



**HAL**  
open science

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

Emeline Bezin, Grégory Ponthière

► **To cite this version:**

Emeline Bezin, Grégory Ponthière. The Tragedy of the Commons and Socialization: Theory and Policy. 2016. halshs-01403244

**HAL Id: halshs-01403244**

**<https://shs.hal.science/halshs-01403244>**

Preprint submitted on 25 Nov 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**PARIS SCHOOL OF ECONOMICS**  
ÉCOLE D'ÉCONOMIE DE PARIS

**WORKING PAPER N° 2016 – 28**

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

**Emeline Bezin  
Grégory Ponthière**

**JEL Codes: C62, D64, Q24, Z1**

**Keywords: Tragedy of the Commons, heterogeneity, Kantian rationality,  
socialization, overlapping generations, fitness**



**PARIS-JOURDAN SCIENCES ÉCONOMIQUES**

48, Bd JOURDAN – E.N.S. – 75014 PARIS  
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10  
[www.pse.ens.fr](http://www.pse.ens.fr)

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

Emeline BEZIN\* Gregory PONTIERE†

November 24, 2016

## Abstract

We revisit the Tragedy of the Commons in a dynamic overlapping generations economy peopled 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, Pure or Impure Kantian players). We examine the dynamics of heterogeneity and land congestion when the prevalence of those different forms of rationality evolves over time following a vertical/oblique socialization process à la Bisin and Verdier (2001). We study the impacts 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 (Pure and Impure), and worsen the Tragedy of Commons with respect to the laissez-faire. Finally, we examine whether a government should promote either a Pure or an Impure Kantian morality, by comparing the relative fitness of Pure/Impure Kantians, and their interactions with the congestion of land.

*Keywords:* Tragedy of the Commons, heterogeneity, Kantian rationality, socialization, overlapping generations, fitness.

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

---

\*Paris School of Economics.

†University Paris East (ERUDITE), Paris School of Economics and Institut universitaire de France (IUF). [corresponding author] Address: Ecole Normale Supérieure, 48 boulevard Jourdan, 75014 Paris, France. E-mail: gregory.ponthiere@ens.fr. Telephone: 0033-1-43136371.

# 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 of the Tragedy of the Commons consists of an economy of shepherds whose sheep graze on the same common (unregulated) parcel of land. When deciding how many sheep they let graze, each shepherd may not fully internalize the negative effect of raising the number of sheep on the output (wool, meat) of other shepherds. As a consequence, the land is congested, the return per sheep is low, and there is an aggregate undersupply of output.

Referring to the game-theoretical literature on Prisoners' Dilemmas, the Tragedy of the Commons can be described as a Nash equilibrium that is not Pareto efficient. To turn back to the example of shepherds, the land congestion that prevails at the *laissez-faire* is not Pareto efficient: each shepherd would be strictly better off than at the *laissez-faire* *provided* there was a way to enforce a global reduction in the number of sheep per shepherd.<sup>1</sup>

The overuse of a common resource has been widely tested in experiments through common pool resource games.<sup>2</sup> In those games, a common, unregulated resource can be extracted by some players at the detriment of others. The theoretical prediction of those games is that each player plays Nash, and extracts a lot. 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 pay-offs would extract.<sup>3</sup>

A possible explanation for that paradoxical result may be that players play Nash, but exhibit inequality aversion, as suggested in Fehr and Schmidt (1999). When players are averse to inequality, they choose to behave in a more cooperative way, to avoid large inequalities in outcomes among players. Another possible explanation is that the concept of Nash equilibrium is not adequate to describe how individuals behave in those experiments. 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, Roemer (2010) introduced the concept of Kantian equilibrium. Kantian rationality consists in choosing not the best strategy for oneself, but the best *generalizable* strategy, i.e. the best strategy in the hypothetical case where a group (possibly everyone) would play it as well.

---

<sup>1</sup>Indeed, such a reduction would, by raising the return per sheep, raise the net output per shepherd, and, as such, this would constitute a Pareto improvement.

<sup>2</sup>See Walker and Gardner (1992), Ostrom, Gardner and Walker (1994), Keser and Gardner (1999), Ostrom et al (2002) and Apesteguia (2006).

<sup>3</sup>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.

The coexistence of alternative forms of rationality raises the question of the *dynamics* of heterogeneity: which form of rationality will survive in the long run? More specifically, when considering the Tragedy of the Commons, one may wonder to what extent Kantian players survive or become extinct in the long-run. Then, turning to an economy with a government, one may also wonder to what extent public intervention can influence the dynamics of heterogeneity, by favoring the survival or the extinction of some form of rationality.

The goal of this paper is to study, in a dynamic model, the survival or the extinction of Kantian behaviors in the context of the Tragedy of the Commons. For that purpose, 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, Pure or Impure Kantian players, the latter being defined as players choosing the best actions generalized to their group only). In line with the literature, the more sheep a shepherd has, and the lower is the available land for each sheep, which reduces the output per sheep.

In order to study the dynamics of heterogeneity in the context of the Tragedy of the Commons, we rely on a vertical/oblique socialization process à la Bisin and Verdier (2001). In that setting, parents, who are motivated by imperfect empathy towards their children, can, through costly socialization efforts, influence the probability to transmit their trait (here their kind of rationality, Nash or Kantian) to their children. The dynamics of heterogeneity thus depends on parental socialization efforts, which are themselves, because of imperfect empathy, functions of the numbers and distributions of sheep among the population.

Our reliance on a socialization model can be motivated as follows. The literature on the management of common pool resources emphasized the existence of strong interactions between, on the one hand, the maintenance of commonly owned resources, and, on the other hand, the enhancement of a community: while the composition of a community can affect the overuse of a common resource, the reciprocal is also true.<sup>4</sup> Obviously this kind of interaction can be rationalized in various ways, including social learning (Gale et al 1993, Bjornerstedt and Weibull 1994). However, the central role of culture in explaining social behaviors motivates the study of models of cultural transmission. Recent empirical evidence emphasized that cultural norms are a key determinant of differences in social behaviors. Focusing on 15 small-scale societies (in Africa, Asia and South America), Henrich et al (2001) show that there exist significant differences across tribes in the outcomes of the ultimatum game.<sup>5</sup> Moreover, there is also some evidence of the role of socialization efforts in the transmission of (non)cooperative behaviors (Knight et al, 1993).

Anticipating our results, we first characterize, at the temporary equilibrium,

---

<sup>4</sup>For example, Ostrom (1987) in McCay and Acheson (1987), taking the case of communal tenure in Swiss Alps, explained that in those rural communities using common land, overgrazing has been kept within control, and that not only has the common resource been protected, but that the community has been enhanced by the absence of overgrazing.

<sup>5</sup>For instance, the mean offer made by the Mapuche (Chile) is larger than the one made by the Machiguenga (Peru) (0.34 against 0.26).

four distinct situations: (1) the symmetric Nash equilibrium (when the population is composed of Nash players only); (2) the symmetric Kantian equilibrium (when the population is composed of Kantian players only); (3) the symmetric  $q$  equilibrium with Pure Kantians (when a fraction  $q$  of the population is Pure Kantian, and chooses the best generalizable strategy independently from the composition of the population, while a fraction  $1 - q$  plays Nash); (4) the symmetric  $q$  equilibrium with Impure Kantians (when a fraction  $q$  of the population is Impure Kantian, and chooses the best strategy generalizable to its group, while a fraction  $1 - q$  plays Nash). We examine the existence and uniqueness of those different types of temporary equilibria. Then, when studying the dynamics of heterogeneity, we examine the conditions for the existence, the uniqueness and the stability of stationary equilibria involving Nash players and (Pure or Impure) Kantian players. We show that (Pure or Impure) Kantians survive in the long-run for any initial partition of the population.

When turning to policy, we first show that introducing a quota on the number of sheep per shepherd may, in some cases, reduce the proportion of Kantians (Pure and Impure), and worsen the Tragedy of the Commons, by reinforcing the congestion of land. We also compare, on the basis of numerical analyses, the impact of a quota with the impact of a Pigouvian tax on animals, and we show that, although a weak quota can worsen the Tragedy, even a low tax can help reducing congestion. Thus, in a third-best world with strong political constraints, the tax performs better than the quota. Finally, we examine whether a government should promote either a Pure or an Impure Kantian morality, by comparing the relative fitness of Nash players, Pure and Impure Kantian players. We show that, for a given partition  $q$ , the total number of animals is lower at the  $q$  equilibrium with Pure Kantians than at the  $q$  equilibrium with Impure Kantians. However, once allowing for the endogeneity of the partition, whether a Pure or an Impure Kantian morality yields a less severe Tragedy varies with the parametrizations, and with the extent of the congestion of land.

Our paper is related to the literature on the dynamics of heterogeneity, and its applications to common resource games. First, our study is especially related to the article by Curry and Roemer (2012), which focuses on the Tragedy of the Commons in the case of 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 Kantian players can observe the type of their opponents, which raises the question of the signaling of types. Our paper complements Curry and Roemer (2012) by considering the survival of Kantian behaviors by means of a socialization approach involving parental efforts. Our paper is also close to Sethi and Somanathan (1996) who study the evolution of social norms in common property resource use by means of an evolutionary game approach. Unlike us, they consider the possibility of punishing defectors. We complement that paper by modeling how parents, by socialization efforts, contribute to the production of behavioral norms in the population. Another related paper, but in the context of climate change, is Van Long (2015), who characterizes the Kantian equilibrium in a dynamic game of climate change mitigations, and studies

the dynamics of the partition of Kantians in the population. Second, 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). Using a general model of evolutionary dynamics, Alger and Weibull (2014) show that Kantian morality can, unlike selfishness, be evolutionarily stable. That general result applies to various environments, including the Tragedy of the Commons. Third, our paper complements also the literature on the economics of culture and socialization (Bisin and Verdier 2001, 2011). While that literature focused mainly on the transmission of preferences traits, we focus here on the transmission of different forms of rationality. Finally, on the policy side, our paper is also related to the public economics literature showing that public policy can sometimes have perverse effects, through its impact on motivations (Brekke et al, 2003; Ponthiere 2013).

The paper is organized as follows. Section 2 presents the model, and characterizes the static Nash equilibrium, Kantian equilibrium,  $q$  equilibrium with Pure Kantians and  $q$  equilibrium with Impure Kantians. Section 3 studies socialization and the long-run dynamics of heterogeneity. The impact of introducing a quota on sheep per shepherd on the dynamics of heterogeneity and on congestion is examined in Section 4. Section 5 studies the influence of a Pigouvian tax on congestion. The relative fitness of the Pure and Impure Kantians (with respect to Nash players) is examined in Section 6. Section 7 concludes.

## 2 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 for the choice of level of input:

- Kantian agents (denoted type- $K$ ): those agents, when choosing the amount of input, choose the best generalizable input level;
- Nash agents (denoted type- $N$ ): those agents, when choosing the amount of input, choose the best input level for himself, conditionally on some beliefs regarding the choice of inputs of others.

At time  $t$ , the proportion of Kantian agents in the population of young adults is equal to  $q_t$ , whereas the proportion of Nash agents is equal to  $1 - q_t$ .

### 2.1 Production

Each individual  $i$  produces an output  $y_t^i$  by means of some input (animals)  $e_t^i$  and land. 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.

The production function for each producer  $i$  is:

$$y_t^i = F(e_t^i, s_t^i) \quad (1)$$

where  $s_t^i \equiv \frac{e_t^i S}{E_t}$  is the amount of land used by the producer  $i$ , while  $E_t \equiv \int_{i=0}^1 e_t^i di$ . Since there is no property right on land, the amount of land used by producer  $i$  is strictly proportional to the share of the number of animals he uses within the total number of animals.

Substituting for  $s_{it}$ , we have:

$$y_t^i = F\left(e_t^i, \frac{e_t^i}{E_t} S\right) \quad (2)$$

Let us normalize  $S$  to 1. Let us assume that  $F(\cdot)$  is homogenous of degree 1. Dividing the LHS and the RHS by  $e_t^i$ , we obtain:

$$\frac{y_t^i}{e_t^i} = F\left(1, \frac{1}{E_t}\right) \quad (3)$$

which the product per producer per animal. Hence the total product per shepherd is:

$$y_t^i = e_t^i F\left(1, \frac{1}{E_t}\right) \quad (4)$$

Defining  $G(E_t) \equiv E_t F\left(1, \frac{1}{E_t}\right)$  as the society's product, we get:

$$y_t^i = \frac{e_t^i}{E_t} G(E_t) \quad (5)$$

which is the functional form used by Curry and Roemer (2012). Note that we have  $G'(E_t) > 0$  and  $G''(E_t) < 0$ .

## 2.2 Preferences

Individuals derive utility from consumption, as well as some disutility from training animals.

Consumption is defined as the difference between the sales and the cost of buying animals:

$$x_t^i = p \frac{e_t^i}{E_t} G(E_t) - ce_t^i \quad (6)$$

where  $p$  is the price of output and  $c$  is the unit cost of animals. In the rest of this paper, we normalize  $p$  to 1 without loss of generality.



Individual utility on consumption  $u(x_t^i)$  is:

$$u(x_t^i) = \frac{(x_t^i)^{1-\sigma}}{1-\sigma} \quad (7)$$

with  $0 < \sigma < 1$ .

Total utility is defined as:

$$U(e_t^i) \equiv \frac{\left(\frac{e_t^i}{\hat{E}_t} G(\hat{E}_t) - ce_t^i\right)^{1-\sigma}}{1-\sigma} - v(e_t^i) \quad (8)$$

where  $v(e_t^i)$  is the disutility of training animals. We assume, for the sake of simplicity, that this disutility takes the following form:

$$v(e_t^i) = \frac{\psi(e_t^i)^2}{2} \quad (9)$$

We have:  $v'(e_t^i) = \psi e_t^i$  and  $v''(e_t^i) = \psi$ .

## 2.3 Decision rule

The two types of agents differ regarding the way in which they choose the number of animals.

### 2.3.1 Nash behavior

Type  $N$  agents solve the standard maximization of utility problem, subject to market prices and to beliefs regarding the choice of animals by other producers  $\hat{e}_t^i$ . This problem can be written as:

$$\begin{aligned} \max_{e_t^N} & \frac{\left(\frac{e_t^N}{\hat{E}_t} G(\hat{E}_t) - ce_t^N\right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^N)^2}{2} \\ \text{s.t. } & \hat{E}_t \equiv \int_{i=0}^1 \hat{e}_t^i di \end{aligned}$$

Note that the marginal impact of  $e_t^N$  on  $\hat{E}_t$  equals 0, due to the fact that there is a continuum of agents. Hence the optimal  $e_t^N$  satisfies:

$$\frac{\left(\frac{G(\hat{E}_t)}{\hat{E}_t} - c\right)}{\left(\frac{e_t^N}{\hat{E}_t} G(\hat{E}_t) - ce_t^N\right)^\sigma} - \psi e_t^N = 0 \quad (10)$$

Given that  $u'(x_t^i) > 0$ , the optimal level of  $e_t^N$  satisfies:

$$e_t^{N*} = \frac{\left(\frac{G(\hat{E}_t)}{\hat{E}_t} - c\right)}{\psi(e_t^{N*})^\sigma \left(\frac{G(\hat{E}_t)}{\hat{E}_t} - c\right)} \quad (11)$$

Hence we have:

$$(e_t^{N*})^{1+\sigma} = \frac{\left(\frac{G(\hat{E}_t)}{\hat{E}_t} - c\right)^{1-\sigma}}{\psi} \quad (12)$$

$$e_t^{N*} = \left[ \frac{\left(\frac{G(\hat{E}_t)}{\hat{E}_t} - c\right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}} \quad (13)$$

Thus the optimal number of animals  $e_t^{N*}$  is decreasing in the expected number of animals of other producers  $\hat{E}_t^{-N}$ , since the derivative of  $\frac{G(\hat{E}_t)}{\hat{E}_t}$  wrt to  $\hat{E}_t^{-N}$  is negative. This is a standard reaction function. Thus we have:

$$\frac{de_t^{N*}}{d\hat{E}_t^{-N}} < 0$$

This depends on the expected levels of animals for all producers, Kantian and non-Kantian, and on the proportions of the two types in the population  $q_t$  and  $1 - q_t$ .

### 2.3.2 Pure Kantians

Kantian rationality is about choosing the best "generalizable" actions. There exist two distinct types of Kantian behaviors, which differ regarding the scope of the "generalizability": generalizability to the whole society (Pure Kantians) or to the group of Kantians only (Impure Kantians).

Pure Kantians (sometimes called Kantians with imperfect information) select the number of animals that leads to the largest generalizable well-being level, that is, the number of animals such that, if chosen by all individuals (without exception), this would lead to the largest level of well-being.

Impure Kantian (sometimes called Kantians with perfect information) select the number of animals that leads to the largest well-being level generalizable to all Kantians, that is, the number of animals such that, if chosen by all Kantian individuals (with exception of the  $N$ -type agents), this would lead to the largest level of well-being for Kantian people. Note that, by definition, Impure Kantian exist only when  $q < 1$ . When  $q = 1$ , Impure Kantians behave like Pure Kantians.

When Pure Kantians, type- $K$  agents solve the following problem. The look for the number of animals that maximize their own well-being provided all other individuals enjoy the same number of animals. The problem can be written as:

$$\begin{aligned} \max_{e_t^K} & \frac{\left(\frac{e_t^K}{E_t} G(E_t) - ce_t^K\right)^{1-\sigma}}{1-\sigma} - \psi \frac{(e_t^K)^2}{2} \\ \text{s.t. } & E_t \equiv e_t^K \end{aligned}$$

The optimal  $e_t^{K*}$  satisfies:

$$\left( \frac{e_t^{K*}}{E_t} G(E_t) - ce_t^{K*} \right)^{-\sigma} (G'(E_t) - c) = \psi e_t^{K*} \quad (14)$$

An interior solution requires  $G'(E_t) - c > 0$ .

Hence:

$$e_t^{K*} = \left[ \frac{(G'(E_t) - c)}{\psi \left( \frac{G(E_t)}{E_t} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}} \quad (15)$$

Given that  $E_t = e_t^{K*}$  with Pure Kantians, it follows that the optimal Kantian number of animals is independent from  $q_t$ , unlike the one of a Nash player.

Let us compare this with the Nash optimal level of animals:

$$e_t^{N*} = \left[ \frac{\left( \frac{G(\hat{E}_t)}{\hat{E}_t} - c \right)}{\psi \left( \frac{G(E_t)}{E_t} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

Given that  $G'(E_t) < \frac{G(\hat{E}_t)}{\hat{E}_t}$ , we obviously have:

$$e_t^{K*} < e_t^{N*} \quad (16)$$

### 2.3.3 Impure Kantians

When Impure Kantians, type- $K$  agents solve the following problem. They look for the number of animals that maximize their own well-being provided all other Kantian individuals enjoy the same number of animals and the  $1 - q$  other individuals play Nash. The problem can be written as:

$$\begin{aligned} \max_{e_t^K} & \frac{\left( \frac{e_t^K}{E_t} G(E_t) - ce_t^K \right)^{1-\sigma}}{1-\sigma} - \psi \frac{(e_t^K)^2}{2} \\ \text{s.t. } & E_t \equiv q_t e_t^K + (1 - q_t) e_t^N \end{aligned}$$

The optimal  $e_t^{K**}$  satisfies:

$$\frac{\left( \frac{E_t - e_t^{K**} q_t}{(E_t)^2} G(E_t) + \frac{e_t^{K**}}{E_t} G'(E_t) q_t - c \right)}{\left( \frac{e_t^{K**}}{E_t} G(E_t) - ce_t^{K**} \right)^\sigma} = \psi e_t^{K**} \quad (17)$$

Hence:

$$e_t^{K**} = \left[ \frac{\left( \frac{E_t - e_t^{K**} q_t}{(E_t)^2} G(E_t) + \frac{e_t^{K**}}{E_t} G'(E_t) q_t - c \right)}{\psi \left( \frac{G(E_t)}{E_t} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}} \quad (18)$$

That condition can be rewritten as:

$$e_t^{K^{**}} = \left[ \frac{\left( \frac{e_t^{K^{**}} q_t}{E_t} \right) \left[ G'(E_t) - \frac{G(E_t)}{E_t} \right] + \frac{G(E_t)}{E_t} - c}{\psi \left( \frac{G(E_t)}{E_t} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}} \quad (19)$$

Note that, when  $q_t = 0$ , this condition collapses to  $\left[ \frac{\left( \frac{G(E_t)}{E_t} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}}$ , which coincides with  $e_t^N$ . When  $q_t = 1$ , the condition collapses to  $e_t^{K^{**}} = \left[ \frac{G'(E_t) - c}{\psi \left( \frac{G(E_t)}{E_t} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}} = e_t^{K^*}$ .

## 2.4 Equilibria

Let us consider now the characterization of the equilibria under the different partitions of the population in terms of Kantian (Pure and Impure) and non-Kantian. For that purpose, we will first consider the existence of an equilibrium when the population is composed exclusively of Kantian individuals or Nash players. Then, in a second stage, we will consider the existence of an equilibrium in mixed societies.

Throughout this paper, we will focus only on symmetric equilibria, i.e. equilibria where each Nash player acts as the other Nash players, and where each Kantian player acts as the other Kantians.

### 2.4.1 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 distribution of animal numbers uniform with  $e_t^i = e_t^{K^*}$  such that:*

$$e_t^{K^*} = \left[ \frac{(G'(e_t^{K^*}) - c)}{\psi \left( \frac{G(e_t^{K^*})}{e_t^{K^*}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

Proposition 1 examines the existence and uniqueness of the Kantian equilibrium.

**Proposition 1** *In an economy with  $q = 1$ , there exists a (symmetric) Kantian equilibrium. A sufficient condition for uniqueness of the Kantian equilibrium is:*

$$\frac{e_t^{K^*} G''(e_t^{K^*})}{G'(e_t^{K^*}) - \frac{G(e_t^{K^*})}{e_t^{K^*}}} > \sigma \quad \forall 0 \leq e_t^{K^*} < G'^{-1}(c)$$

**Proof.** There exists a (symmetric) Kantian equilibrium if and only if the equation

$$e_t^{K*} = \left[ \frac{(G'(e_t^{K*}) - c)}{\psi \left( \frac{G(e_t^{K*})}{e_t^{K*}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

admits a solution.

Remind that the interiority of the optimal number of animals requires:  $G'(e_t^{K*}) - c > 0$ . At  $e_t^{K*} = 0$ , we know that the RHS of the above equation is strictly positive. Moreover, at  $e_t^{K*} = G'^{-1}(c)$ , the RHS of the above condition equals zero. Hence, by continuity, the RHS must necessarily intersect the 45° line at least once.

Regarding uniqueness, note that the derivative of the RHS wrt to  $e_t^{K*}$  is:

$$\frac{1}{1+\sigma} \left[ \frac{(G'(e_t^{K*}) - c)}{\psi \left( \frac{G(e_t^{K*})}{e_t^{K*}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma} - 1} \frac{\left[ \begin{array}{l} (G''(e_t^{K*})) \psi \left( \frac{G(e_t^{K*})}{e_t^{K*}} - c \right)^\sigma \\ - (G'(e_t^{K*}) - c) \psi \sigma \left( \frac{G(e_t^{K*})}{e_t^{K*}} - c \right)^{\sigma-1} \frac{G'(e_t^{K*}) e_t^{K*} - G(e_t^{K*})}{(e_t^{K*})^2} \end{array} \right]}{\psi^2 \left( \frac{G(e_t^{K*})}{e_t^{K*}} - c \right)^{2\sigma}}$$

Hence we have:

$$\frac{1}{1+\sigma} \left[ \frac{(G'(e_t^{K*}) - c)}{\psi \left( \frac{G(e_t^{K*})}{e_t^{K*}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma} - 1} \left[ \frac{G''(e_t^{K*}) - \frac{(G'(e_t^{K*}) - c)\sigma}{G(e_t^{K*}) - ce_t^{K*}} \left( G'(e_t^{K*}) - \frac{G(e_t^{K*})}{e_t^{K*}} \right)}{\psi \left( \frac{G(e_t^{K*})}{e_t^{K*}} - c \right)^\sigma} \right]$$

The sign of this expression is the sign of:

$$G''(e_t^{K*}) - \frac{(G'(e_t^{K*}) - c)\sigma}{G(e_t^{K*}) - ce_t^{K*}} \left( G'(e_t^{K*}) - \frac{G(e_t^{K*})}{e_t^{K*}} \right)$$

or

$$G''(e_t^{K*}) - \frac{(G'(e_t^{K*}) - c)\sigma}{\frac{G(e_t^{K*})}{e_t^{K*}} - c} \frac{1}{e_t^{K*}} \left( G'(e_t^{K*}) - \frac{G(e_t^{K*})}{e_t^{K*}} \right)$$

The negativity of that expression is achieved when:

$$\frac{G''(e_t^{K*})}{\frac{(G'(e_t^{K*}) - c)}{\frac{G(e_t^{K*})}{e_t^{K*}} - c} \frac{1}{e_t^{K*}} \left( G'(e_t^{K*}) - \frac{G(e_t^{K*})}{e_t^{K*}} \right)} > \sigma$$

Given that  $\frac{[G'(e_t^{K^*})-c]}{\frac{G(e_t^{K^*})}{e_t^{K^*}}-c} < 1$ , a sufficient condition for uniqueness is that:

$$\frac{e_t^{K^*} G''(e_t^{K^*})}{\left(G'(e_t^{K^*}) - \frac{G(e_t^{K^*})}{e_t^{K^*}}\right)} > \sigma$$

Hence a sufficient condition for the existence of a unique Kantian equilibrium is:

$$\frac{e_t^{K^*} G''(e_t^{K^*})}{\left(G'(e_t^{K^*}) - \frac{G(e_t^{K^*})}{e_t^{K^*}}\right)} > \sigma \forall 0 \leq e_t^{K^*} < G'^{-1}(c)$$

Clearly, under that condition, the RHS of the initial expression is decreasing, and intersects the  $x$  axis at  $e_t^{K^*} = G'^{-1}(c)$ . Hence, it necessarily intersects the  $45^\circ$  line only once. ■

According to Proposition 1, there always exists at least one Kantian equilibrium. Moreover, a simple condition on the production function  $G(\cdot)$  suffices to guarantee the uniqueness of the Kantian equilibrium. In order to have an idea of the strength of that condition, let us consider the case where  $G(\cdot)$  takes the form of a Cobb-Douglas production function, i.e.  $G(e_t) = Ae_t^\alpha$  with  $0 < \alpha < 1$ .

In that case, the interiority condition implies:

$$A\alpha e^{\alpha-1} > c \iff e < \left(\frac{A\alpha}{c}\right)^{\frac{1}{1-\alpha}}$$

Hence the condition sufficient for uniqueness becomes:

$$\frac{e_t^{K^*} A\alpha(\alpha-1)(e_t^{K^*})^{\alpha-2}}{\left(A\alpha(e_t^{K^*})^{\alpha-1} - A(e_t^{K^*})^{\alpha-1}\right)} > \sigma \forall 0 \leq e_t^{K^*} < G'^{-1}(c) = \left(\frac{A\alpha}{c}\right)^{\frac{1}{1-\alpha}}$$

This can be simplified to:

$$\alpha > \sigma$$

Thus, under  $G(e_t) = Ae_t^\alpha$ , a sufficient condition for uniqueness of the Kantian equilibrium is  $\alpha > \sigma$ . Obviously, in case of quasi-linear preferences (i.e.  $\sigma = 0$ ), that condition is necessarily satisfied, so that the Kantian equilibrium exists and is unique.

## 2.4.2 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 distribution of animal numbers uniform with  $e_t^i = e_t^{N^*}$  such that:

$$e_t^{N^*} = \left[ \frac{\left( \frac{G(e_t^{N^*})}{e_t^{N^*}} - c \right)}{\psi \left( \frac{G(e_t^{N^*})}{e_t^{N^*}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

Proposition 2 states that there exists always a unique symmetric Nash equilibrium in this economy.

**Proposition 2** In an economy with  $q = 0$ , there exists a unique symmetric Nash equilibrium.

**Proof.** There exists a unique symmetric Nash equilibrium if and only if the following equation admits a unique solution:

$$e_t^{N^*} = \left[ \frac{\left( \frac{G(e_t^{N^*})}{e_t^{N^*}} - c \right)}{\psi \left( \frac{G(e_t^{N^*})}{e_t^{N^*}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

Note that interiority requires:  $\frac{G(e_t^{N^*})}{e_t^{N^*}} - c > 0$ .

The equation can be rewritten as:

$$e_t^{N^*} = \left[ \frac{\left( \frac{G(e_t^{N^*})}{e_t^{N^*}} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}}$$

The derivative of the RHS wrt to  $e_t^{N^*}$ :

$$\frac{1}{1+\sigma} \left[ \frac{\left( \frac{G(e_t^{N^*})}{e_t^{N^*}} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}-1} \frac{(1-\sigma)}{\psi} \left( \frac{G(e_t^{N^*})}{e_t^{N^*}} - c \right)^{-\sigma} \frac{\left[ G'(e_t^{N^*}) - \frac{G(e_t^{N^*})}{e_t^{N^*}} \right]}{\psi e_t^{N^*}}$$

Given the interiority condition the first factor and the second factors are positive. Given the concavity of  $G(e_t^N)$ , the last factor is negative. Hence the RHS of the equation is decreasing. Note that when  $\frac{G(e_t^{N^*})}{e_t^{N^*}} - c = 0$ , the RHS of the equation equals 0. Hence, by continuity, the RHS must intersect the 45° line once for an interior  $e_t^{N^*}$ . ■

Having shown that there always exists a unique symmetric Nash equilibrium, let us now compare the optimal levels of animals prevailing, respectively, at the

Nash equilibrium and at the Kantian equilibrium, i.e.,  $e_t^N$  and  $e_t^K$ . As shown in the following corollary, the level of animals prevailing at the Kantian equilibrium is inferior to the one prevailing at the symmetric Nash equilibrium. Moreover, it can be shown that the Kantian equilibrium is Pareto-efficient.

**Corollary 1** • *Given the concavity of  $G(e_t^N)$ , we have:*

$$e_t^{K*} < e_t^{N*}$$

• *Note that the Kantian equilibrium is Pareto-efficient.*

**Proof.** Clearly,  $e_t^{K*}$  maximizes social welfare when  $q = 1$ . ■

That corollary illustrates the welfare consequences of the Tragedy of (unregulated) Commons. At the decentralized equilibrium where all agents choose the number of animals maximizing their own interests, there exists an overuse of animals, leading to an excessive use of land, which implies a welfare loss for each shepherd in comparison with what would prevail at the Kantian equilibrium.

### 2.4.3 $q$ equilibrium with Pure Kantians

Having compared symmetric equilibria in economies that are homogenous in composition, let us now characterize symmetric equilibria when some individuals behave like Nash players, whereas other individuals are Kantians. For that purpose, we will proceed in two stages, and consider first the Pure Kantians.

**Definition 3** *A symmetric  $q$  equilibrium with Pure Kantians (imperfect information) is a distribution of animal numbers where a proportion  $q_t$  of producers has  $e_t^i = e_t^{K*}$  and a proportion  $1 - q_t$  of producers has  $e_t^i = e_t^{N*}$  such that:*

$$e_t^{K*} = \left[ \frac{(G'(e_t^{K*}) - c)}{\psi \left( \frac{G(e_t^{K*})}{e_t^{K*}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

$$e_t^{N*} = \left[ \frac{\left( \frac{G(q_t e_t^{K*} + (1-q_t)e_t^{N*})}{q_t e_t^{K*} + (1-q_t)e_t^{N*}} - c \right)}{\psi \left( \frac{G(q_t e_t^{K*} + (1-q_t)e_t^{N*})}{q_t e_t^{K*} + (1-q_t)e_t^{N*}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

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

**Proposition 3** *In an economy with  $0 < q < 1$ , there exists a unique symmetric  $q$  equilibrium with Pure Kantians.*



**Proof.** There exists a unique symmetric  $q$  equilibrium if and only if the following equation admits a unique solution:

$$e_t^{N^*} = \left[ \frac{\left( \frac{G(q_t e_t^{K^*} + (1-q_t)e_t^{N^*})}{q_t e_t^{K^*} + (1-q_t)e_t^{N^*}} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}}$$

Note that interiority requires:  $\frac{G(q_t e_t^{K^*} + (1-q_t)e_t^{N^*})}{q_t e_t^{K^*} + (1-q_t)e_t^{N^*}} - c > 0$ .

Note that the level of animals chosen by Pure Kantians,  $e_t^{K^*}$ , does not vary with  $e_t^N$ . Hence the derivative of the RHS wrt to  $e_t^{N^*}$ :

$$\frac{1}{1+\sigma} \left[ \frac{\left( \frac{G(E_t^*)}{E_t^*} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}-1} \frac{(1-\sigma) \left( \frac{G(E_t^*)}{E_t^*} - c \right)^{-\sigma} \frac{e_t^{N^*} (1-q_t) [G'(E_t^*)(E_t^*) - G(E_t^*)]}{(E_t^*)^2}}{\psi}$$

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

Given the interiority condition the first factor and the second factors are positive. Given the concavity of  $G(e_t^N)$ , the last factor is negative. Hence the RHS of the equation is decreasing. Note that when  $\frac{G(q_t e_t^{K^*} + (1-q_t)e_t^{N^*})}{q_t e_t^{K^*} + (1-q_t)e_t^{N^*}} - c = 0$ , the RHS of the equation equals 0. Hence, by continuity, the RHS must intersect the 45° line once for an interior  $e_t^{N^*}$ . ■

Thus there exists necessarily a unique symmetric  $q$  equilibrium in an economy composed of a proportion  $q$  of Pure Kantians and a fraction  $1-q$  of Nash players. The following corollary compares the behaviors of Kantians and Nash players with what prevailed, respectively, at the Kantian equilibrium and at the Nash equilibrium.

