Applications of game theory and dynamics to social and biological sciences

Luís Filipe da Silva Martins
Programa Doutoral em Matemática Aplicada
Departamento de Matemática
2018

Orientador

Alberto Adrego Pinto, Professor Catedrático, Faculdade de Ciências da Universidade do Porto

PhD

Applications of game theory and dynamics to social and biological sciences

by

Luís Filipe da Silva Martins

Thesis submitted to obtain the degree of PhD. in Applied Mathematics

Departamento de Matemática
Faculdade de Ciências
Universidade do Porto

October 2018

Acknowledgements

O meu primeiro agradecimento vai endereçado a toda a minha família e mais especificamente aos meus pais.

Obviously, the greatest part of my acknowledgements goes to my supervisor Alberto Pinto, for his knowledge and endless support which enriched this thesis and all the works that we developed together. Most importantly, I am deeply indebted for all the productive (and less productive) discussions we had, for all his advise, and of course for his friendship.

I would like to thank Elvio Accinelli for the very pleasant times we spent during his visits to Porto, for his knowledge of mathematics, economics and the history of economics, which definitely enriched myself, for his friendship and for all the discussions we had about literature, Uruguay and Mexico, among many other things. I would also like to thank Maria Bettini.

I am deeply indebted to Jim M. Cushing, whom I first met on a conference in Portugal when I was still a Master student, where Jim gave a wonderful lecture which introduced me to the topic of structured population dynamics, for his generosity in proposing us some possibilities of research projects, one of which would become chapter 4 of this thesis as well as a joint paper. I hope we can work together again in the future.

I would like to thank Jorge Zubelli for his hospitality when I visited Rio de Janeiro and presented an incipient version of our work at his seminar.

I would like to thank all the other people with whom I worked during this time as a PhD. student: Luis Quintas, Bruno Oliveira, Jorge Oviedo and Amy Veprauskas.

I would like to thank Professor Margarida Brito for all our conversations throughout the years. Many thanks to our research/discussion group whose discussions are, most frequently, the first refereeing of our projects: João Paulo Almeida, José Martins, Bruno Oliveira, Diogo Pinheiro and Renato Soeiro. Many thanks to Saul Mendoza Palacios and Erubiel D'Ordaz for our fruitful and funny discussions.

I would like to thank my friends, for all our long nights filled with long, vivid discussions about almost every imaginable topic, invariably (or less so) ending in a proverbial "we all agree": Melânia Carvalho, Miguel Couto, Margarida Feijó, António Girão, Leonor Guedes de Oliveira, Inês Sampaio and Pedro F. Silva.

I would like to thank each and everyone of my current and former colleagues at the mathematics department for their friendship, and for the pleasant discussions about football, mathematics, books, *etc.*, which we had throughout these years: Atefeh Afsar, Carla Azevedo, Ana Carvalho, Elisa Silveira, Jorge Soares, Pedro Soares and Yusuf Alihu.

Estoy muy agradecido a Karina Peña Rodriguez, Enrique Casares y George por toda su hospitalidad. Gracias por mostrarme México.

Aos meus amigos de sempre (20 e tal anos, pá??): Ana Jesus, Mariana Moura, Guilherme Novais and Inês Pereira.

I finish my acknowledgements in a more immaterial tone by thanking all the novelists, playwrights, film-makers, for all the things I have read or seen in movies or in theatre during my life. There are always those moments where art presents something invaluable. A final acknowledgement to Boavista Football Club. It is always a great pleasure and honour to be a member of this club.

This work had the financial support of FCT - Fundação para a Ciência e a Tecnologia - through a PhD. grant of the MAP-PDMA program with reference PD / BD / 105726 / 2014. I also would like to thank the financial support of IMPA - Instituto de Matemática Pura e Aplicada, at the occasion of the 16th SAET Conference on Current Trends in Economics held in Rio de Janeiro, Brasil, July 2016, where part of chapter 3 of this thesis was developed and presented.

Abstract

In this PhD. thesis several areas of mathematics such as game theory and dynamical systems are applied to social sciences, such as sociology, political science and economics, and to biological sciences.

We analyse a game theoretic model of corruption with two populations: a government (top political elite) and officials appointed by the government that serve the public, and where both can choose between being honest or corrupt. The government, supposed to be the controller of corruption can itself be corrupted and arises the question of how to control the controller. We analyse the importance of citizens in this question through their elective power and when they are active decision makers whose behaviour is either complacent or nor with corruption. We characterize the evolution of corruption using evolutionary dynamical models of the game and we characterize situations where corruption becomes endemic and can become a social trap and perpetuate itself in some cases through cycles of corruption. We also consider a three-population game where the citizens are full players of the game together with the government and officials, and they have two behaviours that are compliance and non-compliance towards corruption. We again study and characterize the evolutionary dynamics of corruption in this setting, taking now into account the influence of citizenship and democracy, represented here by the compliance or not of citizens towards corruption, in the outcomes of the long-term evolution of corruption.

We study the impact of the use of competitive Nash tariffs and social cooperative tariffs in a standard international trade model where firstly countries choose tariffs and secondly firms compete in quantities produced. We study when according to the comparison of competitive and cooperative tariffs the country game is a social equilibrium, is like a prisoner's dilemma or an asymmetrical dilemma where one country is benefited while the other is harmed. Which situation occurs is very important in international trade and in the making of trade agreements. We take into account these situations and we analyse, in terms of the model parameters, other externalities that might arise when the countries choose to enforce cooperative tariffs in a trade agreement, and we explain and interpret these externalities and how they are relevant in the context of international trade.

We study an evolutionary nonlinear matrix model in discrete-time that in addition to a population dynamics component also tracks Darwinian dynamics of the evolution of a suite of mean phenotypic traits subject to natural selection that influence the population dynamics. We prove a bifurcation theorem that characterizes the nature and stability of the fundamental bifurcation that consists of the appearance of a continuum of positive (survival) equilibria that bifurcates from the extinction equilibria that loses stability when the inherent growth rate increases through 1. We obtain that forward bifurcations are stable whereas backward bifurcations are unstable. We apply the results to a modified version of the classic Ricker equation and study other features of the example such as backward bifurcation induce Allee effects where a stable survival equilibrium coexists with a stable extinction equilibrium.

Keywords: Game Theory; dynamics; evolutionary game theory; Nash equilibria; dynamical equilibria; stability; bifurcations; corruption; citizenship; social traps; international trade; tariffs; welfare; trade agreements; externalities; population dynamics; nonlinear matrix models; extinction; survival; persistence.

Resumo

Nesta tese de doutoramento aplicámos diversas áreas da matemática, como teoria de jogos e sistemas dinâmicos às ciências sociais, como sociologia, ciência política e economia e à biologia.

Analisámos um modelo de teoria de jogos, com duas populações, aplicado à corrupção, onde um governo (elite política) e oficiais nomeados pelo governo que servem o público podem ser honestos ou corruptos. O governo é o presumível controlador da corrupção, mas pode também ser corrupto e por isso surge a questão de como controlar o controlador. Analisámos a importância dos cidadãos nesta questão através do seu poder electivo e quando os cidadãos tomam decisões e o seu comportamento é de complacência ou não com a corrupção. Caracterizámos a evolução da corrupção usando modelos de dinâmica evolutiva para o jogo considerado, e caracterizámos situações onde a corrupção se pode tornar endémica e pode tornar-se uma armadilha social e se perpetuar, em alguns casos através de ciclos de corrupção. Também estudámos um jogo com três populações onde os cidadãoes são jogadores plenos, conjuntamente com o governo e os oficiais, e onde têm dois comportamentos: complacência e não-complacência com a corrupção. De novo estudámos a dinâmica evolutiva da corrupção neste contexto, tomando agora em atenção a influência da democracia e cidadania, representada aqui pela complacência e não complacência dos cidadãos face à corrupção, nos resultados da evolução a longo prazo da corrupção.

Estudámos o impacto do uso de tarifas Nash competitivas e tarifas sociais cooperativas num modelo standard the comércio internacional onde primeiramente os países escolhem tarifas e posteriormente as firmas competem em termos de quantidades produzidas. Comparando as tarifas competitivas e cooperativas estudámos quando o jogo entre países é um equilíbrio social, é um dilema do prisioneiro ou um dilema assimétrico onde um país é beneficiado e o outro prejudicado. Analizámos estas situações bem como outras externalidades em termos dos parâmetros do modelo, externalidades essas que podem surgir quando os dois países pretendem aplicar as tarifas cooperativas num acordo comercial, e explicámos e interpretámos estas externalidades e a sua relevância no contexto de comércio internacional.

Estudámos um modelo matricial evolutivo não-linear em tempo discreto que para além da dinâmica populacional inclui também a dinâmica Darwiniana de um conjunto de fenótipos médios na população que estão sujeitos a selecção natural e que influenciam a dinâmica populacional. Demonstrámos um teorema de bifurcação que caracteriza a natureza e estabilidade da bifurcação fundamental que consiste do surgimento de um contínuo de equilibrios positivos (de sobrevivência) que bifurcam do equilíbrio de extinção quando a taxa de crescimento inerente ultrapassa o valor 1 e este último perde estabilidade. Obtivemos que bifurcações para a direita são estáveis enquanto que bifurcações para a esquerda são instáveis. Estudámos uma aplicação dos resultados a uma versão modificada da equação clássica de

Ricker e estudámos outros fenómenos desta aplicação como efeitos de Allee associados a bifurcações para a esquerda, onde um equilibrio de sobrevivência estável e um equilíbrio de extinção estável coexistem.

Palavras-chave: Teoria de jogos; dinâmica; teoria de jogos evolutivos; equilíbrios de Nash; equilíbrios dinâmicos; estabilidade; bifurcações; corrupção; cidadania; armadilhas sociais; comércio internacional; tarifas; bem-estar (welfare); acordos de comércio; externalidades; dinâmica populacional; modelos matriciais não lineares; extinção; sobrevivência; persistência.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 General aims and scope of the thesis

This thesis is the result of different projects I have been involved in. We make use of different areas of mathematics, such as game theory, dynamical systems and evolutionary game theory and we use them with the goal of developing mathematical models to address and study issues in very different and distinct areas of knowledge such as social sciences, namely sociology, political science and economics, and biological sciences.

This work is organized as follows. In chapter 2 we study evolutionary dynamics models of corruption. The first part of this chapter is based on the joint work [5] with E. Accinelli, J. Oviedo, A. A. Pinto and L. Quintas. In chapter 3 we study Nash and social tariffs in international trade and trade agreements. The first part of this chapter is based on the joint work [69] with A. A. Pinto and J. P. Zubelli. In chapter 4 we study the fundamental bifurcation of evolutionary nonlinear matrix models with multiple evolutionary traits. This chapter is based on the joint work [29] with J. M. Cushing, A. A. Pinto and A. Veprauskas.

We may say that the *leitmotiv* of this thesis is the application of mathematics to the modelling of diverse problems in the aforementioned areas of knowledge by means of *game theory*, *dynamics* and *evolutionary game theory*. Chapter 2 addresses the mathematical modelling of corruption using methods from game theory and dynamics and in particular the theory of evolutionary games. In our study of corruption we focus on the interplay between government and officials on one hand, and the citizens of a country and their influence and complacency with corruption. Our main objective is to study how can corruption become a self-enforcing mechanism, and to study the evolution of corruption and how this evolution can be broken and how several feature observed in a society and its institutions can lead to increasing corruption or even degeneration into autocratic political regimes. Citizen intervention is central in our analysis and to allow for the possibility of breaking vicious cycles of corruption. In chapter 3 we study a standard model of international trade paying specific attention to game theoretic aspects, or strategic aspects of tariffs choosing between countries. In our study of tariffs we make a comparison between the results a country obtains in international trade from using the tariffs that yield them their best result competitively and the tariffs that yield the best results when the country cooperates and joins another country. So indeed we compare competition with cooperation. We make a proposal of trade agreement to enhance cooperation and discuss how complicated and difficult its enforcement may be due to externalities related to trade such as the national output in terms of quantities produced by a country and surplus associated to the consumers. In chapter 4 we study an evolutionary version of a population dynamics model. In our study of a population dynamics model we start by considering a variation of the model to include components that model Darwinian evolution. We analyse and characterize when population persistence and possible extinction occurs, and we do a specific study to an evolutionary version of a classical population dynamics model known as the Ricker equation.

We model such problems using different areas of mathematics. More precisely, we use areas/modelling techniques such game theory and dynamics, both in discrete-time and in continuous-time, i.e. using both difference and differential equations.

Game theory is the study of strategic decision making, or the study of rational conflict. This field of applied mathematics is usually considered to have started with the groundbreaking work of mathematician John von Neumann and economist Oskar Morgensten in the book [102]. Several classical references in game theory were used in the making of this thesis. For the interested reader which is not familiar with this area of we recommend for instance [40]. Some of the features from game theory that we used in this work include modelling through normal-form games and multiple-stage games, computation and analysis of Nash equilibria and Perfect Nash equilibria, social optimum equilibria, or Pareto optimum (Pareto efficient) allocation. These classical concepts in game theory allow us to compute Nash equilibria or conditions for a Nash equilibria and so study self-enforcing conditions of some outcomes of the games we model. We apply these concepts to study equilibria in the corruption games that we constructed, to study trade agreements and welfare sharing, that allow us to interpret the results of the models by static analysis.

Another main area of mathematics that was used in this work is dynamical systems. We use dynamical system to interpret dynamical models, using tools such as the study of steady-states, or dynamical equilibria and their stability, or the study of periodic orbits and bifurcations. We also use bifurcation theory, as tools such as Fredholm alternative and Lyapunov-Schmidt reduction to characterize properties and characteristics of bifurcations. For the interest reader in continuous-time dynamics by means of differential equations we recommend the classic reference [46]. For discrete-time dynamics we recommend [37]. For a good introduction to bifurcation theory including the aforementioned tools we recommend [52].

The dynamical models that we consider are not independent from game theory. Indeed, the dynamics that we considered are derived from an extension of classical game theory known as evolutionary game theory, that allows to render the classical, static nature of game theory into a dynamical and evolutionary basis. Evolutionary game theory started in the 1970's with the seminal work of biologists John Maynard Smith and George Robert Price. Contrary to classical game theory which relies heavily on rationality assumptions, evolutionary game theory focuses more on the dynamics of strategy change on a Darwinian competition framework where strategies evolve along time influenced by the frequency of other strategies and with selection of strategies taking place according to their fitness, which measures their ability to survive and reproduce. Several dynamical models have been widely used, the most famous being probably the replicator equation, which has been applied to very wide areas such as economics and other social sciences. For the interested reader in evolutionary game theory we recommend [103] and [47], with the latter also providing a good introduction to the topic of population dynamics. We also recommend [100] for a good reference in evolutionary game theory with more emphasis in biological aspects, such as Darwinian dynamics, evolution and co-evolution, fitness function and natural selection.

In the remainder of this introduction we shall describe with more detail each one the subsequent chapters of this thesis, by framing the work in its related scientific literature and emphasizing its objectives and its main results, and also the prospectives of future works in related topics.

1.2 Detailed description of the thesis

Corruption is one of the most important issues nowadays with the public awareness to the subject increasing even if the phenomena is difficult to overcome. It has been inreasingly studied by sociologists, historians, political scientists and economists. The consequences of corruption are known to be severe. Firstly, for example, corruption is costly for economies, as attested in several studies. The European Union published its first ever anti-corruption report [38] in February 2014, where it concluded that bribery, embezzlement, financial and tax fraud, money laundering, cronyism, *etc.*, "cost the EU economy 120 billion euros per year, just a little less than the annual budget of the European Union". Secondly, apart from the fact that corruption is costly, it deprives citizens of more than money, since it may have very deep sociological consequences. There is a lot of empirical and theoretical evidence showing that high corruption levels are likely to increase income inequality and poverty, actually undermining democracy. A good account of the effect of corruption on these two factors is provided in [44]. Furthermore, in [72] the author presents an empirical account of the effect of corruption in investment, namely that corruption lowers investment and growth. In addition to costs regarding the economy, frequently there are also great political costs associated to corruption. In [91], using data from four Latin American countries, the author asserts that corruption erodes confidence and belief in the legitimacy of the political system.

In general, corruption can be defined as the misuse of public power for private benefit. Different ways of defining corruption and its limitations are discussed in [50] and [48]. In [50] the author divides corruption into three types. First, 'grand corruption' involving political elites. Second, bureaucratic corrupt practices involving appointed bureaucrats or officials that deal more directly with citizens and the corrupt acts that occur when they meet the public. Third, legislative corruption, i.e., when voting behaviour and policy making of the legislator are influenced by private interests and lobbies, eventually benefiting private interest and harming public interest in exchange for benefits to politicians. The three types of corruption there described are obviously interconnected.

There is a profuse economic literature related to the topic of administrative and political corruption. Pioneering works in the area are [85] and [86]. A relevant insight that emerges from many studies is the self-reinforcing nature of corruption: in an environment where corruption is the norm, corruption tends to be imitated and to persist. See for instance [67], [68], [88] and [76]. In much of this literature, the conditions under which people are willing to be corrupted are analysed. For instance, public servants must ensure compliance with the law, payment of taxes by citizens, compliance with rules aimed at preventing pollution, rules aimed at avoiding illegal rent-seeking activities (see for instance [56]), *etc.*, but may be willing to follow a corrupt behaviour in some cases where they may obtain personal gain. In recent works, the evolution of the corruption in a given society is modelled using evolutionary game theory and dynamics such as the replicator dynamics or imitative dynamics. In the evolutionary approach, strategies in a population arise by a trial-and-error and learning process where the best performing strategies will tend to be adopted or imitated. Under this evolutionary approach and under given social conditions, corruption can become a dominant strategy and may be a dynamically stable outcome, in other words, all-pervasive corruption in a country. In [3] the authors consider a model where individuals may have motivations to follow a corrupt behaviour and study how imitation may lead to increasing corruption. In [4], the authors propose a model where income taxation and tax evasion are considered. In [6] the authors propose a model to analyse the effects of corruption in environmental protection by analysing a game between firms and public officials, where firms may pursue corrupt behaviour by bribing public officials in order to get positive reports regarding the environmental status of the firm to avoid fines.

Furthermore, there is the phenomena of ‘grand corruption’ and/or legislative corruption. The central authority or the government is a benevolent planner trying to maximize the social welfare and should play an important role to deter and to control the evolution of corruption. However, individual members of these political elites or the elite groups as a whole can benefit from the evolution of corruption. In addition to their role as balancing the interests of society, they also want to remain in power and may take decisions that serve their own interests at the cost of the society’s interest. In [78] the author does a theoretical analysis of the appropriation of welfare by an autocratic elite, including how it is so and the relation of security of property and contract rights, which enhances the economy, and democracy itself. The relations between individual corruption and institutional corruption is very well exposed and analysed in the books [62] and [96], which focus attentively in the case of the USA Congress. Political elites’ legislative decisions are also influenced by interest and lobby groups that seek to have gains from government’s policy making. Models for this kind of corruption and political influence are proposed in [11] and [43]. Furthermore, a widely studied guise of corruption is vote buying. A classic work of history that documents vote buying practices in the XIX century and before is [41]. A modern account on voting fraud is [7]. An excellent work on the nature and forms that vote buying can assume is given in [89].

The so called petty corruption, i.e., smaller scale corruption that occurs at the implementation end of public services when officials meet citizens, creates incentives for the development of corrupt behaviour at this level and in this way the society as a whole may become corrupt and a self-reinforcing mechanism might be in place. In [96], the author goes beyond individual corruption to address institutional corruption and the relations and possible collusion between these two forms of corruption and how petty corruption may give rise to corrupt practices at a wider sphere. This possibility raises a question: who and how controls the controller?, i.e., who and how should may we institutionally (in the sense of ’rules of the game’ and institutional arrangements) can this be solved.

An interesting discussion on this point was introduced in a lecture by [49], at the occasion of the Nobel Memorial Prize in Economic Sciences in 2007. In the lecture, Hurwicz retakes a question posed by the Latin author Juvenal: *Quis custodiet ipsos custodes?* . This Latin locution has been variously translated as “Who will guard the guardians?”, “Who can watch the watchmen?”, or “Who will guard the guards themselves?”. The quotation is from one of Juvenalian satires from the $1^{st}/2^{nd}$ century CE (see [51]). In the satire Juvenal refers to the inability of enforcing moral behaviour on women by using an enforcer since they are corruptible as well (Satire VI, verses 346-348). The modern usage of the quotation refers very often to the problematic of how to control political power that may be subject to deviating behaviours. This interpretation resounds a passage of the *Republic* by Plato (see [79]) written in 4^{th} century BCE. In the *Republic*, Socrates describes a society with a class of guards to protect it and that they will take care of themselves and should be trusted by people. Glaucon refers that it would be absurd that a guardian would need another guardian. According to Plato, ideally, it would suffice for the guards to perform their functions with honesty, to make them believe that they are better than those to whom they render their services (this is Plato’s presentation of the political philosophy concept of ‘magnificent myth’, sometimes translated as the ‘noble lie’) and that therefore it is their responsibility to watch and to protect the inferiors in the social hierarchy.

Therefore we can say that we have two different views regarding this question. On one hand, Plato’s answer can be seen as optimistic. On the other hand, taking the *lato sensu* interpretation of the Juvenalian quote we can say that his answer is pessimistic since it would require an infinite regress of guardians. In the quoted lecture, Hurwicz suggests that the infinite regress might be avoided if the guardians hold an elective office, since then citizens can act as a top guardian by means of their elective power.

The aim of our work is to give a partial answer to this question in the spirit of Hurwicz. We propose a model where we include the possibility that citizens act as a top guardian and we conclude that they are key in the fight against corruption, because in a democratic country they have the possibility to exert democratic pressure through elections, thereby demanding the first level of guardians, the government, to fight corruption. This kind of influence of citizens fits in the literature of cyclic games, first introduced in [80] that are part of the broader class of polymatrix games (see [81]). In cyclic games, players play “pairwise”, thus forming a cycle.

We consider some of the forms of expression of corruption previously described in our model. More precisely, in our work, we have petty corruption by public officials that may be more interested in their own profits, rather than fulfilling their duties. We also consider corruption that occurs at a higher level, such as the government or the ruler elite, that may collude with officials corruption and also has gains from being maintained in power and eventual promiscuous relations with, for instance, economic elites, or gains from it might get from their relations with interest groups and lobbyists. Hence, government will be the first level of guardians. The second level of guardians is made up of citizens with elective power. More precisely, we introduce a game with three players (populations): government, officials and citizens, where the first two must choose between a corrupt or non-corrupt behaviour. Citizens have a role influencing the prospective benefits government being maintained in power. This influence results in a quantity that we call the index of intolerance to corruption. The persistence of corrupt behaviour in a democratic country will depend on the degree of intolerance of citizens and other socio-economical quantities of the model. Intolerance to corruption can take many forms and be related to a wide range of causes. In future works, we will consider different versions of this index, for instance depending on citizens perception of corruption by the media and other characteristics.

