with homogeneous populations: $q = 0$ and $q = 1$.

Any stationary equilibrium q is stable if and only if at this value of q we have

$$|\frac{dq_{t+1}}{dq_t}| < 1.$$

We have

$$\begin{aligned}\frac{dq_{t+1}}{dq_t} =& 1 + (1-2q_t)\left[\frac{(1-q_t)}{\kappa}(V^{KK} - V^{KN}) - \frac{q_t}{\kappa}(V^{NN} - V^{NK})\right] \\ &+ q_t(1-q_t)\left(\frac{d\tau^{K*}}{dq_t} - \frac{d\tau^{N*}}{dq_t}\right)\end{aligned}$$

One can perform

$$\frac{dq_{t+1}}{dq_t}|_{q=0} = 1 + \frac{1}{\kappa}(V^{KK} - V^{KN}) > 1.$$

$$\frac{dq_{t+1}}{dq_t}|_{q=1} = 1 + \frac{1}{\kappa}(V^{NN} - V^{NK}) > 1.$$

Hence stationary equilibria $q = 0$ and $q = 1$ are unstable.

Now, let us examine the conditions for the existence and uniqueness of an interior stationary equilibrium.

An interior equilibrium exists if and only if

$$\frac{(1-q_t)}{\kappa}(V^{KK} - V^{KN}) - \frac{q_t}{\kappa}(V^{NN} - V^{NK}) = 0,$$

admits one solution.

Define the function $\Gamma : [0,1] \to [0,1]$ given by

$$\Gamma(q_t) \equiv \frac{(1-q_t)}{\kappa}(V^{KK} - V^{KN}) - \frac{q_t}{\kappa}(V^{NN} - V^{NK}),$$

with

$$\begin{aligned} V^{KK} &= G\left(e_t^{K*}\right) - ce_t^{K*} - v\left(e_t^{K*}\right), \\ V^{KN} &= G\left(e_t^{N*}\right) - ce_t^{N*} - v\left(e_t^{N*}\right), \\ V^{NK} &= \frac{e_t^{K*}}{E_t^*}G\left(E_t^*\right) - ce_t^{K*} - v\left(e_t^{K*}\right), \\ V^{NN} &= \frac{e_t^{N*}}{E_t^*}G\left(E_t^*\right) - ce_t^{N*} - v\left(e_t^{N*}\right), \end{aligned}$$

and

$$E_t^* = e_t^{K*}q_t + (1-q_t)e_t^{N*}.$$

We have $\Gamma(0) > 0$ and $\Gamma(1) < 0$ so that there exists at least one interior equilibrium. Then, the interior equilibrium is unique if $\Gamma^{'}(q_t) < 0 \; \forall q_t$. Let us perform this derivative,

$$\begin{aligned}\Gamma^{'}(q_t) = &-\frac{1}{\kappa}\left((V^{KK} - V^{KN}) + (V^{NN} - V^{NK})\right), \\ &+ \frac{(1-q_t)}{\kappa}\frac{de_t^{N*}}{dq_t}\left(G^{'}\left(e_t^{N*}\right) - c - v^{'}\left(e_t^{N*}\right)\right) \\ &- \frac{q_t}{\kappa}\left[\frac{dE^*}{dq_t}\left[e^{N*} - e^{K*}\right]\frac{G'\left(E^*\right) - \frac{G(E^*)}{E^*}}{E^*}\right] \\ &+ \frac{q_t}{\kappa}\frac{de_t^{N*}}{dq_t}\left(\frac{G\left(E_t^*\right)}{E_t^*} - c - v^{'}\left(e_t^{N*}\right)\right).\end{aligned}$$