**Corollary 2** • *The number of animals chosen by Pure Kantian agents is the same at the Kantian equilibrium and at the  $q$  equilibrium.*

- *The number of animals chosen by Nash agents at the  $q$  equilibrium is strictly larger than at the Nash equilibrium.*

**Proof.** The first part of the proof follows from the definition of Pure Kantians, whose optimal behavior is independent from  $q$ .

Second part: we want to show that  $e_t^{N^*}|_{q=0} < e_t^{N^*}|_{q>0}$ . Let us prove this by reductio ad absurdum. Suppose, by absurdum, that

$$e_t^{N^*}|_{q=0} \geq e_t^{N^*}|_{q>0}$$

In the rest, we denote  $e_t^{N^*}|_{q=0}$  by  $e_t^{N^*0}$  and  $e_t^{N^*}|_{q>0}$  by  $e_t^{N^*}$ .

Let us substitute for  $e_t^{N0}$  and  $e_t^{N*}$  using the FOCs:

$$\left[ \frac{\left( \frac{G(q_t e_t^{K*} + (1-q_t)e_t^{N*0})}{q_t e_t^{K*} + (1-q_t)e_t^{N*0}} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}} \geq \left[ \frac{\left( \frac{G(q_t e_t^{N*} + (1-q_t)e_t^{N*})}{q_t e_t^{N*} + (1-q_t)e_t^{N*}} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}}$$

Given that  $\frac{G(E)}{E}$  is decreasing, with  $e_t^{N*0} \geq e_t^{N*}$ , and  $e_t^{N*} > e_t^{K*}$ , then we have:

$$\left[ \frac{\left( \frac{G(q_t e_t^{K*} + (1-q_t)e_t^{N*0})}{q_t e_t^{K*} + (1-q_t)e_t^{N*0}} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}} < \left[ \frac{\left( \frac{G(q_t e_t^{N*} + (1-q_t)e_t^{N*})}{q_t e_t^{N*} + (1-q_t)e_t^{N*}} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}}$$

a contradiction is reached. Hence we have:

$$e_t^{N*0} < e_t^{N*}$$

■

Thus Pure Kantians behave exactly in the same way in an economy composed only of Pure Kantians, or in an economy where Pure Kantians coexist with a fraction  $1 - q$  of Nash players. Note, however, that this does not imply that they enjoy the same welfare level: since Nash players purchase more animals than Pure Kantians, this pushes the return on animals down for Pure Kantians, who thus enjoy a lower welfare level in comparison to what they enjoyed at the Kantian equilibrium.

Regarding Nash players, these tend, at the  $q$  equilibrium, to purchase more animals than at the Nash equilibrium. The reason is that Nash players' reaction curves are decreasing in the number of animals purchased by others. Hence, in presence of a proportion  $q$  of Pure Kantians, who tend to purchase fewer animals, Nash players react by purchasing *more* animals.

#### 2.4.4 $q$ equilibrium with Impure Kantians

Let us now consider an economy composed of both Nash players and Impure Kantians. This leads us to the definition of a symmetric  $q$  equilibrium with Impure Kantians.

**Definition 4** *A symmetric  $q$  equilibrium with Impure Kantians (perfect information) is a distribution of animal numbers where a proportion  $q_t$  of producers*

has  $e_t^i = e_t^{K**}$  and a proportion  $1 - q_t$  of producers has  $e_t^i = e_t^{N**}$ , such that:

$$e_t^{K**} = \left[ \frac{\frac{e_t^{K**} q_t}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{G(E_t^{**})}{E_t^{**}} - c}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

$$e_t^{N**} = \left[ \frac{\frac{G(E_t^{**})}{E_t^{**}} - c}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

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

In comparison to the  $q$  equilibrium with Pure Kantians, a major difference lies in the fact that the level of animals purchased by Impure Kantians does depend on the behavior of Nash players, unlike what prevailed at the  $q$  equilibrium with Pure Kantians. This additional reaction complicates the study of the existence of a  $q$  equilibrium. Proposition 4 identifies a sufficient condition for the existence of a  $q$  equilibrium in the presence of Impure Kantians.

**Proposition 4** • *In an economy with  $0 < q < 1$ , a sufficient condition for the existence of a symmetric  $q$  equilibrium with Impure Kantians is:*

$$\frac{E_t^{**} G''(E_t^{**})}{\left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right]} > \sigma \quad \forall e_t^{K**} \leq e_t^{N**}$$

- *In an economy with  $0 < q < 1$  and quasi-linear preferences ( $\sigma = 0$ ), the symmetric  $q$  equilibrium with Impure Kantians necessarily exists and is unique.*

**Proof.** See the Appendix. ■

In order to have an idea of the strength of that condition, let us take, here again, the case of the Cobb-Douglas production function. The condition then becomes:

$$\alpha > \sigma$$

In the case of quasi-linear preferences ( $\sigma = 0$ ), that condition is trivially satisfied. Thus, under a Cobb-Douglas production function and quasi-linear preferences, we know for sure that there must exist a symmetric  $q$  equilibrium with Impure Kantians.

## 2.5 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. Assuming that the utility functions are additive in the utility of economic activity and in the utility of socialization, those two activities can be treated separately.

We assume that socialization takes place according to a socialization process *à la* Bisin and Verdier (2001) where vertical and oblique socializations interact. Children are born without type. Each parent has one child. Parents can invest in a purely physical socialization effort  $\tau_t^i \in [0, 1]$ .

The disutility of socialization is given by:

$$\frac{\kappa (\tau_t^i)^2}{2} \quad (20)$$

The socialization takes place in two stages. First, the child is subject to the vertical socialization process through his parent. With a probability  $\tau_t^i$ , a child of a parent of type  $i$  will take directly the trait of his parent. With a probability  $1 - \tau_t^i$ , the direct vertical socialization does not take place, and the child takes the trait of an adult drawn randomly in the cohort of his parent (i.e. the role model).

Let us denote by  $P_t^{ij}$  the probability that a child of a parent of type  $i \in \{K, N\}$  takes type  $j \in \{K, N\}$ . We have:

$$\begin{aligned} P_t^{KK} &= \tau_t^K + (1 - \tau_t^K) q_t \\ P_t^{KN} &= (1 - \tau_t^K) (1 - q_t) \\ P_t^{NK} &= (1 - \tau_t^N) q_t \\ P_t^{NN} &= \tau_t^N + (1 - \tau_t^N) (1 - q_t) \end{aligned}$$

The gains from socialization are due to parental imperfect empathy.

For the parent of type  $i \in \{K, N\}$ , let us define  $V^{ij}$ , which is the gain to have a child of type  $j$ :

For an economy with Pure Kantians, we have:

$$V^{KK} = \frac{(G(e_t^{K*}) - ce_t^{K*})^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K*})^2}{2} \quad (21)$$

$$V^{KN} = \frac{(G(e_t^{N*}) - ce_t^{N*})^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N*})^2}{2} \quad (22)$$

$$V^{NK} = \frac{\left(\frac{e_t^{K*}}{E_t^*} G(E_t^*) - ce_t^{K*}\right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K*})^2}{2} \quad (23)$$

$$V^{NN} = \frac{\left(\frac{e_t^{N*}}{E_t^*} G(E_t^*) - ce_t^{N*}\right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N*})^2}{2} \quad (24)$$

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

For an economy with Impure Kantians, we have:

$$V^{KK} = \frac{\left( \frac{e_t^{K**} G(q_t e_t^{K**} + (1-q_t) e_t^{N**})}{q_t e_t^{K**} + (1-q_t) e_t^{N**}} - c e_t^{K**} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K**})^2}{2} \quad (25)$$

$$V^{KN} = \frac{\left( \frac{e_t^{N**} G(q_t e_t^{N**} + (1-q_t) e_t^{N**})}{q_t e_t^{N**} + (1-q_t) e_t^{N**}} - c e_t^{N**} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N**})^2}{2} \quad (26)$$

$$V^{NK} = \frac{\left( \frac{e_t^{K**} G(q_t e_t^{K**} + (1-q_t) e_t^{N**})}{q_t e_t^{K**} + (1-q_t) e_t^{N**}} - c e_t^{K**} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K**})^2}{2} \quad (27)$$

$$V^{NN} = \frac{\left( \frac{e_t^{N**} G(q_t e_t^{K**} + (1-q_t) e_t^{N**})}{q_t e_t^{K**} + (1-q_t) e_t^{N**}} - c e_t^{N**} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N**})^2}{2} \quad (28)$$

In the present context, imperfect empathy does not consist in parent's tendency to evaluate outcomes in the light of their own preferences, but, instead, in parents's tendency to perfectly anticipate the optimal choices of their children in the light of their own decision rule (i.e. of the parents themselves). This explains why  $V^{KN}$  includes the term  $(q_t e_t^{N*} + (1-q_t) e_t^{N*})$  in  $E_t$ : this comes from the fact that type- $K$  Impure Kantians parents evaluate the outcome of their child while considering the impact of their child's decisions on all other Kantians ("if I were choosing the same level of animals as my child does, which utility level it would bring me").

### 2.5.1 Optimal socialization efforts with Pure Kantians

In the case of pure Kantians and Nash players, the optimal socialization efforts can be obtained by considering the following problems.

Given the separability of the utility function in the "consumption component" and the "imperfect altruistic component", a parent of type  $N$  chooses  $\tau_t^N$  such that:

$$\begin{aligned} \max_{\tau_t^N} & -\frac{\kappa (\tau_t^N)^2}{2} + (\tau_t^N + (1-\tau_t^N)(1-q_t)) \left[ \frac{\left( \frac{e_t^{N*} G(E_t^*)}{E_t^*} - c e_t^{N*} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N*})^2}{2} \right] \\ & + (1-\tau_t^N) q_t \left[ \frac{\left( \frac{e_t^{K*} G(E_t^*)}{E_t^*} - c e_t^{K*} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K*})^2}{2} \right] \end{aligned}$$

The FOC is:

$$\tau_t^{N*} = \frac{q_t}{\kappa} \left[ \begin{array}{c} \left[ \frac{\left( \frac{e_t^{N*} G(E_t^*)}{E_t^*} - ce_t^{N*} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N*})^2}{2} \right] \\ - \left[ \frac{\left( \frac{e_t^{K*} G(E_t^*)}{E_t^*} - ce_t^{K*} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K*})^2}{2} \right] \end{array} \right] \quad (29)$$

A parent of type  $K$  chooses  $\tau_t^K$  such that:

$$\begin{aligned} \max_{\tau_t^K} & -\frac{\kappa (\tau_t^K)^2}{2} + (\tau_t^K + (1 - \tau_t^K) q_t) \left[ \frac{(G(e_t^{K*}) - ce_t^{K*})^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K*})^2}{2} \right] \\ & + ((1 - \tau_t^K)(1 - q_t)) \left[ \frac{(G(e_t^{N*}) - ce_t^{N*})^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N*})^2}{2} \right] \end{aligned}$$

The FOC is:

$$\tau_t^{K*} = \frac{(1 - q_t)}{\kappa} \left[ \frac{(G(e_t^{K*}) - ce_t^{K*})^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K*})^2}{2} - \left[ \frac{(G(e_t^{N*}) - ce_t^{N*})^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N*})^2}{2} \right] \right] \quad (30)$$

The equation describing the dynamics of  $q_t$  is given by:

$$q_{t+1} = q_t + q_t(1 - q_t) \left[ \begin{array}{c} \frac{(1-q_t)}{\kappa} \left[ \frac{(G(e_t^{K*}) - ce_t^{K*})^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K*})^2}{2} \right] \\ - \left[ \frac{(G(e_t^{N*}) - ce_t^{N*})^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N*})^2}{2} \right] \\ - \frac{q_t}{\kappa} \left[ \frac{\left( \frac{e_t^{N*} G(E_t^*)}{E_t^*} - ce_t^{N*} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N*})^2}{2} \right] \\ - \left[ \frac{\left( \frac{e_t^{K*} G(E_t^*)}{E_t^*} - ce_t^{K*} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K*})^2}{2} \right] \end{array} \right] \quad (31)$$

### 2.5.2 Optimal socialization efforts with Impure Kantians

Consider now the situation where Nash players coexist with Impure Kantians. Socialization efforts are now given as follows.

Nash parents choose a socialization effort  $\tau_t^N$  such that:

$$\begin{aligned} \max_{\tau_t^N} & -\frac{\kappa (\tau_t^N)^2}{2} + (\tau_t^N + (1 - \tau_t^N)(1 - q_t)) \left[ \frac{\left( \frac{e_t^{N**} G(E_t^{**})}{E_t^{**}} - ce_t^{N**} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N**})^2}{2} \right] \\ & + (1 - \tau_t^N) q_t \left[ \frac{\left( \frac{e_t^{K*} G(E_t^{**})}{E_t^{**}} - ce_t^{K**} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K**})^2}{2} \right] \end{aligned}$$

The FOC is:

$$\tau_t^{N**} = \frac{q_t}{\kappa} \left[ \begin{array}{c} \left( \frac{e_t^{N**} G(E_t^{**}) - c e_t^{N**}}{E_t^{**}} \right)^{1-\sigma} - \frac{\psi(e_t^{N**})^2}{2} \\ - \left[ \frac{\left( \frac{e_t^{K**} G(E_t^{**}) - c e_t^{K**}}{E_t^{**}} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K**})^2}{2} \right] \end{array} \right] \quad (32)$$

A parent of type  $K$  chooses  $\tau_t^K$  such that:

$$\begin{aligned} \max_{\tau_t^K} & -\frac{\kappa (\tau_t^K)^2}{2} + (\tau_t^K + (1 - \tau_t^K) q_t) \left[ \frac{\left( \frac{e_t^{K**} G(E_t^{**}) - c e_t^{K**}}{E_t^{**}} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K**})^2}{2} \right] \\ & + ((1 - \tau_t^K) (1 - q_t)) \left[ \frac{\left( \frac{e_t^{N**} G(q_t e_t^{N**} + (1 - q_t) e_t^{N**}) - c e_t^{N**}}{q_t e_t^{N**} + (1 - q_t) e_t^{N**}} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N**})^2}{2} \right] \end{aligned}$$

The FOC is:

$$\tau_t^{K**} = \frac{(1 - q_t)}{\kappa} \left[ \begin{array}{c} \left( \frac{e_t^{K**} G(E_t^{**}) - c e_t^{K**}}{E_t^{**}} \right)^{1-\sigma} - \frac{\psi(e_t^{K**})^2}{2} \\ - \left[ \frac{\left( \frac{e_t^{N**} G(e_t^{N**}) - c e_t^{N**}}{e_t^{N**}} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N**})^2}{2} \right] \end{array} \right] \quad (33)$$

The equation describing the dynamics of  $q_t$  is given by:

$$q_{t+1} = q_t + q_t(1 - q_t) \left[ \begin{array}{c} \frac{(1 - q_t)}{\kappa} \left[ \begin{array}{c} \left( \frac{e_t^{K**} G(E_t^{**}) - c e_t^{K**}}{E_t^{**}} \right)^{1-\sigma} - \frac{\psi(e_t^{K**})^2}{2} \\ - \left[ \frac{\left( \frac{e_t^{N**} G(e_t^{N**}) - c e_t^{N**}}{e_t^{N**}} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{N**})^2}{2} \right] \end{array} \right] \\ - \frac{q_t}{\kappa} \left[ \begin{array}{c} \left( \frac{e_t^{N**} G(E_t^{**}) - c e_t^{N**}}{E_t^{**}} \right)^{1-\sigma} - \frac{\psi(e_t^{N**})^2}{2} \\ - \left[ \frac{\left( \frac{e_t^{K**} G(E_t^{**}) - c e_t^{K**}}{E_t^{**}} \right)^{1-\sigma}}{1-\sigma} - \frac{\psi(e_t^{K**})^2}{2} \right] \end{array} \right] \end{array} \right] \quad (34)$$

### 3 Long run dynamics

The previous section aimed at studying the existence and uniqueness of different kinds of equilibria under a given partition of the population into groups of Nash

players and Kantian players (either pure or impure). This section studies the dynamics of heterogeneity.

For that purpose, we will, for the sake of simplification, assume that the preference parameter  $\sigma$  equals 0, that is, that individuals have quasi-linear preferences.

As above, we will consider first the long-run dynamics of an economy with Nash players and Pure Kantians, and, then, of an economy with Nash players and Impure Kantians.

### 3.1 Nash players and Pure Kantians

Under  $\sigma = 0$ , the dynamics of  $q_t$  is given by:

$$q_{t+1} \equiv F(q_t) = q_t + q_t(1 - q_t) \begin{bmatrix} \frac{(1-q_t)}{\kappa} \begin{bmatrix} G(e_t^{K*}) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} \\ - \left[ G(e_t^{N*}) - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} \right] \end{bmatrix} \\ -\frac{q_t}{\kappa} \begin{bmatrix} \frac{e_t^{N*}G(E_t^*)}{E_t^*} - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} \\ - \left[ \frac{e_t^{K*}G(E_t^*)}{E_t^*} - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} \right] \end{bmatrix} \end{bmatrix} \quad (35)$$

Let us study the existence, the uniqueness and the stability of a stationary equilibrium. As stated in Proposition 5, there exist, in that economy, at least three stationary equilibria. Two of those stationary equilibria exhibit a homogeneous population, composed either of Pure Kantians only, or of Nash players only. In addition, there exists at least one intermediate stationary equilibrium, i.e. with a population composed of both Pure Kantians and Nash players.

**Proposition 5** *Consider an economy composed of Pure Kantians and Nash players. There exists at least three stationary equilibria:*

$$\begin{aligned} q^* &= 0 \\ 0 &< q^* < 1 \\ q^* &= 1 \end{aligned}$$

**Proof.** See the Appendix. ■

Whereas Proposition 5 only casts some light on the existence of some stationary equilibria with heterogeneous populations, Proposition 6 proposes a condition that is necessary and sufficient for the uniqueness of a stationary equilibria with heterogeneous populations.

**Proposition 6** *A necessary and sufficient condition for the uniqueness of the*



intermediate stationary equilibrium  $0 < q^* < 1$  is:

$$\begin{aligned}
& \underbrace{\left[ \frac{-1}{\kappa q^*} \left[ G(e^{K^*}) - ce^{K^*} - \frac{\psi(e^{K^*})^2}{2} - \left[ G(e^{N^*}) - ce^{N^*} - \frac{\psi(e^{N^*})^2}{2} \right] \right] \right]}_{-} \\
& + \underbrace{\frac{1-q^*}{\kappa} \left[ -\frac{de^{N^*}}{dq} [G'(e^{N^*}) - c - \psi(e^{N^*})] \right]}_{+} \\
& - \underbrace{\frac{q^*}{\kappa} \left[ -[e^{N^*} - e^{K^*}]^2 \frac{G'(E^*) - \frac{G(E^*)}{E^*}}{E^*} \right]}_{-} \\
& - \underbrace{\frac{q^*}{\kappa} \frac{de^{N^*}}{dq} \left[ (e^{N^*} - e^{K^*})(1-q^*) \frac{G'(E^*) - \frac{G(E^*)}{E^*}}{E^*} \right]}_{+} \\
& < 0
\end{aligned}$$

for all  $q^*$ . When the intermediate equilibrium is not unique, there exists an odd number of intermediate stationary equilibria.

**Proof.** See the Appendix. ■

The condition stated in Proposition 6 is based on the following intuition. We know that a necessary and sufficient condition for the uniqueness of an intermediate stationary equilibrium is that the derivative of the term in brackets in the dynamic equation for  $q_t$  (which is equal to the differential in socialization efforts  $\tau_t^K - \tau_t^N$ ) is strictly negative at any intermediate stationary equilibrium.<sup>6</sup> Using that intuition, and computing the derivative of that term at the intermediate stationary equilibrium, we can decompose the impact of a variation in  $q_t$  on the differential in socialization efforts  $\tau_t^K - \tau_t^N$  into four terms, which are presented on the LHS of the condition in Proposition 6.

The first term, which is negative, captures the effect of cultural substitution on the socialization effort of both players. A rise in  $q$  reduces the socialization efforts of Pure Kantian parents and it raises the socialization efforts of Nash parents. The second term, which is positive, captures the impact of a variation of  $q$  on the relative utility gain to have a Pure Kantian child for Pure Kantians  $\Delta V^K = V^{KK} - V^{KN}$ . The third and fourth terms capture the impact of  $q$  on the relative utility gain of having a Nash child for Nash parents, i.e.  $\Delta V^N = V^{NN} - V^{NK}$ . The third term captures the effect, on  $\Delta V^N$ , of a rise in  $q$  for a *given* number of animals purchased by Nash players ( $e^{N^*}$  being given). Nash parents gain more than Pure Kantian parents when  $q$  is raised for a given  $e^{N^*}$ ,

<sup>6</sup>Necessity: if the slope of the transition function at an intermediate stationary equilibrium were positive, then we would be sure that there would exist other intermediate stationary equilibria. Sufficiency: if the slope of the transition function at any intermediate equilibrium was a negative, then clearly that intermediate equilibrium must be unique.

so that we have  $\Delta V^N > 0$ .<sup>7</sup> As a consequence, the third effect is of negative sign. On the contrary, the fourth term is positive. It captures the effect of adjusting  $e^{N*}$  on  $\Delta V^N$ . When Nash players increase the number of animals, both Pure Kantians and Nash players have a decrease in utility, but the latter lose more, implying  $\Delta V^N < 0$ .<sup>8</sup>

When interpreting that condition for uniqueness, it is important to stress the crucial role played by the derivative  $\frac{de^{N*}}{dq}$ , which reflects the extent to which Nash players adapt the number of animals purchased to the proportion of Pure Kantians in the population. That derivative is equal to:

$$\frac{de^{N*}}{dq} = \frac{-[e^{N*} - e^{K*}] \frac{G'(E^*) - \frac{G(E^*)}{E^*}}{\psi E^*}}{1 - (1 - q^*) \frac{G'(E^*) - \frac{G(E^*)}{E^*}}{\psi E^*}} > 0$$

That derivative is positive, since a rise in the proportion of Kantians in the population induces Nash players to purchase more animals (because Kantians purchase fewer animals than Nash players, so that a rise in  $q$  raises the return on animals for Nash players). The two positive terms on the RHS of the condition depend, unlike the negative terms, on the level of  $\frac{de^{N*}}{dq}$ . The uniqueness of an intermediate stationary equilibrium (i.e. with heterogeneous population) is more likely to arise when the reactivity of Nash shepherds is low. However, when  $\frac{de^{N*}}{dq}$  is large, the multiplicity of intermediate equilibria can arise. In other words, a large reactivity of Nash players can generate a bifurcation.<sup>9</sup>

Finally, let us examine the stability of the stationary equilibria. Proposition 7 states that homogeneous stationary equilibria are unstable, and states the necessary and sufficient condition for the (local) stability of intermediate equilibria. Note that, when the intermediate stationary equilibrium is unique, it must necessarily be locally stable.

**Proposition 7** *Consider an economy composed of Pure Kantians and Nash players.*

- *Stationary equilibria  $q^* = 0$  and  $q^* = 1$  are unstable.*
- *A necessary and sufficient condition for the local stability of an interme-*

<sup>7</sup>Indeed, a higher proportion of Pure Kantians implies more space per animal for all, but as Nash players have more animals, this is more beneficial for them.

<sup>8</sup>The reason is that the rise in  $e^{N*}$  reduces the availability of space, and thus reduces the return on each animal. But as Nash players have more animals than Pure Kantians, they suffer from a larger utility loss.

<sup>9</sup>From that perspective, a key parameter is the elasticity of output to the number of animals.

diating stationary equilibrium  $0 < q^* < 1$  is:

$$\left| 1 + q^*(1 - q^*) \left[ \begin{array}{l} \left[ \frac{-1}{\kappa q^*} \right] \left[ G(e^{K^*}) - ce^{K^*} - \frac{\psi(e^{K^*})^2}{2} - \left[ G(e^{N^*}) - ce^{N^*} - \frac{\psi(e^{N^*})^2}{2} \right] \right] \right. \\ \left. + \frac{(1-q^*)}{\kappa} \left[ -\frac{de^{N^*}}{dq} \left[ G'(e^{N^*}) - c - \psi(e^{N^*}) \right] \right] \right. \\ \left. - \frac{q^*}{\kappa} \left[ - \left[ e^{N^*} - e^{K^*} \right]^2 \frac{G'(E^*) - \frac{G(E^*)}{E^*}}{E^*} \right] \right. \\ \left. - \frac{q^*}{\kappa} \frac{de^{N^*}}{dq} \left[ (e^{N^*} - e^{K^*}) (1 - q^*) \frac{G'(E^*) - \frac{G(E^*)}{E^*}}{E^*} \right] \right] \right| < 1 \end{array} \right.$$

- When the intermediate stationary equilibrium is unique, this is also locally stable. Given the continuity of  $F(q_t)$ , the unique intermediate equilibrium is then also globally stable on  $[0, 1]$ .
- When the intermediate stationary equilibrium is not unique, we necessarily have an odd number of intermediate equilibria. Those for which the above condition is satisfied are locally stable, and the others are unstable.

**Proof.** See the Appendix. ■

In sum, this section shows that, provided the reactivity of Nash players  $\frac{de^{N^*}}{dq}$  is sufficiently small, so that the intermediate stationary equilibrium is unique, we know that any economy starting with some proportion of Pure Kantians will still exhibit, in the long-run, a constant positive proportion of Pure Kantians. On the contrary, when the reactivity of Nash players is large, so that there exists a multiplicity of intermediate equilibria, stability may not prevail.

### 3.2 Nash players and Impure Kantians

Assuming, here again, quasi-linear preferences, the equation for the dynamics of  $q_t$  is now:

$$q_{t+1} \equiv H(q_t) = q_t + q_t(1 - q_t) \left[ \begin{array}{l} \frac{(1-q_t)}{\kappa} \left[ \begin{array}{l} \frac{e_t^{K^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{K^{**}} - \frac{\psi(e_t^{K^{**}})^2}{2} \\ - \left[ \frac{e_t^{N^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} \right] \end{array} \right] \\ - \frac{q_t}{\kappa} \left[ \begin{array}{l} \frac{e_t^{N^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} \\ - \left[ \frac{e_t^{K^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{K^{**}} - \frac{\psi(e_t^{K^{**}})^2}{2} \right] \end{array} \right] \end{array} \right] \quad (36)$$

Let us first study the existence of a stationary equilibrium.

Obviously,  $q = 0$  and  $q = 1$  are stationary equilibria. Proposition 8 summarizes our results for existence.

**Proposition 8** Consider an economy composed of Impure Kantians and Nash

players. There exists at least three stationary equilibria:

$$\begin{aligned} q^{**} &= 0 \\ 0 &< q^{**} < 1 \\ q^{**} &= 1 \end{aligned}$$

**Proof.** See the Appendix. ■

Whereas Proposition 8 states that a heterogeneous stationary equilibrium exists, it does not guarantee that this is unique. Proposition 9 provides a condition that is necessary and sufficient for the uniqueness of the intermediate equilibrium.<sup>10</sup>

**Proposition 9** *A necessary and sufficient condition for the uniqueness of the intermediate stationary equilibrium  $0 < q^{**} < 1$  is:*

$$\begin{aligned} &\underbrace{\frac{-1}{\kappa} [V^{KK} - V^{KN} + (V^{NN} - V^{NK})]}_{-} \\ &+ \frac{(1 - q^{**})}{\kappa} \left[ \underbrace{\frac{e^{K^{**}}}{E^{**}} \left( G'(E^{**}) - \frac{G(E^{**})}{E^{**}} \right)}_{+} \right. \\ &\quad \left. + \frac{de^{N^{**}}}{dq} \left( \underbrace{\frac{(1 - q^{**}) e^{K^{**}}}{E^{**}} \left( G'(E^{**}) - \frac{G(E^{**})}{E^{**}} \right)}_{+} \right. \right. \\ &\quad \left. \left. - \underbrace{[G'(e^{N^{**}}) - c - \psi e^{N^{**}}]}_{-} \right) \right] \\ &- \frac{q^{**}}{\kappa} \left[ \underbrace{- [e^{K^{**}} - e^{N^{**}}]^2 \left( G'(E^{**}) - \frac{G(E^{**})}{E^{**}} \right)}_{+} \right. \\ &\quad \left. + \frac{de^{N^{**}}}{dq} \left[ \underbrace{- \frac{1}{E^{**}} (e^{K^{**}} - e^{N^{**}}) (1 - q^{**}) \left( G'(E^{**}) - \frac{G(E^{**})}{E^{**}} \right)}_{-} \right] \right. \\ &\quad \left. + \frac{de^{K^{**}}}{dq} \left[ \underbrace{q^{**} \frac{e^{N^{**}}}{E^{**}} \left( G'(E^{**}) - \frac{G(E^{**})}{E^{**}} \right)}_{-} \right] \right] \\ &< 0 \end{aligned}$$

for all  $q^{**}$ . When the intermediate equilibrium is not unique, there exists an odd number of intermediate stationary equilibria.

**Proof.** See the Appendix. ■

<sup>10</sup>That condition states that the derivative  $\frac{d[\tau_t^K - \tau_t^N]}{dq_t}$  is strictly negative at any intermediate stationary equilibrium.

In comparison with the condition for uniqueness of the intermediate equilibrium in the case of Pure Kantians, the condition stated in Proposition 9 is more complex, since it involves the reactions of Impure Kantians in terms of animal purchase, i.e.  $\frac{de^{K^{**}}}{dq}$ . On the contrary, in the case of Pure Kantians, there was not such a reaction, since Pure Kantians are always purchasing the same number of animals, independently from the composition and behavior of the rest of the population. This is clearly not the case for Impure Kantians.

Let us now interpret the condition given in Proposition 9. The first term on the RHS of the condition in Proposition 9, which is negative, captures the effect of cultural substitution. A rise in  $q$  reduces the socialization efforts of Impure Kantian parents, since those parents rely more on oblique socialization (i.e. the role models) when they are more numerous in the population. On the contrary, a rise in  $q$  raises the socialization efforts of Nash parents, who can rely less on oblique socialization.

The first term of the first bracket is positive: it is the effect of a change in  $q$  on the utility for an Impure Kantian parent to have an Impure Kantian child for a given number of animals for type  $N$ . The second term in the first bracket is the effect, for a given  $q$ , of a change in the number of animals purchased by Nash players on the utility for an Impure Kantian parent to have an Impure Kantian child.

Let us now interpret the second bracket. The first term in the second bracket is the impact, for a given number of animals for both types, of a change in  $q$  on the utility for a Nash parent to have a Nash child. The second term in the second bracket is the impact, for a given  $q$ , of a change in the number of animals purchased by Nash players on the utility for a Nash parent to have a Nash child. The third term in the second brackets is the impact, for a given  $q$ , of a change in the number of animals purchased by Impure Kantians on the utility for a Nash parent to have a Nash child.

Proposition 10 summarizes our results regarding the stability of the three equilibria.

**Proposition 10** *Consider an economy composed of Impure Kantians and Nash players.*

- *The stationary equilibrium  $q^{**} = 0$  is non hyperbolic and is unstable.*
- *The stationary equilibrium  $q^{**} = 1$  is hyperbolic and is unstable.*
- *A necessary and sufficient condition for the local stability of an intermediate stationary equilibrium  $0 < q^{**} < 1$  is that*

$$\frac{d[\tau_t^{K^{**}} - \tau_t^{N^{**}}]}{dq_t} < 0$$

*holds at the intermediate stationary equilibrium.*

- *When the intermediate stationary equilibrium is unique, this is also locally stable. Given the continuity of  $H(q_t)$ , the unique intermediate equilibrium is then also globally stable on  $[0, 1]$ .*

- *When the intermediate stationary equilibrium is not unique, we necessarily have an odd number of intermediate equilibria. Those for which the above condition is satisfied are locally stable, and the others are unstable.*

**Proof.** See the Appendix. ■

Proposition 10 shows that homogeneous stationary equilibria are not stable. Thus, starting from an economy composed of both Nash players and Impure Kantians, it is not possible that only one type of rationality survives in the long-run. In other terms, it is necessarily the case that both types of rationality will survive in the long-run. Proposition 10 also states that, when the intermediate stationary equilibrium is unique (which is not always guaranteed), it is also locally stable. Thus in the case of uniqueness of the intermediate equilibrium, proportions of the two types of players will stabilize in the long-run.

In comparison with the economy composed of Nash players and Pure Kantians, the main difference lies in the fact, mentioned above, that the uniqueness condition for the intermediate equilibrium is more complex, since it includes additional terms depending on the reaction of Impure Kantians to changes in the composition of the population, i.e.  $\frac{de^{K**}}{dq}$ , something which was absent in the case of Pure Kantians.

## 4 Policy (1): quotas

Up to now, we considered an economy at the laissez-faire, without public intervention. However, it is well known that 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:

$$e^{K*} \leq \bar{e} < \lim_{q_t \rightarrow 0} e_t^{N*} \quad (37)$$

The lower bound of the quota,  $e^{K*}$ , coincides with the social optimum. Because of political constraints, it is likely that the quota will exceed that level. The second part of the inequality requires that Nash players are always strictly constrained by the quota. Indeed, we know that  $e_t^{N*}$  is increasing in  $q_t$ , so that it takes its minimal value when there are no Kantian players.