We also consider a modified game with three populations: the government, officials and citizens. Unlike the game first described, citizens not only have an influence on the payoffs of government and officials but are full players of the game. They have two strategies that we will designate by corrupt and non-corrupt, in order not to cause confusion with the strategies of he government and officials, but that may perhaps be more precisely interpreted as compliance and non-compliance towards corruption. Hence the distribution of citizens among these two groups changes dynamically as the the other players. Citizens participation in the game can be seen as representing citizenship and, to a higher extent democracy itself, as it is known that there are historical periods when citizens compliance with some acts as corruption increases or decreases depending on factors such as their perception, media attention and other aspects.

We consider an evolutionary version of the game by means of the replicator dynamics. This kind of dynamics is also called myopic dynamics, since agents adjust their strategies according to a comparison between their payoff and the average payoff. An strategy performing higher than average will tend to be selected by evolution. For good introductions to the fascinating subjects of game theory and evolutionary game theory we recommend [47] or [103]. We obtain five steady states for the dynamics. Four of them are pure equilibria where the government and officials population either choose to be corrupt or non-corrupt. The other equilibrium corresponds to a mixed strategy where the government and officials choose to be corrupt or not with a certain probability, or equivalently, to a certain ratio representing the proportion between the two strategies within the population of officials and the government's acts. The trajectories of the system are initial condition dependent and we characterize the evolutionary outcomes of the system, i.e., the corruption behaviour patterns that are selected by means of the replicator dynamics. We do the stability analysis of the stationary points of the dynamical system and the Nash equilibrium analysis according to the following quantities: i) the fines practised by the non-corrupt government; ii) the re-election power of the government; iii) the welfare that a corrupted government takes from officials; iv) the costs of a non-corrupt government in fighting corruption. We obtain three different outcomes: i) two different Nash equilibria that are stable coexist; ii) only one Nash equilibrium that is stable; iii) cycles of decreasing and increasing corruption. We show how a sudden change in the evolution of corruption might occur as a consequence of changes in the intolerance index and might lead to the non-corruption equilibrium. However, we observe that in some cases when the index of intolerance to corruption is low, then corruption can persist for instance in the form of corruption cycles. More extreme situations may also occur, where equilibria in which corruption exists are stable. That may be associated with situations of declining democracy or even with dictatorships that might also occur when the index of intolerance is not sufficiently high. Thus, an increase in this index, which can be then regarded as a measure of popularity of the government, is essential to surpass those situations.

Game theory has been increasingly used in economic modelling and in problems related to international trade. The strategic nature of international trade, for instance in the choice of tariffs makes it a good field for the intervention of game theory, in other words, the study of strategic decision-making, or the study of rational conflict.

One good reference of applications of game theory to international economics and international trade is provided in [74]. We also recommend the classic reference [101] from the Austrian School for a theoretical analysis of international trade and commercial policy. There is a vast literature in international trade models using game theoretic framework with both complete and incomplete information (uncertainty) and including or not analysis of trade policy. A good review of stratgic/game theoretic aspects of trade policy is provided in [33]. In [21] and [20] the authors consider a Cournot model with tariffs without and with cost uncertainty respectively. In [94] the authors propose a model including government subsidies to firms in the form of R&D subsidies and in [15] they study a model where governments subsidize firms over the produced quantities to help them in competition against foreign producer and also study the extension of the model to a supra-game between governments. The relation with the tariffs literature is that when a negative optimal level of subsidy occurs, it is interpreted as an export tax. [64] extends this work to study optimal export subsidies (and export taxes) under incomplete information. Regarding the subject of export promotion see also [34] for a model where targeted export promotion policies are studied and [42] for a critique of export promotion policies. In [8] the authors study dynamic patterns of trade policy, namely protection with respect to trade volumes. Also on the subject of trade patterns and gains see [45]. In [39] the authors study price competition (inspired from the Bertrand competition model) between two international firms with tariffs. Other works in multimarket/international trade models under oligopoly are for instance [17], [35], [32]. In [14] a model for intra-industry trade is proposed and analysed and in [55] a oligopolistic model with trade restriction is analysed.

The question of the enforcement of trade agreements, i.e., the enforcement of cooperating strategies, is also a very active research topic (see a review of early contributions to this topic in [95]). The enforcement of trade agreements, i.e., the enforcement of cooperating strategies has two important features. Firstly, is that at least one country may have an incentive to unilaterally deviate from its social (cooperative) tariff to its Nash (competitive) tariff. Consequently, such country will eventually deviate to its competitive tariff if there is no punishment associated to this. Therefore, it is important that trade agreements present a mechanism to punish such deviations. Secondly, one must observe that there is not a supra-national authority to enact the punishment mechanism of the violating country for its eventual deviations from cooperation. This implies the need for international agreements to be self-enforcing.

These characteristics, particularly that of self-enforcing agreements lead to the study of the enforcement issues by means of certain repeated games that possess and present a good deal of important features of cooperative self-enforcing agreements. As a result, several instruments have been proposed and studied with the objective of achieving the (self)-enforcement of international trade agreements, as well as addressing the question of their efficiency, particularly when compared to the threat of tariff retaliation. See for instance [9] for a study of enforcement in the context of the General Agreement on Tariffs and Trade (GATT) and its successor, the World Trade Organization (WTO).

In [65] the authors adopt the repeated game approach and propos alternative instruments in the context of a trade agreement between two symmetric countries. More precisely, the authors compare the effectiveness of retaliatory tariffs with that of a financial compensation by means of a monetary fine to the country that violates cooperation. Unlike the most common tariff retaliation, which is imposed by the injured country and only depends on it, monetary fines have an enforcing problem because they must be voluntarily paid by the country that has deviated from cooperation. They showed that monetary fines yield the same cooperative outcome as tariff retaliation, except when a country deviated from the agreement due to an unanticipated shock in the model parameters (in the case, a political-economical parameter) with monetary fines being preferable to tariff retaliation in that situation. They also studied the possibility of the countries exchanging bonds, and concluded that this yields the same cooperation power as tariff retaliation not yielding a more cooperative outcome. In [66], the authors introduce size inequality between countries, by considering one large country and a region of equal market size with a number of identical small countries. The fact that in the second region countries are individually small generates a coordination externality among themselves, as they cannot credibly threat tariff retaliation, but they would if they act like a group. Thus, in trade agreements based solely on tariff retaliation, coordination externalities generate asymmetric outcomes. They show that improvements in efficiency and more symmetric outcomes can be obtained by including specific financial instruments such as monetary fines and bonds in conventional trade agreements based only on the threat of tariff retaliation. In [53], a model with two asymmetric countries is considered and it is shown that an efficient trade agreement might not lead to free trade. Various types of transfers between countries are studied, such as financial (monetary fines), foreign aid and side payments intended to offset a loss resulting from a trade agreement.

Within the frame of repeated games, in [10] and [71] the authors consider a model with two firms competing in the same country and they study the phenomena of dumping. They interpreted the deviation from collusion by the foreign firm as dumping, with the foreign firm subsequently suffering a period of punishment, where punishment results from the home firm lobby on its government, and the imposition of a tariff that makes the foreign country unable to export during the period of punishment.

The authors study two possibilities after the deviation and punishment periods, with both competing in a Cournot way or the repetition of the deviation and punishment phases. In the former paper they study the symmetrical countries case while in the latter they consider the asymmetrical countries case. In [70] the authors consider a model where a firm has the monopoly in its home market, but divides the market of the foreign country with a firm from that country. They study deviation from collusion in the foreign market by the firm that sells in the two countries either by increasing production solely in its foreign market to lower prices and thus making dumping, or by increasing production in both countries and thus lowering prices in both markets and so deviating without making dumping.

In chapter 3, we consider a classic duopoly international trade model with complete information where there are two countries and a firm in each country that sells in its own country and exports to the other one. The exportation is subject to a tariff fixed by the government of the importing country. The international trade model has two stages: in the first stage, the governments simultaneously choose their tariff rates; and in the second stage, the firms observe the tariff rates and simultaneously choose their quantities for home consumption and for export (see, for instance, [40]).

For the second stage, we consider always the classic competitive (Nash) equilibria that determines uniquely the quantities for home consumption and for export. Now, for the first stage, the decision of the governments to impose or not tariffs can be interpreted as the actions of a game specified by the utilities considered for each country. The utilities (each corresponding to a different game) of the countries that we analyse are the relevant economic quantities of the international trade model for the consumers and firms. In particular, we consider the utilities given by the custom revenue of the countries, the consumer surplus of each country, the profit of the firms and the welfare of the countries at the competitive Nash equilibrium of the second stage game. We show that for each of the above utilities there is a Nash (competitive) equilibrium and a social optimum equilibrium corresponding to the maximization of the joint utility of the two governments. For each one of the utilities, we will compute the Nash and social optimum tariffs and compare these social and Nash equilibria in terms of the economic relevant quantities of the model.

There are three typical game outcomes: the social equilibrium (**SE**), where the social optimum coincides with the Nash equilibrium; the prisoner's dilemma (**PD**), where both utilities are bigger in the social optimum than in the Nash equilibrium; and the lose-win social dilemma (**LW**), where the utility of one of the countries is bigger in the social optimum tariffs and the utility of the other country is bigger in the Nash equilibrium tariffs. For every pair of utilities, we will find which of the three types of outcomes **SE**, **PD** or **LW** occurs in terms of the model parameters.

Which one of the three previously mentioned outcomes (**SE**, **PD** or **LW**) occurs presents qualitatively different scenarios for the involved countries. If the game is of the social equilibrium (**SE**) type, then there is *a priori* no need of a trade agreement, because the two countries are already in the social optimum as the competitive equilibrium coincides with the social optimum that maximizes the joint utilities of the two countries. If the game is of prisoner's dilemma (**PD**) type or of lose-win social strategies (**LW**) type, then at least one of the countries can improve its payoff if they choose to cooperate, and that can be done by means of a trade agreement. In the first case, both governments can make a trade agreement such that they choose the social tariffs, thereby improving the utilities of both countries. However, even if both countries improve their utilities, there might be some externalities associated to the trade agreement. In the second case, both governments can also make a trade agreement such that the countries opt for the social tariffs. However, in this case, the situation is qualitatively different as one of the countries is injured by the change to the social tariffs. So in order to enforce cooperation there is need to compensate that country, for instance, by means of a financial compensation or transfers or in other terms stated in the agreement. In this case, as may happen in the previous case, there might be some externalities associated to the trade agreement.

Indeed, in the first case both countries may improve their welfare but these gains may jeopardize the dominant's country position in international trade, albeit their welfare is improved. This may occur for instance in situations where the country's welfare is improved but other aspects such as their industrial output measured in produced quantities may decrease while other components of the welfare increase, such as, for instance, the consumers surplus due to increase in imports.

We consider a welfare balanced international trade agreement, whose main features are that it increases the welfares of the two countries proportionally to the trade agreement index that is the ratio between the total welfare with social tariffs and the total welfare with the Nash tariffs. With this, the welfare shares of each country relative to the total social welfare is the same as the shares they had at the Nash equilibrium. In other words, the relative welfare share or weight between the two countries in terms of aggregate welfare when the social tariffs are enforced is the same as the relative welfare share or weight between the two countries at a competitive (Nash) situation. So in the welfare balanced trade agreement the balance of forces of each country in terms of welfare remains the same, although each one gets an absolute increase in welfare. So there is an absolute benefit for both countries while the welfare balance is maintained. However, as we argued above this trade agreement might have some difficulty since some externalities may appear throughout the process. They are mainly due to the effects of the enforcing of the social tariffs in some aspects of the country's economy such as the surplus of its consumers, their revenues from tariffs (which tend to be lower or zero when the social cooperative tariffs are enforced), or the profits of its firms and their outputs, which relate to the productive or industrial output of the country. We describe with detail the regions where according to our model, some of these externalities, and hence difficulties in the construction of trade agreements may occur. We do not study the problems of enforcing the trade agreement, and possible solutions to the externalities that arose, that may perhaps be avoided with the addition of other features to the trade agreement.

There are some future work possibilities regarding this and related topics, such as the study of the enforcing of trade agreements, and the study of how the compensation between countries in the welfare balance trade agreement should be enhanced, as well as how to mitigate the externalities that we identified. Some ways to mitigate the effect of the externalities might include agreements in specific issues such as R&D, subsidies or fines to firms, or other kinds of transfer between countries.

Widely used models in ecology and biology are structured population discrete-time dynamical models using matrices. In these models the population is divided into classes, the most common example are age classes. Other examples include division into different sizes or different life stages, for instance in epidemiological models, with different classes for the stages of the disease and their impact in the population. The most classical examples of matrix models are perhaps the so-called Leslie matrix models (see [60] and [61]), that are among the first example of population models using matrices. More recent classical examples include [59] and [98]. A classical one-dimensional discrete-time model is the Ricker equation [82], originally developed in 1954 to study the number o fish at a fishery. For the interested reader in matrix modelling in biology we recommend [19], which surveys the classical results and the effects of periodicity and stochasticity. The original models of Leslie are models with constant coefficient matrices. One possibility to extend these kind of models is to include density effects, effectively rendering the matrix model non-linear (see [19], chapter 16). For a good survey on non-linear matrix models we recommend [28]. For a more thorough study see [23].

One fundamental ecological and biological questions in is population extinction and persistence. This can be mathematically formulated in terms of the stability and instability of the extinction equilibrium and of non-extinction equilibria where the population densities are positive, or stability and instability of non-extinction non-equilibria, such as stable cycles or other kind of attractors. In the case of non-linear matrix models for structured populations, a bifurcation theorem answers this question when the projection matrix is primitive by showing the existence of a continuum of positive equilibria that bifurcates from the extinction equilibrium as the inherent population growth rate passes through 1 (see [23]). This theorem also characterizes the stability properties of the bifurcating equilibria by relating them to the direction of bifurcation, which is forward (backward) if, near the bifurcation point, the positive equilibria exist for inherent growth rates greater (less) than 1. More precisely, in a neighbourhood of the bifurcation point, if the bifurcation is forward then the bifurcating equilibria are stable, and if the bifurcation is backward then the bifurcating equilibria are unstable.

Matrix models can be assumed to be dependent on other features than the density effects only. These might include stochastic effects due to environment and demography, periodicity effects, and natural selection pressure from Darwinian evolution. For the interesting subject of Darwinian evolution and dynamics modelling and its relation to evolutionary game theory we recommend the book [100]. Other good references are [83] and [31].

In chapter 4 we consider an evolutionary game theoretic version of a general nonlinear matrix model that includes the dynamics of a vector of mean phenotypic traits subject to natural selection. We extend the fundamental bifurcation theorem to this evolutionary model. The work in this chapter is the extension of the results in the non-evolutionary setting in [23] and [28] and the one-trait evolutionary results in [24]. We model the evolutionary and population dynamics by assuming that they act on the same time scales, which is in accordance to recent literature arguing on the consistency of the evolutionary and populational dynamics time scales (see [2]). We apply the results to an evolutionary version of the classical Ricker model (see [82]) with an added Allee component inspired by the model studied in [90]. This application illustrates the theoretical results and, in addition, several other interesting dynamic phenomena, such as backward bifurcation induced strong Allee effects and survival when multiple traits evolve. We obtain strong Allee effects associated to backward bifurcations where stable positive equilibria coexist with stable extinction equilibria. We also obtain strong Allee effects where other more complicated non-equilibria attractors, such stable cycles coexist with stable extinction equilibria.

Future work in related topics might include the study of other bifurcations apart from the local study about the fundamental bifurcation that we perform on this chapter. One possibility is the study of backward bifurcation induced Allee effects where a *blue-sky* bifurcation where two equilibria of opposite stability properties arise. Other is the study of evolutionary and non-evolutionary versions of population dynamics of semelparous species, i.e., species with only one reproductive cycle. The population dynamics of such species is more accurately described by irreducible non-primitive matrix models. In this cases dynamical situation is more involved and it is not expected that the direction of the bifurcation is enough to characterize stability as occurs in our work. This is due to the well-known fact that for irreducible matrices a strictly dominant eigenvalue does not exist, and there are several eigenvalues with the same modulus, associated to the h-roots of unity where h is the period of the matrix. So more than one eigenvalue leaves the unit circle, unlike the primitive matrix case where, at least locally, only one eigenvalue (the strictly dominant one) leaves the unit circle simultaneously. The situation becomes combinatorially more complicated as dimension increases. However, for low dimensions it may be able to obtain characterizations of the fundamental bifurcation using methods similar to the ones that we practise in our work.

Chapter 2

Evolutionary dynamical modelling of corruption

The first part of this chapter is based on the joint work:

E. Accinelli, F. Martins, J. Oviedo, A. A. Pinto, and L. Quintas. Who controls the controller? a dynamical model of corruption. *The Journal of Mathematical Sociology*, 41(4):220–247, 2017.

The aim of this chapter is to study the evolution of corruption in a game theory setting. The aim of the first part of this chapter is to try to give at least a partial answer to the question of who will control the the controller supposed to prevent the spreading of corruption, since it can itself be corrupted. In societies, citizens often play an important role that may control the controller since they have voting power and their intolerance to corruption exerts democratic influence in maintenance or not of a government in functions. To analyse this we introduce a normal-form game between a government (that is elected by universal suffrage of citizens) and public officials where both can choose between a corrupt behaviour and an honest behaviour. We build an evolutionary version of the game by means of the replicator dynamics and we analyse and fully characterize the possible trajectories of the system according to the index of intolerance to corruption and other relevant quantities of the model. We also consider a tree-population game between government, officials and citizens who have two behaviours, to be compliant or not with corruption. We show that politically active citizens with high index of intolerance to corruption can effectively prevent the spreading of corruption. However, when intolerance is not high enough it ensures situations where democracy is undermined, as corruption can be sustaining either as a social trap, or in extreme situations such as dictatorships, or through cycles of diminishing and increasing corruption, where corruption is endemic.

This chapter is organized in the following way: in the first part of the chapter, in section 2.1, we address the question of controlling the controller of corruption, i.e., the government, or political elite and the importance of citizens through voting power to, in a way, control the government. In the second part of the chapter, in section 2.2 we extend the game to a situation where citizens are players in the full sense of the word and have two behaviours, to be complacent or not with corruption and we analyse the importance of this citizenship, or democratic intervention in the outcomes of the corruption game. In section 2.3 we present some conclusions.

2.1 Controlling the controller

In the first part of the chapter we try, through the modelling and analysis of a game between government and officials, to answer the question of how is it possible to control the government, or the political elites, which is the supposed controller of corruption, since it can also become corrupted. We introduce citizens influence in the game by means of the probability of re-election of a government, which is related to the intolerance of citizens towards corruption, and is like a measure of popularity of the elites. This influence is crucial to answer the previous question since it allows, in some socio-political situations, to control the controller effectively.

This first part is organized as follows: in section 2.1.1 we introduce a normal-form game where the players are the government and officials, where citizens exert influence by means of their power to maintain a government in power or not in a general election. In section 2.1.2 we analyse when corruption can be a self-reinforcing mechanism. In section 2.1.3 we introduce the index of intolerance to corruption and discuss some of its implications. In section 2.1.4 we consider a dynamical system given by the replicator dynamics to explain the evolution of corruption in a society. In section 2.1.5 we analyse the dynamical equilibria of the system, as well as their stability, and their relations with the Nash equilibria of the game. These relations and the dynamical and Nash outcomes are obtained from the characteristic parameters of the government and officials in the game. In section 2.1.6 we consider some particular cases in which it is possible to describe the evolution of corruption by analytically computing the solution of the dynamical system from given initial conditions. In section 2.1.7 we make some considerations about the role of the index of intolerance of corruption and we mention some possible extensions of this index.

2.1.1 The model

Consider an economy or society where the central authority is elected by universal suffrage of citizens. We identify this central authority with the national government, or in presidential regimes, with the president and his/her administration. They make up the ruling elite. The government can be re-elected or not after each electoral period. The government, in its turn, appoints public officials who may or may not be renewed by the new government. These officials are in charge of carrying out the legal and administrative management of the government and serve directly to the citizens when they require to carry out a specific service.

In general, corruption can be defined as the misuse of public power for private benefit[1]. For instance, officials collect bribes for providing permits, licences, passage through costumers, or avoiding the entrance to competitors in a given market. So when an official is corrupt, it can be regarded as if the agent is selling government corruption for personal gain. At the end of each election period, officials must choose between two different behaviours, namely, properly fulfilling his/her role when his/her participation is required by a citizen or, fulfilling his/her role as long as the citizen pays for it a certain amount of money. We call an honest or non-corrupt official the one that chooses to unconditionally fulfil its functions, otherwise we call the official a dishonest or corrupt official.

Following [92], we define government corruption as the complicity of the government (the ruling elite) with officials that sell government property for personal gain, since sometimes, a dishonest official is colluded with a member of the central authority and both take advantage for this behaviour. We

[1] In [77] the author defines it as a deviation "from the formal duties of a public role (elective or appointive) because of private-regarding (personal, close family, private clique) wealth or status gains."

summarize the activity of the government as follows: it must choose between following a corrupt behaviour or a non-corrupt behaviour, meaning to act in complicity with corrupt officials, or alternatively, punishing them. Corrupt behaviour by the elite may also be seen as the appropriation of some welfare from the officials whether the officials are honest or not[2]. Several examples as well studies of this kind of collusion are considered in [96] and [62]. In [78] the author theorizes over the incentive that a autocrat in doing this welfare appropriation, particularly when in a short-term tenure. The author also argues about the existence of a "stationary bandit", to use the terminology of the author who can somehow stabilize corruption by practising it, thus given an argument for the existence of corruption at higher levels such as a government.

Even when some members of the government can be attracted to act in collusion with dishonest officials, it is necessary to consider that the government is interested in being re-elected for the next period, and they know that this happens only if citizens are satisfied with the performance of the government. Citizens will judge the performance of the central authority through the work of officials who deal directly with them. Citizens prefer a non-corrupt government, but they do not have complete information about the behaviour of the government. They know this information only in an indirect way, and only if they have taken contact with some official. It may be difficult for citizens to refuse to pay that amount in countries where officials corruption is frequent in society, otherwise they will not have their needs fulfilled (see [92]). The fact that citizens have incomplete information results in probabilities of re-electing or not the current government, being it corrupt or not. Since governments have a certain valuation for being re-elected, this mechanism results in an externality for the government caused by election. We will describe below how this externality (related to the index of intolerance to corruption) will appear in the game as well as its main consequences in the following sections of this work. We also assume that a corrupt government can try to corrupt the citizens by means of some kind of payment or royalties in exchange for their vote as we will describe in the game model below [3].