Note that e_t^{N*} maximizes the utility of Nash players, that is

$$\frac{G\left(E_t^*\right)}{E_t^*} - c - v^{'}\left(e_t^{N*}\right) = 0.$$

Therefore, we have

$$\begin{aligned}\Gamma^{'}(q_t) = &-\frac{1}{\kappa}\left((V^{KK} - V^{KN}) + (V^{NN} - V^{NK})\right), \\ &+ \frac{(1-q_t)}{\kappa}\frac{de_t^{N*}}{dq_t}\left(G^{'}\left(e_t^{N*}\right) - c - v^{'}\left(e_t^{N*}\right)\right) \\ &- \frac{q_t}{\kappa}\left[\frac{dE^*}{dq_t}\left[e^{N*} - e^{K*}\right]\frac{G'\left(E^*\right) - \frac{G(E^*)}{E^*}}{E^*}\right].\end{aligned}$$

The first term of this sum is positive since both $V^{KK} - V^{KN} > 0$ and $V^{NN} - V^{NK} > 0$. The FOC for Kantian agents is

$$G^{'}\left(e_t^{K*}\right) - c - v^{'}\left(e_t^{K*}\right) = 0.$$

Due to Proposition 3, we know that at the q equilibrium, $e_t^{K*} < e_t^{N*}$, which implies

$$G^{'}\left(e_t^{N*}\right) - c - v^{'}\left(e_t^{N*}\right) < 0.$$

Hence, the second term of the sum is positive.

Finally, we know from the concavity of G that

$$G'\left(E^*\right) - \frac{G\left(E^*\right)}{E^*} < 0.$$

In addition, Corrolary 1 gives

$$\frac{dE^*}{dq_t} < 0.$$

We deduce that the third term of this sum is negative.

Whenever

$$\begin{aligned}&-\frac{1}{\kappa}\left((V^{KK} - V^{KN}) + (V^{NN} - V^{NK})\right), \\ &+ \frac{(1-q_t)}{\kappa}\frac{de_t^{N*}}{dq_t}\left(G^{'}\left(e_t^{N*}\right) - c - v^{'}\left(e_t^{N*}\right)\right) \\ &- \frac{q_t}{\kappa}\left[\frac{dE^*}{dq_t}\left[e^{N*} - e^{K*}\right]\frac{G'\left(E^*\right) - \frac{G(E^*)}{E^*}}{E^*}\right] < 0, \quad \forall q_t,\end{aligned}$$

there exists a unique $q \in]0,1[$ such that $\Gamma(q) = 0$.

To sum up, we have $\Omega(0) = 0$, $\Omega^{'}(0) > 0$, $\Omega(1) = 1$, $\Omega^{'}(1) < 0$. Given the continuity of Ω, if there exists a unique $q \in]0,1[$ such that $\Omega(q) = q$, then $\Omega^{'}(q) < 0$ so that $q \in]0,1[$ is stable. Finally, since $q \in]0,1[$ is unique it is globally attracting.

A.4 Proof of Proposition 5

First, using the FOC for Nash players we deduce that when $\bar{e}$ tends to $e^{N*}|_{q^{LF}}$ the last term of the sum above equals zero. Both other terms remain positive so that we deduce $\partial\Gamma^Q/\partial\bar{e} > 0$ which implies $dq^Q/d\bar{e} > 0$.

Second, we have $q^Q|_{\bar{e}=e^{N*}} = q^{LF}$ and q^Q is continuous in $\bar{e}$. We can conclude that there is a threshold $\tilde{e}$ such that for any $\bar{e} \in [\max\{\tilde{e}, e^{K*}\}, e^{N*}[$,

$$q^Q < q^{LF}.$$

A.5 Proof of Proposition 6

Remind that

$$\frac{dE}{d\bar{e}} = (1 - q_t) - \frac{dq^Q}{d\bar{e}}(\bar{e} - e^{K^*}).$$

We interest in the condition

$$\begin{aligned} &\frac{dE}{d\bar{e}} < 0, \\ \Leftrightarrow &\frac{dq^Q}{d\bar{e}} > \frac{(\bar{e} - e^{K^*})}{(1 - q_t)}. \end{aligned}$$