Let us now consider the impact of imposing the quota  $\bar{e}$  on social outcomes. For that purpose, we now suppose that  $c = 0$ . Furthermore, we will also impose the following cost of shepherd effort:

$$\psi \frac{2}{3} (e_t^i)^{3/2} \quad (38)$$

Moreover, we will now suppose that the production function  $F(e_t^i, s_t^i)$  has the following Cobb-Douglas form:

$$y_t^i = A (e_t^i)^{1/2} (s_t^i)^{1/2} \quad (39)$$

where  $s_t^i \equiv \frac{e_t^i}{E_t}$  is the amount of land used by the producer  $i$ . Hence, the total product per shepherd becomes:

$$y_t^i = \frac{e_t^i}{E_t} A E_t^{1/2} = \frac{e_t^i}{E_t} G(E_t) \quad (40)$$

Thus this modeling is a special case of the one in the previous section, where  $G(E_t) = A E_t^{1/2}$ .

#### 4.1 Pure Kantians vs Nash players

Does the imposition of the quota  $\bar{e}$  on the number of sheep per shepherd contribute to reduce the total number of animals, and, hence, allow the society to reach a Pareto improvement? Proposition 11 shows that this is not necessarily the case. It may indeed be the case, under some conditions, that introducing a quota makes the Tragedy of the Commons even worse than at the *laissez-faire*.

**Proposition 11** *Let us consider an economy with a quota  $\bar{e} \in [e^{K*}, e^{N*}[$ .*

- *There exists some intervals for  $\bar{e}$  such that:*

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

- *There exists some intervals for  $\bar{e}$  such that:*

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

**Proof.** See in the Appendix. ■

Proposition 11 is quite counterintuitive. One would expect that the introduction of a quota would necessarily reduce the extent of land congestion. The intuition behind Proposition 11 goes as follows. True, in the case where the partition of the population into Nash players and Pure Kantian players were constant, imposing a quota would, by reducing the number of sheep of Nash players, reduce the total number of animals with respect to the *laissez-faire*, and, hence, decrease the extent of land congestion.

However, we consider here an economy where the partition of the population into the two types of shepherds is varying over time. Proposition 11 states that a high quota increases the proportion of Nash players. The rational goes as follows. The quota affects socialization gains of each type of parents. On the one hand, it reduces the gain to have a Pure Kantian child for Pure Kantian parents by decreasing the loss to have a child who plays Nash (as the latter will be closer to the optimum for Pure Kantians). On the other hand, when it is

high, the quota always increases the gain to have a Nash child for Nash parents. Indeed, for a given  $q$ , the quotas reduces the total number of animals which 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.

Furthermore, Proposition 11 states that some level of quota increases the total number of animals purchased. The quota has two opposite effects. First it reduces the total number of animals purchased by reducing the number of animals purchased by Pure Kantians. This is the standard *intensive margin* effect. Second, the quota has also an effect on the proportions of the two types of agents, that is, an *extensive margin* effect. Actually, the quota increases the total number of animals purchased by increasing the proportion of Nash players. When the quota is high the second effect outweighs the first one. The increase in the proportion of Nash players is large when the quota has a high negative impact on socialization gains for Pure Kantian parents. The increase in the proportion of Nash players is low when the cultural substitution effect is large. Indeed, through this effect, any increase in the proportion of Nash players tends to reduce the future proportion of this type of agents. Due to the concavity of utility functions, the cultural substitution effect is low compared to the change in socialization gain so that the increase in the proportion of Nash players is substantial. Hence, the total number of animals purchased increases.

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). It is straightforward to see that the relationship between the total number of animals and the quota is not monotonous. 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. But the total number of animals associated with the quota is increasing with the quota  $\bar{e}$ , up to  $\bar{e} = 0.31$ . When we consider even more restrictive quotas, then the total number of animals falls. The extreme point on the left coincides with the social optimum, where Nash players purchase the same number of animals as Pure Kantians.

Actually, if the government wants to reduce the total number of animals by imposing a quota, Figure 1 shows that quotas that restrict the number of animals purchased by Nash players by less than 12 % 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.

## 4.2 Impure Kantians vs Nash players

Let us now consider the introduction of a quota on the number of animals in the case of an economy composed Nash players and Impure Kantians. The problem



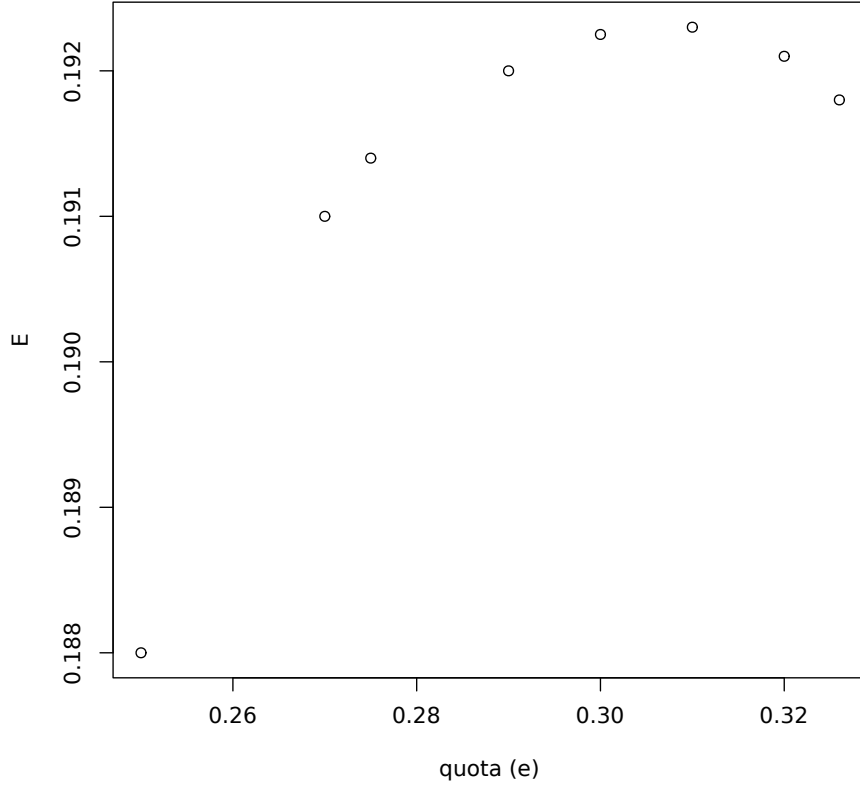


Figure 1: Total number of animals as a function of the quota ( $A = 1/2$ ,  $\psi = 12$ ,  $\kappa = 1/2$ ).

of the Impure Kantian is, under  $\sigma = c = 0$ :

$$\begin{aligned} & \max_{e_t^K} \left( \frac{e_t^K}{E_t} A E_t^{1/2} \right) - \psi \frac{2}{3} (e_t^K)^{3/2} \\ & \text{s.t. } E_t \equiv q_t e_t^K + (1 - q_t) \bar{e} \end{aligned}$$

or alternatively,

$$\max_{e_t^K} \left( e_t^K A (q_t e_t^K + (1 - q_t) \bar{e})^{-1/2} \right) - \psi \frac{2}{3} (e_t^K)^{3/2}$$

The optimal  $e_t^{K**}$  satisfies:

$$A (q_t e_t^{K**} + (1 - q_t) \bar{e})^{-1/2} + \frac{-1}{2} q_t e_t^{K**} A (q_t e_t^{K**} + (1 - q_t) \bar{e})^{-3/2} - \psi (e_t^{K**})^{1/2} = 0$$

Two cases can arise, depending on whether  $e^{K**} \geq \bar{e}$ .

When  $e^{K**} \geq \bar{e}$ , it is necessarily the case that the introduction of a quota reduces the amount of animals. Indeed, we have:

$$q^{LF} e^{K**} + (1 - q^{LF}) e^{N**} > q^Q \bar{e} + (1 - q^Q) \bar{e} = \bar{e}$$

since both  $e^{K**} \geq \bar{e}$  and  $e^{N**} > \bar{e}$ .

Let us now focus on the second case, where  $e^{K**} < \bar{e}$ .

**Proposition 12** *Consider an economy with a quota  $\bar{e} \in [e^{K*}, e^{N**}]$ . Suppose that  $e^{N**} > 6e^{K**}$*

- We have:

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

- There exists some intervals for  $\bar{e}$  such that:

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

**Proof.** See in the Appendix. ■

Again for the case of Impure Kantians, a quota may have a counter-intuitive effect by increasing the total number of animals purchased. As opposed to the case of Pure Kantians, this is not true for all parameters combination. The rationale is as follows. Unlike Pure Kantians, Impure Kantians react to the introduction of a quota by increasing the number of animals. This reaction has two opposite effects. First, it directly increases the total number of animal purchased. However, it also indirectly affects the total number of animal by impacting socialization gains (and thus the proportion of Nash players). In particular, it reduces the gain for Nash parents to have Nash children (letting socialization gains for Impure Kantian parents unchanged). This second effect negatively affects the total number of animal by rising the proportion of Impure Kantians in the economy. When  $e^{N**} > 6e^{K**}$ , implying that the socialization effort of Nash parents is high, then the first effect dominates and the overall impact of the quota is an increase in the total number of animals.

Let us illustrate this counterintuitive case numerically. For that purpose, Figure 2 shows, as the previous figure, the total number of animals as a function of the quota  $\bar{e}$ . Here again, the relationship between the total number of animals and the quota is not monotonous. The extreme point on the right is the total number of animals when the quota is not constraining at all, neither for Impure Kantians, nor for Nash players. Then, when we consider thus lower values for the quota, the quota becomes constraining for Nash players (and not for Impure Kantians). But the total number of animals associated with the quota is increasing with the quota  $\bar{e}$ , up to  $\bar{e} = 0.24$ . When we consider even more restrictive quotas, then the total number of animals falls. The extreme point on the left coincides with the case where  $\bar{e}$  is so low that both Nash players and Impure Kantians are constrained by the quota.

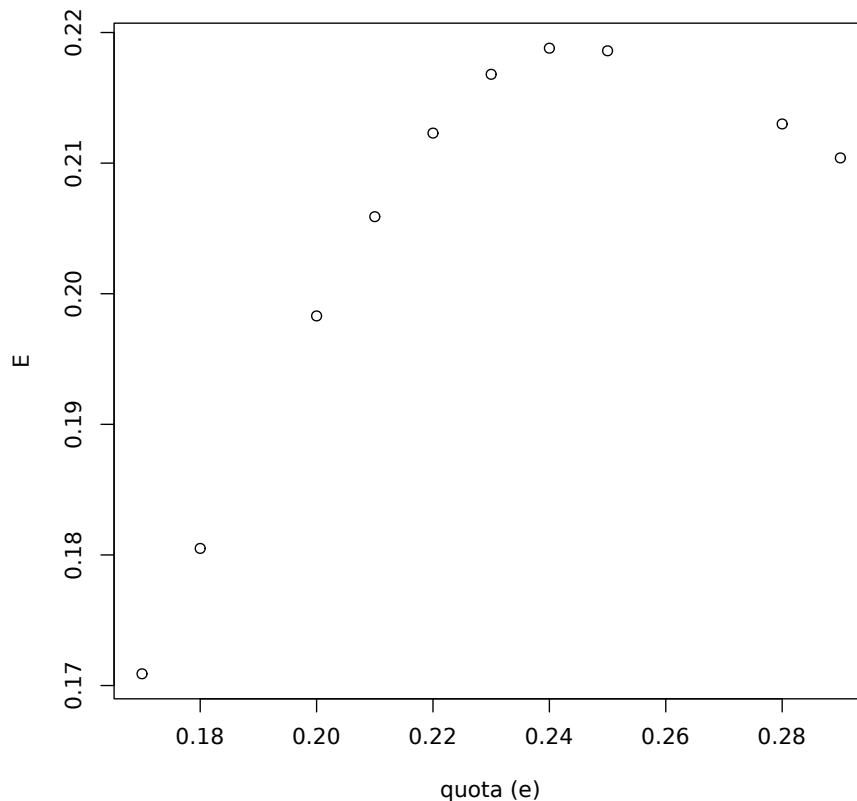


Figure 2: Total number of animals as a function of the quota ( $A = 1/2$ ,  $\psi = 12$ ,  $\kappa = 1/2$ ).

This figure shows that, from a qualitative perspective, the main result obtained when considering Pure Kantians also hold when considering Impure Kantians: in both cases, a soft quota contributes to raise the total number of animals. Note, however, that, from a quantitative perspective, some important differences arise.

## 5 Policy (2): Pigouvian tax

When facing the Tragedy of the Commons, another standard policy tool consists in introducing a Pigouvian tax on the number of animals. The underlying intuition goes as follows. In our economy, Nash players (and, to some extent, Impure Kantians) do not, when choosing the number of animals, internalize

the impact of the number of animals on the society. The negative production externality takes the following form: a higher number of animals tends, by reducing the available space for other animals, to reduce the marginal product of all other animals in the society.

In order to make producers internalize the social consequences of their decisions, a government can introduce a Pigouvian tax on animals. If one denotes such a tax by  $\top$ , the price of each animals then increases from  $c$  to  $c(1 + \top)$ . The introduction of such a tax has two effects on the total number of animals: one in terms of intensive margins and one in terms of extensive margins.

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. But the intensive margin effect is here made ambiguous by the fact that, as Kantians reduce their number of animals, Nash players react by *increasing* their number of animals. Thus the impact of the tax on intensive margins is here more complex than in the case of a quota (where Pure Kantians were not affected at all, but well Impure Kantians).

This ambiguous intensive margin effect may be counterbalanced by another effect, in extensive margins: the tax, by modifying the achieved utility levels by the two types of agents, may modify the composition of the population. It may be the case that the tax, by affecting the incentives to socialize children in different ways across types, raises the proportion of Nash players. This extensive margin effect also affects the influence of the tax at the aggregate level.

## 5.1 Pure Kantians vs Nash players

In the case of Pure Kantians, given that  $E = q(\top)e^{K^*}(\top) + (1 - q(\top))e^{N^*}(\top)$ , the effect of the tax on the total number animals is:

$$\frac{\partial E}{\partial \top} = \frac{\partial q}{\partial \top} e^{K^*}(\top) + q(\top) \frac{\partial e^{K^*}(\top)}{\partial \top} - \frac{\partial q}{\partial \top} e^{N^*}(\top) + (1 - q(\top)) \frac{\partial e^{N^*}(\top)}{\partial \top}$$

This can be rewritten as:

$$\frac{\partial E}{\partial \top} = \underbrace{\frac{\partial q}{\partial \top} [e^{K^*}(\top) - e^{N^*}(\top)]}_{\text{extensive margin effect}} + \underbrace{q(\top) \frac{\partial e^{K^*}(\top)}{\partial \top} + (1 - q(\top)) \frac{\partial e^{N^*}(\top)}{\partial \top}}_{\text{intensive margin effect}}$$

We know that  $e^{K^*}(\top) < e^{N^*}(\top)$ , but since the sign of  $\frac{\partial q}{\partial \top}$  is unknown, the extensive margin effect is ambiguous. Regarding the intensive margin effect, an important difference with respect to the case of the quota is that, in the case of the tax, both types of agents are affected by the policy intervention.<sup>11</sup> The impact on the number of animals for Pure Kantians is unambiguous: Pure

<sup>11</sup>Of course, in a first-best setting, the government would observe the types of agents, and would only tax those who generates negative externalities (i.e. Nash players).

Kantians will, as a consequence of the tax, reduce the number of animals they purchase. We thus have  $\frac{\partial e^{K^*}(\tau)}{\partial \tau} < 0$ .

Concerning the impact of the tax on the number of animals purchased by Nash players, two opposite effects are at work. First, the rise in the price of animals is likely to make Nash players reduce their number of animals. But at the same time, the fall in the number of animals purchased by Pure Kantians pushes Nash players to purchase more animals. Actually, the two effects on  $\frac{\partial e^{N^*}(\tau)}{\partial \tau}$  can be presented as follows:

$$\frac{\partial e^{N^*}(\tau)}{\partial \tau} = \underbrace{\frac{\partial e^{N^*}(\tau)}{\partial \tau} \Big|_{e^{K^*} \text{ constant}}}_{-} + \underbrace{\frac{\partial e^{N^*}(\tau)}{\partial e^{K^*}(\tau)} \frac{\partial e^{K^*}(\tau)}{\partial \tau}}_{+}$$

The impact of the tax on the number of animals purchased by Nash players is negative when the first, direct effect dominates the second, indirect effect (reaction effect). On the contrary, when the reaction effect dominates, Nash players purchase more animals under the tax. Numerically, it can be shown that, when  $q$  is close to 1, it is possible that the second effect dominates the first one, implying that Nash shepherds purchase more animals under the tax than without the tax.<sup>12</sup> However, for lower values of  $q$ , the direct effect generally dominates the indirect effect, so that  $\frac{\partial e^{N^*}(\tau)}{\partial \tau} < 0$ . Thus in general the reaction effect is a second-order effect, which is dominated by the direct effect of the tax. As a consequence, we obtain in that case that the overall intensive margin effect of the tax on the total number of animals is negative. However, given that the impact of the tax on the partition of the population is unknown, the extensive margin effect is not clear. One can hardly know, in theory, what is the total effect of the tax on the number of animals.

The rest of this section examines numerically the impact of introducing a Pigouvian tax on animals on the degree of severity of the Tragedy of the Commons. Given political constraints (lobbying groups, media), policy-makers are constrained in the extent to which they can impose high taxes. Hence, it makes sense to examine the impact, in a third-best world, of small taxes on the extent of the Tragedy of the Commons. This is the task carried out in Table 1 for different values of the ratio of parameters  $A$  (productivity of animals) and  $\psi$  (disutility of the number of animals). To make the link with the previous section, Table 1 also compares the effects of the tax with the impact of quotas of limited sizes.<sup>13</sup>

<sup>12</sup>Note, however, that, since  $q = 1$  is not a stable stationnary equilibrium, one should not pay too much attention to that extreme case.

<sup>13</sup>Table 1 is based on  $\kappa = 0.5$  and  $c = 0.5$ .

$A/\psi$	$E$						
	LF	quotas			tax		
		-10% wrt LF	-20% wrt LF	-50% wrt LF	10%	20%	30%
0.100	0.054	0.059	0.060	0.054	0.052	0.050	0.046
0.125	0.065	0.070	0.072	0.068	0.063	0.060	0.054
0.250	0.112	0.121	0.123	0.118	0.106	0.100	0.090

Table 1: Effects of the quota and the tax on total number of animals (Pure Kantians vs Nash).

Table 1 invites several comments. First, Table 1 shows that, when the ratio  $A/\psi$  goes up, the total number of animals purchased increases in all cases. This result is expected: a higher productivity per animal (i.e. a higher  $A$ ) and a lower disutility of caring for animals (i.e. a lower  $\psi$ ) pushes all individuals, whatever their type, to purchase more animals, implying a rise in  $E$ .

Second, Table 1 also confirms the findings of the previous section: quotas with a limited constraining power do not decrease, but increase the total number of animals purchased. Thus quotas can, as in the previous section, worsen the Tragedy of the Commons in comparison to the laissez-faire.

Third, and more importantly, taxes on animals, even of small sizes, tend to reduce the total number of animals with respect to the laissez-faire. Thus the overall effect of the tax is here to reduce the number of animals, whatever the level of the tax. Contrary to the quota, even a small tax can reduce the Tragedy of the Commons. Thus the tax and the quota have quite different effects in the context of reducing the Tragedy of the Commons. The underlying intuition is that the tax affects here the quantity of animals purchased by *all* individuals, and not only by Nash players (unlike what prevails under the quota).

In sum, it appears that, to reduce the Tragedy of the Commons in a world of political constraints, a small tax may perform better than a small quota (which can make things worse than the laissez-faire), even though in a first-best world the quota performs better (by targetting only the providers of negative externalities, i.e. Nash players). Indeed, as the small quota may raise the total number of animals, this rise is equivalent with what would have been achieved while subsidizing - instead of taxing - the sheep.

## 5.2 Impure Kantians vs Nash players

In the case of Impure Kantians, given that  $E^{**} = q(\tau)e^{K^{**}}(\tau) + (1-q(\tau))e^{N^{**}}(\tau)$ , the effect of the tax on the total number animals can be written as:

$$\frac{\partial E}{\partial \tau} = \frac{\partial q}{\partial \tau} e^{K^{**}}(\tau) + q(\tau) \frac{\partial e^{K^{**}}(\tau)}{\partial \tau} - \frac{\partial q}{\partial \tau} e^{N^{**}}(\tau) + (1 - q(\tau)) \frac{\partial e^{N^{**}}(\tau)}{\partial \tau}$$

This can be rewritten as:

$$\frac{\partial E}{\partial \tau} = \underbrace{\frac{\partial q}{\partial \tau} [e^{K^{**}}(\tau) - e^{N^{**}}(\tau)]}_{\text{extensive margin effect}} + \underbrace{q(\tau) \frac{\partial e^{K^{**}}(\tau)}{\partial \tau} + (1 - q(\tau)) \frac{\partial e^{N^{**}}(\tau)}{\partial \tau}}_{\text{intensive margin effect}}$$

We have  $e^{K^{**}}(\top) < e^{N^{**}}(\top)$ , but the sign of  $\frac{\partial q}{\partial \top}$  is unknown, so that the extensive margin effect remains, here again, ambiguous. Regarding the intensive margin effect, a major difference with respect to the case of Pure Kantians is that Impure Kantians may or may not reduce their number of animals, since they react to what Nash players do (unlike Pure Kantians). We have:

$$\begin{aligned} \frac{\partial e^{K^{**}}(\top)}{\partial \top} &= \underbrace{\frac{\partial e^{K^{**}}(\top)}{\partial \top} \Big|_{e^{N^{**}} \text{ constant}}}_{-} + \underbrace{\frac{\partial e^{K^{**}}(\top)}{\partial e^{N^{**}}(\top)}}_{-} \underbrace{\frac{\partial e^{N^{**}}(\top)}{\partial \top}}_{?} \\ \frac{\partial e^{N^{**}}(\top)}{\partial \top} &= \underbrace{\frac{\partial e^{N^{**}}(\top)}{\partial \top} \Big|_{e^{K^{**}} \text{ constant}}}_{-} + \underbrace{\frac{\partial e^{N^{**}}(\top)}{\partial e^{K^{**}}(\top)}}_{-} \underbrace{\frac{\partial e^{K^{**}}(\top)}{\partial \top}}_{?} \end{aligned}$$

Hence the global intensive margin effect is even more ambiguous in this case than in the case of Pure Kantians. Adding to this the - still unknown - extensive margins effect, we can hardly find some analytical result concerning the impact of the tax.

Turning to numerical analysis, Table 2 compares the impact of quotas and taxes on the total number of animals in the case of Impure Kantians.<sup>14</sup> Table 2 shows that taxes, even when they are low, can reduce the total number of animals with respect to the laissez-faire. On the contrary, weak quotas tend to make things even worse than at the laissez-faire, by implying a rise in the total number of animals. Only quotas restraining Nash players by 50 % can achieve a reduction in the total number of animals, whereas a 10 % tax on animals can achieve such a reduction.

$A/\psi$	$E$						
	LF	quotas			tax		
		-10% wrt LF	-20% wrt LF	-50% wrt LF	10%	20%	30%
0.100	0.059	0.059	0.061	0.050	0.058	0.056	0.050
0.125	0.071	0.072	0.075	0.068	0.067	0.066	0.060
0.250	0.121	0.121	0.125	0.110	0.115	0.109	0.083

Table 2: Effects of the quota and the tax on total number of animals (Impure Kantians vs Nash).

Although non exhaustive, our numerical simulations reveal that taxes and quotas are far from equivalent, in the sense that these have quite different capacities to reduce the severity of the Tragedy of the Commons. We know that, in a first-best world, applying a quota  $\bar{e} = e^{K^*}$  decentralizes the social optimum, by internalizing all congestion externalities. Thus in a first-best world the quota

<sup>14</sup>Table 1 is based on  $\kappa = 0.5$  and  $c = 0.5$ .

is superior to the tax. However, in a third-best world with strong political constraints, governments can hardly implement strong quotas (or high taxes), and in that context, our simulations suggest that the tax performs better than the quota. Whereas weak quotas worsen the Tragedy even more in comparison to the laissez-faire, Pigouvian taxes can, even with low levels, reduce the total number of animals. That result is particularly relevant for governments facing political constraints preventing them from implementing constraining quotas. In such a third-best world, the tax dominates the quota, by leading to a reduction of the severity of the Tragedy.

## 6 Relative fitness of (Im)Pure Kantians

Up to now, we derived separate results for two populations, composed either of Nash players and Pure Kantians, or of Nash players and Impure Kantians. One may be curious to know, from an evolutionary perspective, which trait is the fittest in the context of the Tragedy of the Commons: being a Pure or an Impure Kantian? Put it differently, the question is to know which type of Kantians would be the most numerous in the long-run? At first glance, it is tempting to believe that Impure Kantians are more fit than Pure Kantians, simply because they are better informed, since they take into account, unlike Pure Kantians, that a fraction  $1 - q$  of the population plays Nash. But it is not clear that being more informed implies necessarily to be more fit.

The question of the relative fitness of the two types of Kantian players is also most relevant for policy purposes. Indeed, suppose that a government can, at no cost, convert Kantian agents in either Pure or Impure Kantians. The question is then to know which type of Kantian morality (Pure or Impure) the government should promote, that is, which type of Kantian morality would lead to the lowest congestion of land?

### 6.1 Reaction functions

Let us first come back to the temporary equilibrium, to examine which type of Kantians derives the highest temporal utility for a given partition  $q_t$ . For the sake of simplicity, we focus here on the case of quasi-linear preferences ( $\sigma = 0$ ) and assume  $c = 0$ . As above, we also suppose a Cobb-Douglas production function, which implies:

$$G(E_t) = AE_t^\alpha$$

as well as the disutility from raising animals:

$$\frac{2}{3}\psi (e_t^i)^{3/2}$$

Remind first that the optimal number of animals in the two distinct games



are given by:

$$\begin{aligned}
A(E_t^*)^{\alpha-1} - \psi(e_t^{N*})^{1/2} &= 0 \\
\alpha A e_t^{K*\alpha-1} - \psi(e_t^{K*})^{1/2} &= 0 \iff e_t^{K*} = \left(\frac{\psi}{A\alpha}\right)^{\frac{1}{\alpha-\frac{1}{2}}} \\
A(E_t^{**})^{\alpha-1} - \psi(e_t^{N**})^{1/2} &= 0 \\
A(E_t^{**})^{\alpha-1} + (\alpha-1)q_t e_t^{K**} A(E_t^{**})^{\alpha-2} - \psi(e_t^{K**})^{1/2} &= 0
\end{aligned}$$

Let us now compare those two games. To have analytical results, we consider the case where  $\alpha = 1/2$ . The following proposition summarizes our results.

**Proposition 13** *Suppose  $G(E_t) = AE_t^{1/2}$ ,  $v(e_t^i) = \frac{2}{3}\psi(e_t^i)^{3/2}$  and suppose a given  $q_t$ .*

- *The comparison of the  $q$  equilibrium involving Pure Kantians with the  $q$  equilibrium involving Impure Kantians yields,*  
when  $q_t = 0$  or  $q_t = 1$ :

$$\begin{aligned}
e_t^{K*} &= e_t^{K**} \\
e_t^{N*} &= e_t^{N**}
\end{aligned}$$

when  $0 < q_t < 1$ :

$$\begin{aligned}
e_t^{K*} &< e_t^{K**} \\
e_t^{N*} &> e_t^{N**}
\end{aligned}$$

- *We also have:*

$$\begin{aligned}
\text{if } q_t &= 0 \text{ or } q_t = 1, \text{ then } E_t^* = E_t^{**} \\
\text{if } 0 &< q_t < 1, \text{ then } E_t^* < E_t^{**}
\end{aligned}$$

**Proof.** See the Appendix. ■

Proposition 13 states that, under this analytical case, Impure Kantians tend, for a given partition of the population  $0 < q_t < 1$ , to purchase more animals than Pure Kantians, whereas Nash players purchase more animals when they compete with Pure Kantians than when they compete with Impure Kantians.

Quite interestingly, under an interior  $q_t$ , the total number of animals is lower at the  $q$  equilibrium with Pure Kantians than at the  $q$  equilibrium with Impure Kantians. The underlying intuition goes as follows. At the equilibrium, the effect of the higher number of animals purchased by Impure Kantians dominates the lower number of animals purchased, as a reaction, by Nash players.

## 6.2 Socialization and dynamics

The previous subsection shows that, under particular functional forms, the total number of animals is, for a given partition  $q_t$ , lower when the economy is composed of Pure Kantians and Nash players than when it is composed of Impure Kantians and Nash players. Although interesting, that result does not tell us which Kantian type leads to the smallest aggregate number of animals in the long-run, since that result relies on a *given* partition. The dynamics of heterogeneity is different when the economy is composed of Pure or Impure Kantians, so that there is no obvious reason why one should expect that  $q$  takes the same long-run level in the two cases.

Given the large number of effects at work, the analysis of the relative fitness of the Pure and Impure Kantians has to be carried out numerically. For that purpose, Table 3 shows, under the production function  $G(E_t) = AE_t^{1/2}$  and under the desutility  $\frac{2}{3}\psi(e_t^i)^{3/2}$ , the proportion of Kantian players at the stationary equilibrium for different values of the ratio of parameters  $A/\psi$ , as well as the associated total number of animals  $E$ .<sup>15</sup> To cast light on the relative fitness of Pure and Impure Kantians, we distinguish between two kinds of games: on the one hand, Pure Kantians versus Nash players, and, on the other hand, Impure Kantians versus Nash players.

Parameters	Pure Kantians vs Nash		Impure Kantians vs Nash	
	$q^*$	$E^*$	$q^{**}$	$E^{**}$
0.025	0.14	0.02	0.52	0.02
0.050	0.25	0.05	0.53	0.04
0.100	0.55	0.08	0.58	0.09
0.125	0.62	0.09	0.60	0.11
0.250	0.66	0.19	0.65	0.21
0.500	0.67	0.38	0.66	0.42

Table 3: Relative fitness of Pure and Impure Kantians

Regarding the relative fitness of Pure and Impure Kantians, Table 3 shows that the two types of Kantian agents exhibit quite different degrees of fitness. In all parametrizations considered, Impure Kantians represent more than 50 % of the population, and are thus always more fit than Nash players. This is true whether there is a large number of animals or not, that is, whether the size of negative congestion externalities is large or not. Thus, if one reminds that, in our model, the interior stationary equilibrium  $q$  equals  $\frac{\Delta V^K}{\Delta V^K + \Delta V^N}$ , and depends on the relative intolerance of Kantian and Nash players, we can conclude that Impure Kantians are, under all our parametrizations, more intolerant than Nash

<sup>15</sup>The parameter  $\kappa$  is set to 0.5.

players, independently of the extent of the Tragedy of the Commons.<sup>16</sup>

Turning now to Pure Kantians, we can see that their prevalence at the stationary equilibrium varies much more with the postulated ratio  $A/\psi$ . When  $A/\psi$  is low, that is, when individuals purchase a small number of animals, the proportion of Pure Kantians at the stationary equilibrium is low. They are then clearly dominated, in terms of fitness, by Nash players. However, when the ratio  $A/\psi$  increases (implying that both types of agents purchase more animals), then Pure Kantians become more numerous than Nash players at the stationary equilibrium. Thus, it is only when the Tragedy of the Commons is more severe that Pure Kantians become more intolerant than Nash players.

Comparing the fitness of Pure and Impure Kantians, we can thus conclude that, whereas Impure Kantians are always more intolerant than Nash players, independently of the extent of the Tragedy of the Commons, and are thus always dominant in the long-run, this is not true for Pure Kantians, who are more intolerant than Nash players only when the Tragedy is quite severe. If one now compares directly the prevalence of the two types of Kantians at the long-run equilibrium, we see that, for low levels of the ratio  $A/\psi$ , Impure Kantians are more numerous than Pure Kantians, whereas, for higher levels of  $A/\psi$ , Pure Kantians are more numerous than Impure Kantians. Thus the fitness ranking between Pure and Impure Kantians is not invariant to the parametrizations.

Let us now turn back to the question of the most adequate type of Kantian in the context of the Tragedy of the Commons. Does one type of Kantian lead to lower congestion? When considering that question, a first important thing to stress is that adopting a static perspective may be quite misleading. True, as we showed in the previous section, the total number of animals is, for a *given* partition  $q$ , lower under Pure Kantians than under Impure Kantians. However, thinking in terms of a given partition is quite restrictive here, since the long-run prevalence of Kantians depends strongly on which type of Kantian is considered, and this definitely affects the size of congestion in the long-run. As shown in Table 3, under a low ratio  $A/\psi$ , Pure Kantians are so few at the stationary equilibrium that the total number of animals is larger in the presence of Pure Kantians than under Impure Kantians. Thus the long-run partition of the population definitely matters to see which type of Kantians is the most favorable to minor the Tragedy of the Commons. One cannot make comparisons for a given partition, since the partition definitely varies with the

<sup>16</sup>The underlying intuition goes as follows. For Impure Kantians, we know that the case

$$\underbrace{V^{KK} - V^{KN}}_{\Delta V^K} > \underbrace{V^{NN} - V^{NK}}_{\Delta V^N}$$

is plausible, since we have  $V^{KK} = V^{NK}$ , so that the above inequality becomes:

$$2V^{KK} > V^{NN} + V^{KN}$$

This inequality is plausible, since  $V^{KN}$  is relatively smaller (since an Impure Kantian then generalize  $e^N$  to their group - and thus to the whole society -, so that negative externalities due to congestion reach their maximum.  $V^{KN}$  being relatively smaller, we expect the above inequality to hold, leading to  $q > 1/2$ ).

type of Kantians, whose degrees of relative fitness differ.