The model can be formalized as a normal-form game. The sets of pure strategies are as follows:

1. Officials must choose between two pure strategies: to be corrupt or not, respectively symbolized by O_c and O_{nc}, so that we have $\Gamma_O = \{O_c, O_{nc}\}$.

2. The central authority or government must choose in the set of pure strategies $\Gamma_G = \{G_c, G_{nc}\}$. A corrupt policy is symbolized by G_c while an honest or non-corrupt policy is denoted by G_{nc}. This represents the behaviour of the political elites.

The payoffs for officials and government are represented in the following table. At the end of every period, the ruling elite must choose between to follow a corrupt or a non-corrupt behaviour (rows), and officials must choose between a corrupt or non-corrupt behaviour (columns). In each cell, the first quantity is the officials' payoff and the second quantity is the government's payoff.
where:

- W is the wage of the officials which is paid by the government.

- M is a fine (or welfare punishment) imposed by an honest government to a dishonest official.

[2]This situation may be also interpreted as a dictatorship where there is a small political elite that benefits from corruption. We will explore this interpretation with more detail later.

[3]The vote buying consists in the bribery of a group of citizens with the aim of obtaining their vote in favour of re-election. The voter would be compensated with cash or some bonus. This *modus operandi* is well documented in literature, see for instance [41] and [7].

O|G	G_c	G_{nc}
O_c	$W + M_c - M_g,\ \ M_g - W + V_{G_c} - KP$	$W + M_c - M,\ \ M - W - e + V_{G_{nc}}$
O_{nc}	$W - M'_g,\ \ M'_g - W + V_{G_c} - KP$	$W,\ \ -W + V_{G_{nc}}$

Table 2.1: The payoff table of the government and officials game.

- M_c corresponds to the bribe (the direct benefit) that a dishonest official takes from a citizen when his participation in a certain activity is required.

- M_g is the amount that the dishonest official must pay to his partner in the government when a collusion occurs and both officials and government are corrupt. This quantity may be seen as a fraction of the bribe that the corrupt official charges on citizens: $M_g = \theta M_c$. More generally, it can be seen as the appropriation by the government of an official's welfare.

- M'_g is the amount that an honest official must pay to a dishonest government to keep his position or because they do not want to be punished for breaching the rules of coexistence between the two. We note that M'_g can be negative, i.e., a reward given by the corrupt government to honest officials. When M'_g is positive, it can also be seen as a legal appropriation of the officials' welfare due to ideological reasons.

- e is the cost associated with the capture of corrupt officials[4]. This cost is a measure of the governmental efficiency in the fight against corruption.

- V_{G_c} and $V_{G_{nc}}$ are dichotomous random variables taking, respectively, the values V_{G_c} and 0 with probabilities q_{G_c} and $1 - q_{G_c}$, and the values $V_{G_{nc}}$ and 0 with probabilities $q_{G_{nc}}$ and $1 - q_{G_{nc}}$. For simplicity, we keep the same notation for the values the variables may take and the variables themselves. These two random variables correspond, respectively, to the value that a corrupt government and a non-corrupt government assign to being maintained in power for the next period, or as the valuations attributed to re-election. In the case of a corrupt government this valuation might reflect its potential gains from future relations with, for instance, economic elites and lobbying and interest groups, as well as the government's worries about eventual future prosecutions in the event of an electoral defeat (as observed in countries where two parties alternate in power).

- By KP we symbolize the total amount of money that the ruler elite offers to the citizens to buy their votes, where P is the unitary value paid to each person and K is the number of citizens to which the government pays is such that $0 \leq K \leq H$ where H is the total number of citizens. This may be regarded as the effort that a corrupt government makes to increase its chances of re-election. For more about vote buying and its meaning we refer the reader to [89].

As mentioned previously, citizen are not players in the usual sense in our game, but have a decisive influence on the game by means of the probabilities of re-election of a government.

[4]The costs associated with the capture of a corrupt official may be seen as funded by sanctions that a non-corrupt government obtains from fines to corrupt officials. Certainly if this cost exceeds the total amount of fines collected, the government will have to appeal to other sources to perform this task. This point and questions about the nature of the funding associated to the system of prosecution of corrupt are not considered in this model.

Definition 2.1. *We shall symbolize by* $Q_{G_c} = (q_{G_c}, 1 - q_{G_c})$ *and by* $Q_{G_{nc}} = (q_{G_{nc}}, 1 - q_{G_{nc}})$ *the probability distribution that a corrupt government, (respectively a non-corrupt one) be re-elected. These probabilities can be considered as two different mixed strategies of the citizens.*

We will assume in the first part of this chapter that these probabilities are given. The reason we do this is because in the second part of this chapter we will consider a different game where the citizens will be full players of the game, with two different behaviours, one complacent with corruption and other that is not complacent. Hence the mixed strategy of this citizens will be a probability distribution over these two behaviours. This mixed strategy will have an influence on government and officials and will also dynamically evolve, in a way extending this first framework where the probabilities that are akin to the mixed strategies of the citizens are given and the index of intolerance to corruption to be introduced below is constant.

It is of course possible to endogeneize these probabilities, which can be achieved based on heterogeneous citizens' preferences toward the direct corruption they perceive, that can be measured by the relative amount or probability of finding a corrupt officials, and his/her preference towards the payment P. We explain one possible way of doing this in the appendix 2.A. More precisely we can consider that citizen's preferences are defined by a utility function $u_i : [0,1] \times \Re_+ \to \Re_+$ that we assume to be continuously differentiable, such that $(n_c, P) \hookrightarrow u_i(n_c, P)$ where n_c is the relative amount of corrupt officials and P is the payment offered by the government for the vote of the citizen. In addition we consider that $\frac{\partial u_i(n_c,P)}{\partial n_c} < 0$, $\frac{\partial u_i(n_c,P)}{\partial P} > 0$. Note that under the usual assumptions relative to utilities, the marginal rate of substitution of corruption for money $MRS^i_{n_c,P}$ i.e, the amount of money that a citizen hopes to receive to accept a higher level of corruption, without changing his level of utility, is given (locally) by the expression:

$$MRS^i_{n_c,P} = \frac{\partial P}{\partial n_c} = -\frac{\partial u_i / \partial n_c}{\partial u_i / \partial P} > 0 \; . \tag{2.1}$$

Our analysis mostly refers to the evolution of corruption in countries where citizens have the opportunity to express themselves with relative independence in an election. The probabilities of a corrupt and a non-corrupt government be maintained in power that we have described above may be interpreted as probabilities of re-election, and we will frequently use this interpretation in the remainder of this work. In a way, we can say that they reflect a measurement of the popularity of a government. The case of countries under less democratic regimes or even dictatorial regimes can be considered as extremal cases of our model, in which the probabilities of the government being maintained in power are altered accordingly to the (dictatorial) power of the government, since in most circumstances, a dictatorial regime also has elections that are fraudulent (such as, for instance, Portugal before 1974, and the infamous 1958 Presidential election).

The reasoning for the variables V above is that it is clear that members of a corrupt government have an interest in perpetuating themselves in power, either because of their interest in continuing to be enriched or because of their fear of being penalized by a future government. On the other hand, the interest that a non-corrupt government has in being maintained in power is based on the will its members to fulfil a function of public interest. Eventually, in dictatorships, the government has total or near total probability of being maintained in power unless an outside event such as a coup occurs, with the government always achieving its externality values V_{G_c} or $V_{G_{nc}}$. Thus our model may also explain these situations.

The parameters of the model may be considered as an average of what is observed in the whole society. For instance, it may be that an honest government dismisses some corrupt officials instead of imposing fines on them. This situation may be interpreted as a fine that equals the wage of the official. Another possibility, for instance, is when an honest official works for an honest government, and the government may consider that this has a positive benefit apart from the wage he pays and give the official this positive benefit. Thus, the wage parameter will reflect the average of these situations in the whole society. Similarly for the order of magnitude of fines, briberies and punishments from the part of the officials that may be different across society, with the parameter values representing a societal average. In other words the parameter are for a certain focal or representative individual of a society, and the same for the average/typical behaviour of a government. For the sake of simplicity, we consider that the wages are the same whether officials and government are corrupt or not.

This game fits in the literature of cyclic games introduced in [80] as a 3-person cyclic game. Government acts corruptly or not, and plays a game with public officials. Officials act corrupt or not, thus receiving or not bribes from citizens. Citizens exert influence and may punish the government, thereby closing the cycle.

2.1.2 Corruption as a self-reinforcing mechanism

The Von Neumann-Morgenstern utility theorem shows that, under certain axioms of rational behaviour, a decision-maker faced with risky outcomes of different choices will behave as if he is maximizing the expected values of some function (the von Neumann-Morgenstern utility function) defined over the potential outcomes at some specified point in the future (see [102]). We will follow this point of view to describe the behaviour of the agents involved in our model. We assume that the values of the utility function associated with each choice (for the ruling elite and for the officials) are the potential profits in each state of the world.

Let $n_c(t)$ be the share of corrupt officials at time t and $n_{nc}(t)$ the share of honest officials at time t. We are interested in modelling the evolution of these shares with time by means of the replicator dynamics. We will have that $n_c(t) + n_{nc}(t) = 1$ for all time t. The quantities n_c and n_{nc} can be seen as the probabilities of finding, respectively, a corrupt and a non-corrupt official in the population of officials.

Taking into consideration that q_{G_c} and $q_{G_{nc}}$ are respectively, the probabilities that a corrupt and a non-corrupt government get re-elected, we obtain that the expected payoff of a dishonest government corresponds to

$$E(G_c) = n_c M_g + n_{nc} M'_g - W + R_{G_c} - KP \ , \tag{2.2}$$

and the total payoff of a honest government corresponds to

$$E(G_{nc}) = -W + (M - e)n_c + R_{G_{nc}} \ , \tag{2.3}$$

where R_{G_c} and $R_{G_{nc}}$ are the expected values of governments in case of being maintained in power, i.e., $R_{G_c} = V_{G_c} q_{G_c}$ and analogously for a non-corrupt government $R_{G_{nc}} = V_{G_{nc}} q_{G_{nc}}$.

We denote by $P(G_c)$ the probability that the government follows a corrupt policy. We shall see later that in our model this probability will be endogenously determined. Note that $P(G_{nc}) = 1 - P(G_c)$ is the probability that the government follows a non-corrupt policy. The expected profit of a dishonest official is given by

$$E(O_c) = (W + M_c - M_g)P(G_c) + (W + M_c - M)P(G_{nc}) \ . \tag{2.4}$$

The expected profit of an honest official is given by

$$E(O_{nc}) = (W - M'_g)P(G_c) + WP(G_{nc}) \ . \tag{2.5}$$

Assuming that $M - M_g + M'_g > 0$ and $M_g - M'_g - M + e > 0$, after some algebra we obtain the following statements $E(O_c) > E(O_{nc})$ if and only if:

$$P(G_c) > \frac{M - M_c}{M - M_g + M'_g} \ , \tag{2.6}$$

and $E(G_c) > E(G_{nc})$ if and only if

$$n_c > \frac{(R_{G_{nc}} - R_{G_c}) - M'_g + KP}{M_g - M'_g - M + e} \ . \tag{2.7}$$

The next proposition summarizes these facts:

Proposition 2.1. *Officials prefer to choose a dishonest behaviour if and only if the government corruption is large enough, and reciprocally a high number of corrupt officials encourage governmental corruption.*

Remark 2.1. *Note that if the fines are relatively low with respect to what a corrupt official can obtain by an illegal payment for his services, i.e., if $M_c \geq M$, then even when the government always prefers to follow a non-corrupt behaviour ($P(G_c) = 0$), it is more profitable for the officials to follow a corrupt conduct. So, along time, the amount of corrupt officials increases, and consequently, after some time the government becomes corrupt. More precisely, this will change the governmental behaviour, and it will happen as soon as inequality (2.7) is verified.*

If together with the conditions above we have $M - M_c < 0$ then the threshold on the right hand side is negative and the condition is always verified. Analogously, if $(R_{G_{\bar{c}}} - R_{G_c}) - M'_g - KP < 0$. In these cases the two thresholds are always verified and independently of the proportion of corrupt officials and the probabilities that the government acts corruptly then both players will always prefer to act in a corrupt way. On the contrary if we also have $M - M_c > M - M_g + M'_g > 0$ or $(R_{G_{\bar{c}}} - R_{G_c}) - M'_g - KP > M_g - M'_g - M + e > 0$ then the thresholds above are greater than 1. In these cases it is impossible for the two inequalities to be verified independently of the proportion of corrupt officials and the probability that the government acts corruptly. So that both players will prefer to act corruptly.

A general conclusion can be obtained from proposition (2.1) and summarized in the following way: corruption corrupts. More explicitly, this proposition says that corruption is a self-reinforcing mechanism. The question now is how to break down this process. The answer is in the degree of intolerance of citizens to corruption. It should be considered that even when some elements of the ruling class are willing to follow a corrupt behaviour, if this attitude favours an increase in the number of corrupt officials, their government may not be re-elected, and in that case they do not obtain the value of the re-election externality V_{G_c}. This possibility depends on the interplay between the variables of the model. One particular quantity is the index of intolerance of corruption that we now introduce.

2.1.3 The index of intolerance to corruption

Corruption is wilfully hidden and is not easy to measure directly (see [91]). There have been many attempts to solve this problem but they have all came up with limitations, see for instance [18]. In particular, its perception may be distorted which may influence citizen intolerance towards it.

The intolerance of citizens towards corrupt acts plays a fundamental role regarding the evolution of corruption in society. The possible expressions of this intolerance can take different forms under different regimes. In general, in democratic countries, it manifests itself through citizen's vote. In our analysis we will refer to the expression of this intolerance in the form of citizens voting in an electoral process in democratic countries.

Under this framework, now we define the index of intolerance to corruption as follows:

Definition 2.2. (The index of intolerance to corruption) *Let $q_{G_{nc}}$ be the probability that a corrupt government is re-elected given that the percentage of corrupt officials is n_c and let q_{G_c} be the probability that a non-corrupt government is re-elected. We define the index of intolerance to corruption by the difference:*

$$D_{it} = q_{G_{nc}} - q_{G_c} \ . \tag{2.8}$$

This index captures the social sensibility to the corruption. Note that $-1 \leq D_{it} \leq 1$ and that

$$\begin{aligned} R_{G_{nc}} - R_{G_c} &= V_{G_{nc}} q_{G_{nc}} - V_{G_c} q_{G_c} \\ &= \left(V_{G_{nc}} - V_{G_c}\right) q_{G_{nc}} + V_{G_c} \left(q_{G_{nc}} - q_{G_c}\right) \\ &= \left(V_{G_{nc}} - V_{G_c}\right) q_{G_{nc}} + V_{G_c} D_{it} \ . \end{aligned} \tag{2.9}$$

This expression shows that the difference between the expected value of a non-corrupt government being re-elected and the one that corresponds to a non-corrupt government being re-elected grows with the index of intolerance to corruption.

Substituting (2.9) in inequality (2.7) it follows that government prefers the corrupt strategy if and only if

$$n_c > \frac{\left[\left(V_{G_{nc}} - V_{G_c}\right) q_{G_{nc}} + V_{G_c} D_{it}\right] - M'_g + KP}{M_g - M'_g - M + e}. \tag{2.10}$$

If we consider the additional hypothesis that the valuations of a political group in power can obtain in case of being maintained in power are the same whether it is corrupt or not, i.e. $V_G = V_{G_c} = V_{G_{nc}}$, then equation (2.10) simplifies and the role of the index of intolerance is very clear:

$$n_c > \frac{V_G D_{it} - M'_g + KP}{M_g - M'_g - M + e}. \tag{2.11}$$

In the same conditions as proposition 2.1 the next corollary holds.

Corollary 2.1. *If citizens are sufficiently intolerant with the bad services provided by corrupt officials, then, according to (2.10) or (2.11), it becomes more unlikely that there are enough corrupt officials so that governments prefer to be corrupt, so that the government loses incentives to tolerate or to allow corruption. Insofar as that the degree of tolerance of citizens for the services of corrupt officials decreases or, equivalently, insofar the degree of intolerance for corrupt services increases, the government prefers to punish corrupt officials.*

However, note that the strategy "to be corrupt" can be a dominant strategy for the government if its efficiency to capture corrupt officials is low, or equivalently the cost to catch the corrupt officials is high, i.e., if $e > M - M_g$. The cost to catch the corrupt officials is higher in those countries where the effectiveness of the justice system is low, and in this case, and also when the intolerance index is low, we may be in presence of a negative cycle where an inefficient legal system becomes a cause and a consequence of corruption. We will analyse how exogenous changes in these and other quantities can change the processes of evolution of corruption and revert the spreading of corruption.

If we consider the probabilities of re-election as depending on the number of corrupt officials and the payments, then the index is strongly related with the dis-utility that corruption provides to citizens since the probability q_{G_c} that a corrupt government is re-elected decreases with the share of corrupt officials n_c. If the dis-utility of an increase in corruption for the $i-th$ citizen is very high, i.e., the numerator in (2.1) is very low then the corrupt government must pay a very high price to buy his vote. This could have as a result the impossibility of buying the vote of such citizen, decreasing in this way the probability that the citizen votes for the corrupt government. Analogously for those citizens for whom the marginal utility of money is lower, the denominator in (2.1) is lower, yielding a higher marginal rate of substitution. These citizens are also less likely to vote for a corrupt government. The probability of a non-corrupt government being re-elected should also decrease or do not depend on the number of corrupt officials. In this case we would have that $\frac{\partial}{\partial n_c}(q_{G_{nc}} - q_{G_c}) > 0$, i.e., the derivative of the intolerance index with respect to the percentage of corrupt officials is positive. This means that because of the payments that a corrupt government makes, citizens have an indirect perception of the government corruption, resulting in the probability q_{G_c} decreasing more rapidly than $q_{G_{nc}}$., which is obviously true if $q_{G_{nc}}$ is constant.

2.1.4 The evolution of corruption

To explain the social evolution of corruption we shall follow an evolutionary approach. This approach is based on the fact that strategies that make an agent of the game do better than others will be retained, while strategies that lead to lower payoffs will be abandoned. The success of a strategy is measured by its relative frequency in the population at any given time. Strategies change over time as a function of their relative success in an environment that is made up of other players that keep changing their own strategies adaptively in the same fashion.

Initially, players decide their strategies independently. We assume that players have myopic behaviour, because officials and government can not forecast the consequences that changes in the relative frequency of their strategies can provoke. As such, players analyse the payoffs of his/her strategies in each moment without trying to look at a further horizon, thus the reason why we say the players act myopically. So, in each period, the percentage of individuals that follows a given strategy increases if the expected payoff of such strategy is greater than the average payoff obtained by the population. Otherwise, if the expected payoff is performing worse than average, that strategy becomes less frequent in the population. The dynamical system summarizing these facts is the replicator dynamics (see [103]). In other words, the replicator dynamics considers that the difference between the expected payoff of a strategy and the average payoff of all strategies is the per-capita change in the frequency of that strategy in the population.

As before, we denote by $n_i(t)$ the percentage of officials following strategy $i \in \{O_c, O_{nc}\}$. By $n(t) = (n_c(t), n_{nc}(t))$ we symbolize the distribution of the officials over the set of pure strategies, and by $g(t) =$

$(g_c(t), g_{nc}(t))$ the mixed strategy of the government.

The replicator dynamics is given by the following differential equation:

$$\dot{n}_c = n_c[E(O_c|g) - \bar{E}] \ ,$$

where $\bar{E} = n_c E(O_c|g) + n_{nc} E(O_{nc}|g)$ is the average payoff of the officials given a distribution of officials and $E(O_c|g)$ and $E(O_{nc}|g)$ denote, respectively, the expected value of corrupt and non-corrupt behaviour for an official, given a distribution g over the government behaviour. Since $n_{nc}(t) = 1 - n_c(t)$, the evolution of the percentage of non-corrupt officials is determined by the previous equation, i.e.

$$\dot{n}_{nc} = -\dot{n}_c \ .$$

So for the officials, we need only to consider the equation for the percentage of corrupt officials n_c. Since there are only two pure strategies, after some algebra we obtain the equivalent differential equation:

$$\dot{n}_c = n_c(1 - n_c)(E(O_c|g) - E(O_{nc}|g)) \ . \tag{2.12}$$

To measure the evolution of the governmental corruption we introduce g_c as an index measuring the percentage of corrupt acts committed in public offices regarding the total of acts performed in these government agencies[5] . We endogenize the probability of a government being corrupt by considering the index g_c as the probability that a government follows a corrupt strategy. In other words, this will be the mixed strategy of the government over his set of pure strategies. In a similar way, the evolution of the government policy can be represented by the following differential equation:

$$\dot{g}_c = g_c(1 - g_c)(E(G_c|n, D_{it}) - E(G_{nc}|n, D_{it})) \ . \tag{2.13}$$

where $E(G_c|n, D_{it})$ and $E(G_{nc}|n, D_{it})$ represent, respectively, the expected value of a corrupt behaviour and a non-corrupt behaviour by the government, given a distribution n of the officials over their available strategies and the degree of intolerance to corruption D_{it}. Similarly, as in the case of the officials we have

$$\dot{g}_{nc} = -\dot{g}_c \ ,$$

so that for the government we need only consider the equation for the probability of acting corruption and the probability of acting non-corruptly is determined by the first probability.

Using equalities (2.2), (2.3), (2.4) and (2.5), we obtain that (2.12) and (2.13) may be written.

$$\begin{aligned} \dot{n}_c &= n_c(1 - n_c)\left[(M_c - M_g - M_c + M + M_g')g_c + M_c - M\right] \\ \dot{g}_c &= g_c(1 - g_c)\left[n_c\left(M_g - M_g' - M + e\right) + M_g' + R_{G_c} - R_{G_{nc}} - KP\right] \end{aligned} \tag{2.14}$$

[5]Most indexes measuring corruption actually measure proxies for corruption because corruption is a difficult phenomenon to measure. An example of such an empirical index of the perceived governmental corruption is *Transparent International's* (TI) Corruption Perceptions Index (CPI). This index captures information about administrative and political aspects of corruption. However, its use has not come without criticism (see [18] and [97]). Another example is TI's *Global Corruption Barometer* which asks population directly about the payment of bribes and TI's *Bribe Payers Index* which asks the business sector about willingness to the payment of bribes for the countries operating with the country of their enterprises.