This condition requires $\frac{dq^Q}{d\bar{e}} > 0$. We know the sign of $\frac{dq^Q}{d\bar{e}}$ only in a neighbourhood of e_t^{N*} so that we study the sign of $\frac{dE}{d\bar{e}}$ in a neighbourhood of e_t^{N*}. In a neighborhood of e_t^{N*}, the condition becomes:

$$\begin{aligned} &\frac{dq^Q}{d\bar{e}}|_{e^{N*}} > \frac{(1-q^Q)}{e^{N*} - e^{K*}}, \\ \Leftrightarrow &\frac{-(1-q^Q)\frac{dV^{KN}}{d\bar{e}}|_{e^{N*}} - q^Q \frac{d(V^{NN}-V^{NK})}{d\bar{e}}|_{e^{N*}}}{(V^{KK} - V^{KN} + V^{NN} - V^{NK}) + q^Q \frac{d(V^{NN}-V^{NK})}{dq}|_{e^{N*}}} > \frac{(1-q^Q)}{e^{N*} - e^{K*}}, \\ \Leftrightarrow &- (1-q^Q)\frac{dV^{KN}}{d\bar{e}}|_{e^{N*}} - q^Q \frac{d(V^{NN} - V^{NK})}{d\bar{e}}|_{e^{N*}} \\ &> \\ &\frac{(1-q^Q)}{e^{N*} - e^{K*}} \left((V^{KK} - V^{KN} + V^{NN} - V^{NK}) + q^Q \frac{d(V^{NN} - V^{NK})}{dq}|_{e^{N*}} \right). \end{aligned}$$

We have

$$\frac{d\left(V^{NN} - V^{NK}\right)}{dq}|_{\bar{e}=e^{N*}} = \left(e^{N*} - e^{K*}\right)^2 \frac{1}{E^{Q*}} \left(\frac{G'(E^{Q*})}{E^{Q*}} - G(E^{Q*}) \right)$$

and

$$\frac{d\left(V^{NN} - V^{NK}\right)}{d\bar{e}}|_{\bar{e}=e^{N*}} = \left(e^{N*} - e^{K*}\right) \frac{1}{E^{Q*}} \left(\frac{G'(E^{Q*})}{E^{Q*}} - G(E^{Q*}) \right) (1 - q^Q),$$

so that the condition becomes

$$-(1-q^Q)\frac{dV^{KN}}{d\bar{e}}|_{e^{N*}} > \frac{(1-q^Q)}{e^{N*} - e^{K*}} \left(V^{KK} - V^{KN} + V^{NN} - V^{NK}\right).$$

Let simplify and substitute for $V^{KK} - V^{KN} + V^{NN} - V^{NK}$ as well as for $\frac{dV^{KN}}{d\bar{e}}|_{e^{N*}}$,

$$-G^{'}(e^{N*}) + c + v^{'}(e^{N*}) > \frac{G\left(e^{K*}\right) - G\left(e^{N*}\right)}{e^{N*} - e^{K*}} + \frac{G\left(q^{Q}e^{K*} + (1-q^{Q})e^{N*}\right)}{(q^{Q}e^{K*} + (1-q^{Q})e^{N*})}.$$

Remind that due to the FOC for Nash players, we have

$$\frac{G\left(q^{Q}e^{K*} + (1-q^{Q})e^{N*}\right)}{(q^{Q}e^{K*} + (1-q^{Q})e^{N*})} = c + v^{'}(e^{N*}),$$

so that the above condition is equivalent to

$$- G^{'}(e^{N*}) > \frac{G\left(e^{K*}\right) - G\left(e^{N*}\right)}{e^{N*} - e^{K*}}.$$

$$\Leftrightarrow G^{'}(e^{N*}) < \frac{G\left(e^{N*}\right) - G\left(e^{K*}\right)}{e^{N*} - e^{K*}}.$$

which is true by the concavity of the function $G(e)$ (and since $e^{N*} > e^{K*}$).

We deduce that $dE/d\bar{e} < 0$ in a neighborhood of e^{N*}. In addition, we have $E^{Q}|_{e^{N*}} = E^{LF}$. By continuity of E in $\bar{e}$, we deduce that that there exists $\tilde{e}^{'}$ such that for any $\bar{e} \in [\max\{\tilde{e}^{'}, e^{K*}\}, e^{N*}[$, $E^{Q} > E^{LF}$.