Note, however, that examining the prevalence of Kantians at the stationary equilibrium is, although necessary, not sufficient to be able to know which type of Kantians makes the Tragedy of the Commons less severe. The reason is that the extent of the Tragedy depends not only on extensive margins (proportions of Kantians and Nash players), but, also, on intensive margins (how many animals purchased by each individual). Hence focusing only on the proportion of Kantians may oversimplify the picture. To illustrate this, let us take the case where  $A/\psi$  equals 0.100. In that case, Impure Kantians are, at the stationary equilibrium, more numerous than Pure Kantians (0.58 versus 0.55). However, the total number of animals  $E$  is larger when Nash players face Impure Kantians than when Nash players face Pure Kantians. Thus, one should be cautious before concluding that the presence of the most fit Kantian type leads to the most favorable environmental outcome. Here again, the aggregate number of animals depends both on extensive and intensive margins effects.<sup>17</sup>

## 7 Conclusions

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 (Pure or Impure), and where the partition of the population into those different types of rationalities follows a vertical/oblique socialization process à la Bisin and Verdier (2001). We characterized various types of temporary equilibria (i.e. for a given partition of the population), with either only Nash players, or with only Kantian players, or with both. Then, when studying the long-run dynamics of heterogeneity, we examined the issues of existence, uniqueness and stability of stationary equilibria, i.e. equilibria with a fixed partition of the population into different types of rationality.

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

---

<sup>17</sup>The fact that Impure Kantians purchase more animals than Pure Kantians, and the fact that Nash players internalizing the larger proportion of Kantians react by purchasing more animals, can counterbalance the fact that Impure Kantians are more numerous than Pure Kantians, and lead *in fine* to a larger total number of animals.

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.

We also compared the impact of a quota on animals with the one of a tax on animals, and we showed numerically that, although weak quotas can, because of an extensive margin effect, worsen the Tragedy of the Commons, even a low tax on animals can reduce the congestion of land. Hence, in a third-best world where governments face difficulties in imposing strong quotas/taxes, the tax dominates the quota, by allowing for a reduction of land congestion.

Finally, and still on the policy side, we studied whether a government should, in the context of the Tragedy of the Commons, promote either Pure or Impure Kantian morality, by examining the relative fitness of Pure and Impure Kantian players with respect to Nash players. Here again, thinking for a given partition is misleading. Whereas, for a given partition, the extent of land congestion is lower under Pure Kantians than under Impure Kantians, the opposite may be true once we allow for a varying partition of the population.

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 heterogeneity. Ignoring the dynamics of heterogeneity may lead governments to implement policies that make the Tragedy even worse than at the laissez-faire, even though such policies would have worked well for a given partition of the population.

## 8 References

- Alger, I., Weibull, J. (2014). Evolution leads to Kantian morality. Toulouse School of Economics, IAST Working Papers, 14-10.
- Apestequia, J. (2006). Does information matter in the Commons? Experimental Evidence. *Journal of Economic Behavior and Organization*, 60: 55-69.
- 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.
- Brekke, K., Kverndokk, S., Nyborg, K. (2003). An economic model of moral motivation. *Journal of Public Economics*, 87 (9-10): 1967-1983.
- 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.
- Dannan, F., Elaydi S., Ponomarenko V. (2003), Stability of Hyperbolic and Nonhyperbolic Fixed Points of One-dimensional Maps, *Journal of Difference Equations and Applications*, 9(5), 449-457.

- 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.
- Hardin G. (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.
- Keser, C, Gardner, R. (1999). Strategic behavior of experienced subjects in a common pool resource game, *International Journal of Game Theory*, 28: 241-252.
- Knight, G. P., Cota, M. K., & Bernal, M. E. (1993). The socialization of cooperative, competitive, and individualistic preferences among Mexican American children: The mediating role of ethnic identity. *Hispanic Journal of Behavioral Sciences*, 15(3), 291-309.
- McCay, B., Acheson, J. (1987). *The Question of the Commons. The culture and ecology of communal resources*. University of Arizona Press, Tucson.
- Ostrom, E. (1987). Institutional arrangements for resolving the commons dilemma: some contending approaches. In: McCay, B., Acheson, J. (eds.). *The Question of the Commons. The culture and ecology of communal resources*. University of Arizona Press, Tucson.
- 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.
- Roemer, J. (2010). Kantian equilibrium. *Scandinavian Journal of Economics*, 112: 1-24.
- 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.
- Van Long, N. (2015). Kant-Nash equilibrium in a dynamic game of climate change mitigations. CESifo Area Conference on Energy and Climate Policy.
- 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.

## 9 Appendix

### 9.1 Proof of Proposition 4

#### 9.1.1 Existence

Note first that:

$$\begin{aligned} (e_t^{K^{**}})^{1+\sigma} &= \frac{\left(\frac{e_t^{K^{**}} q_t}{E_t^{**}}\right) \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right]}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} + \frac{\frac{G(E_t^{**})}{E_t^{**}} - c}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} \\ (e_t^{N^{**}})^{1+\sigma} &= \frac{\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} \end{aligned}$$

Hence, from the comparison of those two conditions, and given that the first part of the first equation is negative, we see that  $e_t^{K^{**}} < e_t^{N^{**}}$ .

We know that, from the proof of existence of a  $q$  equilibrium, there exists a unique  $e_t^{N^{**}}$  for each  $e_t^{K^{**}}$  such that:

$$e_t^{N^{**}} = \left[ \frac{\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

Note that  $e_t^{N^{**}}$  is strictly decreasing in  $e_t^{K^{**}}$ . We also have  $e_t^{N^{**}} > 0$  when  $e_t^{K^{**}} = 0$ .

Let us consider the other expression.

$$e_t^{K^{**}} = \left[ \frac{\left(\frac{e_t^{K^{**}} q_t}{E_t^{**}}\right) \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right]}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} + \frac{\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

Hence

$$\begin{aligned} e_t^{K^{**}} &= \left[ \frac{\left(\frac{e_t^{K^{**}} q_t}{E_t^{**}}\right) \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} - c + c\right]}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} + \frac{\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} \right]^{\frac{1}{1+\sigma}} \\ e_t^{K^{**}} &= \left[ \frac{\left(\frac{e_t^{K^{**}} q_t}{E_t^{**}}\right) \left[G'(E_t^{**}) - c - \left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)\right]}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} + \frac{\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} \right]^{\frac{1}{1+\sigma}} \end{aligned}$$

Hence

$$e_t^{K^{**}} = \left[ \left(\frac{e_t^{K^{**}} q_t}{E_t^{**}}\right) \left[ \frac{G'(E_t^{**}) - c}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} - \frac{\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} \right] + \frac{\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

Hence

$$e_t^{K^{**}} = \left( \frac{1}{E_t^{**}} \right)^{\frac{1}{1+\sigma}} \left[ q_t e_t^{K^{**}} \left[ \frac{G'(E_t^{**}) - c}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right] + (1 - q_t) e_t^{N^{**}} \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^{\frac{1}{1+\sigma}}}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right]$$

Let us define

$$e_t^{K^{**}} = \left[ \frac{q_t e_t^{K^{**}}}{E_t^{**}} \left[ \frac{G'(E_t^{**}) - c}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right] + \frac{(1 - q_t) e_t^{N^{**}}}{E_t^{**}} \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^{\frac{1}{1+\sigma}}}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}} \equiv \Theta(e_t^{N^{**}}, e_t^{K^{**}})$$

The term in brackets is a weighted sum of the Kantian and the Nash solutions.

Let us derive  $\Theta(e_t^{N^{**}}, e_t^{K^{**}})$  wrt  $e_t^{N^{**}}$ :

$$\begin{aligned} & \frac{\partial \Theta(e_t^{N^{**}}, e_t^{K^{**}})}{\partial e_t^{N^{**}}} \\ &= \frac{1}{1 + \sigma} \left[ \frac{(e_t^{K^{**}} q_t)}{E_t^{**}} \left[ \frac{G'(E_t^{**}) - c}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma} - 1} \right. \\ & \quad \left. + \frac{(1 - q_t) e_t^{N^{**}}}{E_t^{**}} \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^{\frac{1}{1+\sigma} - 1}}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right] \\ & \quad \left[ \frac{-(e_t^{K^{**}} q_t)(1 - q_t)}{[E_t^{**}]^2} \left[ \frac{G'(E_t^{**}) - c}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right] \right. \\ & \quad \left. + \frac{e_t^{K^{**}} q_t}{E_t^{**}} \left[ \frac{(1 - q_t) G''(E_t^{**}) \psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma - \left[ \frac{[G'(E_t^{**}) - c] \psi^\sigma}{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^{1-\sigma}} \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{(1 - q_t)^{-1} [E_t^{**}]} \right]}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right] \right] \\ & \quad \left. + (1 - q_t) \frac{q_t e_t^{K^{**}}}{[E_t^{**}]^2} \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} + \left[ \frac{(1 - q_t) e_t^{N^{**}}}{E_t^{**}} \frac{(1 - \sigma)(1 - q_t)}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{E_t^{**}} \right] \right] \end{aligned}$$

Let us assume, for the moment, that the second term is negative, and collect the three other terms.

$$\left[ \frac{-(e_t^{K^{**}} q_t)(1 - q_t)}{[E_t^{**}]^2} \left[ \frac{G'(E_t^{**}) - c}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right] + (1 - q_t) \frac{q_t e_t^{K^{**}}}{[E_t^{**}]^2} \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right. \\ \left. + \frac{(1 - q_t) e_t^{N^{**}}}{E_t^{**}} \frac{(1 - \sigma)(1 - q_t)}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{E_t^{**}} \right]$$



Let us factorize:

$$\frac{(1-q_t)e_t^{K^{**}}q_t}{[E_t^{**}]^2} \frac{\left[-[G'(E_t^{**})] + \frac{G(E_t^{**})}{E_t^{**}}\right]}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma} + \frac{((1-q_t)e_t^{N^{**}})(1-\sigma)(1-q_t)}{[E_t^{**}]^2} \frac{\left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right]}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma}$$

Hence

$$\left[\frac{G(E_t^{**})}{E_t^{**}} - G'(E_t^{**})\right] \frac{(1-q_t) \left[e_t^{K^{**}}q_t - ((1-q_t)e_t^{N^{**}})(1-\sigma)\right]}{[E_t^{**}]^2 \psi\left(\frac{G(E_t^{**})}{E_t^{**}} - c\right)^\sigma}$$

Given  $e_t^{K^{**}} < e_t^{N^{**}}$ , we have

$$\begin{aligned} & e_t^{K^{**}}q_t - (1-q_t)e_t^{N^{**}}(1-\sigma) \\ & < e_t^{N^{**}}q_t - (1-q_t)e_t^{N^{**}}(1-\sigma) = e_t^{N^{**}}[\sigma(1-q_t) - 1] < 0 \end{aligned}$$

Hence the sum of the three terms in brackets is unambiguously negative.

We thus have:

$$\frac{\partial \Theta(e_t^{N^{**}}, e_t^{K^{**}})}{\partial e_t^{N^{**}}} < 0$$

Next step: derivation wrt  $e_t^{K^{**}}$ . We have:

$$\begin{aligned} & \frac{\partial \Theta(e_t^{N^{**}}, e_t^{K^{**}})}{\partial e_t^{K^{**}}} \\ &= \frac{1}{1+\sigma} \left[ \frac{(e_t^{K^{**}}q_t)}{E_t^{**}} \left[ \frac{G'(E_t^{**})-c}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}}-c\right)^\sigma} \right]^{\frac{1}{1+\sigma}-1} \right. \\ & \quad \left. + \frac{((1-q_t)e_t^{N^{**}})}{E_t^{**}} \frac{\left(\frac{G(E_t^{**})}{E_t^{**}}-c\right)}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}}-c\right)^\sigma} \right] \\ & \quad \left[ \frac{-(e_t^{N^{**}}(1-q_t))(q_t)}{[E_t^{**}]^2} \left[ \frac{G'(E_t^{**})-c}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}}-c\right)^\sigma} \right] \right. \\ & \quad \left. + \frac{(e_t^{K^{**}}q_t)}{E_t^{**}} \left[ \frac{q_t G''(E_t^{**}) \psi\left(\frac{G(E_t^{**})}{E_t^{**}}-c\right)^\sigma - \left[ \frac{[G'(E_t^{**})-c] \psi^\sigma}{\left(\frac{G(E_t^{**})}{E_t^{**}}-c\right)^{1-\sigma}} q_t \frac{[G'(E_t^{**})] - \frac{G(E_t^{**})}{E_t^{**}}}{E_t^{**}} \right]}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}}-c\right)^\sigma} \right] \right] \\ & \quad \left. + q_t \frac{(1-q_t)e_t^{N^{**}}}{[E_t^{**}]^2} \frac{\left(\frac{G(E_t^{**})}{E_t^{**}}-c\right)}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}}-c\right)^\sigma} + \left[ \frac{(1-q_t)e_t^{N^{**}}}{E_t^{**}} \frac{(1-\sigma)q_t}{\psi\left(\frac{G(E_t^{**})}{E_t^{**}}-c\right)^\sigma} \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{E_t^{**}} \right] \right] \end{aligned}$$

Assuming, for the moment, that the second term in brackets is negative, let us recollect the three other terms and carry out the same manipulation as above.

We have:

$$\left[ \begin{aligned} & \frac{-(e_t^{N**} q_t)(1-q_t)}{[E_t^{**}]^2} \left[ \frac{G'(E_t^{**})-c}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right] + \frac{(1-q_t)q_t e_t^{N**}}{[E_t^{**}]^2} \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \\ & + \frac{(1-q_t)e_t^{N**}}{E_t^{**}} \frac{(1-\sigma)q_t}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{E_t^{**}} \end{aligned} \right]$$

Let us factorize:

$$\frac{(1-q_t)e_t^{N**} q_t}{[E_t^{**}]^2} \left[ -[G'(E_t^{**})] + \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{((1-q_t)e_t^{N**})(1-\sigma)q_t}{[E_t^{**}]^2} \left[ \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right]$$

Hence

$$\frac{\left[ \frac{G(E_t^{**})}{E_t^{**}} - G'(E_t^{**}) \right] (1-q_t) [e_t^{N**} q_t - ((1-q_t)e_t^{N**})(1-\sigma)]}{[E_t^{**}]^2 \psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma}$$

We have

$$e_t^{N**} q_t - (1-q_t)e_t^{N**} (1-\sigma) = e_t^{N**} [q_t - (1-q_t)(1-\sigma)] < 0$$

Hence the sum of the three terms in brackets is unambiguously negative.

We thus have:

$$\frac{\partial \Theta(e_t^{N**}, e_t^{K**})}{\partial e_t^{K**}} < 0$$

Hence, given the two expressions, and given

$$e_t^{K**} = \left[ \frac{(e_t^{K**} q_t)}{E_t^{**}} \left[ \frac{G'(E_t^{**}) - c}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right] + \frac{(1-q_t)e_t^{N**}}{E_t^{**}} \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}} \equiv \Theta(e_t^{N**}, e_t^{K**})$$

We have  $e_t^{K**}$  being an implicit function of  $e_t^{N**}$  defined such that:

$$e_t^{K**} - \Theta(e_t^{N**}, e_t^{K**}) = 0$$

We can write:

$$\frac{de_t^{K**}}{de_t^{N**}} = - \frac{\frac{\partial \Theta(e_t^{N**}, e_t^{K**})}{\partial e_t^{N**}}}{1 - \frac{\partial \Theta(e_t^{N**}, e_t^{K**})}{\partial e_t^{K**}}}$$

Hence we have:

$$\frac{de_t^{K**}}{de_t^{N**}} = \frac{\frac{\partial \Theta(e_t^{N**}, e_t^{K**})}{\partial e_t^{N**}}}{1 - \frac{\partial \Theta(e_t^{N**}, e_t^{K**})}{\partial e_t^{K**}}} < 0$$

Let us denote by  $\Gamma : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  the decreasing function defined such that:

$$e_t^{K**} \equiv \Gamma(e_t^{N**})$$

with  $\Gamma'(e_t^{N**}) < 0$ .

Let us now take the other expression:

$$e_t^{N**} = \left[ \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}} \equiv \Xi(e_t^{K**}, e_t^{N**})$$

We have:

$$\begin{aligned} \frac{\partial \Xi(e_t^{K**}, e_t^{N**})}{\partial e_t^{K**}} &< 0 \\ \frac{\partial \Xi(e_t^{K**}, e_t^{N**})}{\partial e_t^{N**}} &< 0 \end{aligned}$$

We have  $e_t^{K**}$  being an implicit function of  $e_t^{N**}$  defined such that:

$$e_t^{N**} - \Xi(e_t^{K**}, e_t^{N**}) = 0$$

We can write:

$$\frac{de_t^{K**}}{de_t^{N**}} = - \frac{1 - \frac{\partial \Xi(e_t^{K**}, e_t^{N**})}{\partial e_t^{N**}}}{\frac{\partial \Xi(e_t^{K**}, e_t^{N**})}{\partial e_t^{K**}}} = \frac{1 - \frac{\partial \Xi(e_t^{K**}, e_t^{N**})}{\partial e_t^{N**}}}{\frac{\partial \Xi(e_t^{K**}, e_t^{N**})}{\partial e_t^{K**}}} < 0$$

Let us denote by  $\Omega : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  the decreasing function defined such that:

$$e_t^{K**} \equiv \Omega(e_t^{N**})$$

with  $\Omega'(e_t^{N**}) < 0$ .

In order to discuss the existence of an intersection of the two reaction functions in the  $(e_t^{N**}, e_t^{K**})$  space, we will compute  $\Gamma(e_t^{N**})$  and  $\Omega(e_t^{N**})$  when  $e_t^{N**} = e_t^{K**}$ , and  $\Gamma(e_t^{N**})$  and  $\Omega(e_t^{N**})$  when  $e_t^{K**} = 0$ . We have:

$$e_t^{K**} = \left[ \frac{q_t e_t^{K**}}{E_t^{**}} \left[ \frac{G'(E_t^{**}) - c}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right] + \frac{(1 - q_t) e_t^{N**}}{E_t^{**}} \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)}{\psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

Given  $e_t^{N**} = e_t^{K**}$ , we have:

$$\hat{e}_t^{K**} = \frac{\left[ q_t (G'(e_t^{K**}) - c) + (1 - q_t) \left( \frac{G(e_t^{K**})}{e_t^{K**}} - c \right) \right]^{\frac{1}{1+\sigma}}}{\psi^{\frac{1}{1+\sigma}} \left( \frac{G(e_t^{K**})}{e_t^{K**}} - c \right)^{\frac{\sigma}{1+\sigma}}}$$

Take the other expression. We have:

$$\tilde{e}_t^{K^{**}} = \frac{1}{\psi^{\frac{1}{1+\sigma}} \left( \frac{G(e_t^{K^{**}})}{e_t^{K^{**}}} - c \right)^{\frac{\sigma}{1+\sigma}}} \left[ \frac{G(e_t^{K^{**}})}{e_t^{K^{**}}} - c \right]^{\frac{1}{1+\sigma}}$$

We have:

$$\hat{e}_t^{K^{**}} < \tilde{e}_t^{K^{**}} \iff \Gamma(e_t^{K^{**}}) < \Omega(e_t^{K^{**}})$$

We have, at  $e_t^{K^{**}} = 0$ :

$$0 = \left[ 0 + \frac{\left( \frac{G((1-q_t)e_t^{N^{**}})}{(1-q_t)e_t^{N^{**}}} - c \right)}{\psi \left( \frac{G((1-q_t)e_t^{N^{**}})}{(1-q_t)e_t^{N^{**}}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

We thus have  $\hat{e}_t^{N^{**}}$  such that

$$\frac{G((1-q_t)\hat{e}_t^{N^{**}})}{(1-q_t)\hat{e}_t^{N^{**}}} - c = 0$$

Take the other expression.

$$e_t^{N^{**}} = \left[ \frac{\left( \frac{G(E_t^{N^{**}})}{E_t^{N^{**}}} - c \right)}{\psi \left( \frac{G(E_t^{N^{**}})}{E_t^{N^{**}}} - c \right)^\sigma} \right]^{\frac{1}{1+\sigma}}$$

We have, at  $e_t^{K^{**}} = 0$ :

$$e_t^{N^{**}} = \left[ \frac{\left( \frac{G((1-q_t)e_t^{N^{**}})}{(1-q_t)e_t^{N^{**}}} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}}$$

We thus have  $\tilde{e}_t^{N^{**}}$  such that:

$$\tilde{e}_t^{N^{**}} - \left[ \frac{\left( \frac{G((1-q_t)\tilde{e}_t^{N^{**}})}{(1-q_t)\tilde{e}_t^{N^{**}}} - c \right)^{1-\sigma}}{\psi} \right]^{\frac{1}{1+\sigma}} = 0$$

Let us compare  $\hat{e}_t^{N^{**}}$  and  $\tilde{e}_t^{N^{**}}$ . Let us suppose that  $\tilde{e}_t^{N^{**}} \geq \hat{e}_t^{N^{**}}$ . We thus have

$$\frac{G((1-q_t)\tilde{e}_t^{N^{**}})}{(1-q_t)\tilde{e}_t^{N^{**}}} - c \leq \frac{G((1-q_t)\hat{e}_t^{N^{**}})}{(1-q_t)\hat{e}_t^{N^{**}}} - c = 0$$

A contradiction is reached. We thus have:

$$\tilde{e}_t^{N^{**}} < \hat{e}_t^{N^{**}} \iff \Gamma(0) > \Omega(0)$$

In sum, given that  $\Gamma(e_t^{K^{**}}) < \Omega(e_t^{K^{**}})$  and  $\Gamma(0) > \Omega(0)$ , and given that those functions are continuous and strictly decreasing, we can conclude that the two reaction functions intersect at least once.

Let us turn back to the intermediate assumptions according to which:

$$\begin{aligned} & \left[ \begin{array}{l} (1 - q_t) G''(E_t^{**}) \psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma \\ - \left( \frac{[G'(E_t^{**}) - c] \psi \sigma}{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^{1-\sigma}} \right) \\ (1 - q_t) \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{E_t^{**}} \end{array} \right] < 0 \\ & \left[ q_t G''(E_t^{**}) \psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma - \left( \frac{[G'(E_t^{**}) - c] \psi \sigma}{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^{1-\sigma}} q_t \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{E_t^{**}} \right) \right] < 0 \end{aligned}$$

By factorization, we only need:

$$\left[ G''(E_t^{**}) \psi \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^\sigma - \frac{[G'(E_t^{**}) - c] \psi \sigma \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right]}{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)^{1-\sigma} E_t^{**}} \right] < 0$$

for all  $e_t^{K^{**}} \leq e_t^{N^{**}}$ . This can be rewritten as:

$$G''(\cdot) < [G'(E_t^{**}) - c] \sigma \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{G(E_t^{**}) - c E_t^{**}}$$

This can be rewritten as:

$$\begin{aligned} & \frac{G''(E_t^{**})}{[G'(E_t^{**}) - c] \frac{[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}]}{G(E_t^{**}) - c E_t^{**}}} > \sigma \\ & \frac{G''(E_t^{**})}{[G'(E_t^{**}) - c] \frac{[G'(E_t^{**})][E_t^{**}] - G(E_t^{**})}{\left[ \frac{G(E_t^{**})}{E_t^{**}} - c \right] [E_t^{**}]^2}} > \sigma \\ & \frac{G''(E_t^{**})}{\frac{[G'(E_t^{**}) - c]}{[E_t^{**}]^2} \left[ \frac{G'(E_t^{**})[E_t^{**}] - G(E_t^{**})}{\frac{G(E_t^{**})}{E_t^{**}} - c} \right]} > \sigma \end{aligned}$$

Given that  $\frac{[G'(E_t^{**}) - c]}{\frac{G(E_t^{**})}{E_t^{**}} - c} < 1$ , a sufficient condition is:

$$\frac{E_t^{**} G''(E_t^{**})}{\left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right]} > \sigma$$

Note that, under  $\sigma = 0$ , that condition is always satisfied, implying the existence of a symmetric  $q$  equilibrium with Impure Kantians.

### 9.1.2 Uniqueness

Let us derive conditions under which the symmetric  $q$  equilibrium with Impure Kantians is unique in the special case where  $\sigma = 0$ . The FOCs are:

$$\left(\frac{e_t^{K**} q_t}{E_t^{**}}\right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{K**} = 0$$

$$\frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{N**} = 0$$

True, we have, from the previous proof, that:

$$e_t^{K**} \equiv \Omega(e_t^{N**})$$

with  $\Omega'(e_t^{N**}) < 0$  and

$$e_t^{K**} \equiv \Gamma(e_t^{N**})$$

with  $\Gamma'(e_t^{N**}) < 0$ .

Existence always holds at  $\sigma = 0$ . Indeed, we have:

$$\begin{aligned} \Gamma(0) &> \Omega(0) \\ \Gamma(e_t^{K**}) &< \Omega(e_t^{K**}) \end{aligned}$$

Uniqueness unfortunately does not follow merely from  $\Omega'(e_t^{N**}) < 0$  and  $\Gamma'(e_t^{N**}) < 0$ .

A condition for uniqueness of the equilibrium is that the slope of the  $\Omega(\cdot)$  function is higher in absolute value than the slope of the  $\Gamma(\cdot)$  function at any intersection.

Remind that:

$$\frac{G(q_t e_t^{K**} + (1 - q_t) e_t^{N**})}{q_t e_t^{K**} + (1 - q_t) e_t^{N**}} - c - \psi e_t^{N**} = 0$$

defines  $e_t^{K**} = \Omega(e_t^{N**})$ .

We look for  $\Omega'(e_t^{N**})$ . We have:

$$\Omega'(e_t^{N**}) = \frac{de_t^{K**}}{de_t^{N**}} = - \frac{d \left[ \frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{N**} \right]}{d \left[ \frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{N**} \right]}$$

Hence

$$\Omega' (e_t^{N^{**}}) = \frac{de_t^{K^{**}}}{de_t^{N^{**}}} = - \frac{\left[ \frac{G'(E_t^{**})(1-q_t)E_t^{**} - G(E_t^{**})(1-q_t)}{(E_t^{**})^2} - \psi \right]}{\left[ \frac{G'(E_t^{**})q_t E_t^{**} - G(E_t^{**})q_t}{(E_t^{**})^2} \right]}$$

Hence

$$\Omega' (e_t^{N^{**}}) = \frac{de_t^{K^{**}}}{de_t^{N^{**}}} = - \frac{\left[ (1-q_t) \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{E_t^{**}} - \psi \right]}{q_t \left[ \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{E_t^{**}} \right]}$$

Let us now consider the slope of the  $\Gamma (e_t^{N^{**}})$ :

$$\frac{q_t e_t^{K^{**}}}{E_t^{**}} [G'(E_t^{**}) - c] + \frac{((1-q_t)e_t^{N^{**}})}{E_t^{**}} \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right) - \psi e_t^{K^{**}} = 0$$

Hence:

$$\Gamma' (e_t^{N^{**}}) = \frac{de_t^{K^{**}}}{de_t^{N^{**}}} = - \frac{d \left[ \frac{(e_t^{K^{**}} q_t)}{E_t^{**}} [G'(E_t^{**}) - c] + \frac{((1-q_t)e_t^{N^{**}})}{E_t^{**}} \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right) - \psi e_t^{K^{**}} \right]}{d \left[ \frac{(e_t^{K^{**}} q_t)}{E_t^{**}} [G'(E_t^{**}) - c] + \frac{((1-q_t)e_t^{N^{**}})}{E_t^{**}} \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right) - \psi e_t^{K^{**}} \right]}$$

Hence:

$$\Gamma' (e_t^{N^{**}}) = - \frac{\left[ \begin{aligned} & -\frac{(e_t^{K^{**}} q_t)(1-q_t)}{(E_t^{**})^2} [G'(E_t^{**}) - c] \\ & + \frac{(e_t^{K^{**}} q_t)}{E_t^{**}} G''(E_t^{**}) (1-q_t) \\ & + \frac{((1-q_t))E_t^{**} - (1-q_t)e_t^{N^{**}}(1-q_t)}{(E_t^{**})^2} \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right) \\ & + \frac{((1-q_t)e_t^{N^{**}})}{E_t^{**}} \left( \frac{G'(E_t^{**})(1-q_t)E_t^{**} - G(E_t^{**})(1-q_t)}{(E_t^{**})^2} \right) \end{aligned} \right]}{\left[ \begin{aligned} & \frac{q_t E_t^{**} - q_t e_t^{K^{**}} q_t}{(E_t^{**})^2} [G'(E_t^{**}) - c] \\ & + \frac{(e_t^{K^{**}} q_t)}{E_t^{**}} [G''(E_t^{**}) q_t] + \frac{-((1-q_t)e_t^{N^{**}}) q_t}{E_t^{**2}} \left( \frac{G(E_t^{**})}{E_t^{**}} - c \right) \\ & + \frac{((1-q_t)e_t^{N^{**}})}{E_t^{**}} \left( \frac{G'(E_t^{**}) q E_t^{**} - G(E_t^{**}) q}{E_t^{**2}} \right) - \psi \end{aligned} \right]}$$

Hence

$$\Gamma' (e_t^{N^{**}}) = - \frac{\left[ \begin{aligned} & \frac{(e_t^{K^{**}} q_t)}{E_t^{**}} G''(E_t^{**}) (1-q_t) \\ & + \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \frac{(1-q_t)}{(E_t^{**})^2} [-e_t^{K^{**}} q_t + (1-q_t)e_t^{N^{**}}] \end{aligned} \right]}{\left[ \frac{2(1-q)e_t^{N^{**}} q_t}{(E_t^{**})^2} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{(e_t^{K^{**}} q_t)}{E_t^{**}} [G''(E_t^{**}) q_t] - \psi \right]}$$

We now compare this slope with the slope of  $\Omega'(\cdot)$ :

$$\begin{aligned}\Omega'(e_t^{N^{**}}) &= -\frac{\left[(1-q_t)\left(G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right) - \psi E_t^{**}\right]}{q_t \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right]} \\ \Gamma'(e_t^{N^{**}}) &= -\frac{\left[\left(e_t^{K^{**}} q_t\right) G''(E_t^{**}) (1-q_t) + \left(G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right) \frac{(1-q_t)}{E_t^{**}} \left[-e_t^{K^{**}} q_t + (1-q_t) e_t^{N^{**}}\right]\right]}{\frac{2(1-q) e_t^{N^{**}} q_t}{E_t^{**}} \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right] + q_t e_t^{K^{**}} [G''(E_t^{**}) q_t] - \psi E_t^{**}}\end{aligned}$$

Remind that, at the equilibrium, we have:  $\Omega(e_t^{N^{**}}) = \Gamma(e_t^{N^{**}})$ . Uniqueness is achieved when  $|\Omega'(e_t^{N^{**}})| > |\Gamma'(e_t^{N^{**}})|$ .

Note that, at  $q = 0$ , we have  $|\Omega'(e_t^{N^{**}})| = +\infty > |\Gamma'(e_t^{N^{**}})|$ , we then have uniqueness.

When  $q = 1$ , we have:  $|\Omega'(e_t^{N^{**}})| = \frac{\psi E_t^{**}}{\left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right]} > |\Gamma'(e_t^{N^{**}})| = 0$ .

We thus also have uniqueness.

Is this true for all  $q > 0$ ?

Let us prove that  $\Omega'(e_t^{N^{**}}) - \Gamma'(e_t^{N^{**}}) < 0$ .