To simplify the notation we write

$$\begin{aligned} A &= -M_g + M + M'_g, & B &= M_c - M \\ A' &= M_g - M'_g - M + e, & B' &= M'_g + R_{G_c} - R_{G_{nc}} - KP \end{aligned} \tag{2.15}$$

The dynamical system (2.14) then takes the form:

$$\begin{aligned} \dot{n}_c &= n_c(1-n_c)\,(Ag_c + B) \\ \dot{g}_c &= g_c(1-g_c)\,(A'n_c + B')\ . \end{aligned} \tag{2.16}$$

2.1.5 Dynamical equilibria, stability and Nash equilibria

In this section we analyse the equilibria of the dynamical system given by equations (2.16) and analyse their stability. In a neighbourhood of an equilibrium point we consider the values of the parameters of the model and the value of the intolerance index as given. In a small neighbourhood of the equilibrium they can be considered as constant, not affecting the stability analysis, which is a local property, i.e., it only concerns a neighbourhood of each equilibrium point. Note that in dynamical system 2.14 the index of intolerance affects only the differential equation of the government. It is possible to consider that the ruler elite doesn't have exact knowledge of this index and considers it as an average value on that neighbourhood. With this study, we characterize the long-term evolution of corruption.

Government and officials characteristics

To analyse the stability of the equilibria, we will use the following characterizations of corrupt and non-corrupt governments.

(i) The non-corrupt government practices appropriate fines if $M > M_c$, $(B < 0)$, and it practices inadequate fines if $M < M_c$, $(B > 0)$.

(ii) Let the re-election power threshold be

$$T_1 = M'_g - KP\ .$$

The non-corrupt government has high re-election power if $R_{G_{nc}} - R_{G_c} = V_{G_{nc}}q_{G_{nc}} - V_{G_c}q_{G_c} > T_1$, $(B' < 0)$, and it has low re-election power if $R_{G_{nc}} - R_{G_c} = V_{G_{nc}}q_{G_{nc}} - V_{G_c}q_{G_c} < T_1, (B' > 0)$. It may also be interpreted as saying that the corrupt government has low (respectively high) re-election power.

(iii) Let the non-corrupt government efficiency threshold be

$$T_2 = R_{G_{nc}} - R_{G_c} + M - M_g + KP\ .$$

The non-corrupt government is cost efficient in fighting against corruption if $e < T_2, (A' + B' < 0)$ and it is cost inefficient in fighting against corruption if $e > T_2, (A' + B' > 0)$.

(iv) The corrupt government penalizes honest officials more than dishonest officials if $M'_g > M_g - M_c, (A + B > 0)$, and penalizes dishonest officials more than honest officials if $M'_g < M_g - M_c, (A + B < 0)$. It may also be interpreted as saying that the corrupt government takes more welfare from non-corrupt officials than from corrupt officials.

Remark 2.2. *If the degree of intolerance to corruption is high enough then the non-corrupt government has high re-election power. In other words, when the expected value R_{G_c} of the valuation associated with a corrupt government being maintained in power increases, then it is less likely that a non-corrupt government has high re-election power. We see that the re-election power threshold decreases when the total amount KP paid by the corrupt government to buy votes increases. Thus it is more likely that a non-corrupt government has high re-election power when the amount of the payment increases and/or the number of people that are paid increases. We may see this as a positive popularity effect on a non-corrupt government, yielding a higher re-election power. Similarly, the threshold increases when the value M'_g of the punishment that a corrupt government puts on honest officials increases. So it gets more likely that a corrupt government has high re-election power since the welfare of a corrupt government has increased due to an increase in M'_g.*

Remark 2.3. *If we interpret M_g as a fraction $\theta < 1$ of M_c, then $A + B$ is positive. However, it could happen that this quantity is negative if a corrupt government is charging corrupt officials (M_g) more than the bribe that the corrupt officials receive (M_c), i.e., $\theta > 1$, meaning that the government takes the full bribe, plus some amount from the welfare of the official, or equivalently, that officials are paying government a portion of their salaries to keep their jobs. Another possible interpretation is that the amount M'_g is negative meaning that an honest official receives money from a dishonest government to keep his position. This kind of corruption of the ruling elites can be considered like 'legal' corruption associated to self-imposed laws charging severe fines to officials that are corrupt by breaking these laws and so breaking the (ideological) solidarity with the government. We shall see that this condition may lead to the equilibrium corresponding to a corrupt government with honest officials being stable.*

Dynamical equilibria and Nash equilibria

The dynamical system (2.16) has the following four dynamic equilibria (or steady-states) corresponding to pure strategies of the game.

1. The corruption ('bad') equilibrium $(n_c^1, g_c^1) = (1, 1)$. This equilibrium may be interpreted a country where the law is not respected and with general high levels of corrupt both at the officials' and at government's level. This equilibrium is a Nash equilibrium if and only if

$$E(O_c|g_c = 1) \geq E(O_{nc}|g_c = 1) \text{ and } E(G_c|n_c = 1) \geq E(G_{nc}|n_c = 1) \ .$$

   Equivalently

$$A + B \geq 0 \text{ and } A' + B' \geq 0 \ .$$

2. The corrupt officials equilibrium $(n_c^2, g_c^2) = (1, 0)$. The interpretation of this equilibrium is the case of a *de facto* government of the officials. It corresponds to the case of a weak *de jure* government that is unable to control the corruption of the officials and that it is maintained in power by them

only to masque corruption. This equilibrium is a Nash equilibrium if and only if

$$E(O_c|g_c = 0) \geq E(O_{nc}|g_c = 0) \text{ and } E(G_c|n_c = 1) \leq E(G_{nc}|n_c = 1) \ .$$

Equivalently

$$B \geq 0 \text{ and } A' + B' \leq 0 \ .$$

3. The corrupt government equilibrium $(n_c^3, g_c^3) = (0, 1)$. The interpretation for this equilibrium is the case of a strong ideological dictatorship that in fact does not allow (illegal) corruption. In other words, where government corruption (that is, corruption by a usually small political elite) may be seen as legal form of corruption. However, the government imposes its power by force with high penalties for officials that deviate from honest behaviour. This equilibrium is a Nash equilibrium if and only if

$$E(O_c|g_c = 1) \leq E(O_{nc}|g_c = 1) \text{ and } E(G_c|n_c = 0) \geq E(G_{nc}|n_c = 0) \ .$$

Equivalently

$$A + B \leq 0 \text{ and } B' \geq 0 \ .$$

4. The non-corruption ('good') equilibrium $(n_c^1, g_c^1) = (0, 0)$. It corresponds to the case in which government and officials fulfil their functions properly and general levels of corruption are low. This equilibrium is a Nash equilibrium if and only if

$$E(O_c|g_c = 0) \leq E(O_{nc}|g_c = 0) \text{ and } E(G_c|n_c = 0) \leq E(G_{nc}|n_c = 0) \ .$$

Equivalently

$$B \leq 0 \text{ and } B' \leq 0 \ .$$

The corrupt officials equilibrium $(n_c^2, g_c^2) = (1, 0)$ and the corrupt government equilibrium $(n_c^3, g_c^3) = (0, 1)$ will be designated in some situations as *asymmetric* pure equilibria, since both players choose opposite pure strategies.

The dynamical system (2.16) also has another dynamical equilibrium that may correspond to a mixed Nash equilibrium of the game. If A and A' are not equal to zero, then the point (n_c^T, g_c^T) is a steady state, where

$$\bar{n}_c^T = -\frac{B'}{A'} = \frac{(R_{G_{nc}} - R_{G_c}) - M_g' + KP}{M_g - M_g' - M + e} \text{ and } \bar{g}_c^T = -\frac{B}{A} = \frac{M - M_c}{M - M_g + M_g'} \ .$$

In our framework, this equilibrium makes sense if $0 < \bar{n}_c^T < 1$ and $0 < \bar{g}_c^T < 1$ are satisfied. Then the steady state $(\bar{n}_c^T, \bar{g}_c^T)$ is also a mixed Nash equilibria for the game. We note that if $\bar{n}_c^T$ is equal to 0 or 1, then $(\bar{n}_c^T, g_c)$ is a steady state for any g_c; and if $\bar{g}_c^T$ is equal to 0 or 1, then $(n_c, \bar{g}_c^T)$ is a steady state for any n_c.

Stability of dynamical equilibria

In this subsection we see how the stability of the steady-states mentioned above depends on the thresholds characterizing government and officials.

The Hartman-Grobman theorem states that the orbit structure of a dynamical system in a neighbourhood of a hyperbolic equilibrium point is topologically equivalent to the orbit structure of the linearised dynamical system. (See [46] for more on the Hartman-Grobman theorem and stability).

Assuming that A and A' are non-zero, then the point $(n_c^T, g_c^T) = (-B'/A',\ -B/A)$ is a steady state for the dynamical system. The linearisation at this point is given by the matrix:

$$J\left(-\frac{B'}{A'}, -\frac{B}{A}\right) = \begin{bmatrix} 0 & -\frac{A'B}{A^2}(B+A) \\ -\frac{AB'}{A'^2}(B'+A') & 0 \end{bmatrix}.$$

The eigenvalues of this matrix are:

$$\lambda = \pm\sqrt{\frac{B'B}{AA'}(B'+A')(B+A)}\ .$$

Thus, if $\frac{B'B}{AA'}(B'+A')(B+A) > 0$ then this point is a saddle point for the dynamics. In other cases the Hartman-Grobman theorem does not apply since the matrix above has eigenvalues with zero real part, meaning that the point is not hyperbolic.

For the four pure equilibria we will see that the characterization of its asymptotic stability by means of the linearisation is the same as the characterization of it being a strict Nash equilibria, i.e., its strategies constitute strict best responses against each other, or in other words, the inequalities characterizing pure Nash equilibria presented in section 2.1.5 are strict.

For the corruption equilibrium $(n_c^B, g_c^B) = (1,1)$ the linearisation is

$$J(1,1) = \begin{bmatrix} -(A+B) & 0 \\ 0 & -(A'+B') \end{bmatrix}.$$

The eigenvalues are $\lambda_1 = -(A+B)$ and $\lambda_2 = -(A'+B')$. Hence, the corruption equilibrium $(n_c^B, g_c^B) = (1,1)$ is stable if the non-corrupt government is cost inefficient in fighting against corruption and the corrupt government penalizes honest officials more than dishonest officials. This is the same as saying that to be corrupt is a strict best response against corrupt behaviour, both for the government and the officials: i) when government is corrupt, officials are also corrupt because of the higher penalties for corrupt officials; ii) when officials are corrupt, government is also corrupt due to its inefficiency.

For the corrupt officials equilibrium $(n_c^2, g_c^2) = (1,0)$ the linearisation is

$$J(1,0) = \begin{bmatrix} -B & A' \\ 0 & A'+B' \end{bmatrix}.$$

The eigenvalues are $\lambda_1 = -B$ and $\lambda_2 = (A'+B')$. Hence, the equilibrium $(n_c^2, g_c^2) = (1,0)$ is stable if the non-corrupt government is cost efficient in fighting against corruption but practices inadequate fines. In terms of best responses this means that: i) when government is honest, officials are corrupt because of inadequate fines; ii) when officials are corrupt government is honest since it is cost efficient and so it is not costly to enhance anti-corruption mechanisms.

For the corrupt government equilibrium $(n_c^3, g_c^3) = (0, 1)$ the linearisation is

$$J(0,1) = \begin{bmatrix} A+B & A' \\ 0 & -B' \end{bmatrix}.$$

The eigenvalues are $\lambda_1 = (A+B)$ and $\lambda_2 = -B'$. Hence, the corrupt government equilibrium $(n_c^3, g_c^3) = (0, 1)$ is stable if the corrupt government penalizes dishonest officials more than honest officials and the non-corrupt government has low re-election power. In terms of best responses this means that: i) when government is corrupt officials are honest because of lower penalties; ii) when officals are honest government is corrupt since acting non-corruptly has low re-election power.

For the non-corruption equilibrium $(n_c^4, g_c^4) = (0, 0)$ the linearisation is

$$J(0,0) = \begin{bmatrix} B & 0 \\ 0 & B' \end{bmatrix}.$$

The eigenvalues are $\lambda_1 = B$ and $\lambda_2 = B'$. Hence, the non-corruption equilibrium $(n_c^4, g_c^4) = (0, 0)$ is stable if the non-corrupt government practices appropriate fines and has high re-election power. In terms of best responses this means that: i) when government is honest officials are honest because of the fines they have for corrupt behaviour; ii) when officials are honest government is also honest due to the high re-election power of acting non-corruptly.

As we hinted before we note that a given pure dynamical equilibria of the above is asymptotically stable if and only if it is a strict Nash equilibrium for the game. Recall that a strict Nash equilibrium is a Nash equilibrium such that it is strict best response in the sense that every other strategy yields a lower payoff that the choice in the specified mixed Nash equilibrium. This means that a pure Nash equilibrium is strict if its component strategies are strict best responses against each other. Furthermore, since a mixed Nash equilibrium is obtained when the player is indifferent towards two or more pure strategies, i.e., they yield the same payoff, we have that a mixed Nash equilibrium is never a strict Nash equilibrium. Furthermore, the mixed equilibrium is never asymptotically stable as we have seen above. So we conclude that a dynamical equilibrium is asymptotically stable for our dynamics if and only if it is a strict Nash equilibrium.

We observe that the corruption and non-corruption equilibria can be simultaneously stable, as well as the other two asymmetric pure equilibria. We observe that when either the non-corruption equilibrium or the corruption equilibrium are stable, then the two asymmetric equilibria can not be stable, and *vice-versa*, if one of the asymmetric equilibria are stable then the non-corruption and corruption equilibria can not be stable. In other words, a stable pure equilibrium immediately destabilizes the two adjacent pure equilibria. Hence, we conclude that only three situations are possible: there is no stable pure equilibrium; there is exactly one stable pure equilibrium; there are exactly two stable pure equilibria.

When only one of the four pure equilibria is stable, then it is globally asymptotically stable, meaning that all trajectories inside the unit cube tend towards that equilibrium. In this case there is no 'mixed' equilibrium, or more correctly, it lies outside the unit square.

When there are exactly two stable pure equilibria, they must not be adjacent in the unit square, so they must be opposite vertexes of the square. This situation is when the corruption and non-corruption equilibria are both stable, or when the two asymmetric equilibria are stable. When the corruption and non-corruption equilibria are both stable we have that $-(A+B) < 0$, $-(A'+B') < 0$, $B < 0$ and $B' < 0$. It is easy to see that these inequalities imply that $0 < \bar{n}_c^T < 1$ and $0 < \bar{g}_c^T < 1$. Similarly, when the two asymmetric equilibria are stable then the mixed equilibrium lies inside the unit cube. So when here are exactly two stable pure equilibria, the mixed equilibrium lies inside the unit cube.

Hence the existence of two stable pure equilibria forces the existence of the mixed Nash equilibrium. In this case it is possible to show that dynamically the mixed Nash equilibrium must be a saddle whose stable manifold is the separatrix between the basins of attraction of the two stable pure equilibria. The picture in this case is like figure 2.1 in the case where the two stable pure equilibria are the corruption and non-corruption equilibria. For the case where the two stable pure equilibria are the two asymmetric equilibria the figure is like figure 2.1 but with the arrows reverse and this time the stable manifold of the mixed equilibrium connects the two asymmetric equilibria. We will see this with detail in the following section in some examples. The other situation where a mixed equilibrium exists is when none of the four pure equilibria are stable. In this case it is expected that, because there is no stable pure equilibria, some oscillatory behaviour might occur. We will see in what follows that this is indeed the case, with the appearance of periodic orbits.

Non-corrupt government with high re-election power and appropriate fines

In this subsection, we assume that $B < 0$ and $B' < 0$. Hence the non-corrupt government has high re-election power and uses appropriate fines. In this case the 'good' equilibrium $(n_c^4, g_c^4) = (0, 0)$ is always asymptotically stable.

(1) When $A > -B$ and $A' > -B'$ then $A + B > 0$ and $A' + B' > 0$. From these conditions the following inequalities are verified: $A > 0$, $A' > 0$, $0 < -\frac{B}{A} < 1$, and $0 < -\frac{B'}{A'} < 1$, implying the existence of a mixed equilibrium in the interior of the unit square. In this case we also have that $\frac{B'B}{AA'}(B' + A')(B + A) > 0$, so the Hartman-Grobman theorem can be applied to the mixed equilibrium, yielding a saddle point. In this case the 'bad' equilibrium and the 'good' equilibrium without corruption are asymptotically stable. See figure (2.1) for the general picture of the dynamics in this case.

Note that this case corresponds to a social situation where:

1. The amount M of the fine imposed by a non-corrupt government to a corrupt official is relatively high, meaning that it is greater than the bribe M_c the official takes from citizens, i.e., $M > M_c$.
2. The inequality $M'_g > M_g - M_c$ is verified. Recall that M'_g is the amount that an honest official must pay to a dishonest government to keep his place. This means that a corrupt government punishes honest behaviour more than dishonest behaviour.
3. The government is inefficient to catch corrupts officials, or equivalently, e is relatively high (relatively high costs to combat corruption).
4. The non-corrupt government has high re-election power. This may be written as $V_{G_c} D_{it} > M'_g - KP - (V_{G_{nc}} - V_{G_c}) q_{G_{nc}}$. This occurs if the index of intolerance is high enough, and the corrupt government has a high valuation for being re-elected.

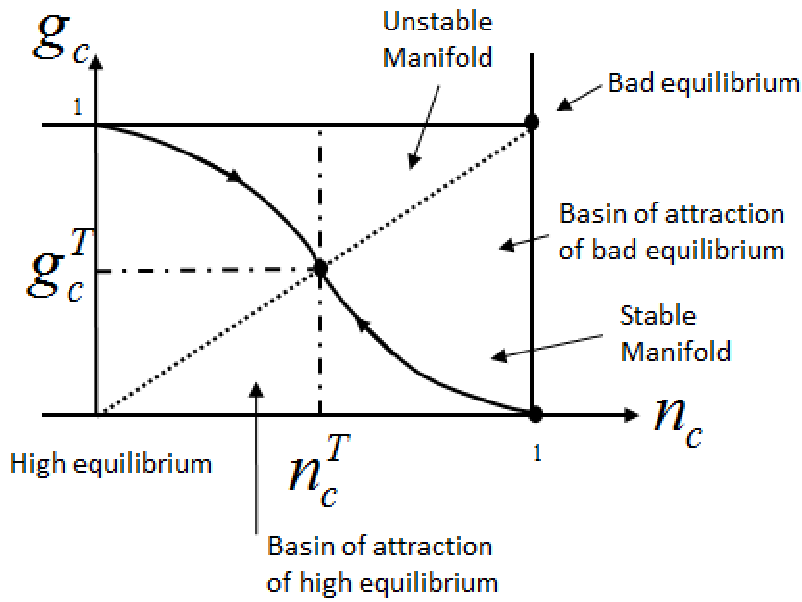

Figure 2.1: The dynamics of the system, with the basin of attraction of the two stable pure equilibria and the mixed Nash equilibrium in case (1).

This is a good example of ongoing spontaneous coordination because the non-corrupt equilibrium is stable. However, if the initial distributions of officials and government's actions correspond to a point in the basin of attraction of the 'bad' equilibrium, then officials and government have incentives to act in a corrupt way. Thus, the general levels of corruption will increase, and corruption becomes a self-enforcing mechanism over time. In this case, as in the previously described case of corruption as a self-reinforcing mechanism, we can regarded it as a social trap.

However, the basin of attraction of the 'bad' equilibrium $(n_c^1, g_c^1) = (1, 1)$ decreases when the interior equilibria gets closer to the 'bad' equilibrium (n_c^1, g_c^1), i.e., when $A + B$ and $A' + B'$ tend to zero. Hence, the basin of attraction of the 'good' equilibrium $(n_c^4, g_c^4) = (0, 0)$ is larger when the non-corrupt government's costs in fighting corruption are close to the non-corrupt government efficiency threshold, i.e. e is closer to T_2, and the corrupt government penalizes honest officials similarly to dishonest officials, i.e. M_g' is closer to $M_g - M_c$.

The cost efficiency e can get closer to the threshold T_2 because of different reasons: (a) non-corrupt government is able to decrease the value of costs to capture corrupt officials, which can occur for instance if the justice system becomes more reliable; (b) the efficiency threshold T_2 rises due for instance to an increase in the index of intolerance, an increase in the valuation of re-election by a corrupt government, an increase in the probability of a non-corrupt government being re-elected, to an increase of the fine imposed by a non-corrupt government to a dishonest official, or an increase in the vote buying by part of the corrupt government.

Hence, the levels of corruption that were increasing can change if the degree of intolerance of citizens increases. If the government believes that this change in intolerance can take place then (depending also on the value that the government assigns to be re-elected), it may result in a change in the basin of attractions of the 'good' and 'bad' equilibria, making some paths that would initially evolve towards the 'bad' equilibrium now evolve towards the 'good' equilibrium. This possibility is summarized in the following fact:

Remark 2.4. *The basin of attraction of the 'bad' equilibrium* $(n_c^B, g_c^B) = (1,1)$ *decreases when the degree of intolerance increases and/or the cost to capture the corrupt officials decreases.*

Thus, the index of intolerance of citizens with respect to corruption, if high enough, and if the government is interested in being re-elected can play an important role at the time to control the controller acting as a servomechanism correcting the evolution of corruption. It acts as a barrier stopping corruption, since, under several circumstances, it can reverse a process of growing corruption. The higher it is, the more difficult it gets that corruptions grows and develops within the government. In figure (2.2) we plot some trajectories of the system that exemplify the previous remark. For the same initial conditions with different model parameters, corresponding to an increase in the degree of intolerance, we see that initial conditions originally in the basin of attraction of the 'bad' equilibrium are instead converging to the 'good' equilibrium. This illustrates the shrinking of the basin of attraction of the corruption equilibrium as the degree of intolerance grows.

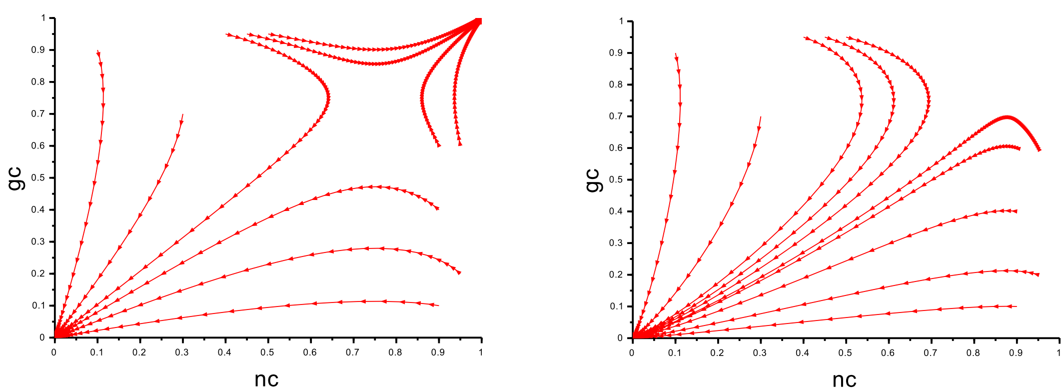

Figure 2.2: Some trajectories of the system for the same initial conditions with different parameters. Left-hand side: lower degree of intolerance. Right-hand side: higher degree of intolerance.

(2) Assuming that $A > -B$ and $A' < -B'$ it follows that $(A+B) > 0$, $(A'+B') < 0$ then there is no mixed Nash equilibrium because either $-\frac{B'}{A'} > 1$ or $-\frac{B'}{A'} < 0$. The 'bad' equilibrium $(n_c^1, g_c^1) = (1,1)$ is a saddle point, as well as the equilibrium $(n_c^2, g_c^2) = (1,0)$, and the equilibrium $(n_c^3, g_c^3) = (0,1)$ is a repulsor. In this case there is a unique asymptotically stable dynamic equilibrium and this is the Nash equilibrium without corruption, i.e., $(n_c^4, g_c^4) = (0,0)$, with all the interior initial conditions being attracted to this point. See figure (2.3). This conditions correspond to a well ruled society.

(3) Assuming that $A < -B$ and $A' < -B'$ it follows that $(A+B) < 0$, $(A'+B') < 0$ then the 'bad' equilibrium is a repulsor, there is no mixed equilibrium, and there is a unique equilibrium that is asymptotically stable, that is the 'good' equilibrium $(n_c^4, g_c^4) = (0,0)$, with all interior initial conditions being attracted to this point. The equilibria $(n_c^2, g_c^2) = (1,0)$ and $(n_c^3, g_c^3) = (0,1)$ are saddle points.

(4) Assuming that $A < -B$ and $A' > -B'$ it follows that $(A+B) < 0$, $(A'+B') > 0$. Then, there is no mixed equilibrium and the 'bad' equilibrium $(n_c^1, g_c^1) = (1,1)$ is a saddle point as well as

the equilibrium $(n_c^3, g_c^3) = (0,1)$, and the equilibrium $(n_c^2, g_c^2) = (1,0)$ is a repulsor. The only asymptotically stable equilibrium is the 'good' equilibrium $(n_c^4; g_c^4) = (0,0)$, with all interior initial conditions being attracted to this point. This corresponds to the case of a well ruled society.

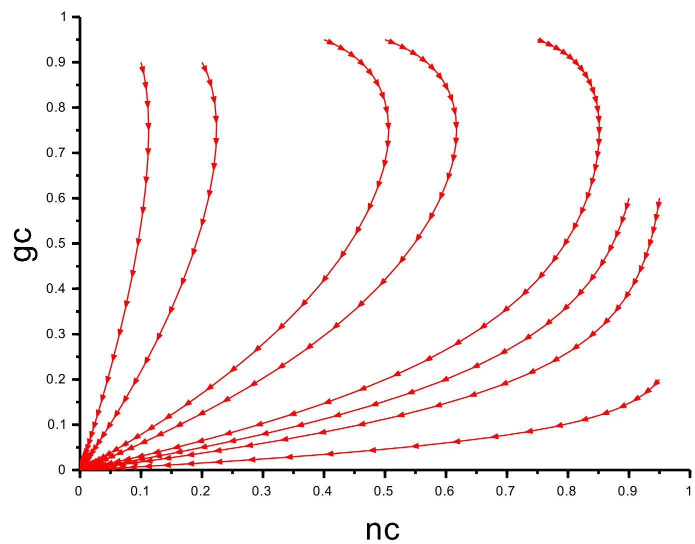

Figure 2.3: Some trajectories of the system for case (2): a well ruled society.

In cases (2), (3) and (4) the outcome of the dynamics is similar since there is no mixed equilibrium and the 'good' equilibrium is globally asymptotically stable. However, in cases (3) and (4) the quantity $A+B = M_g' + (M_c - M_g)$ is negative, that is $M_g' < M_g - M_c$, i.e., the penalties are higher for dishonest officials (see remark (2.3)).

In case (2) the assumptions are describing a socio-political situation corresponding to: (a) a corrupt government that penalizes honest officials more than dishonest officials; (b) an index of intolerance relatively high; and/or (c) a governmental elite with a high interest in being re-elected; and/or (d) the non-corrupt government is highly efficient in fighting corruption, i.e., low values of e. In this case the punishments are higher for honest officials than for corrupt officials, but then government efficiency plays an important role since government's costs are low so that it is better for the government to fight corruption at the officials' level rather than complying with them.

In case (3) we have that in addition to appropriate fines and high re-election power for the non-corrupt government the non-corrupt government has cost efficiency and higher punishments for corrupt officials than for honest officials. We may refer to this case as the ideal parameters case. As one should expect, the 'good' equilibrium is globally asymptotically stable.

In case (4) we have that the non-corrupt government is cost inefficient in fighting corruption, but since the punishments by a corrupt government for corrupt officials are high when compared to honest officials, so there will be less corrupt officials and the government inefficiency plays a lesser role, so that the 'good' equilibrium is still globally stable.

Non-corrupt government with low re-election power and inappropriate fines

In this subsection, we assume that $B > 0$ and $B' > 0$. Hence the non-corrupt government has low re-election power and uses inappropriate fines. We have the following four cases.

(5) Assuming that $A < -B, A' < B'$ then the inequalities $(A + B) < 0$, $(A' + B') < 0$ hold. In this situation both the 'good' non-corruption equilibrium $(n_c^4, g_c^4) = (0, 0)$ and the 'bad' corrupt equilibrium $(n_c^1, g_c^1) = (1, 1)$ are repulsors and the mixed Nash equilibrium is a saddle point. The equilibria $(n_c^2, g_c^2) = (1, 0)$ and $(n_c^3, g_c^3) = (0, 1)$, i.e., the equilibria where government prefers to be honest but officials prefer to be corrupt, and reciprocally, where government prefers to be corrupt but officials prefer to be honest, are local attractors. Which one of these two situations occurs is initial condition dependent. In figure (2.4) we plot some transition paths of the system. Depending on the initial condition, the transition path approaches either $(n_c^2, g_c^2) = (1, 0)$ or $(n_c^3, g_c^3) = (0, 1)$. The exception is one initial condition that approaches the mixed equilibrium, since that initial condition lies on the stable manifold of the mixed equilibrium. We observe that the stable manifold of the mixed equilibrium is a curve passing through the mixed equilibrium that connects the 'bad' and the 'good' equilibrium.

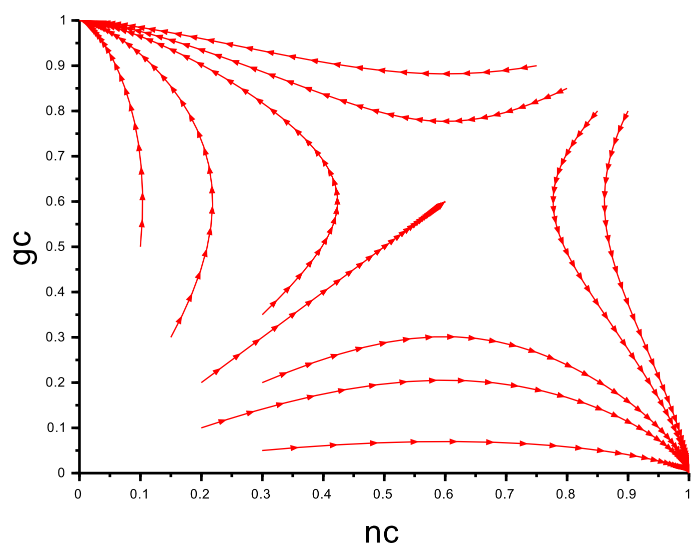

Figure 2.4: Some trajectories of the system for case (5).

(6) Assuming that $A > -B, A' < -B'$ then $A + B > 0$ and $A' + B' < 0$. There is no mixed Nash equilibrium and the equilibrium $(n_c^2, g_c^2) = (1, 0)$ is the only equilibrium point that is asymptotically stable, and all initial conditions in the interior of the unit square are attracted to this equilibrium. We plot some trajectories of the system for this case in figure (2.5).

(7) Assuming that $A < -B,\ A' > -B'$ then $A + B < 0$ and $A' + B' > 0$. There is no mixed Nash equilibrium and the equilibrium $(n_c^3, g_c^3) = (0, 1)$ is the only equilibrium point that is asymptotically stable, and all initial conditions in the interior of the unit square are attracted to this equilibrium.

(8) Assuming that $A > -B$, $A' > -B'$ then $A + B > 0$ and $A' + B' > 0$. In this case the 'good' equilibrium $(n_c^4, g_c^4) = (0, 0)$ is a repulsor and the corner equilibria $(n_c^3, g_c^3) = (0, 1)$ and $(n_c^2, g_c^2) = (1, 0)$ are saddle points. The only equilibrium point that is asymptotically stable is the 'bad' equilibrium $(n_c^1, g_c^1) = (1, 1)$, with all interior initial conditions being attracted to this point. We plot some trajectories of the system in figure (2.6).

In case (6) society is evolving to an equilibrium where officials prefer to be corrupt, even with an honest government. As we argued above this can be seen as a *de facto* government of the officials. Our assumptions imply that governmental fines to punish corrupt behaviour are relatively low, and that the government has low re-election power, because citizens perceive this government as a corrupt one. However, the government is unable to diminish officials' corruption because of the government: (a) being focused on re-election; (b) being inefficient; or (c) practising low fines on corrupt officials.

In order to have long term evolution to the equilibrium where officials are corrupt but government is honest, i.e., to have asymptotic stability of the corrupt officials equilibrium $(n_c^2, g_c^2) = (1, 0)$ (cases (5) and (6)), it is necessary that $B > 0$ and $A' + B' < 0$, i.e., that the non-corrupt government practices inappropriate fines, so that officials will want to be corrupt since their fines are not very high, and that the non-corrupt government is cost efficient in fighting corruption, and so prefers to be honest. In this situation the fines that a non-corrupt government practices are low, so a large number of officials will choose to be corrupt. The government makes little effort to fight corruption at the officials level since they have a low cost e or equivalently, they are efficient in fighting corruption. So they prefer to do it and act non-corruptly and fight corruption. If they are not efficient, i.e., the cost e is high, then the government would get higher utility by acting corruptly and henceforth destabilize the equilibrium.

Cases (5) and (7) may occur only when $A + B < 0$ (as in cases (3) and (4)). This means that in order to have long term evolution to an equilibrium where government is corrupt but officials prefer to be honest it is necessary that a corrupt government penalizes corrupt officials more than honest officials. See remark (2.3).

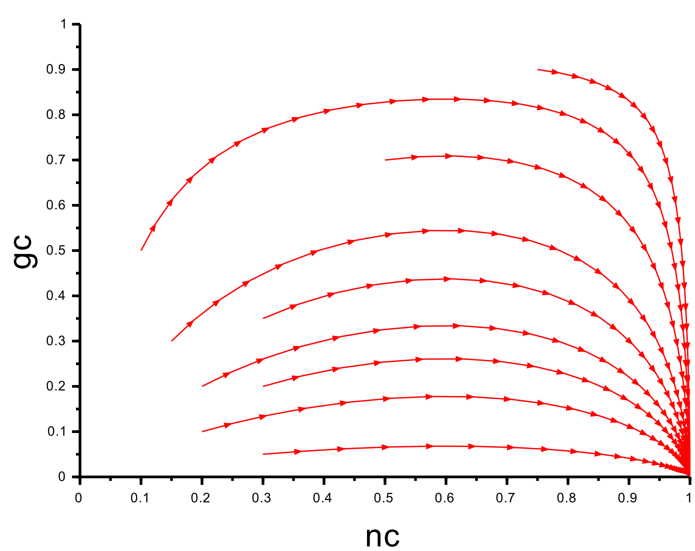

Figure 2.5: Some trajectories of the system for case (6): *de facto* government of the officials.

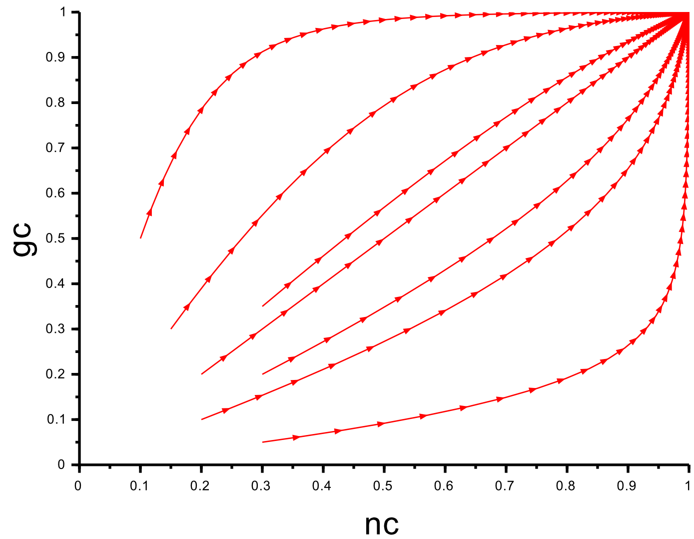

Figure 2.6: Some trajectories of the system for case (8): badly ruled society.

In case (8) we have an unruled society where the corruption equilibrium is stable thus ensuing a social trap. The society is evolving towards full corruption both on the governmental level and on the officials' level, due to low fines to punish corrupt officials, higher penalties for honest officials than for corrupt officials, high costs to capture corrupt officials and because of low re-election power, for instance because of a culture of impunity due to low intolerance index. This extreme situation may occur in a dictatorship. However, and most importantly it may occur in democratic societies where, for instance, citizens' perception of corruption is not clear, and where corruption is deeply and culturally rooted in society, making it very difficult to eradicate. The term *endemic or systemic corruption* has been used to designate this kind of corruption phenomena. In this case, corruption becomes a cause of several social and economic ills, such as increasing levels of inequality, poverty and lack of social mobility.

Corruption cycles

Recent works show that alternation in power can delay or stop processes of increasing corruption. However, this alternation can give place to a cyclical process of corruption, in which periods of increasing corruption alternate with periods in which it decreases. As previously mentioned, this represents a situation of *endemic* corruption, in this case where periods with high levels of corruption repeatedly occur. These periods are limited by the change of the ruler elite. Such is the case of Mexico in the period in which the alternation in power took place between the PRI (*Partido Revolucionario Institucional*) and the opposition party, PAN (*Partido Acción Nacional*). While we are witnessing a first period in which corruption descends, then this process accelerates, growing until the moment when the PAN is replaced in power by its competitor, the PRI. From the new triumph of the PRI we witness a new process of restraint and subsequent expansion of the corruption process. See [93] for a good empirical study of cycles in Mexican politics and its relation to the phenomena of corruption. In [13] the authors propose a game between politicians and contractors with the assumptions of material compensations to voters, and obtain corruption cycles in a discrete setting by analysing the repeated game. As we shall show the possibility of such cyclical process appear as a particular case of our model.

Periodic orbits appear when the radicand in the expression for the eigenvalues of the mixed Nash equilibrium is negative, i.e. $\frac{B'B}{AA'}(B'+A')(B+A) < 0$. Note that in this case Hartman-Grobman's theorem is not applicable, because the eigenvalues of the mixed equilibrium are purely imaginary numbers, i.e., have zero real part. We will now analyse and interpret the circumstances where this value happens to be negative.

The reasoning behind the existence of cycles of corruption is basically a restatement of the fact that in these two situations there is no pure strategy Nash equilibrium.

Let us consider the case where $B < 0,\ B' > 0$ and $A > -B,\ A' < -B'$. These inequalities imply that the corruption equilibrium $(n_c^1, g_c^1) = (1,1)$ and the non-corruption equilibrium $(n_c^4, g_c^4) = (0,0)$ are saddle points. The asymmetric equilibria $(n_c^2, g_c^2) = (1,0)$ and $(n_c^3, g_c^3) = (0,1)$ are also saddle points. These inequalities imply the existence of a mixed Nash equilibrium in the interior of the unit square, and by the previous formula, the eigenvalues of its linearisation are purely imaginary numbers, so that Hartman-Grobman theorem does not apply. In this case the dynamics are given by periodic orbit oscillating around the mixed interior Nash equilibrium, which is the focus of such periodic orbits. This situation corresponds to cycles of growth and decline of corruption. Recall that in this case the assumptions imply that there are low costs to capture corrupt officials, resulting in high efficiency, and there are high fines to punish corrupt officials, but the intolerance index is low. This interplay between these quantities results in the appearance of periodic orbits, as shown in figure (2.7). The periodic orbits in the figure rotate clockwise. The rationale behind this situation is the following. When general levels of corruption are low, because of low intolerance to corruption, a non-corrupt government has low re-election power, which causes an increase in government corruption, which in turn causes more officials to prefer to be corrupt since a corrupt government punishes honest officials more than corrupt officials. So the general levels of corruption have risen. As such, facing an increasingly number of corrupt officials, government becomes less corrupt, taking advantage of low costs to capture corrupt officials, and subsequently, because of the high fines the non-corrupt government practices, officials have a disincentive to be corruption, thus increasing the number of honest officials. Hence, the overall levels of corruption in society have declined to the original levels and the cycle restarts again.

A similar situation occurs if $B > 0,\ B' < 0$ and $A < -B,\ A' > -B'$. This situation corresponds to high re-election power, but the costs of the non-corrupt government in fighting corruption are high and fines practised are low. This again results in the appearance of periodic orbits, as shown in figure (2.7), but this time the periodic orbits in the figure rotate counter-clockwise. The rationale in this situation is the following. When corruption is generally high at both government and officials level, officials have an incentive to be honest, since the corrupt government penalizes corrupt behaviour more than honest behaviour since $A + B < 0$. But when more officials are honest, because of high re-election power by the non-corrupt government (or low re-election power by the corrupt government), governmental corruption decreases. On the other hand, while the general levels of corruption are low, since the fines practised are low, officials have an incentive to be corrupt, and the non-corrupt government, facing an increasing number of corrupt officials, and since it has low efficiency in prosecuting them subsequently becomes more corrupt. So, the general levels of corruption are high again and the cycle restarts.

Periodic orbits may appear naturally if the intolerance index is a function of the percentage of corrupt agents. In this situation the index increases when the number of corrupt officials grows, and decreases as does the percentage of corrupt officials. The corrupt political elite feels the pressure of a high index

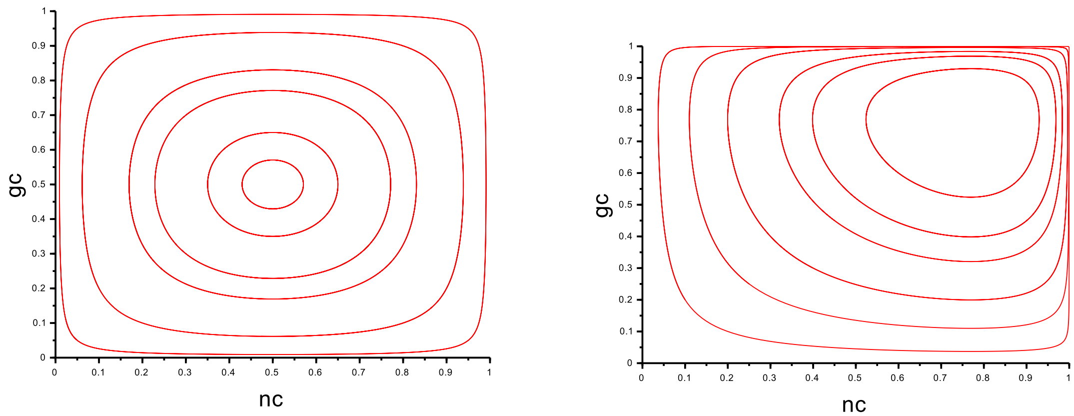

Figure 2.7: Some trajectories of the system (2.16) showing periodic orbits.

of intolerance, possibly reducing its expected value in this case of re-election, because the probability of being re-elected is reduced. As a result, the government corruption is reduced, and government will seek to punish corrupt officials more severely. But by reducing the amount of corrupt officials, the index of intolerance decreases, and therefore the pressure on the government declines, again permitting an increase in governmental corruption and allowing for an increase in the number of corrupt officials, thus restarting the cycle of corruption.

Game outcomes

Depending on the characteristics of the society, the dynamical equilibria may be or not Nash equilibria of the game.

Analysing the best responses of the players of the game, or using formulas (2.2),(2.3),(2.4),(2.5) instead, we have that each one of the four dynamical equilibria corresponding to pure strategies of the game is stable if and only if it is a (strict) Nash equilibrium. The other dynamical equilibrium, when it is interior, is always a Nash equilibrium of the game, but since it is mixed, it is never strict, so it can be a saddle or stable but is not asymptotically stable.

Hence, regarding the Nash/dynamical equilibria of the game, we have the following possibilities:

1. There is only one pure Nash equilibrium that is also asymptotically stable for the dynamics. This situation corresponds to cases (2),(3),(4),(6),(7) and (8).

2. There are two pure Nash equilibria that are both asymptotically stable for the dynamics. There are two possibilities for this case: either the two Nash equilibria are the non-corruption and the corruption equilibrium, or the government corruption equilibrium and the officials' corruption equilibrium. In this situation there is also a third mixed Nash equilibrium which is a saddle point of the dynamics whose stable manifold separates between the two pure Nash equilibria. This occurs in cases (1) and (5).

3. There is no pure Nash equilibrium. In this case the only Nash equilibrium is the mixed strategy that dynamically corresponds to a focus point for periodic orbits. This is the case of corruption cycles.

2.1.6 Some particular cases

For each time t we say that the pair $(n_c(t), g_c(t))$ defines the state of corruption of the society in time t. Thus, given the dynamical system (2.16) and an initial condition in time $t = t_0$ (i.e., an initial state of corruption), $(n_c(t_0), g_c(t_0)) = (n_{c_0}, g_{c_0})$, we say that $\xi(\cdot, (n_{c_0}, g_{c_0})) \to \Re^2$ is a solution of the dynamical system with such initial condition if and only if $\xi(t, (n_{c_0}, g_{c_0}))$ verifies the system (2.16), and $\xi(t_0, (n_{c_0}, g_{c_0})) = (n_{c_0}, g_{c_0})$. Classic theorems in the theory of differential equations show that once an initial condition is fixed, there is a unique solution for the differential equation and that the function $\xi(t, \cdot) : \Re^2 \to \Re^2$ is smooth, i.e., the solution of the dynamical system (2.16) is smooth with respect to initial conditions (see, for instance [46]).