We have  $\Omega'(e_t^{N^{**}}) - \Gamma'(e_t^{N^{**}})$  equal to:

$$\begin{aligned}& -\frac{\left[(1-q_t)\left(G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right) - \psi E_t^{**}\right]}{q_t \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right]} \\ & + \frac{\left[\left(e_t^{K^{**}} q_t\right) G''(E_t^{**}) (1-q_t) + \left(G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right) \frac{(1-q_t)\left[-e_t^{K^{**}} q_t + (1-q_t) e_t^{N^{**}}\right]}{E_t^{**}}\right]}{\frac{2(1-q) e_t^{N^{**}} q_t}{E_t^{**}} \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right] + q_t e_t^{K^{**}} [G''(E_t^{**}) q_t] - \psi E_t^{**}}\end{aligned}$$

Let us rewrite this with a common denominator:

$$\begin{aligned}& \frac{\left[(1-q_t)\left(G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right) - \psi E_t^{**}\right] \left[\frac{2(1-q) e_t^{N^{**}} q_t}{E_t^{**}} \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right] + q_t e_t^{K^{**}} [G''(E_t^{**}) q_t] - \psi E_t^{**}\right]}{q_t \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right] \left[\frac{2(1-q) e_t^{N^{**}} q_t}{E_t^{**}} \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right] + q_t e_t^{K^{**}} [G''(E_t^{**}) q_t] - \psi E_t^{**}\right]} \\ & + \frac{\left[\left(e_t^{K^{**}} q_t\right) G''(E_t^{**}) (1-q_t) + \left(G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right) \frac{(1-q_t)}{E_t^{**}} \left[-e_t^{K^{**}} q_t + (1-q_t) e_t^{N^{**}}\right]\right] q_t \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right]}{\left[\frac{2(1-q) e_t^{N^{**}} q_t}{E_t^{**}} \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right] + q_t e_t^{K^{**}} [G''(E_t^{**}) q_t] - \psi E_t^{**}\right] q_t \left[G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}\right]}\end{aligned}$$

The denominator is positive. Hence the whole expression is positive if and



only if:

$$\begin{aligned}
& - \left[ (1 - q_t) \left( G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right) - \psi E_t^{**} \right] \left[ \frac{2(1-q)e_t^{N^{**}} q_t}{E_t^{**}} \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] \right. \\
& \left. + q_t e_t^{K^{**}} [G'' (E_t^{**}) q_t] - \psi E_t^{**} \right] \\
& + \left[ \left[ (e_t^{K^{**}} q_t) G'' (E_t^{**}) (1 - q_t) + \left( G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right) \frac{(1-q_t)[-e_t^{K^{**}} q_t + (1-q_t)e_t^{N^{**}}]}{E_t^{**}} \right] \right. \\
& \left. q_t \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] \right] \\
& < 0
\end{aligned}$$

Let us now factorize and develop:

$$\begin{aligned}
& \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] \left[ \begin{aligned} & - \frac{2(1-q)^2 e_t^{N^{**}} q_t}{E_t^{**}} \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] + \psi E_t^{**} \frac{2(1-q)e_t^{N^{**}} q_t}{E_t^{**}} \\ & - (1 - q_t) q_t e_t^{K^{**}} [G'' (E_t^{**}) q_t] + (e_t^{K^{**}} q_t) G'' (E_t^{**}) q_t (1 - q_t) \\ & + \left( G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right) \frac{(1-q_t)q_t}{E_t^{**}} [-e_t^{K^{**}} q_t + (1 - q_t)e_t^{N^{**}}] \end{aligned} \right] \\
& + \psi E_t^{**} \left[ q_t e_t^{K^{**}} [G'' (E_t^{**}) q_t] - \psi E_t^{**} + (1 - q_t) \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] \right] \\
& < 0
\end{aligned}$$

Hence

$$\begin{aligned}
& \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] \left[ \begin{aligned} & \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] (1 - q_t) q_t \left( - \frac{(1-q)e_t^{N^{**}} + q_t e_t^{K^{**}}}{E_t^{**}} \right) \\ & + \psi E_t^{**} \frac{2(1-q)e_t^{N^{**}} q_t}{E_t^{**}} \end{aligned} \right] \\
& + \psi E_t^{**} \left[ q_t e_t^{K^{**}} [G'' (E_t^{**}) q_t] - \psi E_t^{**} + (1 - q_t) \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] \right] \\
& < 0
\end{aligned}$$

Hence

$$\begin{aligned}
& \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] \left[ \begin{aligned} & \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] (1 - q_t) q_t (-1) + \psi E_t^{**} \frac{2(1-q)e_t^{N^{**}} q_t}{E_t^{**}} \\ & + \psi E_t^{**} \left[ q_t e_t^{K^{**}} [G'' (E_t^{**}) q_t] - \psi E_t^{**} + (1 - q_t) \left[ G' (E_t^{**}) - \frac{G (E_t^{**})}{E_t^{**}} \right] \right] \end{aligned} \right] \\
& < 0
\end{aligned}$$

The first term is negative. The second one is also negative.

Hence we always have, for any  $q_t$ , that  $\Omega' (e_t^{N^{**}}) - \Gamma' (e_t^{N^{**}}) < 0$  so that  $|\Omega' (e_t^{N^{**}})| - |\Gamma' (e_t^{N^{**}})| > 0$ . We thus have the uniqueness of the  $q$  equilibrium with Impure Kantians.

## 9.2 Proof of Proposition 5

The dynamics of  $q_t$  is given by:

$$q_{t+1} = q_t + q_t(1-q_t) \left[ \begin{array}{l} \frac{(1-q_t)}{\kappa} \left[ G(e_t^{K*}) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} - \left[ G(e_t^{N*}) - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} \right] \right] \\ - \frac{q_t}{\kappa} \left[ \frac{e_t^N G(E_t^*)}{E_t^*} - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} - \left[ \frac{e_t^{K*} G(E_t^*)}{E_t^*} - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} \right] \right] \end{array} \right]$$

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

Regarding the existence of an intermediate stationary equilibrium, let us first notice that, at such an intermediate equilibrium, we have:

$$\begin{aligned} & \frac{(1-q_t)}{\kappa} \left[ G(e_t^{K*}) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} - \left[ G(e_t^{N*}) - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} \right] \right] \\ = & \frac{q_t}{\kappa} \left[ \frac{e_t^N G(E_t^*)}{E_t^*} - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} - \left[ \frac{e_t^{K*} G(E_t^*)}{E_t^*} - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} \right] \right] \end{aligned}$$

In other words, socialization efforts are equal for the two types. Some simplifications yield:

$$\begin{aligned} & G(e_t^{K*}) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} - G(e_t^{N*}) + ce_t^{N*} + \frac{\psi(e_t^{N*})^2}{2} \\ & - q_t G(e_t^{K*}) + q_t ce_t^{K*} + q_t \frac{\psi(e_t^{K*})^2}{2} + q_t G(e_t^{N*}) - q_t ce_t^{N*} - q_t \frac{\psi(e_t^{N*})^2}{2} \\ = & \frac{q_t e_t^{N*} G(E_t^*)}{E_t^*} - q_t ce_t^{N*} - \frac{q_t \psi(e_t^{N*})^2}{2} - \frac{q_t e_t^{K*} G(E_t^*)}{E_t^*} + q_t ce_t^{K*} + q_t \frac{\psi(e_t^{K*})^2}{2} \end{aligned}$$

Hence

$$\frac{[G(e_t^{K*}) - G(e_t^{N*})](1-q_t)}{[e_t^{N*} - e_t^{K*}]} + c + \frac{\psi[e_t^{N*} + e_t^{K*}]}{2} = \frac{q_t G(q_t e_t^{K*} + (1-q_t)e_t^{N*})}{q_t e_t^{K*} + (1-q_t)e_t^{N*}}$$

Let us now substitute for optimal animal levels. Under  $\sigma = 0$ , we have:

$$\begin{aligned} e_t^{K*} &= \frac{(G'(e_t^{K*}) - c)}{\psi} \\ e_t^{N*} &= \frac{\left(\frac{G(E_t^*)}{E_t^*} - c\right)}{\psi} \end{aligned}$$

Hence

$$\frac{[G(e_t^{K*}) - G(e_t^{N*})](1-q_t)}{\psi[e_t^{N*} - e_t^{K*}]} + \frac{[e_t^{N*} + e_t^{K*}]}{2} = \frac{q_t \frac{G(E_t^*)}{E_t^*} - c}{\psi}$$

Remind that  $e_t^{K^*}$  does not depend on  $q_t$ , and can thus be treated as a parameter. Thus we have two equations with two unknowns. Whether a stationary equilibrium exists or not depends on whether there exists some pair  $(q_t, e_t^{N^*})$  satisfying the equations:

$$e_t^{N^*} = \frac{\left(\frac{G(E_t^*)}{E_t^*} - c\right)}{\psi}$$

$$\frac{[G(e_t^{K^*}) - G(e_t^{N^*})](1 - q_t)}{\psi [e_t^{N^*} - e_t^{K^*}]} + \frac{[e_t^{N^*} + e_t^{K^*}]}{2} = \frac{q_t \frac{G(E_t^*)}{E_t^*} - c}{\psi}$$

From the first equation, we know that  $e_t^{N^*}$  is an implicit function of  $q_t$ :

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

We use the implicit function theorem to obtain the derivative of  $e_t^{N^*}$  wrt  $q_t$ :

$$\frac{de_t^{N^*}}{dq_t} = - \frac{-\frac{G'(E_t^*)[e_t^{K^*} - e_t^{N^*}]E_t^* - G(E_t^*)[e_t^{K^*} - e_t^{N^*}]}{\psi[E_t^*]^2}}{1 - \frac{G'(E_t^*)(1 - q_t)E_t^* - G(E_t^*)(1 - q_t)}{\psi[E_t^*]^2}}$$

Further simplifications yield:

$$\frac{de_t^{N^*}}{dq_t} = \frac{-[e_t^{N^*} - e_t^{K^*}] \frac{G'(E_t^*) - \frac{G(E_t^*)}{E_t^*}}{\psi[E_t^*]}}{1 - (1 - q_t) \frac{G'(E_t^*) - \frac{G(E_t^*)}{E_t^*}}{\psi[E_t^*]}}$$

Note that, because of the concavity of  $G(\cdot)$ , we know that  $\frac{G'(E_t^*) - \frac{G(E_t^*)}{E_t^*}}{\psi[E_t^*]} < 0$ . Moreover, given that  $e_t^{N^*} > e_t^{K^*}$ , we know that the numerator is positive. Furthermore, we also know that the denominator is also positive. Hence, without ambiguity, we obtain:

$$\frac{de_t^{N^*}}{dq_t} > 0$$

The more numerous the Pure Kantians are, and the more animals Nash players purchase. We thus have the following function:

$$e_t^N \equiv e(q_t)$$

with  $e'(q_t) > 0$ .

Let us now take the other equation, and substitute for  $e(q_t)$ .

$$\frac{[G(e_t^{K^*}) - G(e(q_t))](1 - q_t)}{\psi [e(q_t) - e_t^{K^*}]} + \frac{[e(q_t) + e_t^{K^*}]}{2} - \frac{q_t \frac{G(q_t e_t^{K^*} + (1 - q_t)e(q_t))}{q_t e_t^{K^*} + (1 - q_t)e(q_t)} - c}{\psi} = 0$$

Let us now consider the case where  $q_t = 0$ . We know, from the equations for optimal socialization efforts, that:

$$\begin{aligned}\tau_t^N &= 0 \\ \tau_t^K &= \frac{1}{\kappa} \left[ G(e_t^{K*}) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} - \left[ G(e(0)) - ce(0) - \frac{\psi(e(0))^2}{2} \right] \right]\end{aligned}$$

Pure Kantians behave like a social planner, and chooses the number of animals that, generalized to all players, will bring the highest individual utility. On the contrary, when  $q = 0$ , Nash players all chose the same number of animals, but suffer from a coordination failure. As a consequence, they tend to have too many animals. As a consequence, the utility of each Nash player when there is no Pure Kantian player is necessarily *inferior* to the utility obtained by Pure Kantians. This explains why, at  $q = 0$ , the socialization effort chosen by Pure Kantians is higher than the one chosen by Nash players. Hence the term in brackets

$$G(e_t^{K*}) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} - \left[ G(e(0)) - ce(0) - \frac{\psi(e(0))^2}{2} \right] > 0$$

Therefore  $\tau_t^K(0) > 0$ . Hence, Pure Kantians survive when they are in ultra minority. The survival of Pure Kantians when  $q = 0$  is due to two things. First, the standard cultural substitution effect: when  $q = 0$ , Nash parents rely on the society for the transmission of their trait, so that  $\tau_t^N = 0$ . Second, the effort chosen by Pure Kantians parents is strictly positive  $\tau_t^K > 0$ .

Let us now consider the case where  $q_t = 1$ . The socialization effort for the Kantian is now:

$$\tau_t^K = 0$$

The Pure Kantian does not invest any effort in socialization, because he relies entirely on the rest of the society. This is again the standard cultural substitution effect.

Regarding the socialization investment of the Nash player, note that, for any  $q_t$ , his socialization effort is:

$$\tau_t^N = \frac{q_t}{\kappa} \left[ \frac{e_t^{N*}}{E_t^*} G(E_t^*) - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} - \left[ \frac{e_t^{K*}}{E_t^*} G(E_t^*) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} \right] \right]$$

Since  $e_t^N$  maximizes, by definition, the function  $U(x) \equiv \frac{xG(E_t^*)}{E_t^*} - cx - \frac{\psi(x)^2}{2}$  for any level of  $e_t^{K*}$  and  $q_t$ , we have, for any  $e_t^{K*} \neq e_t^{N*}$ , that the first term in brackets exceeds strictly the second term in brackets. Hence we have:

$$\tau_t^N > 0$$

To be sure, we can rewrite the socialization effort of the Nash parent as:

$$\tau_t^N = \frac{q_t}{\kappa} \left[ (e_t^{N*} - e_t^{K*}) \left( \frac{G(E_t^*)}{E_t^*} - c \right) + \frac{\psi}{2} \left[ (e_t^{K*})^2 - (e_t^{N*})^2 \right] \right]$$

Let us now substitute for

$$e_t^{N*} = \frac{G(E_t^*) - c}{\psi}$$

this yields:

$$\tau_t^N = \frac{q_t}{\kappa} \left[ \frac{\psi}{2} (e_t^{N*} - e_t^{K*})^2 \right] > 0$$

Thus Nash parents always invest a positive effort in socialization when  $q_t > 0$ .

In the case where  $q_t \rightarrow 1$ , we have:

$$\tau_t^N = \frac{1}{\kappa} \left[ \frac{\psi}{2} (e_t^{N*} - e_t^{K*})^2 \right]$$

Given that

$$\begin{aligned} \tau_t^K - \tau_t^N &> 0 \text{ when } q_t = 0 \\ \tau_t^K - \tau_t^N &< 0 \text{ when } q_t = 1 \end{aligned}$$

we know, by continuity, that there exists at least one level of  $q_t \in ]0, 1[$  such that  $\tau_t^K - \tau_t^N = 0$ , that is, at least one intermediate stationary equilibrium.

### 9.3 Proof of Proposition 6

The dynamic equation for  $q_t$  is:

$$q_{t+1} = q_t + q_t(1 - q_t) \begin{bmatrix} \frac{(1-q_t)}{\kappa} \left[ \begin{array}{l} G(e_t^{K*}) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} \\ - \left[ G(e_t^{N*}) - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} \right] \end{array} \right] \\ - \frac{q_t}{\kappa} \left[ \begin{array}{l} \frac{e_t^{N*}G(E_t^*)}{E_t^*} - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} \\ - \left[ \frac{e_t^{K*}G(E_t^*)}{E_t^*} - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} \right] \end{array} \right] \end{bmatrix}$$

Regarding the uniqueness of the intermediate stationary equilibrium, it should be stressed that a necessary and sufficient condition for the uniqueness of an intermediate stationary equilibrium is that the derivative of the term in brackets is negative at any intermediate stationary equilibrium. Indeed, if it were positive, then there would be at least 3 intermediate equilibria.

Let us thus differentiate the bracket (equal to  $\tau_t^K - \tau_t^N$ ) with respect to  $q_t$ .

This yields:

$$\begin{aligned}
& \frac{-1}{\kappa} \left[ G(e_t^{K*}) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} - \left[ G(e_t^{N*}) - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} \right] \right] \\
& + \frac{(1-q_t)}{\kappa} \left[ -\frac{de_t^{N*}}{dq_t} [G'(e_t^{N*}) - c - \psi(e_t^{N*})] \right] \\
& - \frac{1}{\kappa} \left[ \frac{e_t^{N*} G(E_t^*)}{E_t^*} - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} - \left[ \frac{e_t^{K*} G(E_t^*)}{E_t^*} - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} \right] \right] \\
& - \frac{q_t}{\kappa} \left[ \frac{\begin{aligned} & \left[ \frac{de_t^{N*}}{dq_t} G(E_t^*) \right. \\ & \left. + e_t^{N*} G'(E_t^*) \left( e_t^{K*} - e_t^{N*} + (1-q_t) \frac{de_t^{N*}}{dq_t} \right) \right] (E_t^*) \\ & - e_t^{N*} G(E_t^*) \left( e_t^{K*} - e_t^{N*} + (1-q_t) \frac{de_t^{N*}}{dq_t} \right) \end{aligned}}{[E_t^*]^2} \right. \\
& \left. - c \frac{de_t^{N*}}{dq_t} - \psi(e_t^{N*}) \frac{de_t^{N*}}{dq_t} \right. \\
& \left. - \left[ \frac{\begin{aligned} & e_t^{K*} G'(E_t^*) \left( e_t^{K*} - e_t^{N*} + (1-q_t) \frac{de_t^{N*}}{dq_t} \right) (E_t^*) \\ & - \left( e_t^{K*} - e_t^{N*} + (1-q_t) \frac{de_t^{N*}}{dq_t} \right) e_t^{K*} G(E_t^*) \end{aligned}}{[E_t^*]^2} \right] \right]
\end{aligned}$$

A necessary and sufficient condition for the uniqueness of the intermediate stationary equilibrium is that this derivative is negative at the intermediate stationary equilibrium.

Remind that, at the intermediate stationary equilibrium, socialization efforts are equal. Hence we have:

$$\begin{aligned}
& \frac{(1-q_t)}{\kappa} \left[ G(e_t^{K*}) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} - \left[ G(e_t^N) - ce_t^N - \frac{\psi(e_t^N)^2}{2} \right] \right] \\
& = \frac{q_t}{\kappa} \left[ \frac{e_t^N G(E_t^*)}{E_t^*} - ce_t^N - \frac{\psi(e_t^N)^2}{2} - \left[ \frac{e_t^{K*} G(E_t^*)}{E_t^*} - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} \right] \right]
\end{aligned}$$

Hence, the previous expression can be simplified to:

$$\begin{aligned}
& \left[ \frac{-1}{\kappa q} \right] \left[ G(e_t^{K*}) - ce_t^{K*} - \frac{\psi(e_t^{K*})^2}{2} - \left[ G(e_t^{N*}) - ce_t^{N*} - \frac{\psi(e_t^{N*})^2}{2} \right] \right] \\
& + \frac{(1-q_t)}{\kappa} \left[ -\frac{de_t^{N*}}{dq_t} [G'(e_t^{N*}) - c - \psi(e_t^{N*})] \right] \\
& - \frac{q_t}{\kappa} \left[ \begin{aligned} & \left[ \begin{aligned} & \frac{de_t^{N*}}{dq_t} G(E_t^*) \\ & + e_t^{N*} G'(E_t^*) \left( e_t^{K*} - e_t^{N*} + (1-q_t) \frac{de_t^{N*}}{dq_t} \right) \end{aligned} \right] (E_t^*) \\ & - e_t^{N*} G(E_t^*) \left( e_t^{K*} - e_t^{N*} + (1-q_t) \frac{de_t^{N*}}{dq_t} \right) \\ & \frac{[E_t^*]^2}{\left( \begin{aligned} & e_t^{K*} G'(E_t^*) \left( e_t^{K*} - e_t^{N*} + (1-q_t) \frac{de_t^{N*}}{dq_t} \right) (E_t^*) \\ & - \left( e_t^{K*} - e_t^{N*} + (1-q_t) \frac{de_t^{N*}}{dq_t} \right) e_t^{K*} G(E_t^*) \end{aligned} \right)} \right] \\ & - \left[ \begin{aligned} & -c \frac{de_t^{N*}}{dq_t} - \psi(e_t^{N*}) \frac{de_t^{N*}}{dq_t} \end{aligned} \right] \end{aligned} \right]
\end{aligned}$$

The first term, which is negative, captures the effect of cultural substitution on the socialization effort of both players. It reduces the effort of Pure Kantian and it raises the effort of Nash. The second term captures the impact of  $q$  on the relative utility gain to have a Pure Kantian child for Pure Kantians. That second term is positive. The third term captures the impact of  $q$  on the relative utility gain of having a Nash child for Nash parents. The sign of that third term is ambiguous, but we can decompose and simplify that third term to see the opposite effects at work.

We know that a rise in  $q$  will necessarily increase  $\Delta V^K$ , but its impact on  $\Delta V^N$  is ambiguous. The ambiguity comes from the fact that the change in  $e^{N*}$  (i.e. a rise when  $q$  rises) affects both the utility of having a Pure Kantian child and of having a Nash child. Two effects are at work.

To isolate the effect of a rise in  $q$  for a given  $e^{N*}$ , let us set  $\frac{de_t^{N*}}{dq_t} = 0$  in the third term. We then obtain that, if  $e^{N*}$  was given, the rise in  $q$  would raise  $\Delta V^N = V^{NN} - V^{NK}$  :

$$\left[ - [e_t^{N*} - e_t^{K*}]^2 \frac{[G'(E_t^*)] - \frac{G(E_t^*)}{E_t^*}}{E_t^*} \right] > 0$$

Thus Nash parents gain more than Pure Kantian parents when  $q$  is raised for a given  $e^{N*}$ . Indeed, a higher proportion of Kantians implies more space per animal for all, but as Nash players have more animals, this is more beneficial for them.

The second effect arises when  $q$  is given but Nash players adapt the number of animals  $e^{N*}$ . To isolate that effect, let us now delete all terms of the third

term that are not in  $\frac{de_t^{N^*}}{dq_t}$ . We then obtain:

$$\frac{q_t}{\kappa} \frac{de_t^{N^*}}{dq_t} \left[ \frac{G(E_t^*)}{E_t^*} - c - \psi e_t^{N^*} + (e_t^{N^*} - e_t^{K^*}) (1 - q_t) \frac{G'(E_t^*) - \frac{G(E_t^*)}{E_t^*}}{E_t^*} \right]$$

Using the FOC for Nash players:  $\frac{G(E_t^*)}{E_t^*} - c - \psi e_t^{N^*} = 0$ , we have:

$$\frac{q_t}{\kappa} \frac{de_t^{N^*}}{dq_t} \left[ (e_t^{N^*} - e_t^{K^*}) (1 - q_t) \frac{G'(E_t^*) - \frac{G(E_t^*)}{E_t^*}}{E_t^*} \right] < 0$$

Hence, when Nash players increase the number of animals, both Pure Kantians and Nash players have a decrease in utility, but the latter seem to lose more. Why? Again, the rise in  $e_t^{N^*}$  reduces the availability of space, and thus reduces the return on each animal. But as Nash players have more animals, they suffer from a larger utility loss.

Using that decomposition for the third term, we obtain the following condition for uniqueness of the intermediate stationary equilibrium:

$$\begin{aligned} & \underbrace{\left[ \frac{-1}{\kappa q^*} \left[ G(e^{K^*}) - ce^{K^*} - \frac{\psi(e^{K^*})^2}{2} - \left[ G(e^{N^*}) - ce^{N^*} - \frac{\psi(e^{N^*})^2}{2} \right] \right] \right]}_{-} \\ & + \underbrace{\frac{(1 - q^*)}{\kappa} \left[ -\frac{de^{N^*}}{dq^*} [G'(e^{N^*}) - c - \psi(e^{N^*})] \right]}_{+} \\ & - \underbrace{\frac{q^*}{\kappa} \left[ -[e^{N^*} - e^{K^*}]^2 \frac{G'(E^*) - \frac{G(E^*)}{E^*}}{E^*} \right]}_{-} \\ & - \underbrace{\frac{q^*}{\kappa} \frac{de^{N^*}}{dq^*} \left[ (e^{N^*} - e^{K^*}) (1 - q^*) \frac{G'(E^*) - \frac{G(E^*)}{E^*}}{E^*} \right]}_{+} \\ & < 0 \end{aligned}$$

## 9.4 Proof of Proposition 7

The dynamic equation is:

$$q_{t+1} = q_t + q_t(1 - q_t) \left[ \begin{array}{l} \frac{(1 - q_t)}{\kappa} \left[ \begin{array}{l} G(e_t^{K^*}) - ce_t^{K^*} - \frac{\psi(e_t^{K^*})^2}{2} \\ - \left[ G(e_t^{N^*}) - ce_t^{N^*} - \frac{\psi(e_t^{N^*})^2}{2} \right] \end{array} \right] \\ - \frac{q_t}{\kappa} \left[ \begin{array}{l} \frac{e_t^{N^*} G(E_t^*)}{E_t^*} - ce_t^{N^*} - \frac{\psi(e_t^{N^*})^2}{2} \\ - \left[ \frac{e_t^{K^*} G(E_t^*)}{E_t^*} - ce_t^{K^*} - \frac{\psi(e_t^{K^*})^2}{2} \right] \end{array} \right] \end{array} \right]$$



Differentiating this with respect to  $q_t$  yields:

$$\begin{aligned}
& 1 + (1 - 2q_t) \left[ \begin{aligned} & \frac{(1-q_t)}{\kappa} \left[ \begin{aligned} & G(e_t^{K^*}) - ce_t^{K^*} - \frac{\psi(e_t^{K^*})^2}{2} \\ & - \left[ G(e_t^{N^*}) - ce_t^{N^*} - \frac{\psi(e_t^{N^*})^2}{2} \right] \end{aligned} \right] \\ & - \frac{q_t}{\kappa} \left[ \frac{e_t^{N^*} G(E_t^*)}{E_t^*} - ce_t^{N^*} - \frac{\psi(e_t^{N^*})^2}{2} - \left[ \frac{e_t^{K^*} G(E_t^*)}{E_t^*} - ce_t^{K^*} - \frac{\psi(e_t^{K^*})^2}{2} \right] \right] \end{aligned} \right] \\
& + q_t(1 - q_t) \left[ \begin{aligned} & \left[ \frac{-1}{\kappa q} \right] \left[ G(e_t^{K^*}) - ce_t^{K^*} - \frac{\psi(e_t^{K^*})^2}{2} - \left[ G(e_t^{N^*}) - ce_t^{N^*} - \frac{\psi(e_t^{N^*})^2}{2} \right] \right] \\ & + \frac{(1-q)}{\kappa} \left[ -\frac{de_t^{N^*}}{dq} \left[ G'(e_t^{N^*}) - c - \psi(e_t^{N^*}) \right] \right] \\ & - \frac{q}{\kappa} \left[ - \left[ e_t^{N^*} - e_t^{K^*} \right]^2 \frac{G'(E_t^*) - \frac{G(E_t^*)}{E_t^*}}{E_t^*} \right] \\ & - \frac{q}{\kappa} \frac{de_t^{N^*}}{dq} \left[ \left( e_t^{N^*} - e_t^{K^*} \right) \left( 1 - q \right) \frac{G'(E_t^*) - \frac{G(E_t^*)}{E_t^*}}{E_t^*} \right] \end{aligned} \right]
\end{aligned}$$

When  $q = 0$ , the derivative is:

$$1 + \frac{1}{\kappa} \left[ G(e_t^{K^*}) - ce_t^{K^*} - \frac{\psi(e_t^{K^*})^2}{2} - \left[ G(e_t^N) - ce_t^N - \frac{\psi(e_t^N)^2}{2} \right] \right] > 1$$

Hence the equilibrium  $q = 0$  is unstable.

Take now the equilibrium  $q = 1$ . The derivative is:

$$\begin{aligned}
& 1 + \frac{1}{\kappa} \left[ \frac{e_t^{N^*} G(E_t^*)}{E_t^*} - ce_t^{N^*} - \frac{\psi(e_t^{N^*})^2}{2} - \left[ \frac{e_t^{K^*} G(E_t^*)}{E_t^*} - ce_t^{K^*} - \frac{\psi(e_t^{K^*})^2}{2} \right] \right] \\
= & 1 + \frac{1}{\kappa} \left[ e_t^{N^*} \left[ \frac{G(E_t^*)}{E_t^*} - c - \frac{\psi e_t^{N^*}}{2} \right] - \left[ \frac{e_t^{K^*} G(E_t^*)}{E_t^*} - ce_t^{K^*} - \frac{\psi(e_t^{K^*})^2}{2} \right] \right]
\end{aligned}$$

Using the FOC for Nash players:  $\frac{G(E_t^*)}{E_t^*} - c - \psi e_t^{N*} = 0$ , we have:

$$\begin{aligned}
& 1 + \frac{1}{\kappa} \left[ e_t^{N*} \frac{\psi e_t^{N*}}{2} - \left[ \frac{e_t^{K*} G(E_t^*)}{E_t^*} - c e_t^{K*} - \frac{\psi (e_t^{K*})^2}{2} \right] \right] \\
= & 1 + \frac{1}{\kappa} \left[ \frac{\psi (e_t^{N*})^2}{2} - \left[ \frac{e_t^{K*} G(E_t^*)}{E_t^*} - c e_t^{K*} - \frac{\psi (e_t^{K*})^2}{2} \right] \right] \\
= & 1 + \frac{1}{\kappa} \left[ \frac{\psi (e_t^{N*})^2}{2} - e_t^{K*} \left[ \frac{G(E_t^*)}{E_t^*} - c - \frac{\psi e_t^{K*}}{2} \right] \right] \\
= & 1 + \frac{1}{\kappa} \left[ \frac{\psi (e_t^{N*})^2}{2} - e_t^{K*} \left[ \psi e_t^{N*} - \frac{\psi e_t^{K*}}{2} \right] \right] \\
= & 1 + \frac{1}{\kappa} \left[ \frac{\psi (e_t^{N*})^2}{2} + \frac{\psi (e_t^{K*})^2}{2} - \psi e_t^{N*} e_t^{K*} \right] \\
= & 1 + \frac{1}{\kappa} \frac{\psi}{2} \left[ (e_t^{N*} - e_t^{K*})^2 \right] > 1
\end{aligned}$$

Hence the equilibrium  $q = 1$  is also unstable.

Finally, at the intermediate equilibrium, the derivative is:

$$1+q(1-q) \left[ \begin{aligned} & \left[ \frac{-1}{\kappa q} \right] \left[ G(e_t^{K*}) - c e_t^{K*} - \frac{\psi (e_t^{K*})^2}{2} - \left[ G(e_t^{N*}) - c e_t^{N*} - \frac{\psi (e_t^{N*})^2}{2} \right] \right] \\ & + \frac{(1-q)}{\kappa} \left[ -\frac{de_t^{N*}}{dq} \left[ G'(e_t^{N*}) - c - \psi (e_t^{N*}) \right] \right] \\ & - \frac{q}{\kappa} \left[ -[e_t^{N*} - e_t^{K*}]^2 \frac{G'(E_t^*) - \frac{G(E_t^*)}{E_t^*}}{E_t^*} \right] - \frac{q}{\kappa} \frac{de_t^{N*}}{dq} \left[ (e_t^{N*} - e_t^{K*}) (1-q) \frac{G'(E_t^*) - \frac{G(E_t^*)}{E_t^*}}{E_t^*} \right] \end{aligned} \right]$$

Local stability requires that expression to be less than 1 in absolute value. Note that, when the intermediate stationary equilibrium is unique, the factor in brackets is negative (by Proposition 6), and so the stability condition is necessarily satisfied.

In the case where the intermediate stationary equilibrium is unique, we know that, by the continuity of  $F(q_t)$ , the basin of attraction of the intermediate stationary equilibrium is equal to  $[0, 1]$ , implying that the equilibrium is globally stable.

## 9.5 Proofs of Propositions 8, 9 and 10

### 9.5.1 Toolbox

Let us first compute the derivatives of each term  $V^{KK}$ ,  $V^{KN}$ ,  $V^{NK}$  and  $V^{NN}$ .

$$\begin{aligned} V^{KK} &= \frac{e_t^{K**} G(E_t^{**})}{E_t^{**}} - ce_t^{K**} - \frac{\psi(e_t^{K**})^2}{2} \\ V^{KN} &= G(e_t^{N**}) - ce_t^{N**} - \frac{\psi(e_t^{N**})^2}{2} \\ V^{NK} &= \frac{e_t^{K**} G(E_t^{**})}{E_t^{**}} - ce_t^{K**} - \frac{\psi(e_t^{K**})^2}{2} \\ V^{NN} &= \frac{e_t^{N**} G(E_t^{**})}{E_t^{**}} - ce_t^{N**} - \frac{\psi(e_t^{N**})^2}{2} \end{aligned}$$

where

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

with derivatives

$$\begin{aligned} \frac{dE(q_t)}{dq_t} &= e_t^{K**} + q_t \frac{de_t^{K**}}{dq_t} - e_t^{N**} + (1 - q_t) \frac{de_t^{N**}}{dq_t} \\ \frac{d^2 E(q_t)}{dq_t^2} &= \frac{de_t^{K**}}{dq_t} + \frac{de_t^{K**}}{dq_t} + q_t \frac{de_t^{K**2}}{dq_t^2} - \frac{de_t^{N**}}{dq_t} - \frac{de_t^{N**}}{dq_t} + (1 - q_t) \frac{d^2 e_t^{N**}}{dq_t^2} \end{aligned}$$

The derivatives of  $V^{KK}$ ,  $V^{KN}$ ,  $V^{NK}$  and  $V^{NN}$  are:

$$\begin{aligned} \frac{dV^{KK}}{dq_t} &= \frac{de_t^{K**}}{dq_t} \left( \frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{K**} \right) \\ &\quad + \frac{dE(q_t)}{dq_t} \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \end{aligned}$$

Hence

$$\begin{aligned} \frac{dV^{KK}}{dq_t} &= \frac{de_t^{K**}}{dq_t} \left( \frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{K**} + q_t \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right) \\ &\quad + \left( e_t^{K**} - e_t^{N**} + (1 - q_t) \frac{de_t^{N**}}{dq_t} \right) \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \end{aligned}$$

Using the Envelope Theorem, we know that the term in brackets is equal to zero (it is the FOC of the Impure Kantians).

Hence:

$$\frac{dV^{KK}}{dq_t} = \left( e_t^{K**} - e_t^{N**} + (1 - q_t) \frac{de_t^{N**}}{dq_t} \right) \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)$$

Hence, at  $q = 0$ , we have:

$$\frac{dV^{KK}}{dq_t} = \left( \frac{de_t^{N^{**}}}{dq_t} \right) \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)$$

Let us now compute the second-order derivative:

$$\begin{aligned} \frac{d^2V^{KK}}{dq_t^2} &= \frac{d^2e_t^{K^{**}}}{dq_t^2} \left( \frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{K^{**}} + qt \frac{e_t^{K^{**}}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right) \\ &\quad + \frac{de_t^{K^{**}}}{dq_t} \left( \frac{\frac{dE(q_t)}{dq_t} \frac{1}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] - \psi \frac{de_t^{K^{**}}}{dq_t}}{\left( \frac{e_t^{K^{**}}}{E_t^{**}} + qt \frac{\frac{de_t^{K^{**}}}{dq_t} E_t^{**} - e_t^{K^{**}} \frac{dE(q_t)}{dq_t}}{E_t^{**2}} \right) \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)} \right. \\ &\quad \left. + qt \frac{e_t^{K^{**}}}{E_t^{**}} \frac{dE(q_t)}{dq_t} \left( G''(E_t^{**}) - \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right) \right) \\ &\quad + \left( \frac{de_t^{K^{**}}}{dq_t} - \frac{de_t^{N^{**}}}{dq_t} - \frac{de_t^{N^{**}}}{dq_t} + (1-qt) \frac{d^2e_t^{N^{**}}}{dq_t^2} \right) \left[ \frac{e_t^{K^{**}}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right] \\ &\quad + \left( e_t^{K^{**}} - e_t^{N^{**}} + (1-qt) \frac{de_t^{N^{**}}}{dq_t} \right) \left[ \frac{\frac{de_t^{K^{**}}}{dq_t} E_t^{**} - e_t^{K^{**}} \frac{dE(q_t)}{dq_t}}{E_t^{**2}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right. \\ &\quad \left. + \frac{e_t^{K^{**}}}{E_t^{**}} \left( G''(E_t^{**}) - \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right) \right] \end{aligned}$$

Using, again, the Envelope Theorem, we have:

$$\begin{aligned} \frac{d^2V^{KK}}{dq_t^2} &= \frac{de_t^{K^{**}}}{dq_t} \left( \frac{\frac{dE(q_t)}{dq_t} \frac{1}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] - \psi \frac{de_t^{K^{**}}}{dq_t}}{\left( \frac{e_t^{K^{**}}}{E_t^{**}} + qt \frac{\frac{de_t^{K^{**}}}{dq_t} E_t^{**} - e_t^{K^{**}} \frac{dE(q_t)}{dq_t}}{E_t^{**2}} \right) \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)} \right) \\ &\quad + \left( \frac{de_t^{K^{**}}}{dq_t} - \frac{de_t^{N^{**}}}{dq_t} - \frac{de_t^{N^{**}}}{dq_t} + (1-qt) \frac{d^2e_t^{N^{**}}}{dq_t^2} \right) \left[ \frac{e_t^{K^{**}}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right] \\ &\quad + \left( e_t^{K^{**}} - e_t^{N^{**}} + (1-qt) \frac{de_t^{N^{**}}}{dq_t} \right) \left[ \frac{\frac{de_t^{K^{**}}}{dq_t} E_t^{**} - e_t^{K^{**}} \frac{dE(q_t)}{dq_t}}{E_t^{**2}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right. \\ &\quad \left. + \frac{e_t^{K^{**}}}{E_t^{**}} \left( G''(E_t^{**}) - \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right) \right] \end{aligned}$$

Hence

$$\begin{aligned}
\frac{d^2V^{KK}}{dq_t^2} = & \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \left[ \begin{aligned} & \frac{de_t^{K**}}{dq_t} \frac{dE(q_t)}{dq_t} \frac{1}{E_t^{**}} \\ & + \frac{de_t^{K**}}{dq_t} \left( \frac{e_t^{K**}}{E_t^{**}} + qt \frac{\frac{de_t^{K**}}{dq_t} E_t^{**} - e_t^{K**} \frac{dE(q_t)}{dq_t}}{E_t^{**2}} \right) \\ & + \frac{de_t^{K**}}{dq_t} qt \frac{e_t^{K**}}{E_t^{**}} \frac{dE(q_t)}{dq_t} \frac{1}{E_t^{**}} \\ & + \frac{e_t^{K**}}{E_t^{**}} \left( \frac{de_t^{K**}}{dq_t} - \frac{de_t^{N**}}{dq_t} - \frac{de_t^{N**}}{dq_t} + (1-qt) \frac{d^2e_t^{N**}}{dq_t^2} \right) \\ & + \left[ \begin{aligned} & \left( e_t^{K**} - e_t^{N**} + (1-qt) \frac{de_t^{N**}}{dq_t} \right) \\ & \left( \frac{de_t^{K**}}{dq_t} E_t^{**} - e_t^{K**} \frac{dE(q_t)}{dq_t} + \frac{e_t^{K**}}{(E_t^{**})^2} \right) \end{aligned} \right] \end{aligned} \right] \\
& + \frac{de_t^{K**}}{dq_t} \left( -\psi \frac{de_t^{K**}}{dq_t} + qt \frac{e_t^{K**}}{E_t^{**}} \frac{dE(q_t)}{dq_t} (G''(E_t^{**})) \right) \\
& + \left( e_t^{K**} - e_t^{N**} + (1-qt) \frac{de_t^{N**}}{dq_t} \right) \left[ \frac{e_t^{K**}}{E_t^{**}} (G''(E_t^{**})) \right]
\end{aligned}$$

Note that  $q = 0$ , this simplifies to:

$$\begin{aligned}
\frac{d^2V^{KK}}{dq_t^2} = & \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \left[ \begin{aligned} & \frac{de_t^{K**}}{dq_t} \frac{dE(q_t)}{dq_t} \frac{1}{E_t^{**}} + \frac{de_t^{K**}}{dq_t} \left( \frac{e_t^{K**}}{E_t^{**}} \right) \\ & + \frac{e_t^{K**}}{E_t^{**}} \left( \frac{de_t^{K**}}{dq_t} - 2 \frac{de_t^{N**}}{dq_t} + \frac{d^2e_t^{N**}}{dq_t^2} \right) \\ & + \frac{de_t^{N**}}{dq_t} \left( \frac{de_t^{K**}}{dq_t} E_t^{**} - e_t^{K**} \frac{dE(q_t)}{dq_t} + \frac{e_t^{K**}}{(E_t^{**})^2} \right) \end{aligned} \right] \\
& - \psi \left( \frac{de_t^{K**}}{dq_t} \right)^2 + \frac{de_t^{N**}}{dq_t} \left[ \frac{e_t^{K**}}{E_t^{**}} G''(E_t^{**}) \right]
\end{aligned}$$

Let us now compute  $\frac{de_t^{N**}}{dq_t}$ :

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

Note that, at the equilibrium, both  $e_t^{K**}$  and  $e_t^{N**}$  are functions of  $q_t$  only. In other words, for each level of  $q_t$ , we know from Proposition 4 that there exist only one unique vector  $(e_t^{K**}, e_t^{N**})$ .

We have:

$$\left( \frac{e_t^{K**} q_t}{E_t^{**}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{K**} = 0$$

Thus, using the Theorem of implicit function, we have:

$$\frac{de_t^{K**}}{dq_t} = - \frac{d \left[ \left( \frac{e_t^{K**} q_t}{E_t^{**}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{K**} \right]}{d \left[ \left( \frac{e_t^{K**} q_t}{E_t^{**}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{K**} \right]} \frac{1}{de_t^{K**}}$$

Hence the numerator is:

$$\begin{aligned}
& \left( \frac{e_t^{K^{**}} E_t^{**} - e_t^{K^{**}} q_t \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst}}{E_t^{**2}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \\
& + \left( \frac{e_t^{K^{**}} q_t}{E_t^{**}} \right) \left[ G''(E_t^{**}) \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst} - \frac{G'(E_t^{**}) \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst} E_t^{**} - G(E_t^{**}) \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst}}{E_t^{**2}} \right] \\
& + \frac{G'(E_t^{**}) \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst} E_t^{**} - G(E_t^{**}) \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst}}{E_t^{**2}} \\
& \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \left( \frac{e_t^{K^{**}} E_t^{**}}{E_t^{**2}} \right) + \left( \frac{e_t^{K^{**}} q_t}{E_t^{**}} \right) G''(E_t^{**}) \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst} \\
& + \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst} \frac{1}{E_t^{**2}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] [E_t^{**} - 2e_t^{K^{**}} q_t]
\end{aligned}$$

Hence

$$\begin{aligned}
& \frac{e_t^{K^{**}}}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \\
& + \left( \frac{e_t^{K^{**}} q_t}{E_t^{**}} \right) \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst} G''(E_t^{**}) \\
& + \left( \frac{(1-q_t)e_t^{N^{**}} - e_t^{K^{**}} q_t}{E_t^{**}} \right) \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst} \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)
\end{aligned}$$

The denominator is:

$$\begin{aligned}
& q_t \left( \frac{(1-q_t)e_t^{N^{**}}}{E_t^{**2}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \\
& + \frac{e_t^{K^{**}} q_t q_t}{E_t^{**}} \left[ G''(E_t^{**}) - \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right] \\
& + \frac{q_t}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] - \psi
\end{aligned}$$

Hence

$$q_t \left( \frac{2(1-q_t)e_t^{N^{**}}}{E_t^{**2}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{e_t^{K^{**}} q_t q_t}{E_t^{**}} [G''(E_t^{**})] - \psi$$

Hence we have:

$$\frac{de_t^{K^{**}}}{dq_t} = - \frac{\left[ \frac{(e_t^{K^{**}})}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \left( \frac{e_t^{K^{**}} q_t}{E_t^{**}} \right) \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst} G''(E_t^{**}) \right.}{\left. + \left( \frac{(1-q_t)e_t^{N^{**}} - e_t^{K^{**}} q_t}{E_t^{**}} \right) \frac{dE_t^{**}}{dq_t} \Big|_{e_t^{K^{**}} = cst} \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right]}{q_t \left( \frac{2(1-q_t)e_t^{N^{**}}}{E_t^{**2}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{e_t^{K^{**}} q_t q_t}{E_t^{**}} [G''(E_t^{**})] - \psi}$$

Let us now substitute for  $\left. \frac{dE_t^{**}}{dq_t} \right|_{e_t^{K^{**}}=cst} = e_t^{K^{**}} - e_t^{N^{**}} + (1-q_t) \frac{de_t^{N^{**}}}{dq_t}$ :

$$\frac{de_t^{K^{**}}}{dq_t} = - \frac{\left[ \frac{e_t^{K^{**}}}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \left( \frac{e_t^{K^{**}} q_t}{E_t^{**}} \right) \left( e_t^{K^{**}} - e_t^{N^{**}} + (1-q_t) \frac{de_t^{N^{**}}}{dq_t} \right) G''(E_t^{**}) \right] + \left( \frac{(1-q_t)e_t^{N^{**}} - e_t^{K^{**}} q_t}{E_t^{**}} \right) \left( e_t^{K^{**}} - e_t^{N^{**}} + (1-q_t) \frac{de_t^{N^{**}}}{dq_t} \right) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}{qt \left( \frac{2(1-q_t)e_t^{N^{**}}}{E_t^{**2}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{e_t^{K^{**}} q_t}{E_t^{**}} [G''(E_t^{**})] - \psi}$$

Hence

$$\frac{de_t^{K^{**}}}{dq_t} = - \frac{\left[ \frac{e_t^{K^{**}}}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \left( \frac{e_t^{K^{**}} q_t}{E_t^{**}} \right) \left( e_t^{K^{**}} - e_t^{N^{**}} + (1-q_t) \frac{de_t^{N^{**}}}{dq_t} \right) G''(E_t^{**}) \right] + \left( \frac{(1-q_t)e_t^{N^{**}} - e_t^{K^{**}} q_t}{E_t^{**}} \right) \left( e_t^{K^{**}} - e_t^{N^{**}} + (1-q_t) \frac{de_t^{N^{**}}}{dq_t} \right) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}{qt \left( \frac{2(1-q_t)e_t^{N^{**}}}{E_t^{**2}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{e_t^{K^{**}} q_t}{E_t^{**}} [G''(E_t^{**})] - \psi}$$

Hence, at  $q = 0$ , we have:

$$\frac{de_t^{K^{**}}}{dq_t} = \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \frac{\left[ 1 + \left( \frac{de_t^{N^{**}}}{dq_t} \right) \frac{1}{E_t^{**}} \right]}{\psi}$$

Let us now compute  $\frac{de_t^{N^{**}}}{dq_t}$ . We have:

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

Hence

$$\frac{de_t^{N^{**}}}{dq_t} = - \frac{\frac{d \left[ e_t^{N^{**}} - \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)}{\psi} \right]}{dq_t}}{\frac{d \left[ e_t^{N^{**}} - \frac{\left( \frac{G(E_t^{**})}{E_t^{**}} - c \right)}{\psi} \right]}{de_t^{N^{**}}}}$$

Hence

$$\frac{de_t^{N^{**}}}{dq_t} = - \frac{\left[ - \frac{\left( \left. \frac{dE_t^{**}}{dq_t} \right|_{e_t^{N^{**}}=cst} \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}{\left[ 1 - \frac{\left( (1-q_t) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}$$

Hence, replacing by  $\left. \frac{dE_t^{**}}{dq_t} \right|_{e_t^{N^{**}}=cst} = \left( e_t^{K^{**}} + q_t \frac{de_t^{K^{**}}}{dq_t} - e_t^{N^{**}} \right)$ :

$$\frac{de_t^{N^{**}}}{dq_t} = - \frac{\left[ - \frac{\left( \left( e_t^{K^{**}} + q_t \frac{de_t^{K^{**}}}{dq_t} - e_t^{N^{**}} \right) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}{\left[ 1 - \frac{\left( (1-q_t) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}$$

At  $q = 0$ , we have  $e_t^{K^{**}} - e_t^{N^{**}} = 0$  and:

$$\frac{de_t^{N^{**}}}{dq_t} = 0$$

Hence, noting that:

$$\frac{de_t^{K^{**}}}{dq_t} = \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \frac{\left[ 1 + \left( \frac{de_t^{N^{**}}}{dq_t} \right) \frac{1}{E_t^{**}} \right]}{\psi}$$

We have:

$$\frac{de_t^{K^{**}}}{dq_t} = \frac{\left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right]}{\psi}$$

Let us now compute the second-order derivative:

$$\frac{d^2 e_t^{N^{**}}}{dq_t^2} = - \frac{\left[ \begin{aligned} & \left[ -\frac{1}{\psi} \left( \begin{aligned} & \left( \frac{de_t^{K^{**}}}{dq} + \frac{de_t^{K^{**}}}{dq_t} + qt \frac{d^2 e_t^{K^{**}}}{dq_t^2} - \frac{de_t^{N^{**}}}{dq_t} \right) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right. \right. \\ & + \left. \left[ \begin{aligned} & \left( e_t^{K^{**}} + qt \frac{de_t^{K^{**}}}{dq_t} - e_t^{N^{**}} \right) \\ & \left( G''(E_t^{**}) \frac{dE_t^{**}}{dq} - \frac{G'(E_t^{**}) \frac{dE_t^{**}}{dq} E_t^{**} - G(E_t^{**}) \frac{dE_t^{**}}{dq}}{E_t^{**2}} \right) E_t^{**} \\ & - \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \frac{dE_t^{**}}{dq} \end{aligned} \right] \frac{1}{E_t^{**2}} \end{aligned} \right] \right]}{\psi} \\ & \left[ 1 - \frac{\left( (1-qt) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right] \\ & - \frac{\left[ \begin{aligned} & \left[ \begin{aligned} & \left( \left( e_t^{K^{**}} + qt \frac{de_t^{K^{**}}}{dq_t} - e_t^{N^{**}} \right) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right) \\ & \left( -\frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right) \\ & + \left( 1-qt \right) \left[ \begin{aligned} & \left( G''(E_t^{**}) \frac{dE_t^{**}}{dq} - \frac{G'(E_t^{**}) \frac{dE_t^{**}}{dq} E_t^{**} - G(E_t^{**}) \frac{dE_t^{**}}{dq}}{E_t^{**2}} \right) E_t^{**} \\ & - \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \frac{dE_t^{**}}{dq} \end{aligned} \right] \end{aligned} \right] \right]}{\psi} \end{aligned} \right]}{\left[ 1 - \frac{\left( (1-qt) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]^2}$$



Hence, at  $q = 0$  :

$$\frac{d^2 e_t^{N^{**}}}{dq_t^2} = - \frac{\left[ \begin{array}{l} -\frac{1}{\psi} \left( \left( \frac{de_t^{K^{**}}}{dq} + \frac{de_t^{K^{**}}}{dq_t} - \frac{de_t^{N^{**}}}{dq_t} \right) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right) \\ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \end{array} \right]}{\left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]^2}$$

Hence

$$\frac{d^2 e_t^{N^{**}}}{dq_t^2} = \frac{\frac{1}{\psi} \left( 2 \frac{de_t^{K^{**}}}{dq} \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}$$

Hence, given  $\frac{de_t^{K^{**}}}{dq} = \frac{\left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right]}{\psi}$  at  $q = 0$ :

$$\frac{d^2 e_t^{N^{**}}}{dq_t^2} = \frac{\left( 2 \frac{1}{E_t^{**}} \left[ \frac{\left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}{\psi} \right]^2 \right)}{\left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}$$

Hence, in the general case where  $q \neq 0$ , we now have a system with two equations and two unknowns:

$$\frac{de_t^{N^{**}}}{dq_t} = \frac{\left[ -\frac{\left( \left( e_t^{K^{**}} + q_t \frac{de_t^{K^{**}}}{dq_t} - e_t^{N^{**}} \right) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}{\left[ 1 - \frac{\left( (1-q_t) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}$$

$$\frac{de_t^{K^{**}}}{dq_t} = \frac{\left[ \frac{e_t^{K^{**}}}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \left( \frac{e_t^{K^{**}} q_t}{E_t^{**}} \right) \left( e_t^{K^{**}} - e_t^{N^{**}} + (1-q_t) \frac{de_t^{N^{**}}}{dq_t} \right) G''(E_t^{**}) \right. \\ \left. + \left( \frac{(1-q_t)e_t^{N^{**}} - e_t^{K^{**}} q_t}{E_t^{**}} \right) \left( e_t^{K^{**}} - e_t^{N^{**}} + (1-q_t) \frac{de_t^{N^{**}}}{dq_t} \right) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right]}{q_t \left( \frac{2(1-q_t)e_t^{N^{**}}}{E_t^{**2}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{e_t^{K^{**}} q_t}{E_t^{**}} [G''(E_t^{**})] - \psi}$$

Substituting the first equation in the second one, we obtain:

$$\frac{de_t^{K^{**}}}{dq_t} = \frac{\left[ \begin{aligned} & \frac{e_t^{K^{**}}}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \\ & + \frac{e_t^{K^{**}} q_t}{E_t^{**}} \left( e_t^{K^{**}} - e_t^{N^{**}} - \frac{(1-q_t) \left[ -\frac{(e_t^{K^{**}} + q_t \frac{de_t^{K^{**}}}{dq_t} - e_t^{N^{**}}) \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}{\psi E_t^{**}} \right]}{\left[ 1 - \frac{(1-q_t) \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}{\psi E_t^{**}} \right]} \right) \\ & + \left( \frac{(1-q_t) e_t^{N^{**}} - e_t^{K^{**}} q_t}{E_t^{**}} \right) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \\ & + \left( e_t^{K^{**}} - e_t^{N^{**}} - \frac{(1-q_t) \left[ -\frac{(e_t^{K^{**}} + q_t \frac{de_t^{K^{**}}}{dq_t} - e_t^{N^{**}}) \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}{\psi E_t^{**}} \right]}{\left[ 1 - \frac{(1-q_t) \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}{\psi E_t^{**}} \right]} \right) \right]}{q_t \left( \frac{2(1-q_t) e_t^{N^{**}}}{E_t^{**2}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{e_t^{K^{**}} q_t q_t}{E_t^{**}} [G''(E_t^{**})] - \psi} \right]} G''(E_t^{**}) \end{aligned}$$

This expression will be used later on.

Let us now compute the other derivatives.

$$V^{KN} = G(e_t^{N^{**}}) - ce_t^{N^{**}} - \frac{\psi (e_t^{N^{**}})^2}{2}$$

Let us first compute the first derivative wrt  $q_t$ :

$$\begin{aligned} \frac{dV^{KN}}{dq_t} &= G'(e_t^{N^{**}}) \frac{de_t^{N^{**}}}{dq_t} - c \frac{de_t^{N^{**}}}{dq_t} - \psi e_t^{N^{**}} \frac{de_t^{N^{**}}}{dq_t} \\ &= \frac{de_t^{N^{**}}}{dq_t} [G'(e_t^{N^{**}}) - c - \psi e_t^{N^{**}}] \end{aligned}$$

The second-order derivative is:

$$\begin{aligned} \frac{d^2 V^{KN}}{dq_t^2} &= \frac{d^2 e_t^{N^{**}}}{dq_t^2} [G'(e_t^{N^{**}}) - c - \psi e_t^{N^{**}}] + \frac{de_t^{N^{**}}}{dq_t} \left[ G''(e_t^{N^{**}}) \frac{de_t^{N^{**}}}{dq_t} - \psi \frac{de_t^{N^{**}}}{dq_t} \right] \\ &= \frac{d^2 e_t^{N^{**}}}{dq_t^2} [G'(e_t^{N^{**}}) - c - \psi e_t^{N^{**}}] + \left( \frac{de_t^{N^{**}}}{dq_t} \right)^2 [G''(e_t^{N^{**}}) - \psi] \end{aligned}$$

Let us now compute the last derivative. We have:

$$V^{NN} = \frac{e_t^{N^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{N^{**}} - \frac{\psi (e_t^{N^{**}})^2}{2}$$

Hence

$$\begin{aligned}
\frac{dV^{NN}}{dq_t} &= \frac{\left[ \frac{de_t^{N**}}{dq_t} G(E_t^{**}) + e_t^{N**} G'(E_t^{**}) \frac{dE_t^{**}}{dq_t} \right] E_t^{**} - e_t^{N**} G(E_t^{**}) \frac{dE_t^{**}}{dq_t}}{E_t^{**2}} - c \frac{de_t^{N**}}{dq_t} - \psi e_t^{N**} \frac{de_t^{N**}}{dq_t} \\
&= \frac{de_t^{N**}}{dq_t} \left[ \frac{G(E_t^{**})}{E_t^{**}} - c - \psi e_t^{N**} \right] + \frac{e_t^{N**} G'(E_t^{**}) \frac{dE_t^{**}}{dq_t} E_t^{**} - e_t^{N**} G(E_t^{**}) \frac{dE_t^{**}}{dq_t}}{E_t^{**2}} \\
&= 0 + \frac{dE_t^{**}}{dq_t} \frac{e_t^{N**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)
\end{aligned}$$

Hence

$$\frac{dV^{NN}}{dq_t} = \left( e_t^{K**} + qt \frac{de_t^{K**}}{dq_t} - e_t^{N**} + (1-qt) \frac{de_t^{N**}}{dq_t} \right) \frac{e_t^{N**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)$$

Let us now compute  $\frac{d\Delta V^N}{dq_t}$ :

$$\frac{d\Delta V^N}{dq_t} = \left[ \frac{\left[ \frac{de_t^N}{dq_t} G(E_t^{**}) + e_t^N G'(E_t^{**}) \frac{dE_t^{**}}{dq_t} \right] E_t^{**} - (e_t^N G(E_t^{**})) \frac{dE_t^{**}}{dq_t}}{[E_t^{**}]^2} - c \frac{de_t^N}{dq_t} - \psi e_t^N \frac{de_t^N}{dq_t} \right. \\
\left. - \frac{\left[ \frac{de_t^{K**}}{dq_t} G(E_t^{**}) + e_t^{K**} G'(E_t^{**}) \frac{dE_t^{**}}{dq_t} \right] E_t^{**} - (e_t^{K**} G(E_t^{**})) \frac{dE_t^{**}}{dq_t}}{[E_t^{**}]^2} - c \frac{de_t^{K**}}{dq_t} - \psi e_t^{K**} \frac{de_t^{K**}}{dq_t} \right]$$

Further simplifications yield, at  $q = 0$ :

$$\frac{d\Delta V^N}{dq_t} = \left[ \frac{\left[ \frac{de_t^{N**}}{dq_t} G(E_t^{**}) \right]}{E_t^{**}} - c \frac{de_t^{N**}}{dq_t} - \psi \frac{de_t^{N**}}{dq_t} \right. \\
\left. - \left[ \frac{\left[ \frac{de_t^{K**}}{dq_t} G(E_t^{**}) \right]}{e_t^{N**}} - c \frac{de_t^{K**}}{dq_t} - \psi e_t^{K**} \frac{de_t^{K**}}{dq_t} \right] \right]$$

Hence, given the FOCs for Nash players, we have:

$$\frac{d\Delta V^N}{dq_t} = 0$$

## 9.5.2 Existence and stability of homogeneous stationary equilibria

The equation for the dynamics of  $q_t$  is now:

$$q_{t+1} \equiv H(q_t) = q_t + q_t(1-q_t) \left[ \begin{array}{l} \frac{(1-q_t)}{\kappa} \left[ \begin{array}{l} \frac{e_t^{K**} G(E_t^{**})}{E_t^{**}} - ce_t^{K**} - \frac{\psi(e_t^{K**})^2}{2} \\ - \left[ G(e_t^{N**}) - ce_t^{N**} - \frac{\psi(e_t^{N**})^2}{2} \right] \end{array} \right] \\ - \frac{q_t}{\kappa} \left[ \begin{array}{l} \frac{e_t^{N**} G(E_t^{**})}{E_t^{**}} - ce_t^{N**} - \frac{\psi(e_t^{N**})^2}{2} \\ - \left[ \frac{e_t^{K**} G(E_t^{**})}{E_t^{**}} - ce_t^{K**} - \frac{\psi(e_t^{K**})^2}{2} \right] \end{array} \right] \end{array} \right]$$

Let us first study the existence of a stationary equilibrium.

Obviously,  $q = 0$  and  $q = 1$  are stationary equilibria.

Let us now consider the stability of those homogeneous equilibria.

At  $q_t = 1$ , we know that

$$\begin{aligned}\tau_t^K &= 0 \\ \tau_t^N &= \frac{1}{\kappa} \left[ \begin{array}{l} \frac{e_t^{N^{**}}}{e_t^{K^{**}}} G(e_t^{K^{**}}) - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} \\ - \left[ \frac{e_t^{K^{**}}}{e_t^{K^{**}}} G(e_t^{K^{**}}) - ce_t^{K^{**}} - \frac{\psi(e_t^{K^{**}})^2}{2} \right] \end{array} \right]\end{aligned}$$

But remind that, when  $q = 1$ ,  $e_t^{K^{**}} = e_t^{K^*}$ . Hence,

$$\tau_t^N = \frac{1}{\kappa} \left[ \frac{e_t^{N^{**}}}{e_t^{K^*}} G(e_t^{K^*}) - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} - \left[ G(e_t^{K^*}) - ce_t^{K^*} - \frac{\psi(e_t^{K^*})^2}{2} \right] \right] > 0$$

Hence,

$$\begin{aligned}H'(q_t) &= 1 + (1 - 2q_t) \left[ \begin{array}{l} \frac{(1-q_t)}{\kappa} \left[ \begin{array}{l} \frac{e_t^{K^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{K^{**}} - \frac{\psi(e_t^{K^{**}})^2}{2} \\ - \left[ \frac{e_t^{N^{**}} G(e_t^{N^{**}})}{e_t^{N^{**}}} - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} \right] \end{array} \right] \\ - \frac{q_t}{\kappa} \left[ \begin{array}{l} \frac{e_t^{N^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} \\ - \left[ \frac{e_t^{K^*} G(E_t^{**})}{E_t^{**}} - ce_t^{K^{**}} - \frac{\psi(e_t^{K^{**}})^2}{2} \right] \end{array} \right] \end{array} \right] \\ &\quad + q_t(1 - q_t) \frac{d[\tau_t^K - \tau_t^N]}{dq_t}\end{aligned}$$

Hence

$$\begin{aligned}H'(q_t) &= 1 - 1 \left[ \begin{array}{l} \frac{(1-q_t)}{\kappa} \left[ \begin{array}{l} \frac{e_t^{K^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{K^{**}} - \frac{\psi(e_t^{K^{**}})^2}{2} \\ - \left[ \frac{e_t^{N^{**}} G(e_t^{N^{**}})}{e_t^{N^{**}}} - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} \right] \end{array} \right] \\ - \frac{q_t}{\kappa} \left[ \begin{array}{l} \frac{e_t^{N^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} \\ - \left[ \frac{e_t^{K^*} G(E_t^{**})}{E_t^{**}} - ce_t^{K^{**}} - \frac{\psi(e_t^{K^{**}})^2}{2} \right] \end{array} \right] \end{array} \right] \\ &= 1 - 1 [0 - \tau_t^N] > 1\end{aligned}$$

hence  $q = 1$  is unstable.

Let us now consider the stability of  $q = 0$ . The derivative of  $H(q_t)$  at  $q_t = 0$  is:

$$\begin{aligned}H'(q_t) &= 1 + (1 - 2q_t) [\tau_t^K - \tau_t^N] \\ &\quad + q_t(1 - q_t) \frac{d[\tau_t^K - \tau_t^N]}{dq_t}\end{aligned}$$

Remind that, at  $q = 0$ , we have:

$$e_t^{K**} = \frac{\frac{G(e_t^{N**})}{e_t^{N**}} - c}{\psi} = e_t^{N**}$$

Hence we have:

$$\tau_t^K = \frac{1}{\kappa} \left[ \frac{e_t^{N**} G(e_t^{N**})}{e_t^{N**}} - ce_t^{N**} - \frac{\psi (e_t^{N**})^2}{2} - \left[ G(e_t^{N**}) - ce_t^{N**} - \frac{\psi (e_t^{N**})^2}{2} \right] \right] = 0$$

Thus Impure Kantian parents invest 0 in socialization when  $q_t = 0$ . The reason is simple: when  $q_t = 0$ , Impure Kantians behave exactly as Nash players, and thus, from the perspective of parents, there is no difference between having a Nash player and an Impure Kantian.

Moreover, at  $q = 0$ , we have:

$$\tau_t^N = \frac{0}{\kappa} \left[ \frac{\frac{e_t^{N**} G(E_t^{**})}{E_t^{**}} - ce_t^{N**} - \frac{\psi (e_t^{N**})^2}{2}}{\left[ \frac{e_t^{K**} G(E_t^{**})}{E_t^{**}} - ce_t^{K**} - \frac{\psi (e_t^{K**})^2}{2} \right]} \right] = 0$$

Thus both socialization efforts are 0.

At  $q_t = 0$  we have:

$$H'(0) = 1$$

Second-order derivative is:

$$H''(q_t) = -2[\tau_t^K - \tau_t^N] + 2(1 - 2q_t) \frac{d[\tau_t^K - \tau_t^N]}{dq_t} + q_t(1 - q_t) \frac{d^2[\tau_t^K - \tau_t^N]}{dq_t^2}$$

At  $q_t = 0$  we have:

$$H''(q_t) = 2 \frac{d[\tau_t^K - \tau_t^N]}{dq_t}$$

which is to be computed.

From the toolbox, we have, at  $q = 0$ :

$$\frac{dV^{KK}}{dq} - \frac{dV^{KN}}{dq} = 0$$

Back to the initial expression, and reminding that at  $q = 0$  we have  $\Delta V^K = \Delta V^N = 0$ , it follows that

$$\begin{aligned} \frac{d[\tau_t^K - \tau_t^N]}{dq_t} &= -\Delta V^K + (1 - q) \frac{d\Delta V^K}{dq_t} - \left[ \Delta V^N + q \frac{d\Delta V^N}{dq_t} \right] \\ &= 0 + \frac{d\Delta V^K}{dq_t} - \left[ 0 + 0 \frac{d\Delta V^N}{dq_t} \right] \\ &= \frac{d\Delta V^K}{dq_t} \end{aligned}$$

Let us compute  $\frac{d\Delta V^K}{dq_t} = \frac{dV^{KK}}{dq} - \frac{dV^{KN}}{dq}$ .  
From the toolbox, we have, at  $q = 0$ :

$$\begin{aligned}\frac{dV^{KN}}{dq} &= \frac{de_t^N}{dq_t} [G'(e_t^N) - c - \psi e_t^N] \\ \frac{dV^{KK}}{dq} &= \left( \frac{de_t^N}{dq_t} \right) \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)\end{aligned}$$

Noting, that, from the toolbox, we have  $\frac{de_t^N}{dq_t} = 0$ .  
Hence we have:

$$\begin{aligned}\frac{d\Delta V^K}{dq_t} &= \frac{dV^{KK}}{dq} - \frac{dV^{KN}}{dq} \\ &= 0\end{aligned}$$

Hence

$$H''(0) = 0$$

Hence, following Dannan, Elaydi and Ponomarenko (2003, Theorem 2.3), we know that only two cases are possible: either  $H'''(q_t)$  at the fixed point is strictly negative, implying that the equilibrium is asymptotically stable, or  $H'''(q_t) > 0$  at the fixed point is strictly positive, implying that the equilibrium is unstable.