Definition 2.3. (The trajectory of corruption) *Given the dynamical system (2.16) and an initial condition in time $t = t_0$, we define the trajectory of the corruption, as the set $\Gamma \subset \Re^2$ given by:*

$$\Gamma = \{(n_c(t), g_c(t)) = \xi(t, (n_c(t_0), g_c(t_0))),\ \forall\, t \geq t_0\}\,.$$

Note that each trajectory defines a set of possible future states of corruption, i.e., for each initial condition, there is only one set of possible future states (since we do not consider shocks and stochastic effects in this work). So, the corruption in a given society, once the initial condition is fixed, evolves along a trajectory.

Definition 2.4. (The transition path) *Given the dynamical system (2.16) and an initial condition the set of possible states for all $t > t_0$ will be called the* ***transition path****.*

This transition path is given by the set of possible states $(n_c(t), g_c(t))$ that represents the evolution of the corruption, from a fixed initial time $t = t_0$ until the system rests in a dynamical equilibrium. To obtain the possible transition path explicitly as a function is equivalent to obtain the analytical solutions of the dynamical system (2.13).

In general, for a dynamical system, it is not possible to obtain a complete analytical solution, however, like we did in the previous sections in some cases it is possible to analyse the behaviour of the solution close to each dynamical equilibrium. Certainly, to use this approach we need to assume that the parameters are given, or that they can be considered in a neighbourhood of each one of this points as constant.

However, in some particular cases it is possible to obtain analytical solutions. We will now consider some of these cases.

We consider here the case where $A = 0$ or $A' = 0$. Note that $A = 0$ is equivalent to $M_g = M + M'_g$ and $A' = 0$ is equivalent to $M_g + e = M + M'_g$. To simplify the notation consider the case $V_{G_c} = V_{G_{nc}} = V_G$. In these cases, since the system (2.16) turns out to be uncoupled and its solution is relatively simple, and once the initial conditions $n_c(0)$ and $g_c(0)$ in $t = 0$ are fixed we obtain the classical logistic solution.

$$n_c(t) = \frac{n_c(0)e^{Bt}}{(1 - n_c(0)) + n_c(0)e^{Bt}}, \ \text{ and } \ g_c(t) = \frac{g_c(0)e^{(M'_g - KP)t + \int_0^t D_{it}V_G dt}}{(1 - g_c(0)) + g_c(0)e^{(M'_g - KP)t - \int_0^t D_{it}V_G dt}}.$$

The evolution depends on the signs of B and $(M'_g - KP)t + \int_0^t D_{it}V_G dt$.

Recall that $B = M_c - M$ and $B'(t) = M'_g - KP - V_g D_{it}(n_c(t))$. Then, $B < 0$ if and only the value M of the fine is high enough, i.e., if and only if $M > M_c$. Moreover, $B'(t) < 0$ if and only if the intolerance index $D_{it}(n_c(t))$ is for all time $t > 0$ high enough i.e,; $D_{it}V_g(n_c(t)) > M'_g - KP$.

More in detail, the following cases are possible and particularly interesting:

- If $B < 0$ and for all time $B'(t) = M'_g - KP - V_g D_{it}(n_c(t)) < 0$ i.e, if $D_{it}(n_c(t)) > \frac{1}{V_G}(M'_g - KP)$ for all the $t > 0$. The 'good' equilibrium $(n_c^4, g_c^4) = (0, 0)$ is globally asymptotically stable, and independently of the initial condition inside the unit square, society is evolving to a situation where officials and government prefer to follow an honest behaviour, i.e., $g_c(t) \to 0$ and $n_c(t) \to 0$,, so $\xi(t, n_{c_0}, g_{c_0}) \to (0, 0)$, when $t \to +\infty$.

- If $B > 0$ and is for al time $B'(t) = M'_g - KP - V_g D_{it}(n_c(t)) > 0$ i.e, if $D_{it}(n_c(t)) < \frac{1}{V_G}(M'_g - KP)$ for all the $t > 0$. the 'bad' equilibrium $(n_c^1, g_c^1) = (1, 1)$ is globally asymptotically stable, and independently of the initial condition inside the unit square, society is evolving to a situation where officials and government prefer to follow a corrupt behaviour, i.e, $g_c(t) \to 1$, and $n_c(t) \to 1$, so $\xi(t, n_{c_0}, g_{c_0}) \to (1, 1)$, when $t \to +\infty$.

- If $B > 0$ and for all time t $B'(t) < 0$ then the equilibrium $(n_c^2, g_c^2) = (1, 0)$ is globally asymptotically stable, i.e., $g_c(t) \to 0$, and $n_c(t) \to 1$, so $\xi(t, n_{c_0}, g_{c_0}) \to (1, 0)$, when $t \to \infty$.

- If $B < 0$ and for all t $B'(t) > 0$ then the equilibrium $(n_c^3, g_c^3) = (0, 1)$ is asymptotically stable, i.e., $g_c(t) \to 1$, and $n_c(t) \to 0$, so $\xi(t, n_{c_0}, g_{c_0}) \to (0, 1)$, when $t \to \infty$.

Note that the index of intolerance and the value that the current government assign to the re-election play a central role in the possible evolution of corruption. The greater his interest in re-election, the more he will tend to control corruption, at least that which directly affects voters. In all the above cases, the basin of attraction of the asymptotically stable equilibria is the whole interior of the unit square.

In the degenerate cases when $B = 0$ or $B' = 0$, we have that, respectively, $n_c(t)$ or $g_c(t)$ is constant.

2.1.7 The role of the index of intolerance revisited

In some cases, corruption can be considered as a social trap [87]. Under several circumstances, the 'bad' equilibrium is asymptotically stable. In this case, if the initial distribution of corrupt officials and government's corrupt acts are in the basin of attraction of this equilibrium, neither official nor the government have incentives to act in a non-corrupt way. It is in this sense that we consider the corruption as a self- reinforcing mechanism and may say that a social trap is in place. In other words, an equilibrium that is Pareto dominated, is a Nash equilibrium and is dynamically stable. Corrupt actions by a party encourage corrupt actions by the other. If everybody is corrupt, nobody wants to be honest. To be corrupt is the rational way, because under these initial conditions, the expected value of this behaviour is higher than the expected value of the non-corrupt behaviour. Under this prospect, corruption looks like a sticky problem that can not be changed internally by agents. This grim prospect is analysed in several works such as in the cited work about social traps or in situations where practices such as gratitude payments (that can be regarded as a kind of 'black' market rent) are deeply rooted (see [54] for a very interesting study on the case of medical doctors gratitude payments in Hungary). In cases where corruption has advanced in different areas of society, neither alternation in power may be a guarantee against corruption, since a corrupt ruler elite can buy the vote of hopeless citizens and alternation in power might not really signify any deep change in the economic/societal structure where corruption develops. However, the degree of intolerance of citizens to corruption plays an important role to deter corruption. In these cases, when the annoyance of citizens over corruption is high enough, the process can be reversed as we have seen above. This is precisely what the Index of Intolerance measures. The recent events in Brazil show the evolution of society towards a low equilibrium, in which corruption seems to reach the different political elites, as well as officials, and where the high degree of nonconformity of citizens restrains a given party elite who appears as an accomplice of corrupt acts and ends up imposing another political group in government. The high degree of corruption of the different political sectors calls into question the possible curbing of the evolution of corruption. However, the high degree of intolerance of corruption demonstrated by citizens patent in massive demonstration that took place in Brasil in the last few years may force the political sectors now in the government to take precaution, if they intend to remain in power or even in some cases avoid imprisonment or prosecution for corruption offences, after some possible (or even probable) subsequent changes in the forces in power (see Brazil Corruption Report - Business Anti-Corruption Portal in [16] and [104]). See also the discussion about Mexican politics in [93] and in the preceding section on corruption cycles. There are examples of success in deterrence of corruption, for instance the cases of Singapore and Honk Kong (see [84]). These examples show that only when the rate of intolerance to corruption is high these processes of increasing corruption can be stopped.

Now consider that the intolerance index is an increasing and convex function of the number of corrupt officials and assume that $V_{G_{nc}} = V_{G_c} = V_G$. Then there exist n_{cmax} such that for all $n_c(t) > n_{cmax}$ the inequality

$$V_G D_{it}(n_c(t)) > \left[n_c(t)\left(M_g - M'_g - M + e\right) + M'_g - KP\right] \ ,$$

or equivalently

$$D_{it}(n_c(t)) > \frac{1}{V_G}\left[n_c(t)\left(M_g - M'_g - M + e\right) + M'_g - KP\right]$$

holds, then taking in to account equation (2.10) it follows that $E(G_{nc}) > E(G_c)$. Note that this n_{cmax} will be the maximum percentage of corrupt officials that a government interested in being re-elected can tolerate, and corresponds to the point where the convex function $y = D_{it}(n_c)$ intercepts the straight line $y = \frac{1}{V_G}(M_g - M'_g - M + e)n_c + \frac{1}{V_G}(M'_g - KP)$. So, a rational elite ruler must be receptive to the unhappiness of citizens if she has some interest in being re-elected.

Note that under the hypothesis of our model, it is natural to assume that the Intolerance Index grows with the amount of corrupt officials, because citizens perceive the corruption through the actions of the officials. When the amount of corrupts officials increases, the perception of corruption increases, increasing the intolerance of the population. However in future works it will be necessary to complete this index, considering other sources of information for citizens, for instance the press, rumours and investigations about vote buying, about government corruption and other aspects.

2.2 Democracy and citizenship in the evolution of corruption

In the second part of the chapter we consider that citizens are full players of the game. They have two behaviours or strategies, to be complacent with corruption, which we will call corrupt, or not be complacent, which we will call non-corrupt. We will analyse the evolution of corruption by means of the replicator dynamics.

This second part is structured in the following way: in section 2.2.1 we introduce the model as a normal-form game between three players or populations: the government, the officials and the citizens. Successively, in the next sections we shall describe the main characteristics of these different players of the game. In section 2.2.2 we discuss the characteristics of citizens, its strategies or behaviours and the associated payoffs. The description of the participation of the government is introduced in 2.2.3 and in section 2.2.4 we describe the action of the officials. In section 2.2.5 we introduce a dynamical system, consisting of the replicator equation describing the evolution of the corruption on society. In section 2.2.6 we discuss the socio-political characterizations of the players of the game according to the parameters of our model. The dynamical equilibria of the dynamical system are described in section 2.2.7. We study the stability of the pure equilibria and the stability and existence of the other mixed equilibria. In section 2.2.8 we analyse the dynamics on the edges of the cube, how they are enough to characterize the asymptotic stability of the vertexes and their relation to the socio-political characterization of the game. In section 2.2.9 we analyse when the pure strategy profiles are dynamical attractors and in section 2.2.10 we provide some graphical examples.

2.2.1 The model

We consider a three-population, two-strategy normal-form game where the populations are: the citizens, the government i.e., the political elite (elected directly by universal suffrage of citizens), and the officials named directly by the current government to carry out the administrative tasks of the state. The government may be regarded as a first level ruler elite. The officials may be regarded as a second level elite. For simplicity we assume that the group of officials is small while compared to the citizens group. In other words, since the group of officials is small, it has a negligible influence in elections since they are heavily outnumbered by citizens. At the end of each electoral period, citizens re-elect the current government or not. The officials must choose between to accomplish their task in an honest way or to require citizens an illegal payment for the purpose of performing some task. It may be for the fulfillment of some service or the payment of some bribe after some infraction by a citizen. It is up to the ruler elite or government to choose between punishing or being an accomplice of corrupt officials. In the first case we say that the government acts in an honest or non-corrupt way and in the second one we say that the government is acting in a corrupt way. We say that a citizen acts in a corrupt way (or briefly we will say that it is corrupt) if he is willing to sell his vote to a corrupt government, which can also be interpreted *latu sensu* as being complacent with corruption, in other case we will say that he is an honest or non-corrupt citizen.

Summarizing, we have a game with three populations: citizens, government, and officials, where each one must choose between two pure strategies. This game is symbolized by the formal expression: $\Gamma = \{P, S, U\}$, where

- $P = \{O, G, C\}$ represents the players or populations of the game. By C we denote the citizens, by O the officials and by G the government.

- $S = S_O \times S_G \times S_C$ is the strategy space where $S_O = \{O_c, O_{nc}\}$, $S_G = \{G_c, G_{nc}\}$ and $S_C = \{C_c, C_{nc}\}$ represents the set of pure strategies for each player, namely, to be corrupt or not-corrupt.
- U represents utility, or payoff obtained by each player depending on his/her own strategy and on the strategies followed by the other players. We will describe the payoffs in the next subsections.

Furthermore, as usual, a mixed strategy for each player, in our case for each population, is a probability distribution over his/her set of pure strategies. As previously, we will denote by $n = (n_c, n_{nc})$ a distribution of the officials over the two strategies described above, by $g = (g_c, g_{nc})$ represents the percentage of corrupt and non-corrupt acts of the government and by $c = (c_c, c_{nc})$ we represent a distribution of probabilities over the set of pure strategies of the citizens. In this case c_c corresponds to the percentage of citizens in the first group, or equivalently the probability that the average citizen follows a compliant behaviour and c_{nc} the percentage of citizens in the second group, i.e., the probability that the average citizen follows a non- compliant behaviour.

In the next three section we shall describe the payoffs of the game Γ.

2.2.2 The citizens

We assume that there are two types of citizens. Those of the first group are willing to give their vote to a corrupt government in exchange for receiving a monetary compensation. The citizens of the second group are not willing to sell their vote to a corrupt government. This may also be interpreted as saying that one group has some compliance with corruption while the other has not.

It may of course be argued that under rationality no citizen would like to be compliant with a corrupt government. However, for some reasons, he/she might be compelled to have some degree of compliance towards a corrupt government. Indeed, we may re-interpret social inaction by citizens as a a kind of silent compliance.

The number of citizens in one group or another may change according to the amount that a corrupt government is willing to pay for the vote. To simplify, we assume that this amount is fixed and equal to P. We also assume that citizens who are not willing to accept government corruption by selling their vote will receive benefits only when the government is an honest government with which they feel represented. These citizens have preferences for an honest government, to which they assign a value of satisfaction $\alpha > 0$. They also assign a level of dissatisfaction $\beta > 0$ to a corrupt government. In addition, we assume that when a citizen meets a corrupt official, the citizen has to make a payment, that we assume normalized as equal to -1. This payment is caused by the necessity of bribing the official in order to make him fulfill or accelerate the fulfillment of a given service. Furthermore, we also assume that when the citizen is honest and has to make such a payment to a corrupt official we consider that he has a level of dissatisfaction measured by $\gamma > 0$.

We consider that citizens can choose between two possible strategies by choosing to which one of the previously described groups they belong: the one of complicity with a corrupt government, or the rejection of this type of government. Hence, the citizens in the first group may be seen as having a compliant behaviour with corruption, while the citizens in the second group do not. The profits corresponding to each type of citizen are represented in the following table.

Regarding the parameters that feature in the payoff table we have the following:

- $\alpha > 0$ is the utility or satisfaction that a honest citizen assigns to a non-corrupt government.

	G_c, O_c	G_c, O_{nc}	G_{nc}, O_c	G_{nc}, O_{nc}
C_c	$P-1$	P	-1	0
C_{nc}	$-\beta-1-\gamma$	$-\beta$	$\alpha-1-\gamma$	α

Table 2.2: The payoff table of the citizens in the three-population game.

- $\beta > 0$ is the disutility or dissatisfaction that a honest citizen assigns to a corrupt government.
- $+1$ corresponds to the normalized payment that a citizen must make when he encounters a corrupt official.
- $P > 0$ is the payment that a corrupt government makes to buy the vote of a citizens. It may be interpreted *latu sensu* as a welfare gain of a corrupt or compliant citizen with governmental corruption.
- $\gamma > 0$ is the dis-utility or dissatisfaction that a honest citizen assigns when he/she encounters a corrupt official. This value adds to the payment that he has to make to the official and represents the dislike of corruption by a honest citizen.

Instead of assigning the payoff 0 to a corrupt citizen when the government and officials are non-corrupt we could have assigned the payoff α_c, as a satisfaction that the citizen attains, and the satisfaction α_{nc} when the citizen is not compliant. In order to simplify notation we consider the first value as 0, so that α may be seen as $\alpha = \alpha_{nc} - \alpha_c$. It is clear that $\alpha_{nc} > \alpha_c$.

We observe that the payoffs in the previous matrix may be thought as being in units of dis-utility obtained from the payment of a bribe, since the payment of the bribe to a corrupt official is -1. It can also be argued that a citizen, even if compliant with corruption also gets a dis-utility when meeting a corrupt official. In order to simplify notation we consider that this dis-utility is already included in the -1, so that when the citizen is honest he also gets an increment $\gamma > 0$ of dis-utility to the -1. As in the case of the satisfaction α we may regard γ as $\gamma = \gamma_{nc} - \gamma_c$. It is clear that one should have $\gamma_{nc} > \gamma_c$ since the dissatisfaction for meeting a corrupt official should be bigger for the non-compliant citizen.

To simplify we assumed that to every corrupt citizen is offered the same amount P of monetary compensation for the vote in a corrupt government, which certainly limits the possible actions of a corrupt government, a fact that, as we shall see, may not be minor. This amount P corresponds to a simplified version of possible royalties granted by corrupt governments to possible electors. In the case of entrepreneurs, it may correspond to public concessions or, in the case of poor families of developing countries to consumption bundles. More generically it corresponds to a welfare gain that a citizen that is compliant to corruption has when a corrupt government is in charge, even if this welfare gain is unexpected by the citizen or inadvertently obtained. We can thus say that these payoffs are implicit when a citizen is compliant with corruption. The same occurs for the factors of dissatisfaction with corrupt government β and satisfaction with honest government α.

Different groups or individuals can typically receive different personal profits from a corrupt government. These individual profits are far from being favourable to the development of social welfare, and moreover, they are antagonistic to the social interest. Again, given this disparity, the amount P can be regarded as the average of such welfare gain in a country, or in other words the gain of a representative, or focal, compliant citizen when a corrupt government is in charge. The same with the payments of bribes which may be different so that we consider an average value that we normalize to 1. This is just a simplification of the notation M_c we used in the first part of the chapter in order to avoid cumbersome expressions and facilitate interpretation.

The expected payoffs of each strategy of the citizens depends on the strategy followed by the other players of the game, the officials and the government, so they are the conditional expected values $E(C_j|n,g)$ of the pure strategies $j \in \{c, nc\}$, given that the distribution of the officials is n and the mixed strategy of the government is g. To simplify the notation we shall denote this by $E(C_j),\ j \in \{c, nc\}$. Then we have:

$$\begin{aligned} &E(C_c) = (P-1)g_c n_c + P g_c(1-n_c) - (1-g_c)n_c \\ &E(C_{nc}) = -(\beta+1+\gamma)g_c n_c - \beta g_c(1-n_c) + (\alpha-1-\gamma)(1-g_c)n_c + \alpha(1-g_c)(1-n_c)\ . \end{aligned} \tag{2.17}$$

where $n = (n_c, 1-n_c)$ is a distribution of the official over the set of pure strategies or a mixed strategy for the officials. Similarly $g = (g_c, 1-g_c)$ represents a mixed strategy for the government.

If we consider that citizens' preferences over the result of their actions are given by a Von Neumann-Morgentern utility function then they will chose to accept the governmental corruption if $E(C_c) > E(C_{nc})$ and in this case they will prefer to be corrupt. In the case where the reciprocal inequality is verified, citizens will choose a non corrupt behaviour. They are indifferent in the case where the equality is verified. This equality defines a threshold such that if the corruption of the officials and government is over this threshold, citizens will also choose a corrupt behaviour.

Again it can be argued that even if they can act compliantly towards corruption, citizens tend to act irrationally in this situation since they would not like to act compliantly even if this yields them a better payoff. Indeed, as we explained in the introduction to this work, evolutionary game theory allows to relax the rationality assumption and focus on the dynamics of change in behaviours even if they are no rational. The insight of evolutionary game theory is that having more than one possible strategies, or behaviour, through some fitness and evaluation of performance, selection will act to select the best options, which will survive in the population.

According to the replicator dynamics, the evolution of the citizens strategies over time is given by the following the differential equations system:

$$\begin{aligned} \dot{c}_c &= c_c(1-c_c)[E(C_c) - E(C_{nc})] \\ \dot{c}_{nc} &= -\dot{c}_c\ . \end{aligned} \tag{2.18}$$

2.2.3 The government

The payoffs for the government are similar to the ones in the first part of this chapter. The payoff table is as follows:

More precisely we have that:

- W is the wage paid by the government to an official.

	O_c, C_c	O_c, C_{nc}	O_{nc}, C_c	O_{nc}, C_{nc}
G_c	$M_g - W - P + V_{c,c}$	$M_g - W + V_{nc,c}$	$M'_g - W - P + V_{c,c}$	$M'_g - W + V_{nc,c}$
G_{nc}	$M - W - e + V_{c,nc}$	$M - W - e + V_{nc,nc}$	$-W + V_{c,nc}$	$-W + V_{nc,nc}$

Table 2.3: The payoff table of the government in the three population game.

- P is the amount paid by government to buy votes. It may be interpreted as the amount paid by government to influence and increase its chances of remaining in power.

- M_g is the punishment that a corrupt government imposes to a corrupt official (for instance for the governmental complicity with corruption of the official).

- M'_g is the punishment that a corrupt government imposes to an honest official (for instance for the honest official to keep his place).

- M is the fine imposed by an honest government to a corrupt official.

- e is the cost of a non-corrupt government in fighting corruption. This is a measure of a non-corrupt government's efficiency in fighting corruption.

- $V_{i,j}$ are the valuations of a government of type j when encountering a citizen of type i, or in other words, when citizens of type i votes for it. This represents a voting externality of the government related with both the valuation that a government assigns to being maintained in power for one period more and the probability of that occurring according to the vote in an election, and an externality parameter related with citizen compliance.

As usual, we consider for simplicity that the values of wage, fines and P correspond to average values observed in society. In other words it represent the average/typical payoffs for government in a society.