Computing the third order derivative, we obtain:

$$\begin{aligned}H'''(q_t) &= -2 \frac{d[\tau_t^K - \tau_t^N]}{dq_t} + 2(1-2q_t) \frac{d^2[\tau_t^K - \tau_t^N]}{dq_t^2} - 4 \frac{d[\tau_t^K - \tau_t^N]}{dq_t} \\ &\quad + (1-2q_t) \frac{d^2[\tau_t^K - \tau_t^N]}{dq_t^2} + q_t(1-q_t) \frac{d^3[\tau_t^K - \tau_t^N]}{dq_t^3}\end{aligned}$$

At  $q = 0$ , we have:

$$H'''(q_t) = 3 \frac{d^2[\tau_t^K - \tau_t^N]}{dq_t^2}$$

Let us now calculate  $\frac{d^2[\tau_t^K - \tau_t^N]}{dq_t^2}$  with general forms. Remember that we have:

$$\frac{d[\tau_t^K - \tau_t^N]}{dq_t} = -\Delta V^K + (1-q) \frac{d\Delta V^K}{dq_t} - \left[ \Delta V^N + q \frac{d\Delta V^N}{dq_t} \right]$$

Hence the second-order derivative is:

$$\frac{d^2[\tau_t^K - \tau_t^N]}{dq_t^2} = -2 \frac{d\Delta V^K}{dq_t} + (1-q) \frac{d^2\Delta V^K}{dq_t^2} - 2 \frac{d\Delta V^N}{dq_t} - q \frac{d^2\Delta V^N}{dq_t^2}$$

At  $q = 0$ , this simplifies to:

$$\frac{d^2[\tau_t^K - \tau_t^N]}{dq_t^2} = -2 \frac{d\Delta V^K}{dq_t} + \frac{d^2\Delta V^K}{dq_t^2} - 2 \frac{d\Delta V^N}{dq_t}$$

We already know that  $\frac{d\Delta V^K}{dq_t} = 0$  at  $q = 0$ :

$$\frac{d^2 [\tau_t^K - \tau_t^N]}{dq_t^2} = \frac{d^2 \Delta V^K}{dq_t^2} - 2 \frac{d\Delta V^N}{dq_t}$$

Hence, given that, at  $q = 0$ , we have  $\frac{d\Delta V^N}{dq_t} = 0$ , it follows that:

$$\frac{d^2 [\tau_t^K - \tau_t^N]}{dq_t^2} = \frac{d^2 \Delta V^K}{dq_t^2}$$

From the toolbox, we have, at  $q = 0$ :

$$\begin{aligned} \frac{d^2 V^{KK}}{dq_t^2} &= \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \left[ \begin{aligned} &\frac{de_t^{K**}}{dq_t} \frac{dE(q_t)}{dq_t} \frac{1}{E_t^{**}} + \frac{de_t^{K**}}{dq_t} \left( \frac{e_t^{K**}}{E_t^{**}} \right) \\ &+ \frac{e_t^{K**}}{E_t^{**}} \left( \frac{de_t^{K**}}{dq_t} - 2 \frac{de_t^{N**}}{dq_t} + \frac{d^2 e_t^{N**}}{dq_t^2} \right) \\ &+ \frac{de_t^{N**}}{dq_t} \left( \frac{\frac{de_t^{K**}}{dq_t} E_t^{**} - e_t^{K**} \frac{dE(q_t)}{dq_t}}{E_t^{**2}} + \frac{e_t^{K**}}{E_t^{**}} \frac{1}{E_t^{**}} \right) \end{aligned} \right] \\ &\quad - \psi \left( \frac{de_t^{K**}}{dq_t} \right)^2 + \frac{de_t^{N**}}{dq_t} \left[ \frac{e_t^{K**}}{E_t^{**}} (G''(E_t^{**})) \right] \end{aligned}$$

$$\begin{aligned} \frac{d^2 V^{KN}}{dq_t^2} &= \frac{d^2 e_t^{N**}}{dq_t^2} [G'(e_t^{N**}) - c - \psi e_t^{N**}] + \frac{de_t^{N**}}{dq_t} \left[ G''(e_t^{N**}) \frac{de_t^{N**}}{dq_t} - \psi \frac{de_t^{N**}}{dq_t} \right] \\ &= \frac{d^2 e_t^{N**}}{dq_t^2} [G'(e_t^{N**}) - c - \psi e_t^{N**}] + \left( \frac{de_t^{N**}}{dq_t} \right)^2 [G''(e_t^{N**}) - \psi] \end{aligned}$$

as well as,

$$\frac{de_t^{N**}}{dq_t} = \left[ - \frac{\left[ \frac{\left( (e_t^{K**} + q_t \frac{de_t^{K**}}{dq_t} - e_t^{N**}) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}{\left[ 1 - \frac{\left( (1-q_t) \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]} \right]$$

$$\frac{de_t^{K**}}{dq_t} = - \frac{\left[ \frac{e_t^{K**}}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \left( \frac{e_t^{K**} q_t}{E_t^{**}} \right) \left( e_t^{K**} - e_t^{N**} + (1-q_t) \frac{de_t^{N**}}{dq_t} \right) G''(E_t^{**}) \right]}{q_t \left( \frac{2(1-q_t)e_t^{N**}}{E_t^{**2}} \right) \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] + \frac{e_t^{K**} q_t}{E_t^{**}} [G''(E_t^{**})] - \psi}$$

when  $q = 0$ , those last two expressions vanish to:

$$\frac{de_t^{N**}}{dq_t} = 0$$

$$\frac{de_t^{K**}}{dq_t} = \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{\psi}$$

Hence

$$\frac{de_t^{K**}}{dq_t} = \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{\psi} < 0$$

Let us now substitute for those two expressions in  $\frac{d^2V^{KK}}{dq_t^2}$  and  $\frac{d^2V^{KN}}{dq_t^2}$ . We have, given  $\frac{dE(q_t)}{dq_t} = e_t^{K**} + qt \frac{de_t^{K**}}{dq_t} - e_t^{N**} + (1-qt) \frac{de_t^{N**}}{dq_t} = 0$  at  $q = 0$ :

$$\begin{aligned} \frac{d^2V^{KK}}{dq_t^2} &= \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \left[ 2 \frac{de_t^{K**}}{dq_t} + \frac{d^2e_t^{N**}}{dq_t^2} \right] - \psi \left( \frac{de_t^{K**}}{dq_t} \right)^2 \\ \frac{d^2V^{KN}}{dq_t^2} &= \frac{d^2e_t^{N**}}{dq_t^2} \left[ G'(e_t^{N**}) - c - \psi e_t^{N**} \right] \end{aligned}$$

Remind that, from the toolbox, we have:

$$\frac{d^2e_t^{N**}}{dq_t^2} = \frac{\left( 2 \frac{1}{E_t^{**}} \left[ \frac{(G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}})}{\psi} \right]^2 \right)}{\left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}$$

Hence, substituting in  $\frac{d^2V^{KK}}{dq_t^2}$  and  $\frac{d^2V^{KN}}{dq_t^2}$ , we have:

$$\begin{aligned} \frac{d^2V^{KK}}{dq_t^2} &= \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \left[ 2 \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{\psi} + \frac{\left( 2 \frac{1}{E_t^{**}} \left[ \frac{(G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}})}{\psi} \right]^2 \right)}{\left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]} \right] \\ &\quad - \psi \left( \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{\psi} \right)^2 \end{aligned}$$



Hence

$$\frac{d^2 V^{KK}}{dq_t^2} = \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \left[ \frac{2 \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] \left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]}{1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi}} \right]}{\psi \left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]} \right] - \psi \left( \frac{G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}}}{\psi} \right)^2$$

Hence

$$\begin{aligned} \frac{d^2 V^{KK}}{dq_t^2} &= \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right]^2 \frac{1}{\psi} \left[ \frac{2}{1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi}} - 1 \right] \\ &= \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right]^2 \frac{1}{\psi} \left[ \frac{1 + \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi}}{1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi}} \right] \end{aligned}$$

Moreover,

$$\frac{d^2 V^{KN}}{dq_t^2} = \frac{d^2 e_t^{N**}}{dq_t^2} [G'(e_t^{N**}) - c - \psi e_t^{N**}]$$

Hence

$$\frac{d^2 V^{KN}}{dq_t^2} = \frac{\left( 2 \frac{1}{E_t^{**}} \left[ \frac{\left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}{\psi} \right]^2 \right)}{\left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]} [G'(e_t^{N**}) - c - \psi e_t^{N**}]$$

Hence

$$\begin{aligned}
& \frac{d^2 V^{KK}}{dq_t^2} - \frac{d^2 V^{KN}}{dq_t^2} \\
&= \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right]^2 \frac{1}{\psi} \left[ \frac{1 + \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi}}{1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi}} \right] \\
& \quad - \frac{\left( 2 \frac{1}{E_t^{**}} \left[ \frac{\left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}{\psi} \right]^2 \right)}{\left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]} [G'(e_t^{N**}) - c - \psi e_t^{N**}]
\end{aligned}$$

Hence

$$\frac{\left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right]^2}{\left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]} \left[ \frac{\frac{1}{\psi} + \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi^2}}{-[G'(e_t^{N**}) - c - \psi e_t^{N**}] \frac{2}{E_t^{**} \psi^2}} \right]$$

Hence

$$\frac{\left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right]^2}{\left[ 1 - \frac{\left( \frac{1}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}{\psi} \right]} \frac{1}{E_t^{**} \psi^2} \left[ \left[ E_t^{**} \psi + \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) - 2 [G'(e_t^{N**}) - c - \psi e_t^{N**}] \right] \right]$$

Hence

the stability condition for  $q = 0$  is:

$$E_t^{**} \psi + \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) - 2 [G'(e_t^{N**}) - c - \psi e_t^{N**}] < 0$$

Hence

$$e_t^{N**} \psi + \left( G'(e_t^{N**}) - \frac{G(e_t^{N**})}{e_t^{N**}} \right) - 2 [G'(e_t^{N**}) - c - \psi e_t^{N**}] < 0$$

Let us now replace by  $e_t^{N**} = \frac{\left( \frac{G(e_t^{N**})}{e_t^{N**}} - c \right)}{\psi}$  when  $q = 0$ . We have:

$$\frac{\left( \frac{G(e_t^{N**})}{e_t^{N**}} - c \right)}{\psi} \psi + \left( G'(e_t^{N**}) - \frac{G(e_t^{N**})}{e_t^{N**}} \right) - 2 \left[ G'(e_t^{N**}) - c - \psi \frac{\left( \frac{G(e_t^{N**})}{e_t^{N**}} - c \right)}{\psi} \right] < 0$$

Hence

$$\left[ -c - G'(e_t^{N^{**}}) + 2 \left( \frac{G(e_t^{N^{**}})}{e_t^{N^{**}}} \right) \right] > 0$$

Hence, given the interior condition  $e_t^{N^{**}} > 0$ :

$$\left[ \frac{G(e_t^{N^{**}})}{e_t^{N^{**}}} - G'(e_t^{N^{**}}) + \frac{G(e_t^{N^{**}})}{e_t^{N^{**}}} - c \right] > 0$$

Hence, 0 is unstable.

### 9.5.3 Intermediate equilibrium: existence, uniqueness and stability

Regarding the existence, let us consider the space  $(q_t, \tau_t^K - \tau_t^N)$ . We know that, at  $q = 1$ , we have  $\tau_t^K(1) - \tau_t^N(1) < 0$ .

Indeed, we have, at  $q = 1$ :

$$\begin{aligned} \tau_t^K &= 0 \\ \tau_t^N &= \frac{1}{\kappa} \left[ \begin{aligned} &\frac{e_t^{N^{**}}}{e_t^{K^{**}}} G(e_t^{K^{**}}) - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} \\ &- \left[ \frac{e_t^{K^{**}}}{e_t^{K^{**}}} G(e_t^{K^{**}}) - ce_t^{K^{**}} - \frac{\psi(e_t^{K^{**}})^2}{2} \right] \end{aligned} \right] \end{aligned}$$

Note also that, at  $q = 0$ , we have:

$$\begin{aligned} \tau_t^K &= \frac{1}{\kappa} \left[ \frac{e_t^{N^{**}} G(e_t^{N^{**}})}{e_t^{N^{**}}} - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} - \left[ G(e_t^{N^{**}}) - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} \right] \right] = 0 \\ \tau_t^N &= \frac{0}{\kappa} \left[ \frac{e_t^{N^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{N^{**}} - \frac{\psi(e_t^{N^{**}})^2}{2} - \left[ \frac{e_t^{K^{**}} G(E_t^{**})}{E_t^{**}} - ce_t^{K^{**}} - \frac{\psi(e_t^{K^{**}})^2}{2} \right] \right] = 0 \end{aligned}$$

Hence  $\tau_t^K(0) - \tau_t^N(0) = 0$ .

We also know, from the analysis of stability at  $q = 0$ :

$$\begin{aligned} \frac{d[\tau_t^K(0) - \tau_t^N(0)]}{dq_t} &= 0 \\ \frac{d^2[\tau_t^K(0) - \tau_t^N(0)]}{dq_t^2} &> 0 \end{aligned}$$

Hence we know that the function  $\tau_t^K(q) - \tau_t^N(q)$  is above the  $x$  axis in the  $(q_t, \tau_t^K - \tau_t^N)$  in the neighborhood of 0 and below the  $x$  axis in the  $(q_t, \tau_t^K - \tau_t^N)$  in the neighborhood of 1. Hence, by continuity, there exists at least one intersection of the function  $\tau_t^K(q) - \tau_t^N(q)$  with the  $x$  axis. Therefore there exists at least one intermediate stationary equilibrium.

Regarding uniqueness, a necessary and sufficient for uniqueness of the intermediate stationary equilibrium is that the slope of the function  $\tau_t^K(q) - \tau_t^N(q)$  is strictly negative at any intermediate stationary equilibrium. We have:

$$\tau_t^K - \tau_t^N = \frac{(1 - q_t)}{\kappa} [V^{KK} - V^{KN}] - \frac{q_t}{\kappa} [V^{NN} - V^{NK}]$$

Hence the derivative:

$$\begin{aligned} \frac{d[\tau_t^K - \tau_t^N]}{dq_t} &= \frac{-1}{\kappa} [V^{KK} - V^{KN} + (V^{NN} - V^{NK})] \\ &\quad + \frac{(1-q_t)}{\kappa} \left[ \frac{dV^{KK}}{dq_t} - \frac{dV^{KN}}{dq_t} \right] \\ &\quad - \frac{q_t}{\kappa} \left[ \frac{dV^{NN}}{dq_t} - \frac{dV^{NK}}{dq_t} \right] \end{aligned}$$

The first term is negative. It captures the cultural substitution effect. The second and third terms will be decomposed as follows, using the toolbox for expressions  $\frac{dV^{KK}}{dq_t}$ ,  $\frac{dV^{KN}}{dq_t}$ ,  $\frac{dV^{NN}}{dq_t}$  and  $\frac{dV^{NK}}{dq_t} = \frac{dV^{KK}}{dq_t}$ :

$$\begin{aligned} \frac{d[\tau_t^K - \tau_t^N]}{dq_t} &= \frac{-1}{\kappa} [V^{KK} - V^{KN} + (V^{NN} - V^{NK})] \\ &\quad + \frac{(1-q_t)}{\kappa} \left[ \begin{aligned} &\left( \left( e_t^{K**} - e_t^{N**} + (1-q_t) \frac{de_t^{N**}}{dq_t} \right) \right) \\ &\frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \\ &- \left( \frac{de_t^{N**}}{dq_t} [G'(e_t^{N**}) - c - \psi e_t^{N**}] \right) \end{aligned} \right] \\ &\quad - \frac{q_t}{\kappa} \left[ \begin{aligned} &\left[ \begin{aligned} &\left( e_t^{K**} + q_t \frac{de_t^{K**}}{dq_t} - e_t^{N**} + (1-q_t) \frac{de_t^{N**}}{dq_t} \right) \\ &\frac{e_t^{N**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \end{aligned} \right] \\ &- \left( \left( e_t^{K**} - e_t^{N**} + (1-q_t) \frac{de_t^{N**}}{dq_t} \right) \right) \\ &\quad \left. - \left( \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right) \right] \end{aligned}$$

Let us now decompose the last two terms into those associated with variations in the number of animals purchased by each type (for a given  $q$ ), and into terms associated with a variation in  $q$  for given number of animals:

$$\begin{aligned} \frac{d[\tau_t^K - \tau_t^N]}{dq_t} &= \frac{-1}{\kappa} [V^{KK} - V^{KN} + (V^{NN} - V^{NK})] \\ &\quad + \frac{(1-q_t)}{\kappa} \left[ \begin{aligned} &\left( e_t^{K**} - e_t^{N**} \right) \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \\ &+ \left( (1-q_t) \frac{de_t^{N**}}{dq_t} \right) \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \\ &- \left( \frac{de_t^{N**}}{dq_t} [G'(e_t^{N**}) - c - \psi e_t^{N**}] \right) \end{aligned} \right] \\ &\quad - \frac{q_t}{\kappa} \left[ \begin{aligned} &-\frac{1}{E} (e_t^{K**} - e_t^{N**}) \left( (1-q_t) \frac{de_t^{N**}}{dq_t} \right) \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \\ &+ \left( q_t \frac{de_t^{K**}}{dq_t} \right) \frac{e_t^{N**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \\ &- [e_t^{K**} - e_t^{N**}]^2 \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \end{aligned} \right] \end{aligned}$$

Let us now try to sign each of those terms.

$$\begin{aligned}
\frac{d[\tau_t^K - \tau_t^N]}{dq_t} &= \underbrace{\frac{-1}{\kappa} [V^{KK} - V^{KN} + (V^{NN} - V^{NK})]}_{-} \\
&+ \frac{(1-q_t)}{\kappa} \left[ \underbrace{\left( e_t^{K**} - e_t^{N**} \right) \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}_{+} \right. \\
&\quad \left. + \frac{de_t^{N**}}{dq_t} \left( \underbrace{\left( (1-q_t) \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}_{+} \right. \right. \\
&\quad \left. \left. - \underbrace{[G'(e_t^{N**}) - c - \psi e_t^{N**}]}_{-} \right) \right] \\
&- \frac{q_t}{\kappa} \left[ \underbrace{- [e_t^{K**} - e_t^{N**}]^2 \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)}_{+} \right. \\
&\quad \left. + \frac{de_t^{N**}}{dq_t} \left[ - \frac{1}{E_t^{**}} (e_t^{K**} - e_t^{N**}) (1-q_t) \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right] \right] \\
&\quad \left. + \frac{de_t^{K**}}{dq_t} \left[ q_t \frac{e_t^{N**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right] \right]
\end{aligned}$$

The first term of the first bracket is positive: it is the effect of a change in  $q$  on the utility for a Kantian parent to have a Kantian child for a given number of animals for type  $N$ . The second term in the first bracket is the effect, for a given  $q$ , of a change in the number of animals purchased by Nash players on the utility for a Kantian parent to have a Kantian child.

Let us now interpret the second bracket. The first term in the second bracket is the impact, for a given number of animals for both types, of a change in  $q_t$  on the utility for a Nash parent to have a Nash child. The second term in the second bracket is the impact, for a given  $q_t$ , of a change in the number of animals purchased by Nash players on the utility for a Nash parent to have a Nash child. The third term in the second brackets is the impact, for a given  $q_t$ , of a change in the number of animals purchased by Impure Kantians on the utility for a Nash parent to have a Nash child.

Having carried out this decomposition, let us go further in the calculations:

$$\begin{aligned}
\frac{d[\tau_t^K - \tau_t^N]}{dq_t} &= \frac{-1}{\kappa} \left[ \frac{e_t^{N**} G(E_t^{**})}{E_t^{**}} - G(e_t^{N**}) \right] \\
&+ \frac{1}{\kappa} \left( (e_t^{K**} - e_t^{N**}) \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right) \\
&+ \frac{de_t^{N**}}{dq_t} \frac{(1 - q_t)}{\kappa} \left( \begin{aligned} &- [G'(e_t^{N**}) - c - \psi e_t^{N**}] \\ &+ \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \\ &- q_t \frac{e_t^{N**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \end{aligned} \right) \\
&- \frac{q_t}{\kappa} \left[ \left( e_t^{K**} + q_t \frac{de_t^{K**}}{dq_t} - e_t^{N**} \right) \frac{e_t^{N**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right]
\end{aligned}$$

Hence:

$$\begin{aligned}
\frac{d[\tau_t^K - \tau_t^N]}{dq_t} &= \frac{-1}{\kappa} \underbrace{\left[ \frac{e_t^{N**} G(E_t^{**})}{E_t^{**}} - G(e_t^{N**}) \right]}_{-} \\
&+ \frac{1}{\kappa} \underbrace{\left( (e_t^{K**} - e_t^{N**}) \frac{e_t^{K**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right)}_{+} \\
&+ \frac{de_t^{N**}}{dq_t} \frac{(1 - q_t)}{\kappa} \left( \begin{aligned} &- [G'(e_t^{N**}) - c - \psi e_t^{N**}] \\ &+ \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \left( \frac{e_t^{K**}}{E_t^{**}} - q_t \frac{e_t^{N**}}{E_t^{**}} \right) \end{aligned} \right) \\
&- \frac{q_t}{\kappa} \underbrace{\left[ (e_t^{K**} - e_t^{N**}) \frac{e_t^{N**}}{E_t^{**}} \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right) \right]}_{-} \\
&- \frac{de_t^{K**}}{dq_t} \frac{e_t^{N**}}{E_t^{**}} \frac{q_t}{\kappa} q_t \left( G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right)
\end{aligned}$$

Remind that the FOC for Impure Kantians is:

$$\begin{aligned}
\left( \frac{E_t^{**} - e_t^{K**} q_t}{(E_t^{**})^2} G(E_t^{**}) + \frac{e_t^{K**}}{E_t^{**}} G'(E_t^{**}) q_t - c \right) &= \psi e_t^{K**} \\
\frac{e_t^{K**} q_t}{E_t^{**}} \left[ G'(E_t^{**}) - \frac{G(E_t^{**})}{E_t^{**}} \right] - c + \frac{G(E_t^{**})}{E_t^{**}} &= \psi e_t^{K**}
\end{aligned}$$

Hence, given that we are here at an intermediate stationary equilibrium, the

condition is:

$$\begin{aligned}
\frac{d[\tau_t^K - \tau_t^N]}{dq_t} &= \frac{-1}{\kappa} \underbrace{[V^{KK} - V^{KN} + (V^{NN} - V^{NK})]}_{-} \\
&+ \frac{(1 - q^{**})}{\kappa} \left[ \underbrace{\left( e^{K^{**}} - e^{N^{**}} \right) \frac{e^{K^{**}}}{E^{**}} \left( G'(E^{**}) - \frac{G(E^{**})}{E^{**}} \right)}_{+} \right. \\
&\quad \left. + \frac{de^{N^{**}}}{dq^{**}} \left( \frac{e^{K^{**}}}{E^{**}} G'(E^{**}) + \frac{G(E^{**})}{E^{**}} \left( 1 - \frac{e^{K^{**}}}{E^{**}} \right) \right) \right. \\
&\quad \left. - \frac{q^{**}}{\kappa} \left[ \underbrace{-[e^{K^{**}} - e^{N^{**}}]^2 \left( G'(E^{**}) - \frac{G(E^{**})}{E^{**}} \right)}_{+} \right. \right. \\
&\quad \left. \left. + \frac{de^{N^{**}}}{dq^{**}} - \frac{1}{E^{**}} (e^{K^{**}} - e^{N^{**}}) (1 - q^{**}) \left( G'(E^{**}) - \frac{G(E^{**})}{E^{**}} \right) \right. \right. \\
&\quad \left. \left. + \frac{de^{K^{**}}}{dq^{**}} q^{**} \frac{e^{N^{**}}}{E^{**}} \left( G'(E^{**}) - \frac{G(E^{**})}{E^{**}} \right) \right. \right. \\
&\quad \left. \left. \underbrace{\left( G'(E^{**}) - \frac{G(E^{**})}{E^{**}} \right)}_{-} \right] \right]
\end{aligned}$$

## 9.6 Proof of Proposition 11

Existence and uniqueness of the stationary equilibrium with quotas. A stationary equilibrium under quota  $q^Q$ , if it exists, satisfies

$$(1 - q^Q)(V^{KK} - V^{KN}) - q^Q(V^{NN} - V^{NK}) = 0,$$

with

$$\begin{aligned}
V^{KK} &= A(e^{K^*})^{1/2} - \frac{2\psi(e^{K^*})^{3/2}}{3}, \\
V^{KN} &= A(\bar{e})^{1/2} - \frac{2\psi(\bar{e})^{3/2}}{3}, \\
V^{NK} &= e^{K^*} A(q^Q e^{K^*} + (1 - q^Q)\bar{e})^{-1/2} - \frac{2\psi(e^{K^*})^{3/2}}{3}, \\
V^{NN} &= \bar{e} A(q^Q e^{K^*} + (1 - q^Q)\bar{e})^{-1/2} - \frac{2\psi(\bar{e})^{3/2}}{3}.
\end{aligned}$$

Proof for existence follows from proof of Proposition 6 and is due to the

cultural substitution effect. The uniqueness condition writes as

$$\begin{aligned}
& - (V^{KK} - V^{KN} + V^{NN} - V^{NK}) \\
& + (1 - q^Q) \frac{d(V^{KK} - V^{KN})}{dq} - q^Q \frac{d(V^{NN} - V^{NK})}{dq} < 0, \\
\Leftrightarrow & - (V^{KK} - V^{KN} + V^{NN} - V^{NK}) - q^Q \frac{d(V^{NN} - V^{NK})}{dq} < 0, \\
\Leftrightarrow & - (V^{KK} - V^{KN} + V^{NN} - V^{NK}) \\
& - q^Q \left( \frac{1}{2} (\bar{e} - e^{K*})^2 A (q^Q e^{K*} + (1 - q^Q) \bar{e})^{-3/2} \right) < 0 \\
\Leftrightarrow & - A e^{K*} \left[ (e^{K*})^{-\frac{1}{2}} - (q^Q e^{K*} + (1 - q^Q) \bar{e})^{-\frac{1}{2}} \right] \\
& - A \bar{e} \left[ (q^Q e^{K*} + (1 - q^Q) \bar{e})^{-\frac{1}{2}} - (\bar{e})^{-\frac{1}{2}} \right] \\
& - q^Q \left( \frac{1}{2} (\bar{e} - e^{K*})^2 A (q^Q e^{K*} + (1 - q^Q) \bar{e}) \right)^{-3/2} < 0
\end{aligned}$$

which is true since  $e^{K*} < q^Q e^{K*} + (1 - q^Q) \bar{e} < \bar{e}$ .

We shall prove that for some value of  $\bar{e} \in [e^{K*}, e^{N*}]$ , the introduction of the quota increases the total number of animal, i.e. has a negative impact on global provision of public good. To do so, we proceed in two steps. In a first step we show that the introduction of a quota may reduce the number of Pure Kantians in the economy, i.e.  $q^Q(e^{N*}) \equiv q^{LF} > q^Q(\bar{e})$ , for some  $\bar{e} \in [e^{K*}, e^{N*}]$ .

We have

$$\frac{dq^Q}{d\bar{e}} = - \frac{(1 - q^Q) \frac{d(V^{KK} - V^{KN})}{d\bar{e}} - q^Q \frac{d(V^{NN} - V^{NK})}{d\bar{e}}}{- (V^{KK} - V^{KN} + V^{NN} - V^{NK}) - q^Q \frac{d(V^{NN} - V^{NK})}{dq}}.$$

The denominator is negative so that  $\frac{dq^Q}{d\bar{e}}$  is of the sign of  $(1 - q^Q) \frac{d(V^{KK} - V^{KN})}{d\bar{e}} - q^Q \frac{d(V^{NN} - V^{NK})}{d\bar{e}}$ . We have

$$\begin{aligned}
\frac{dV^{KK}}{d\bar{e}} &= 0, \\
\frac{dV^{KN}}{d\bar{e}} &= \frac{1}{2} (\bar{e})^{-\frac{1}{2}} - \psi(\bar{e})^{1/2}, \\
\frac{dV^{NK}}{d\bar{e}} &= \frac{-1}{2} (1 - q) e^{K*} A (q e^{K*} + (1 - q) \bar{e})^{-3/2}, \\
\frac{dV^{NN}}{d\bar{e}} &= A (q e^{K*} + (1 - q) \bar{e})^{-1/2} + \bar{e} (1 - q) \frac{-1}{2} A (q e^{K*} + (1 - q) \bar{e})^{-3/2} - \psi(\bar{e})^{1/2}.
\end{aligned}$$



In particular,

$$\begin{aligned}\frac{dV^{KN}}{d\bar{e}}\Big|_{\bar{e}=e^{N^*}} &= \frac{1}{2}(e^{N^*})^{-\frac{1}{2}} - \psi(e^{N^*})^{1/2} < 0, \\ \frac{d(V^{NN} - V^{NK})}{d\bar{e}}\Big|_{\bar{e}=e^{N^*}} &= \frac{-1}{2}(1-q)(e^{N^*} - e^{K^*})A(qe^{K^*} + (1-q)e^{N^*})^{-3/2} < 0,\end{aligned}$$

using the FOC for Nash and Pure Kantians and given that  $e^{K^*} < qe^{K^*} + (1-q)e^{N^*} < e^{N^*}$ . We deduce that  $\frac{dq^Q}{d\bar{e}}\Big|_{\bar{e}=e^{N^*}} > 0$ . Since we also have  $q^Q(e^{N^*}) = q^{LF}$ , we deduce that there exists some value  $\bar{e} \in [e^{K^*}, e^{N^*}[$  such that  $q^Q(\bar{e}) < q^{LF}$ .

In a second step we interest in the sign of  $\Delta E \equiv E^Q - E^{LF}$ .

Let us now compare the total number of animals at the laissez-faire and under the quota.

We have:

$$\begin{aligned}E^{LF} &= q^{LF}e^{K^*} + (1-q^{LF})e_t^{N^*} \\ E^Q &= q^Qe^{K^*} + (1-q^Q)\bar{e}\end{aligned}$$

We define  $\Delta E \equiv E^{LF} - E^Q$ ,

$$\Delta E = q^{LF}e^{K^*} + (1-q^{LF})e_t^{N^*} - [q^Qe^{K^*} + (1-q^Q)\bar{e}]$$

Hence

$$\Delta E = e^{K^*}(q^{LF} - q^Q) + (e_t^{N^*} - \bar{e}) - (q^{LF}e_t^{N^*} - q^Q\bar{e})$$

Hence

$$\Delta E = e^{K^*}(q^{LF} - q^Q) + (e_t^{N^*} - \bar{e}) - (q^{LF}e_t^{N^*} - q^Q\bar{e} - q^{LF}\bar{e} + q^{LF}\bar{e})$$

Hence

$$\Delta E = e^{K^*}(q^{LF} - q^Q) + (e_t^{N^*} - \bar{e}) - (q^{LF}(e_t^{N^*} - \bar{e}) + \bar{e}(q^{LF} - q^Q))$$

Let us also define:

$$\begin{aligned}\Delta q &\equiv q^{LF} - q^Q > 0 \\ \Delta e &\equiv e_t^{N^*} - \bar{e} > 0\end{aligned}$$

and re-express  $\Delta E$  as

$$\Delta E = e^{K^*}\Delta q + \Delta e - (q^{LF}\Delta e + \bar{e}\Delta q)$$

Hence,

$$\Delta E = -(\bar{e} - e^{K^*})\Delta q + (1 - q^{LF})\Delta e$$

When the quota raises the number of sheep, we have  $\Delta E < 0$ .  $\Delta E$  is negative if and only if

$$-(\bar{e} - e^{K^*})\Delta q + (1 - q^{LF})\Delta e < 0$$

which is equivalent to

$$(1 - q^{LF})\Delta e < (\bar{e} - e^{K*})\Delta q$$

that is

$$(1 - q^{LF})/(\bar{e} - e^{K*}) < \Delta q/\Delta e$$

This condition certainly does not hold for any value of  $\bar{e}$  (in particular it is violated at  $\bar{e} = e^{K*}$ ). Let us study this condition in a neighborhood of  $e_t^{N*}$ . To do so, take the limit when  $\bar{e}$  tends to  $e_t^{N*}$  for both sides of the inequality. For the left-hand side, we have

$$\lim_{\bar{e} \rightarrow e_t^{N*}} (1 - q^{LF})/(\bar{e} - e^{K*}) = (1 - q^{LF})/(e_t^{N*} - e^{K*})$$

For the right-hand side, note that

$$\lim_{\bar{e} \rightarrow e_t^{N*}} \Delta q/\Delta e = \frac{dq}{d\bar{e}} \Big|_{\bar{e}=e_t^{N*}}$$

since  $q^{LF}$  and  $q^Q$  are values of the function  $q(\bar{e})$  (i.e. the values of the stationary equilibrium for any  $\bar{e}$ ) respectively taken at  $e_t^{N*}$  and  $\bar{e}$ .

Therefore in a neighborhood of  $e_t^{N*}$ , the condition becomes:

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

We have

$$\frac{d(V^{NN} - V^{NK})}{dq} \Big|_{\bar{e}=e^{N*}} = \frac{1}{2} (e^{N*} - e^{K*})^2 A (qe^{K*} + (1 - q)e^{N*})^{-3/2}$$

and

$$\frac{d(V^{KK} - V^{KN})}{d\bar{e}} \Big|_{\bar{e}=e^{N*}} = Ae^{K*} \frac{-1}{2} (q^Q e^{K*} + (1 - q^Q)e^{N*})^{-3/2} (1 - q^Q),$$

so that the condition for  $\Delta E < 0$  becomes

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

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

$$-\frac{1}{2}A(e^{N^*})^{-1/2} + \psi(e^{N^*})^{1/2} > \frac{A(e^{K^*})^{1/2} - A(e^{N^*})^{1/2}}{e^{N^*} - e^{K^*}} + A(q^Q e^{K^*} + (1 - q^Q)e^{N^*})^{-1/2}.$$

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

$$A(q^Q e^{K^*} + (1 - q^Q)e^{N^*})^{-1/2} = \psi(e^{N^*})^{1/2},$$

so that the above condition is equivalent to

$$\begin{aligned} & -\frac{1}{2}A(e^{N^*})^{-1/2} + \psi(e^{N^*})^{1/2} > \frac{A(e^{K^*})^{1/2} - A(e^{N^*})^{1/2}}{e^{N^*} - e^{K^*}} + \psi(e^{N^*})^{1/2} \\ \Leftrightarrow & -\frac{1}{2}A(e^{N^*})^{-1/2} > \frac{A(e^{K^*})^{1/2} - A(e^{N^*})^{1/2}}{e^{N^*} - e^{K^*}} \\ \Leftrightarrow & \frac{1}{2}A(e^{N^*})^{-1/2} < \frac{A(e^{N^*})^{1/2} - A(e^{K^*})^{1/2}}{e^{N^*} - e^{K^*}}, \end{aligned}$$

which is true by the concavity of the function  $f(x) = Ax^{\frac{1}{2}}$  (and since  $e^{N^*} > e^{K^*}$ ).

## 9.7 Proof of Proposition 12

Existence and uniqueness of the stationary equilibrium with quotas. A stationary equilibrium under quota  $q^Q$ , if it exists, satisfies

$$(1 - q^Q)(V^{KK} - V^{KN}) - q^Q(V^{NN} - V^{NK}) = 0,$$

with

$$\begin{aligned} V^{KK} &= e^{K^{**}} A(q^Q e^{K^{**}} + (1 - q^Q)\bar{e})^{-1/2} - \frac{2\psi(e^{K^{**}})^{3/2}}{3}, \\ V^{KN} &= A(\bar{e})^{1/2} - \frac{2\psi(\bar{e})^{3/2}}{3}, \\ V^{NK} &= V^{KK} = e^{K^{**}} A(q^Q e^{K^{**}} + (1 - q^Q)\bar{e})^{-1/2} - \frac{2\psi(e^{K^{**}})^{3/2}}{3}, \\ V^{NN} &= \bar{e} A(q^Q e^{K^{**}} + (1 - q^Q)\bar{e})^{-1/2} - \frac{2\psi(\bar{e})^{3/2}}{3}. \end{aligned}$$

Proof for existence follows from proof of Proposition. Uniqueness requires

$$\begin{aligned} & -(V^{KK} - V^{KN} + V^{NN} - V^{NK}) \\ & + (1 - q^Q) \frac{d(V^{KK} - V^{KN})}{dq} - q^Q \frac{d(V^{NN} - V^{NK})}{dq} \\ & + \frac{de^{K^{**}}}{dq} \left( (1 - q^Q) \frac{d(V^{KK} - V^{KN})}{de^{K^{**}}} - q^Q \frac{d(V^{NN} - V^{NK})}{de^{K^{**}}} \right) \equiv -B < 0. \end{aligned}$$

which is assumed in what follows.

Let us first express the derivative. We have

$$\begin{aligned} \frac{dq^Q}{d\bar{e}} &= \left[ (1-q^Q) \frac{d(V^{KK} - V^{KN})}{d\bar{e}} - q^Q \frac{d(V^{NN} - V^{NK})}{d\bar{e}} \right. \\ &\quad \left. + \frac{de^{K**}}{d\bar{e}} \left( (1-q^Q) \frac{d(V^{KK} - V^{KN})}{de^{K**}} - q^Q \frac{d(V^{NN} - V^{NK})}{de^{K**}} \right) \right] / \\ &\quad \left[ (V^{KK} - V^{KN} + V^{NN} - V^{NK}) - (1-q^Q) \frac{d(V^{KK} - V^{KN})}{dq} + q^Q \frac{d(V^{NN} - V^{NK})}{dq} \right. \\ &\quad \left. - \frac{de^{K**}}{dq} \left( (1-q^Q) \frac{d(V^{KK} - V^{KN})}{de^{K**}} - q^Q \frac{d(V^{NN} - V^{NK})}{de^{K**}} \right) \right]. \end{aligned}$$

Find out the sign of this derivative is not so easy. Therefore, we directly show that, for some parameters combination we have  $\Delta E \equiv E^Q - E^{LF} < 0$  since it implies  $\frac{dq^Q}{d\bar{e}} > 0$ . Using similar arguments than the case of Pure Kantians, we find that the condition for  $\Delta E \equiv E^Q - E^{LF} < 0$  when  $\bar{e}$  close to  $e^{N**}$  is

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

Note that  $\frac{dV^{KK}}{de^{K**}} = 0 = \frac{dV^{NK}}{de^{K**}}$  and  $\frac{dV^{KN}}{de^{K**}} = 0$  so that the above condition

re-writes as

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

When studying introduction of a quota in the case of Pure Kantians, we already performed  $\frac{d(V^{NN} - V^{NK})}{d\bar{e}} \Big|_{e^{N^{**}}}$  and  $\frac{d(V^{NN} - V^{NK})}{dq} \Big|_{e^{N^{**}}}$  and found that

$$\frac{d(V^{NN} - V^{NK})}{d\bar{e}} \Big|_{e^{N^{**}}} = -\frac{(1 - q^Q)}{e^{N^{**}} - e^{K^*}} \frac{d(V^{NN} - V^{NK})}{dq} \Big|_{e^{N^{**}}}.$$

Computing each of the derivatives involved in the above condition one can check that

$$\frac{d(V^{KK} - V^{KN})}{d\bar{e}} \Big|_{e^{N^{**}}} = -\frac{(1 - q^Q)}{e^{N^{**}} - e^{K^*}} \frac{d(V^{KK} - V^{KN})}{dq} \Big|_{e^{N^{**}}},$$

also,

$$\begin{aligned}
-\frac{de^{K^{**}}}{dq} &= \frac{de^{K^{**}}}{d\bar{e}} \frac{e^{N^{**}} - e^{K^*}}{(1 - q^Q)} + \\
& \frac{\frac{1}{2} A e^{K^{**}} (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-3/2}}{\left[ \frac{q^Q A (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-3/2} - (3/4) (q^Q)^2 e^{K^{**}} A (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-5/2}}{+(1/2) \Psi(e^{K^{**}})^{-1/2}} \right]}
\end{aligned}$$

and,

$$\frac{dV^{NN}}{de^{K^{**}}} \Big|_{e^{N^{**}}} = -\frac{1}{2} q^Q \bar{e} A (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-3/2}$$

Therefore, the condition above re-writes as,

$$\begin{aligned}
& - (1 - q^Q) \frac{dV^{KN}}{d\bar{e}} \Big|_{e^{N^{**}}} > \\
& \frac{(1 - q^Q)}{e^{N^{**}} - e^{K^*}} \left[ V^{KK} - V^{KN} + V^{NN} - V^{NK} + \right. \\
& \quad \left. \frac{(1/4)(q^Q)^2 A^2 e^{K^{**}} e^{N^{**}} \left[ (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-3/2} \right]^2}{\left[ q^Q A (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-\frac{3}{2}} - (3/4)(q^Q)^2 e^{K^{**}} A (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-\frac{5}{2}} + (1/2) \Psi(e^{K^{**}})^{-\frac{1}{2}} \right]} \right] \\
& + q^Q \frac{de^{K^{**}}}{d\bar{e}} \frac{1}{e^{N^{**}} - e^{K^*}} B.
\end{aligned}$$

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

$$\begin{aligned}
& - \frac{1}{2} A (e^{N^{**}})^{-1/2} + \psi(e^{N^{**}})^{1/2} > \frac{Ae^{K^*} (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-1/2} - A (e^{N^{**}})^{1/2}}{e^{N^{**}} - e^{K^*}} \\
& + A (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-1/2} \\
& + \frac{(1/4) \frac{1}{e^{N^{**}} - e^{K^*}} (q^Q)^2 A^2 e^{K^{**}} e^{N^{**}} \left[ (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-3/2} \right]^2}{q^Q A (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-\frac{3}{2}} - (3/4)(q^Q)^2 e^{K^{**}} A (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-\frac{5}{2}} + (1/2) \Psi(e^{K^{**}})^{-\frac{1}{2}}} \\
& + \frac{q^Q}{(1 - q^Q)} \frac{de^{K^{**}}}{d\bar{e}} \frac{1}{e^{N^{**}} - e^{K^*}} B.
\end{aligned}$$

Since,  $Ae^{K^*} (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-1/2} < Ae^{K^*} (e^{K^*})^{-1/2} = A (e^{K^*})^{1/2}$ , a sufficient condition is

$$\begin{aligned}
& - \frac{1}{2} A (e^{N^{**}})^{-1/2} + \psi(e^{N^{**}})^{1/2} > \frac{A (e^{K^*})^{1/2} - A (e^{N^{**}})^{1/2}}{e^{N^{**}} - e^{K^*}} \\
& + A (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-1/2} \\
& + \frac{(1/4)(q^Q)^2 A^2 e^{K^{**}} e^{N^{**}} \left[ (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-3/2} \right]^2 \frac{1}{e^{N^{**}} - e^{K^*}}}{q^Q A (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-\frac{3}{2}} - (3/4)(q^Q)^2 e^{K^{**}} A (q^Q e^{K^{**}} + (1 - q^Q) \bar{e})^{-\frac{5}{2}} + (1/2) \Psi(e^{K^{**}})^{-\frac{1}{2}}} \\
& + \frac{q^Q}{(1 - q^Q)} \frac{de^{K^{**}}}{d\bar{e}} \frac{1}{e^{N^{**}} - e^{K^*}} B.
\end{aligned}$$

Using similar reasoning than for the case of Pure Kantians, we have

$$- \frac{1}{2} A (e^{N^{**}})^{-1/2} + \psi(e^{N^{**}})^{1/2} > \frac{A (e^{K^*})^{1/2} - A (e^{N^{**}})^{1/2}}{e^{N^{**}} - e^{K^*}} + A (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-1/2},$$

so that a sufficient condition for the inequality to hold is

$$\begin{aligned} & \frac{(1/4)(q^Q)^2 A^2 e^{K^{**}} e^{N^{**}} \left[ (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-3/2} \right]^2 \frac{1}{e^{N^{**}} - e^{K^{**}}}}{\left[ \frac{q^Q A (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-\frac{3}{2}} - (3/4)(q^Q)^2 e^{K^{**}} A (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-\frac{5}{2}}}{+(1/2)\Psi(e^{K^{**}})^{-\frac{1}{2}}} \right]} \\ & + \frac{q^Q}{(1 - q^Q)} \frac{de^{K^{**}}}{d\bar{e}} \frac{1}{e^{N^{**}} - e^{K^{**}}} B < 0. \end{aligned}$$

Let us substitute for  $\frac{de^{K^{**}}}{d\bar{e}}$ ,

$$\begin{aligned} & \frac{(1/4)(q^Q)^2 A^2 e^{K^{**}} e^{N^{**}} \left[ (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-3/2} \right]^2 \frac{1}{e^{N^{**}} - e^{K^{**}}}}{\left[ \frac{q^Q A (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-\frac{3}{2}} - (3/4)(q^Q)^2 e^{K^{**}} A (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-\frac{5}{2}}}{+(1/2)\Psi(e^{K^{**}})^{-\frac{1}{2}}} \right]} \\ & \frac{\frac{1}{2} q^Q A (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-\frac{3}{2}} \left( -1 + \frac{3}{2} q^Q \frac{e^{K^{**}}}{q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}}} \right) \frac{1}{e^{N^{**}} - e^{K^{**}}}}{\left[ \frac{q^Q A (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-\frac{3}{2}} - (3/4)(q^Q)^2 e^{K^{**}} A (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-\frac{5}{2}}}{+(1/2)\Psi(e^{K^{**}})^{-\frac{1}{2}}} \right]} B \\ & < 0, \\ & \Leftrightarrow (1/2)(q^Q) A e^{K^{**}} e^{N^{**}} (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-3/2} + B \left( -1 + \frac{3}{2} q^Q \frac{e^{K^{**}}}{q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}}} \right) < 0. \end{aligned}$$

We consider two cases.

(i) Suppose that  $B > \frac{1}{2} A q^Q (e^{N^{**}} - e^{K^{**}})^2 (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-3/2}$ . A sufficient condition is now,

$$\begin{aligned} & (1/2)(q^Q) A e^{K^{**}} e^{N^{**}} (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-3/2} \\ & + \left[ \frac{\frac{1}{2} A q^Q (e^{N^{**}} - e^{K^{**}})^2 (q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}})^{-3/2}}{\left( -1 + \frac{3}{2} q^Q \frac{e^{K^{**}}}{q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}}} \right)} \right] \\ & < 0 \end{aligned}$$

Since  $q^Q < \frac{1}{2}$  and  $\frac{e^{K^{**}}}{q^Q e^{K^{**}} + (1 - q^Q) e^{N^{**}}} < 1$ , a sufficient condition is

$$-\frac{1}{4}(e^{N^{**}} - e^{K^{**}})^2 - e^{K^{**}} e^{N^{**}} < 0.$$

This is a polynomial function of  $e^{N^{**}}$  which is convex and positive whenever  $e^{N^{**}} > 6e^{K^{**}}$ . Hence, the condition holds when parameters are such that  $e^{N^{**}} > 6e^{K^{**}}$ .

(ii) Suppose that  $B < \frac{1}{2}Aq^Q(e^{N^{**}} - e^{K^{**}})^2 (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2}$ . Remind that the condition for  $\Delta E < 0$  is

$$\begin{aligned} & \left[ (1 - q^Q) \frac{d(V^{KK} - V^{KN})}{d\bar{e}} \Big|_{e^{N^{**}}} - q^Q \frac{d(V^{NN} - V^{NK})}{d\bar{e}} \Big|_{e^{N^{**}}} \right. \\ & \left. + \frac{de^{K^{**}}}{d\bar{e}} \left( (1 - q^Q) \frac{d(V^{KK} - V^{KN})}{de^{K^{**}}} \Big|_{e^{N^{**}}} - q^Q \frac{d(V^{NN} - V^{NK})}{de^{K^{**}}} \Big|_{e^{N^{**}}} \right) \right] \\ & > \\ & B \left[ \frac{(1 - q^Q)}{e^{N^{**}} - e^{K^{**}}} + q^Q \frac{de^{K^{**}}}{d\bar{e}} \frac{1}{e^{N^{**}} - e^{K^{**}}} \right] \end{aligned}$$

A sufficient condition for this inequality to hold is then

$$\begin{aligned} & \left[ (1 - q^Q) \frac{d(V^{KK} - V^{KN})}{d\bar{e}} \Big|_{e^{N^{**}}} - q^Q \frac{d(V^{NN} - V^{NK})}{d\bar{e}} \Big|_{e^{N^{**}}} \right. \\ & \left. + \frac{de^{K^{**}}}{d\bar{e}} \left( (1 - q^Q) \frac{d(V^{KK} - V^{KN})}{de^{K^{**}}} \Big|_{e^{N^{**}}} - q^Q \frac{d(V^{NN} - V^{NK})}{de^{K^{**}}} \Big|_{e^{N^{**}}} \right) \right] \\ & > \\ & \left[ \frac{1}{2}Aq^Q(e^{N^{**}} - e^{K^{**}})^2 (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \right. \\ & \left. \frac{(1 - q^Q)}{e^{N^{**}} - e^{K^{**}}} + q^Q \frac{de^{K^{**}}}{d\bar{e}} \frac{1}{e^{N^{**}} - e^{K^{**}}} \right] \end{aligned}$$

That is,

$$\begin{aligned} & \left[ (1 - q^Q) \frac{d(V^{KK} - V^{KN})}{d\bar{e}} \Big|_{e^{N^{**}}} - q^Q \frac{d(V^{NN} - V^{NK})}{d\bar{e}} \Big|_{e^{N^{**}}} \right. \\ & \left. + \frac{de^{K^{**}}}{d\bar{e}} \left( (1 - q^Q) \frac{d(V^{KK} - V^{KN})}{de^{K^{**}}} \Big|_{e^{N^{**}}} - q^Q \frac{d(V^{NN} - V^{NK})}{de^{K^{**}}} \Big|_{e^{N^{**}}} \right) \right] \\ & > \\ & \frac{1}{2}Aq^Q(e^{N^{**}} - e^{K^{**}}) (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \left[ 1 - q^Q + q^Q \frac{de^{K^{**}}}{d\bar{e}} \right] \end{aligned}$$

Let substitute for  $\frac{d(V^{KK} - V^{KN})}{d\bar{e}} \Big|_{e^{N^{**}}}$ ,  $\frac{d(V^{NN} - V^{NK})}{d\bar{e}} \Big|_{e^{N^{**}}}$ ,  $\frac{d(V^{KK} - V^{KN})}{de^{K^{**}}} \Big|_{e^{N^{**}}}$ , and rearrange the terms,

$$\begin{aligned} & (1 - q^Q) \left[ \begin{aligned} & -\frac{1}{2}A(1 - q^Q)e^{K^{**}} (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \\ & - \left( \frac{1}{2}A(e^{N^{**}})^{-\frac{1}{2}} - \Psi(e^{N^{**}})^{\frac{1}{2}} \right) \\ & + \frac{1}{2}q^Q(e^{N^{**}} - e^{K^{**}})A (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \end{aligned} \right] \\ & + (q^Q)^2 \frac{1}{2}Ae^{N^{**}} (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \frac{de^{K^{**}}}{d\bar{e}} \\ & > \\ & \frac{1}{2}Aq^Q(e^{N^{**}} - e^{K^{**}}) (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \left[ 1 - q^Q + q^Q \frac{de^{K^{**}}}{d\bar{e}} \right]. \end{aligned}$$



This is equivalent to

$$\begin{aligned}
& (1 - q^Q) \begin{bmatrix} -\frac{1}{2}A(1 - q^Q)e^{K^{**}} (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \\ -\left(\frac{1}{2}A(e^{N^{**}})^{-\frac{1}{2}} - \Psi(e^{N^{**}})^{\frac{1}{2}}\right) \\ +\frac{1}{2}q^Q(e^{N^{**}} - e^{K^{**}})A (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \\ -\frac{1}{2}Aq^Q(e^{N^{**}} - e^{K^{**}}) (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \end{bmatrix} \\
& > \\
& q^Q \frac{de^{K^{**}}}{d\bar{e}} \begin{bmatrix} \frac{1}{2}Aq^Q(e^{N^{**}} - e^{K^{**}}) (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \\ -q^Q \frac{1}{2}Ae^{N^{**}} (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \end{bmatrix}
\end{aligned}$$

or,

$$\begin{aligned}
& (1 - q^Q) \begin{bmatrix} -\frac{1}{2}Ae^{K^{**}} (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} \\ -\left(\frac{1}{2}A(e^{N^{**}})^{-\frac{1}{2}} - \Psi(e^{N^{**}})^{\frac{1}{2}}\right) \\ +\frac{1}{2}Aq^Q (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} e^{K^{**}} \end{bmatrix} \\
& > \\
& q^Q \frac{de^{K^{**}}}{d\bar{e}} \begin{bmatrix} \frac{1}{2}Aq^Q (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} (-e^{K^{**}}) \end{bmatrix}
\end{aligned}$$

We can show that

$$-\frac{1}{2}Ae^{K^{**}} (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} - \left(\frac{1}{2}A(e^{N^{**}})^{-\frac{1}{2}} - \Psi(e^{N^{**}})^{\frac{1}{2}}\right) > 0$$

and a sufficient condition is

$$\begin{aligned}
& (1 - q^Q) \begin{bmatrix} \frac{1}{2}Aq^Q (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} e^{K^{**}} \end{bmatrix} \\
& > \\
& q^Q \frac{de^{K^{**}}}{d\bar{e}} \begin{bmatrix} \frac{1}{2}Aq^Q (q^Q e^{K^{**}} + (1 - q^Q)e^{N^{**}})^{-3/2} (-e^{K^{**}}) \end{bmatrix}
\end{aligned}$$

which is equivalent to

$$q^Q \frac{de^{K^{**}}}{d\bar{e}} + (1 - q^Q) > 0,$$

which condition always holds.

## 9.8 Proof of Proposition 13

When  $\alpha = 1/2$ , the four equations become:

$$\begin{aligned}
 A (q_t e_t^{K^*} + (1 - q_t) e_t^{N^*})^{-1/2} - \psi (e_t^{N^*})^{1/2} &= 0 \\
 \alpha A e_t^{K^* - 1/2} - \psi (e_t^{K^*})^{1/2} &= 0 \iff e_t^{K^*} = \frac{A}{2\psi} \\
 A (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-1/2} - \psi (e_t^{N^{**}})^{1/2} &= 0 \\
 \left[ A (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-1/2} \right. \\
 \left. + \frac{-1}{2} q_t e_t^{K^{**}} A (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-3/2} \right] - \psi (e_t^{K^{**}})^{1/2} &= 0
 \end{aligned}$$

We have:

$$A (q_t e_t^{K^*} + (1 - q_t) e_t^{N^*})^{-1/2} - \psi (e_t^{N^*})^{1/2} = 0$$

Hence

$$(1 - q_t) (e_t^{N^*})^2 + q_t e_t^{K^*} e_t^{N^*} - \frac{A^2}{\psi^2} = 0$$

Hence:

$$\Delta = (q_t e_t^{K^*})^2 + 4(1 - q_t) \frac{A^2}{\psi^2}$$

Hence:

$$e_t^{N^*} = \frac{-q_t e_t^{K^*} + \sqrt{(q_t e_t^{K^*})^2 + 4(1 - q_t) \frac{A^2}{\psi^2}}}{2(1 - q_t)}$$

We also have:

$$e_t^{N^{**}} = \frac{-q_t e_t^{K^{**}} + \sqrt{(q_t e_t^{K^{**}})^2 + 4(1 - q_t) \frac{A^2}{\psi^2}}}{2(1 - q_t)}$$

Hence, substituting for  $e_t^{K^*}$  in  $e_t^{N^*}$ , we obtain:

$$e_t^{N^*} = \frac{A - \frac{q_t}{2} + \sqrt{\left(\frac{q_t}{2}\right)^2 + 4(1 - q_t)}}{\psi \cdot 2(1 - q_t)}$$

Let us now differentiate  $e_t^{N^{**}}$  with respect to  $e_t^{K^{**}}$ . We obtain:

$$\begin{aligned}
 \frac{\partial e_t^{N^{**}}}{\partial e_t^{K^{**}}} &= \frac{-q_t + \frac{1}{2} \left[ (q_t e_t^{K^{**}})^2 + 4(1 - q_t) \frac{A^2}{\psi^2} \right]^{-1/2} 2 (q_t e_t^{K^{**}}) q_t}{2(1 - q_t)} \\
 &= \frac{-1 + \left[ (q_t e_t^{K^{**}})^2 + (1 - q_t) \frac{4A^2}{\psi^2} \right]^{-1/2} q_t e_t^{K^{**}}}{2(1 - q_t)}
 \end{aligned}$$

That expression is negative iff:

$$\begin{aligned} \left[ (q_t e_t^{K^{**}})^2 + (1 - q_t) \frac{4A^2}{\psi^2} \right]^{-1/2} q_t e_t^{K^{**}} &< 1 \\ (q_t e_t^{K^{**}})^{-1/2} \left( q_t e_t^{K^{**}} + (1 - q_t) \frac{4A^2}{\psi^2 q_t e_t^{K^{**}}} \right)^{-1/2} q_t e_t^{K^{**}} &< 1 \\ \left[ \frac{q_t e_t^{K^{**}}}{q_t e_t^{K^{**}} + (1 - q_t) \frac{4A^2}{\psi^2 q_t e_t^{K^{**}}}} \right]^{1/2} &< 1 \end{aligned}$$

which is true.

Let us now take the FOC for optimal  $e_t^{K^{**}}$  and differentiate with respect to  $e_t^{K^{**}}$ . We obtain:

$$\begin{aligned} A \frac{-1}{2} (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-3/2} \left( q_t + (1 - q_t) \frac{\partial e_t^{N^{**}}}{\partial e_t^{K^{**}}} \right) + \frac{-1}{2} q_t A (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-3/2} \\ + \frac{-1}{2} q_t e_t^{K^{**}} A \frac{-3}{2} (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-5/2} \left( q_t + (1 - q_t) \frac{\partial e_t^{N^{**}}}{\partial e_t^{K^{**}}} \right) - \frac{1}{2} \psi (e_t^{K^{**}})^{-1/2} \end{aligned}$$

Let us first compute  $\left( q_t + (1 - q_t) \frac{\partial e_t^{N^{**}}}{\partial e_t^{K^{**}}} \right)$ :

$$\begin{aligned} q_t + (1 - q_t) q_t \frac{-1 + \left[ (q_t e_t^{K^{**}})^2 + (1 - q_t) \frac{4A^2}{\psi^2} \right]^{-1/2} q_t e_t^{K^{**}}}{2(1 - q_t)} \\ = q_t \frac{1}{2} \left[ 1 + \frac{q_t e_t^{K^{**}}}{\left[ (q_t e_t^{K^{**}})^2 + (1 - q_t) \frac{4A^2}{\psi^2} \right]^{1/2}} \right] > 0 \end{aligned}$$

Let factorize the expression:

$$\begin{aligned} A \frac{-1}{2} (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-3/2} \left( q_t + (1 - q_t) \frac{\partial e_t^{N^{**}}}{\partial e_t^{K^{**}}} \right) \left[ 1 - \frac{3}{2} \frac{q_t e_t^{K^{**}}}{(q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})} \right] \\ + \frac{-1}{2} q_t A (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-3/2} - \frac{1}{2} \psi (e_t^{K^{**}})^{-1/2} \\ < 0 \end{aligned}$$

The rationale goes as follows. Suppose that  $1 - \frac{3}{2} \frac{q_t e_t^{K^{**}}}{(q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})} > 0$ .

Then, given  $q_t + (1 - q_t) \frac{\partial e_t^{N^{**}}}{\partial e_t^{K^{**}}} > 0$ , the three terms of the sum are negative.

Suppose now that  $1 - \frac{3}{2} \frac{q_t e_t^{K^{**}}}{(q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})} < 0$ . Let us rewrite the expression

as:

$$\begin{aligned}
& A \frac{-1}{2} (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-3/2} q_t \left[ 1 - \frac{3}{2} \frac{q_t e_t^{K^{**}}}{(q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})} \right] \\
& + \frac{-1}{2} q_t A (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-3/2} - \frac{1}{2} \psi (e_t^{K^{**}})^{-1/2} \\
& + A \frac{-1}{2} (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-3/2} (1 - q_t) \frac{\partial e_t^{N^{**}}}{\partial e_t^{K^{**}}} \left[ 1 - \frac{3}{2} \frac{q_t e_t^{K^{**}}}{(q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})} \right]
\end{aligned}$$

We know from the SOC that the sum of the first three terms is negative. But as  $\frac{\partial e_t^{N^{**}}}{\partial e_t^{K^{**}}} < 0$  and  $1 - \frac{3}{2} \frac{q_t e_t^{K^{**}}}{(q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})} < 0$ , the last term is also negative.

Let us define:

$$F(e_t^{K^{**}}) = A (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-1/2} + \frac{-1}{2} q_t e_t^{K^{**}} A (q_t e_t^{K^{**}} + (1 - q_t) e_t^{N^{**}})^{-3/2} - \psi (e_t^{K^{**}})^{1/2}$$

$$\text{where } e_t^{N^{**}} = \frac{-q_t e_t^{K^{**}} + \sqrt{(q_t e_t^{K^{**}})^2 + 4(1 - q_t) \frac{A^2}{\psi^2}}}{2(1 - q_t)}.$$

We know that

$$\begin{aligned}
F(e_t^{K^{**}}) &= 0 \\
F'(e_t^{K^{**}}) &< 0
\end{aligned}$$

We want to compare  $e_t^{K^*}$  and  $e_t^{K^{**}}$ . For that purpose, given the two previous equations, we have that if  $F(e_t^{K^*}) > 0$  then  $e_t^{K^*} < e_t^{K^{**}}$ .

Let us now substitute for  $e_t^{K^*}$  in that expression. We obtain:

$$\begin{aligned}
F(e_t^{K^*}) &= A (q_t e_t^{K^*} + (1 - q_t) e_t^{N^*})^{-1/2} + \frac{-1}{2} q_t e_t^{K^*} A (q_t e_t^{K^*} + (1 - q_t) e_t^{N^*})^{-3/2} - \psi (e_t^{K^*})^{1/2} \\
&= A \left( q_t \frac{A}{2\psi} + (1 - q_t) \frac{A}{\psi} \frac{-\frac{q_t}{2} + \sqrt{\left(\frac{q_t}{2}\right)^2 + 4(1 - q_t)}}{2(1 - q_t)} \right)^{-1/2} \\
&\quad + \frac{-1}{2} q_t \frac{A}{2\psi} A \left( q_t \frac{A}{2\psi} + (1 - q_t) \frac{A}{\psi} \frac{-\frac{q_t}{2} + \sqrt{\left(\frac{q_t}{2}\right)^2 + 4(1 - q_t)}}{2(1 - q_t)} \right)^{-3/2} - \psi \left( \frac{A}{2\psi} \right)^{1/2}
\end{aligned}$$

Hence we obtain:

$$\left[ \begin{aligned}
& A \left( \frac{A}{2\psi} \right)^{-1/2} \left( \frac{q_t}{2} + \sqrt{\left(\frac{q_t}{2}\right)^2 + 4(1 - q_t)} \right)^{-1/2} \\
& + \frac{-1}{2} q_t A \left( \frac{A}{2\psi} \right)^{-1/2} \left( \frac{q_t}{2} + \sqrt{\left(\frac{q_t}{2}\right)^2 + 4(1 - q_t)} \right)^{-3/2} - \psi \left( \frac{A}{2\psi} \right)^{1/2}
\end{aligned} \right]$$

Let us now check the sign of that expression. It is positive iff:

$$\left[ A \left( \frac{A}{2\psi} \right)^{-1/2} \left( \frac{q_t}{2} + \sqrt[2]{\left( \frac{q_t}{2} \right)^2 + 4(1-q_t)} \right)^{-1/2} + \frac{-1}{2} q_t A \left( \frac{A}{2\psi} \right)^{-1/2} \left( \frac{q_t}{2} + \sqrt[2]{\left( \frac{q_t}{2} \right)^2 + 4(1-q_t)} \right)^{-3/2} - \psi \left( \frac{A}{2\psi} \right)^{1/2} \right] > 0$$

Let us multiply the whole expression by  $\left( \frac{A}{2\psi} \right)^{1/2}$ . Then:

$$\left( \frac{q_t}{2} + \sqrt[2]{\left( \frac{q_t}{2} \right)^2 + 4(1-q_t)} \right)^{-1/2} \left[ 1 - \frac{\frac{1}{2} q_t}{\frac{1}{2} q_t + \sqrt[2]{\left( \frac{q_t}{2} \right)^2 + 4(1-q_t)}} \right] - \frac{1}{2} > 0$$

Hence

$$\frac{\sqrt[2]{\left( \frac{q_t}{2} \right)^2 + 4(1-q_t)}}{\left( \frac{q_t}{2} + \sqrt[2]{\left( \frac{q_t}{2} \right)^2 + 4(1-q_t)} \right)^{3/2}} > \frac{1}{2}$$

We know that the LHS equals  $\frac{1}{\sqrt[2]{2}} > \frac{1}{2}$  when  $q_t = 0$  and that the LHS equals  $1/2$  when  $q_t = 1$ . Plotting the LHS for  $q_t \in [0, 1]$  shows that the above inequality is always satisfied.

Thus we have  $F(e_t^{K^*}) > 0$  implying  $e_t^{K^*} < e_t^{K^{**}}$ .

Note that, since Nash players have, for a given  $q_t$ , the same reaction functions, given by:

$$e_t^{N^*} = \frac{-q_t e_t^{K^*} + \sqrt[2]{(q_t e_t^{K^*})^2 + 4(1-q_t) \frac{A^2}{\psi^2}}}{2(1-q_t)}$$

$$e_t^{N^{**}} = \frac{-q_t e_t^{K^{**}} + \sqrt[2]{(q_t e_t^{K^{**}})^2 + 4(1-q_t) \frac{A^2}{\psi^2}}}{2(1-q_t)}$$

and since the reaction functions for Nash players are decreasing in respectively  $e_t^{K^*}$  and  $e_t^{K^{**}}$ , the inequality  $e_t^{K^*} < e_t^{K^{**}}$  implies  $e_t^{N^*} > e_t^{N^{**}}$ .

Let us now check whether, under our parametrization, we have  $E_t^* = q_t e_t^{K^*} + (1-q_t) e_t^{N^*} \geq E_t^{**} = q_t e_t^{K^{**}} + (1-q_t) e_t^{N^{**}}$ .

We have

$$E_t^* = q_t \frac{A}{2\psi} + (1-q_t) \frac{A - \frac{q_t}{2} + \sqrt[2]{\left( \frac{q_t}{2} \right)^2 + 4(1-q_t)}}{\psi \cdot 2(1-q_t)}$$

$$E_t^{**} = q_t e_t^{K^*} + (1-q_t) \frac{-q_t e_t^{K^*} + \sqrt[2]{(q_t e_t^{K^*})^2 + 4(1-q_t) \frac{A^2}{\psi^2}}}{2(1-q_t)}$$

We have

$$E_t^{**} = q_t e_t^{K^{**}} + (1-q_t) \frac{-q_t e_t^{K^{**}} + \sqrt[2]{(q_t e_t^{K^{**}})^2 + 4(1-q_t) \frac{A^2}{\psi^2}}}{2(1-q_t)}$$

Let us simplify further:

$$E_t^* = \frac{1}{2} \left( q_t e_t^{K^*} + \sqrt{(q_t e_t^{K^*})^2 + 4(1-q_t) \frac{A^2}{\psi^2}} \right)$$

$$E_t^{**} = \frac{1}{2} \left( q_t e_t^{K^{**}} + \sqrt{(q_t e_t^{K^{**}})^2 + 4(1-q_t) \frac{A^2}{\psi^2}} \right)$$

We have:

$$E_t^* < E_t^{**} \iff \sqrt{(q_t e_t^{K^*})^2 + 4(1-q_t) \frac{A^2}{\psi^2}} - \sqrt{(q_t e_t^{K^{**}})^2 + 4(1-q_t) \frac{A^2}{\psi^2}} < q_t (e_t^{K^{**}} - e_t^{K^*})$$

Obviously, when  $q_t = 0$  or  $q_t = 1$ , the LHS is equal to the RHS.

The function  $\sqrt{(q_t x)^2 + 4(1-q_t) \frac{A^2}{\psi^2}}$  being increasing in  $x$ , and  $e_t^{K^*} < e_t^{K^{**}}$ , we have that the LHS is negative. But the RHS is always positive, since,  $e_t^{K^{**}} > e_t^{K^*}$ . Hence the inequality is always satisfied.