Furthermore, observe that as in the interpretation of α and γ we can think of the valuation $V_{i,j}$ analogously. Indeed we can think of $V_{c,nc}$ and $V_{nc,c}$ as being zero, which is the same as seeing $V_{c,c}$ and $V_{nc,nc}$ as the differences $V_{c,c} - V_{c,nc}$ and $V_{nc,nc} - V_{nc,c}$ respectively, so that we would only have the valuation of an honest government towards a non-compliant citizen and the valuation of a corrupt government towards a compliant citizen, and the other valuations being zero. A simplification of this would be o consider that the valuations o the government only depend on the government type and not on the citizen's type. This would mean that $V_{c,c} = V_{nc,c} = \tilde{V}_c$ and $V_{c,nc} = V_{nc,nc} = \tilde{V}_{nc}$.

Remark 2.5 (A consideration on the punishments of the corrupt government)**.** *It is often observed that whenever a corrupt ruler elite has a sufficiently short time horizon, it is in his interest to confiscate the property of his subjects, or to abrogate previously signed contracts and generally to ignore the long-run economic consequences of his choices. This point is widely discussed in [78]. A corrupt elite with such short time horizon takes welfare from all aspects of society, even from those who could act as his accomplices, in our case the corrupt officials. This welfare is measured by M_g and M'_g.*

The expected payoffs for pure strategies of the government are:

$$
\begin{aligned}
E(G_c) &= (M_g - W - P + V_{c,c})n_c c_c + (M_g - W + V_{nc,c})n_c(1-c_c) + \\
&+ (M'_g - W - P + V_{c,c})(1-n_c)c_c + (M'_g - W + V_{nc,c})(1-n_c)(1-c_c) \\
E(G_{nc}) &= (M - W - e + V_{c,nc})n_c c_c + (M - W - e + V_{nc,nc})n_c(1-c_c) + \\
&+ (-W + V_{c,nc})(1-n_c)c_c + (-W + V_{nc,nc})(1-n_c)(1-c_c)\ .
\end{aligned}
\tag{2.19}
$$

Considering the expected utilities of the government we have that it chooses to follow a corrupt strategy if $E(G_c) > E(G_{nc})$. Otherwise, if $E(G_c) < E(G_{nc})$ the ruler elite prefers to follow a non-corrupt behaviour. In the case of equality, i.e; $E(G_c) = E(G_{nc})$ the government is indifferent between the two. This threshold is surpassed if the citizens don't have a high level of rejection of corruption, i.e., if the group C_c is large and if for example punishing corrupt officials has a high cost for the government, because it is inefficient in the fight against corruption.

As in the case of citizens, following the replicator dynamics, the evolution of the strategic decision of the government will be given by the differential equations system:

$$
\begin{aligned}
&\dot{g}_c = g_c(1-g_c)[E(G_c) - E(G_{\bar{c}})] \\
&\dot{g}_{\bar{c}} = -\dot{g}_c\ .
\end{aligned}
\tag{2.20}
$$

2.2.4 The officials

The payoffs of the officials only depend on the strategy followed by the government and their own strategy, and not on the citizens behaviour, which only interferes with the government in the way we have seen before in the previous section. Their payoffs are similar to those we discussed in the first part of this chapter. The payoffs of the officials are given in the following table.

	G_c, C_c	G_c, C_{nc}	G_{nc}, C_c	G_{nc}, C_{nc}
O_c	$W - M_g + 1$	$W - M_g + 1$	$W - M + 1$	$W - M + 1$
O_{nc}	$W - M'_g$	$W - M'_g$	W	W

Table 2.4: The payoff table of the officials in the three population game.

Note that regardless of whether the citizen is honest or not, a corrupt officer will always demand a payment to the citizen for performing the required management. This assumption is represented by the +1, which always appears in the row corresponding to the official's corrupt strategy.

As before, officials are in direct contact with the citizens whom they bribe in the case officials act corruptly, and they can be punished by the ruling elite. Eventually, such punishment may be zero, i.e. the official may be fired, so as previously, we consider the parameters in the table as averages among the populations in consideration in the game.

The expected payoffs associated with each strategy are given by:

$$\begin{aligned} E(O_c) &= (W - M_g + 1)g_c + (W - M + 1)(1 - g_c) \\ E(O_{nc}) &= (W - M_g')g_c + W(1 - g_c) \ . \end{aligned} \tag{2.21}$$

They will choose a corrupt strategy if and only if $E(O_c) > E(O_{nc})$. Similar considerations that in the case of citizens and government can be given for the choice of officials.

The dynamics for the choice of the officials is given by the replicator dynamics, which is the the following system of differential equations

$$\begin{aligned} &\dot{n}_c = n_c(1 - n_c)[E(O_c) - E(O_{\bar{c}})] \\ &\dot{n}_{\bar{c}} = -\dot{n}_c \ . \end{aligned} \tag{2.22}$$

2.2.5 The evolutionary dynamics of corruption in society

The tables above show that different groups of officials typically benefit in different way from different types of government. In turn different types of government can be promoted for different kind of citizens and institutions. Citizens obviously have different preferences towards the government. There is generally a conflict between individual or political interest and social interest, and this conflict is aggravated by corrupt governments and institutions that pursue spurious interests such as the officials. According to the interests of the ruler elite, i.e. a corrupt government may try to appear honest by punishing officials that have been found in acts of corruption, although this punishment has only a propaganda end to potential voters. In other words, under certain conditions it can be difficult to decrease government corruption.

Returning now to each of the systems of differential equations previously considered, we will analyse the evolution of corruption in society. Precisely the possible solutions of a system conformed by the three differential equations above introduced will represent the evolution of corruption in the society in the three levels corresponding to each one of the players or populations of the game. Taking into account the classification of various types of corruption presented in [50], we have that our dynamical system tracks the evolution of *grand corruption* at the level of the political elites, such as the government, tracks the dynamics of *petty corruption* that occurs at the level of officials, and finally the dynamics of the citizen's willingness to sell their vote, which contrary to the first part of this chapter will now dynamically evolve. In this case, this last part corresponds to the willingness of citizen's to engage in vote buying which may also be seen as the dynamics of citizen's compliance with governmental corruption. So that the role of citizen's intolerance that we described in section 2.1 is now in the three population game a dynamically evolving mixed strategy, or distribution of the action, or influence/pressure that citizen's and democracy exert in corruption at the government's level. The valuations $V_{i,j}$ of the government are also related to the citizen type, and can be extended to situations where they do depend on other parameters, or in the number of corrupt officials, for instance.

The evolution of corruption in the society is given by the three differential equation systems (2.18, 2.20, 2.22). Each one of these systems of differential equations has only one free equation since the second one is determined by the first. So we obtain a system with three differential equations.

$$\begin{cases} \dot{n}_c = n_c(1-n_c)(E(O_c)-E(O_{nc})) \\ \dot{g}_c = g_c(1-g_c)(E(G_c)-E(G_{nc})) \\ \dot{c}_c = c_c(1-c_c)(E(C_c)-E(C_{nc})) \end{cases} \tag{2.23}$$

Taking into account equations (2.17), (2.19) and (2.21), after some algebra we have

$$\begin{cases} \dot{n}_c = n_c(1-n_c)\left[(M+M'_g-M_g)g_c+1-M\right] \\ \dot{g}_c = g_c(1-g_c)\left[(e+M_g-M-M'_g)n_c+ (V_{c,c}-V_{c,nc}-V_{nc,c}+V_{nc,nc}-P)c_c+V_{nc,c}-V_{nc,nc}-M'_g\right] \\ \dot{c}_c = c_c(1-c_c)\left[(P+\beta+\alpha)g_c+\gamma n_c-\alpha\right] \end{cases} \tag{2.24}$$

These equations may be written in the following form:

$$\begin{cases} \dot{n}_c = n_c(1-n_c)\left(A_o g_c + D_o\right) \\ \dot{g}_c = g_c(1-g_c)\left(E_g n_c + B_g c_c + D_g\right) \\ \dot{c}_c = c_c(1-c_c)\left[A_c g_c + \gamma n_c - \alpha\right] \end{cases} \tag{2.25}$$

where

$$A_o = M + M'_g - M_g$$

$$D_o = 1 - M$$

$$E_g = e + M_g - M - M'_g$$

$$B_g = V_{c,c} - V_{c,nc} - V_{nc,c} + V_{nc,nc} - P$$

$$D_g = V_{nc,c} - V_{nc,nc} + M'_g$$

$$A_c = P + \beta + \alpha$$

If the parameters of the model are known, then the dynamical system is well defined and if in addition the initial conditions are given, this system has an unique solution. Thus the possible solutions of this dynamical system will reflect the possible trajectories of evolution of corruption in the society. Since in general it is not possible to compute the solutions of a dynamical system we will have to rely heavily in stability analysis of the equilibria, or stationary states of the system.

So if our interest is to analyse the evolution of corruption from some time $t = t_0$ we need to know the state of the system at the moment t_0, i.e., the initial distributions $c(t_0), n(t_0)$ and $g(t_0)$. The time t_0 represents the initial moment from which we are interested in reflecting on the future development of corruption. Let assume that in time t_0 $c(t_0) = c_0$, $n(t_0) = n_0$ and $g(t_0) = g_0$ we denote these initial conditions by $d_0 = (c_0, n_0, g_0)$. Fixed the initial conditions, the solution of the system (2.23) is a unique function $\xi(t; t_0, d_0) : [t_0, \infty) \to \mathbb{R}^3$ representing the states of the system in each time t after t_0. So $\xi(t; t_0, d_0) = (n(t), g(t), c(t))$ represents for each $t \geq t_0$ the state of the system. In our case that corresponds to the instantaneous distributions over the set of possible behaviours of the players.

It is worth to say that a change in the parameters of the model can give place to a change in the evolution of corruption, qualitatively changing the trajectories of corruption and other dynamical feature such as the attractor. Societies with similar or even equal initial conditions can evolve to very different steady states. This means that changes in public policies with repercussion in the values of parameters of the model can give place to changes in the future evolution of the corruption. The dynamical system shows the role of each of these parameters and its relevance in the fight against corruption or in the perpetuation of corruption. We will precisely discuss in the remainder of this chapter the effect of the model parameters in such outcomes and interpret the diverse situations corresponding to different socio-political conditions that may arise and what the limit behaviour of the dynamics would be.

As we said, in general it is not possible to obtain the analytic solutions of the system, but we can analyse the behaviour of solutions in a neighbourhood of each stationary state, by means of Hartman Grobman's theorem when it is applicable, i.e., when the equilibrium is hyperbolic. With this objective let us consider the dynamical equilibria, or steady-states of the system.

2.2.6 Socio-political characteristics of government, officials and citizens

We now write the characteristics of each player of the game according to the parameters of model. The inequalities that follow and which characterize the players will be crucial in the subsequent analysis that we will do regarding the stability ff the equilibria of the system and the evolution of corruption and outcomes that occur since they will allow us to interpret each scenario corresponding to different socio-political conditions.

Each one of the inequalities corresponds to one of the pure strategy best-responses of the game. Hence *a priori* there would be 12 different inequalities. However, since we observed that the payoffs of the officials are independent of the citizens behaviour then we only have 10 different inequalities regarding the best-responses of the game.

Of these 10 inequalities, 3 of them are mandatory with respect to our assumptions, and we will always assume they occur. We start with them.

(i) $\alpha > 0$. This is the utility or satisfaction that a honest citizen drives from a non-corrupt government. This means that the best response of citizens when facing honest government and officials is to be honest. Recall that writing α as a difference as we explained before this is equivalent to $\alpha_{nc} > \alpha_c$, meaning that the valuation towards a non-corrupt government by a honest citizen is greater than that of a compliant citizen.

(ii) $A_c - \alpha > 0$ since $A_c - \alpha = P + \beta > 0$. This case corresponds to the best response when the government is corrupt and officials are honest.

(iii) $A_c + \gamma - \alpha > 0$ since $A_c + \gamma - \alpha = P + \beta + \gamma > 0$. This case corresponds to the best response when the government and officials are corrupt.

The last two inequalities mean that the citizens best-response when government is corrupt is to be corrupt, or in other words to be compliant with it. This is because since the government is corrupt, it is always willing to lose some welfare P corresponding to vote buying. Furthermore, a non-corrupt citizen, or non-compliant, always has a dis-utility associated to a corrupt government, and in the case the officials are corrupt citizens have a further dis-utility associated to this. However this does not mean that a corrupt outcome is inevitable since the valuations $V_{i,j}$ (or externalities) in the government payoffs may force corruption to decrease. Indeed, whether or not the government and officials have these incentives is characterized by the following inequalities.

(i) $\gamma - \alpha > 0$. This inequality has a clear interpretation which is that the dissatisfaction or dis-utility that a honest citizen gets from encountering a corrupt official is smaller than preference of the honest citizen to have a honest government. Interpreting this in another way as we suggested before we have that $\alpha_c - \gamma_c > \alpha_{nc} - \gamma_{nc}$ meaning that the net satisfaction of a compliant citizen is greater than that of a non-compliant citizen. Conversely, if $\gamma - \alpha < 0$ the non-compliant citizen gets a higher satisfaction.

(ii) $D_o > 0$, or equivalently $1 > M$. This means that the non-corrupt government practices inappropriate fines towards a corrupt official, since the fine is lower than the value of the bribe a corrupt official gets from a citizen, i.e., $+1$. Respectively, $D_o < 0$, equivalently $1 < M$ means that the no-corrupt government practices appropriate fines.

(iii) $A_o + D_o > 0$, or equivalently $M'_g > M_g - 1$. This means that a corrupt government penalizes honest officials more than dishonest officials. Conversely, if $A_o + D_o < 0$, or equivalently $M'_g < M_g - 1$ mean that a corrupt government penalizes dishonest officials more than honest officials. We may say that in this case the corrupt government takes more welfare from non-corrupt officials than from corrupt officials.

(iv) $E_g + D_g > 0$, or equivalently $e + M_g - M + V_{nc,c} - V_{nc,nc} > 0$. This inequality concerns the choice of the government when officials are corrupt and citizens are non-compliant with corruption. It reflects the efficiency of the non-corrupt government in fighting corruption when citizens are non-compliant. In the previous case the parameter e is high so that the costs of fighting corruptions are high, which means low efficiency. If $E_g + D_g < 0$ the non-corrupt government is cost efficient in fighting officials corruption when citizens are non-compliant.

(v) $E_g + B_g + D_g > 0$, or equivalently $e + M_g - M - P + V_{c,c} - V_{c,nc} > 0$. This inequality concerns the choice of the government when officials are corrupt and citizens are compliant with corruption. It reflects the efficiency of the non-corrupt government in fighting corruption when citizens are compliant. In this case the non-corrupt government is cost inefficient in fighting corruption. If $E_g + B_g + D_g < 0$ the non-corrupt government is cost efficient in fighting officials' corruption when citizens are compliant.

(vi) $D_g > 0$, or equivalently $V_{nc,c} + M'_g > V_{nc,nc}$. This inequality concerns the choice of the government when officials are non-corrupt and citizens are non-compliant with corruption. In this case the valuation that non-compliant citizens give to a corrupt government plus the value of the punishment to the non-corrupt officials is higher than the valuation that non-compliant citizens give to a non-corrupt government. It reflects the re-election power of a government when citizens are non-compliant. In this case the non-corrupt government has low re-election power, or in other words, the corrupt government has high re-election power when citizens are compliant. If $D_g < 0$ the non-corrupt government has high re-election power.

(vii) $B_g + D_g > 0$, or equivalently $V_{c,c} + M'_g - P > V_{c,nc}$. This inequality concerns the choice of the government when officials are non-corrupt and citizens are compliant with corruption. In this case the valuation that compliant citizens give to a corrupt government plus the value of the punishment to the non-corrupt officials and minus the welfare loss P corresponding to vote buying is higher than the valuation that compliant citizens give to a non-corrupt government. It reflects the re-election power of a government when citizens are compliant. In this case the non-corrupt government has low re-election power, or in other words, the corrupt government has high re-election power when citizens are compliant. If $B_g + D_g < 0$ the non-corrupt government has high re-election power.

Inequality (i) concerns the preference of the citizen when facing an honest government and corrupt officials. His preference thus depends of the satisfaction drawn from the non-corrupt government and the dis-utility of facing a corrupt official.

We observe that inequalities (ii) and (iii) are similar to the ones we had when considering the two population game in the first part of this chapter. These two inequalities refer to best responses of the officials when facing a corrupt government and a non-corrupt government. Citizens choice does not influence the best response of the officials. What officials do is: when the government is non-corrupt, they compare the fines the he incurs to the payment he receives from the bribe; when the government is corrupt he compares the penalties he would get when the government is corrupt to the penalty when the government is not corrupt.

We observe that we have two inequalities, (iv) and (v), including the parameter e which measures the cost of a non-corrupt government in fighting corruption, instead of only one we had in the first part of this chapter. This is due to the fact that there are now two possible citizen behaviours, compliance or not. Observe that (iv) and (v) are very similar but that in (v), the inequality for the efficiency of the government when citizens are compliant with corruption, there is an extra term P that corresponds to the vote buying parameter. If we consider the special case where the valuations of the government only depend on the government strategy and not on citizens type, then the term P becomes the only difference between (iv) and (v). Indeed, if we define the efficiency threshold of a non-corrupt government:

$$T_e = \tilde{V}_{nc} - \tilde{V}_c + M - M_g$$

the two inequalities become respectively $e > T_e$ and $e > T_e + P$. So we have that $P > 0$ pushes the efficiency threshold to the right, which means that an efficient non-corrupt government when citizens are non-compliant with corruption is also efficient when citizens are compliant. In other words, a non-corrupt government that is inefficient when citizens are compliant is also inefficient when citizens are non-compliant. The reason is that while a non-corrupt government might not be sufficiently efficient when citizens are not-compliant, because the cost e is too high, and hence prefers to act corruptly, it may be efficient when citizens are compliant since in that case there is a loss in government's welfare of magnitude $P > 0$ if the government strategy is to be corrupt, so that the cost e is not so high, and so the best choice would be to be non-corrupt. Actually, in this case, even with a slightly higher cost, say $e_2 \gtrapprox e$ such that $e_2 < T_e + P$, the government would still be efficient when citizens are compliant. If $P = 0$ then there is no difference between the two inequalities.

Similarly for inequalities (vi) and (vii) which can be interpreted as characterizing the re-election power of a government. We also had one inequality that we interpreted as measuring the re-election power of a government in the first part of this chapter. This time we have two inequalities regarding this characteristic of the government since one of them regards the re-election power of the government under citizen compliance with corruption while the other regards the re-election power of the government when citizens are not compliant with corruption. Inequalities (vi) and (vii) are relatively similar apart from the fact that the vote buying parameter P appears in (vii). To better understand these two inequalities again consider the special case where the valuations of the government only depend on the government strategy. Now observe that (vi) can be written as

$$\tilde{V}_{nc} - \tilde{V}_c < M'_g$$

and that (vii) can be written as

$$\tilde{V}_{nc} - \tilde{V}_c < M'_g - P \ .$$

So we have that the value $P > 0$ pushes the threshold to the left, and so we have that if a non-corrupt has high re-election power (or that a corrupt government government has low re-election power) when citizens are not-compliant with corruption, then the non-corrupt government also has high re-election power when citizens are compliant with corruption. So it is more likely that a non-corrupt government has high re-election when citizens are compliant with corruption since while the government might prefer to be corrupt because its re-election power is low when citizens are compliant, the government might prefer to be honest when citizens are not compliant while his re-election power is high. This is due to the fact that in this last case government has a welfare loss of $P > 0$ if he decides to be corrupt, and if P is high enough then the damage to the government's payoff would increase so that it would be better to chose the strategy non-corrupt, which means high re-election power for the non-corrupt government. In other words, a given re-election power of a non-corrupt government might be too low when citizens are non-compliant, but because when citizen's are compliant the government incurs in a loss of welfare equal to $P > 0$ if it decides to be corrupt, then the same re-election power is enough to make the government prefer to be non-corrupt, i.e. it then has high re-election power.

2.2.7 The steady-states of the system

It is clear that the vertices of the three-dimensional unit cube $\mathcal{C} = [0, 1]^3$ are equilibria of the system (2.25). The vertexes correspond to pure strategy profiles of the game. Depending on the values of the parameters considered in tables 2.2, 2.3 and 2.4, these points could be also pure Nash equilibria for the game.

Apart from these equilibria there could be others corresponding to situations where some or all the players of the game randomize between the two pure strategies when they are indifferent between them. For instance the point $N^* = (n_c^*, g_c^*, c_c^*)$ where the identities

$$E(O_c^*) - E(O_{nc}^*) = E(G_c^*) - E(G_{nc}^*) = E(C_c^*) - E(C_{nc}^*) = 0$$

are verified, can be an interior point of the cube $\mathcal{C}$, so that the system would have an interior stationary state.

In such case, the corresponding distributions

$$n^* = (n_c^*, n_{nc}^*); \ \ g^* = (g_c^*, g_{nc}^*); c^* = (c_c^*, c_{nc}^*)$$

that solve the previous equations constitute a mixed Nash equilibria for the game. We will refer to this equilibrium as the strictly mixed equilibrium of the game. If the system is in this stationary state then the government, officials and citizens are indifferent between acting in a corrupt or honest manner so that they randomize between the two behaviours in the proportion given by the distribution of the steady state.

Furthermore, we may have situation where not all players randomize but only one or two of them, or in other words, at least one of the population has only one type of individuals, since they all choose the same strategy while the other populations randomize. In our case this yields equilibria that lie on faces of the unit cube $\mathcal{C}$ or on edges. More specifically, the stationary states of the system are characterized by the following relationships:

$$\dot{n}_c = 0 \iff n_c = 0 \textbf{ or } n_c = 1 \textbf{ or } A_o g_c + D_o = 0$$
$$\dot{g}_c = 0 \iff g_c = 0 \textbf{ or } g_c = 1 \textbf{ or } E_g n_c + B_g c_c + D_g = 0$$
$$\dot{c}_c = 0 \iff c_c = 0 \textbf{ or } c_c = 1 \textbf{ or } A_c g_c + \gamma n_c - \alpha = 0$$

The vertexes of the cube readily solve these three equations. Furthermore we obtain the completely mixed equilibrium N^* given by:

$$\begin{aligned} n_c^* &= \frac{A_o\alpha + A_c D_o}{A_o\gamma} \\ g_c^* &= -\frac{D_o}{A_o} \\ c_c^* &= -\frac{A_o D_g\gamma + A_o E_g\alpha + A_c E_g D_o}{A_o B_g\gamma} \end{aligned} \tag{2.26}$$

As we said above if the parameters of the model are such that the inequalities $0 < \ n_c^*, \ g_c^*, \ c_c^*, < 1$ hold, then N^* is a steady state for the dynamical system that is completely mixed and so it is at the same time a Nash equilibrium for the game. Of course, the parameters may be such that this equilibrium is outside of the unit cube, in which case it makes no sense in our context and we will say that the completely mixed equilibrium does not exist.

The other solutions of the three previously equations correspond to situations where not every population of the game mixes between the two pure strategies. These solutions lie on faces of the unit cube. As in the case of the completely mixed Nash equilibrium we have that these face equilibria might be outside the unit cube in which case they do not make sense to our analysis and we will say that the face equilibria do not exist. These face equilibria are given by the following expressions:

$$(n_c^*, g_c^*, c_c^*) = \left(-\frac{D_g}{E_g}, -\frac{D_o}{A_o}, 0\right)$$

$$(n_c^*, g_c^*, c_c^*) = \left(-\frac{D_g + B_g}{E_g}, -\frac{D_o}{A_o}, 1\right)$$

$$(n_c^*, g_c^*, c_c^*) = \left(0, \frac{\alpha}{A_c}, -\frac{D_g}{B_g}\right)$$

$$(n_c^*, g_c^*, c_c^*) = \left(1, -\frac{\gamma - \alpha}{A_c}, -\frac{E_g + D_g}{B_g}\right)$$

Let us now consider the stability of the pure equilibria as well as the stability and existence of the mixed equilibria.

Stability of the pure equilibria

Using Hartman-Grobman's theorem it is possible to obtain some conclusions about the behaviour of the dynamical system (2.25) in neighbourhoods of the stationary states corresponding to the vertices of the cube.

We begin by obtaining the eigenvalues of the linearisations corresponding to each one of the pure equilibria.

$$(n_c^1, g_c^1, c_c^1) = (0,0,0),\ \lambda_1 = D_g, \lambda_{2,3} = \pm\sqrt{-\alpha D_o}$$

$$(n_c^2, g_c^2, c_c^2) = (0,1,0),\ \lambda_1 = -D_g, \lambda_{2,3} = \pm\sqrt{(A_o + D_o)(A_c - \alpha)}$$

$$(n_c^3, g_c^3, c_c^3) = (1,0,0),\ \lambda_1 = E_g + D_g, \lambda_{2,3} = \pm\sqrt{-D_o(\gamma - \alpha)}$$

$$(n_c^4, g_c^4, c_c^4) = (1,1,0),\ \lambda_1 = -(E_g + D_g), \lambda_{2,3} = \pm\sqrt{-(A_o + D_o)(A_c + \gamma - \alpha)}$$

$$(n_c^5, g_c^5, c_c^5) = (0,0,1),\ \lambda_1 = B_g + D_g, \lambda_{2,3} = \pm\sqrt{\alpha Do}$$

$$(n_c^6, g_c^6, c_c^6) = (0,1,1),\ \lambda_1 = -(B_g + D_g), \lambda_{2,3} = \pm\sqrt{-(A_o + D_o)(A_c - \alpha)}$$

$$(n_c^7, g_c^7, c_c^7) = (1,0,1),\ \lambda_1 = E_g + B_g + D_g, \lambda_{2,3} = \pm\sqrt{D_o(\gamma - \alpha)}$$

$$(n_c^8, g_c^8, c_c^8) = (1,1,1),\ \lambda_1 = -(E_g + B_g + D_g), \lambda_{2,3} = \pm\sqrt{(A_o + D_o)(A_c + \gamma - \alpha)}$$

Hence we have that the pure equilibria are always saddle points, which occurs when the radicand in the square root is positive, or have a pair of pure imaginary numbers, which occurs when the radicand in the square root is negative, in which case Hartman-Grobman's theorem is inconclusive.

The quantities that appear in the eigenvalues above are deeply related with the behaviour of the dynamics along the three edges that are incident on that vertex as we will see later.

Stability and existence of the completely mixed equilibrium

Let us now study the stability and existence of the completely mixed equilibrium of the system.

The mixed equilibrium exists if and only if $(n^*, g^*, c^*) \in (0,1)^3$. Let us consider the following notation:

$$N^* = n_c^*(1 - n_c^*) > 0\ ;\ G^* = g_c^*(1 - g_c^*) > 0\ ;\ C^* = c_c^*(1 - c_c^*) > 0\ .$$

To analyse the stability we will consider the linear approximation of the system (2.25). The Jacobian matrix of this linearisation at the mixed equilibrium is equal to

$$\begin{bmatrix} 0 & N^* A_o & 0 \\ G^* E_g & 0 & G^* B_g \\ C^* \gamma & C^* A_c & 0 \end{bmatrix}$$

The eigenvalues are the roots of the characteristic polynomial:

$$\lambda^3 - \lambda(N^* G^* E_g A_o - G^* C^* B_g A_c) - N^* G^* C^* B_g A_o \gamma = 0\ .$$

Recall that Vieta's formulas provide some relations between the roots of a polynomial which may allow us to derive some knowledge of the sign of the solutions, which is precisely what we want to know the stability of equilibrium. Consider a generic polynomial

$$p(\lambda) = \lambda^3 + a_2 \lambda^2 + a_1 \lambda + a_0\ .$$

Vieta's formulas say that if x, y, z are roots of the polynomial then

$$x + y + z = -a_2$$

$$xy + xz + yz = a_1$$

$$xyz = -a_0$$

Vieta's formulas imply that the sum of the three roots of the above polynomial is zero. So there is a root with positive real part and a root with negative real part, or otherwise the three roots have zero real part, in which case the roots are a pair of pure imaginary numbers and a root that equals zero. We must consider two cases:

1. $B_g A_o \gamma = 0$ then one of the solutions is zero. The other two are either two pure imaginary numbers that are conjugate, or two symmetric real numbers. In this case the point is not hyperbolic and Hartman-Grobman's is not applicable.

2. $B_g A_o \gamma \neq 0$. Then zero is not a solution. So the polynomial has a real solution with non-zero part, and so must have another with non-zero real part with opposed sign, and so the mixed equilibrium is a saddle. The roots may have non-zero imaginary part, in which case we have a spiralling saddle.

When $B_g A_o \gamma = 0$ we are in a degenerate case in which either $A_o = 0$ or $B_g = 0$ (since we assumed $\gamma \neq 0$). It can be readily seen from the expressions (2.26) that we computed for the mixed equilibrium that when $A_o = 0$ or $B_g = 0$ the equilibrium does not exist since it is outside the unit cube.

So we conclude that when the completely mixed equilibrium exists, i.e. it is in the interior of the unit cube, then it is a saddle and in some cases it may be a spiralling saddle.

We finish by observing that when $N^*G^*C^* = 0$ then at least one of these is zero which means that the mixed equilibrium is on the boundary of the cube, which may be a face, and edge, or even in a vertex. When $N^*G^*C^* = 0$, then zero is again a solution of the characteristic polynomial, so that Hartman-Grobman's theorem is not applicable.

Stability and existence of the face mixed equilibria

Let us now study the stability and existence of the mixed equilibria in the faces of the unit cube. We will make use the same notation as before for each one of the four face equilibria.

$$N^* = n_c^*(1-n_c^*) > 0 \; ; \; G^* = g_c^*(1-g_c^*) > 0 \; ; \; C^* = c_c^*(1-c_c^*) > 0 \; .$$

For each equilibria in the faces one of this quantities is zero.

We begin by observing that there are no mixed equilibria in the faces corresponding to $g_c = 0$ and $g_c = 1$.

We consider each one of the four remaining faces separately:

1. Consider the face such that $n_c = 0$. In this face we have that $N^* = 0$. The Jacobian matrix is

$$J_{n_c=0} = \begin{bmatrix} A_oG^* + D_o & 0 & 0 \\ G^*E_g & 0 & G^*B_g \\ C^*\gamma & C^*A_c & 0 \end{bmatrix}$$

The characteristic polynomial is

$$(A_oG^* + D_o - \lambda)(\lambda^2 - (G^*C^*A_cB_g)$$

and the eigenvalues are:

$$\lambda = A_oG^* + D_o, \text{ and } \lambda = \pm\sqrt{G^*C^*A_cB_g)} \; .$$

2. Consider the face such that $n_c = 1$. In this face we have that $N^* = 0$. The Jacobian matrix is

$$J_{n_c=1} = \begin{bmatrix} -(A_oG^* + D_o) & 0 & 0 \\ G^*E_g & 0 & G^*B_g \\ C^*\gamma & C^*A_c & 0 \end{bmatrix}$$

The characteristic polynomial is

$$(-(A_oG^* + D_o) - \lambda)(\lambda^2 - (G^*C^*A_cB_g)$$

and the eigenvalues are:

$$\lambda = -(A_oG^* + D_o), \text{ and } \lambda = \pm\sqrt{G^*C^*A_cB_g}) \;.$$

3. Consider the face such that $c_c = 0$. In this face we have that $C^* = 0$. The Jacobian matrix is

$$J_{c_c=0} = \begin{bmatrix} 0 & N^*A_o & 0 \\ G^*E_g & 0 & G^*B_g \\ 0 & 0 & A_cG^* + \gamma N^* - \alpha \end{bmatrix}$$

The characteristic polynomial is

$$(A_cG^* + \gamma N^* - \alpha - \lambda)(\lambda^2 - N^*G^*E_gA_o)$$

and the eigenvalues are:

$$\lambda = A_cG^* + \gamma N^* - \alpha, \text{ and } \lambda = \pm\sqrt{N^*G^*E_gA_o} \;.$$

4. Consider the face such that $c_c = 1$. In this face we have that $C^* = 0$. The Jacobian matrix is

$$J_{c_c=1} = \begin{bmatrix} 0 & N^*A_o & 0 \\ G^*E_g & 0 & G^*B_g \\ 0 & 0 & -(A_cG^* + \gamma N^* - \alpha) \end{bmatrix}$$

The characteristic polynomial is

$$(-(A_cG^* + \gamma N^* - \alpha) - \lambda)(\lambda^2 - N^*G^*E_gA_o)$$

and the eigenvalues are:

$$\lambda = -(A_cG^* + \gamma N^* - \alpha), \text{ and } \lambda = \pm\sqrt{N^*G^*E_gA_o} \;.$$

We conclude that every face mixed equilibria is either a saddle point or else has a pair of conjugate pure imaginary numbers. In this second case Hartman-Grobman's theorem is not applicable because the point is not hyperbolic.

2.2.8 The dynamics on the edges of the cube

Each edge of the unit cube has two fixed coordinates and one free coordinate. The dynamics along an edge is obtained by looking at the relevant equation of the dynamical system (2.25), i.e. the equation of the free coordinate and substituting the values of the fixed coordinates. After substitution, the value of the free coordinate will increase if the obtained quantity is positive and it will decrease if the quantity is negative. A special feature of the replicator dynamics is that this value is constant along the edge, so what matters is the sign of this quantity. These quantities that tell the behaviour of the vector field that defines the differential equation along the edges of the unit cube are precisely the quantities that appear in the socio-political characterization of government, officials and citizens that we have done in section 2.2.6, which, we recall, also characterize the best-responses against each pure strategy profile. After doing this study of the vector field in each edge of the unit cube we obtain the information that is summarized in the next figure (2.8).

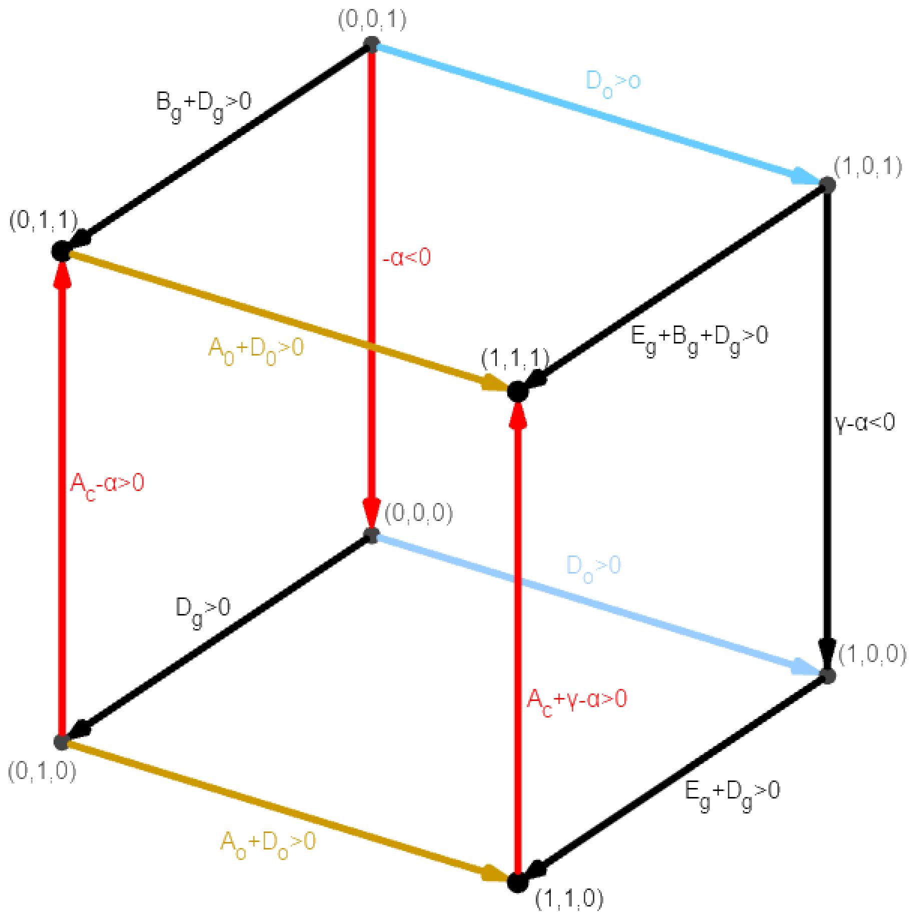

Figure 2.8: The behaviour of the dynamics along the edges of the unit cube.

The edges in red correspond to the three mandatory inequalities that we have identified in section 2.2.6. These three edges always point in the directions that are shown in the figure. The edges with direction $(0, 1, 0)$ that we have painted in blue and yellow correspond to the choices of the officials, and we observe that the values in such edges do not depend on the height, i.e., they do not depend on the citizens distribution. The other edges that are painted in black correspond to the other inequalities identified in section 2.2.6. The edges in blue, yellow and black are free in the sense that according to the model parameters they may go in either direction, although as we observed the two blue edges and the two yellow edges need to agree.

From figure 2.8 we also see that the quantities that appeared in the eigenvalues of the linearisation of each pure equilibrium above are deeply related with the behaviour of the dynamics along the three edges that are incident on that pure equilibrium.

As we have seen above the vertexes of the cube are either saddle points or have a pair of pure imaginary numbers. It is possible for a given vertex to be an attractor along its three incident edges without contradicting the fact that it is a saddle. Indeed, for a vertex to be an attractor it is necessary that the trajectories along the three incident edges approach the vertex but this is not sufficient, since for it to be an attractor it should be so in a small sphere in $\mathbb{R}^3$ around the point. However, in our case, only the intersection of this sphere with the unit cube matters. So using a continuity argument, if there are no other equilibria in the edges and if the edges are attracted towards the vertex, then it can be shown that the vertex is asymptotically stable for initial conditions in a neighbourhood that are in the unit cube, that is, for initial conditions that matter to our analysis. So we have that the dynamical behaviour in the incident edges is also sufficient to characterize the local stability of the vertex. We prove this in the appendix 2.B. The reason for this is that the edges incident on a vertex are not eigenvectors of the linearisation, so those eigenvectors that represent the unstable directions of the vertex "exit" the cube, and so they will not be relevant to our analysis.

The simple eigenvalue associated to a pure equilibria, i.e., the one that is not a square root always has an eigenvector which is $(0, 1, 0)$. This eigenvalue has modulus equal to the quantity that corresponds to the edge with that direction that is incident on the vertex.

The other two eigenvalues that are a pair of square roots are related with the values along edges with directions $(1, 0, 0)$ and $(0, 0, 1)$, although these directions are not eigenvectors. The expression in the radicand is precisely the product of the quantities relative to these two edges with a signal. The eigenvectors associated to these eigenvalues are of the form $(1, 0, \pm z)$. In the case where the three incident edges are attracted to the vertex, the eigenvector that is associated to the positive eigenvalue, hence unstable, has a direction that "leaves" the unit cube as expected.

It is also worth noting that a stable equilibrium, since all its incident edges are attracted towards it immediately implies the instability of all its neighbours (adjacent) vertexes. This, together with the fact that, as we discussed, not all combinations of dynamics on the edges are possible, implies that there are, at most, two pure attractors.

We also have the following assertions whose proofs can be seen in [103]:

Theorem 2.1. *For the replicator dynamics we have the following:*

1. *A dominated pure strategy are asymptotically eliminated. In other words, pure strategies that are dominated will vanish in the limit.*
2. *A pure strategies that is iteratively eliminated will vanish in the limit.*
3. *A Nash equilibrium is Lyapunov stable.*
4. *A strategy is asymptotically stable if and only if it is a strict Nash equilibrium.*

The first assertion in this theorem implies that when there exists a dominated strategy then the trajectories approach a face of the cube, so that the limit, if it exists, would lie on that face. This implies that there cannot be a completely mixed equilibrium when there is a dominated strategy, since the trajectory starting at that point would always stay at the same point so that the dominated strategy would not vanish. However, this does no imply that the limit exists in that face since the interior trajectory might approach a cycle in the face. From the second assertion we have that if one can find dominated strategies, successively, then the remaining pure vertex would be a global attractor. From the third assertion we have that a mixed strategy cannot be an attractor. So only the vertexes of the cube can be asymptotically stable. When a vertex is Nash but not strict Nash, then it is Lyapunov stable but not an attractor, in which case we conclude that there must be an edge made of equilibria, i.e., one where the respective quantity showed in figure 2.8 vanishes.

So we now have that the stability of the vertexes may be characterized by the behaviour of the incident edges. Hence, each one of the possible cases that we mentioned corresponding to a specific behaviour along the edges, and hence corresponding to a specific socio-political characterization allows us to know the stability of the pure equilibria. We may also know the nature of the other equilibria and characterize the evolution of corruption. We will present some examples of these cases in the following sections.

2.2.9 Pure strategy attractors

We now discuss the three inequalities that need to verify to have asymptotic stability for each of the pure strategies of the game. This can be readily done by looking at figure 2.8.

1. $(n_c^1, g_c^1, c_c^1) = (0, 0, 0)$. This is the 'good' equilibrium where general levels of corruption are low and citizens are non-compliant. Inequality $\alpha < 0$ is always verified so that a non-corrupt government and officials will make more citizens less compliant with corruption. Furthermore, one needs $D_o < 0$, which is appropriate fines to punish corrupt officials and $D_g < 0$ which means efficiency, or low costs in fighting corruption.

2. $(n_c^2, g_c^2, c_c^2) = (0, 1, 0)$. This is the equilibrium where only government is corrupt. This equilibrium is never stable since the citizen compliance will increase. This may be seen as if the citizens have no other possibility than being compliant, which may even be inadvertently or unintentionally since it will be a more profitable strategy because of the welfare gain P citizens have plus the dis-utility β he has because the government is corrupt.

3. $(n_c^3, g_c^3, c_c^3) = (1, 0, 0)$. This is the equilibrium where only the officials are corrupt. This equilibrium is stable if the fines for corrupt officials are not very high and since there are a lot of corrupt officials because of low fines, the government prefers to act honestly since it is efficient and so makes little effort to fight corruption. Furthermore, for this equilibrium to be stable citizens need to value the dis-utility of finding corrupt officials less than the valuation of their satisfaction with an honest government.

4. $(n_c^4, g_c^4, c_c^4) = (1, 1, 0)$. This is the equilibrium where both government and officials are corrupt. This equilibrium is unstable since in this case, again citizen will be, even if inadvertently or unintentionally, compliant, since it will be a more profitable strategy because because of the welfare gain P citizens have plus the dis-utilities β and
gamma he has because the government is corrupt and officials are corrupt respectively.

5. $(n_c^5, g_c^5, c_c^5) = (0, 0, 1)$. This is the equilibrium where only citizens are compliant with corruption and officials and government are honest. This equilibrium is always unstable since citizens will always change strategy because of the fact that the government is honest, which means their (even if implicit) valuation $\alpha > 0$ will make them want to be non-compliant.

6. $(n_c^6, g_c^6, c_c^6) = (0, 1, 1)$. This is the equilibrium where government is corrupt and citizens compliant. The stability of this equilibrium depends on the re-election power of the government and on the penalties imposed on corrupt officials, which would need to be high.

7. $(n_c^7, g_c^7, c_c^7) = (1, 0, 1)$. This is the equilibrium where only officials are corrupt. For this equilibrium to be stable one would need to have low costs for the government to fight corruption, but still unable to lower officials corruption because the fines practised are low. Furthermore one would need citizens to have lower dis-utility caused by facing corrupt officials than the valuation they give to having a honest government.

8. $(n_c^8, g_c^8, c_c^8) = (1, 1, 1)$. This is the equilibrium where all agents are corrupt and citizens are also compliant. It can be stable when the government punishes honest behaviour more than corrupt behaviour by the officials, in which case officials prefer to be corrupt and when government has high inefficiency in fighting corruption. In this situation the citizens will also be compliant.

2.2.10 Some examples

We now provide some graphical examples of the dynamics in some cases.

Let us consider the case where the following inequalities are verified: the fines from a non-corrupt government are appropriate, meaning that for a non-corrupt government, officials will prefer to be honest. This is $D_o < 0$. Corrupt government punishments are higher for corrupt officials, i.e. $D_o + A_o < 0$. The dissatisfaction γ is lower than α so that $\gamma - \alpha < 0$. Furthermore, the re-election power of a non-corrupt government is higher, meaning that $D_g < 0$ and $B_g + D_< 0$ and the efficiency of a non-corrupt government is high, meaning that $E_g + D_g < 0$ and $E_g + B_g + D_g < 0$.

In this case the corrupt strategy is strictly dominated for the government, and after this strategy is eliminated, the strategy to be corrupt is also eliminated for the officials because of the high fines. After this, citizens will be non-compliant, so from theorem 2.1 we conclude that the 'good' equilibrium $(n_c^1, g_c^1, c_c^1) = (0, 0, 0)$ is the unique global attractor of the dynamics. We plot some trajectories of the system in this case in figure 2.9.

Let us now consider the case where the following inequalities are verified: the fines from a non-corrupt government are appropriate, meaning that for a non-corrupt government, officials will prefer to be honest. This is $D_o < 0$. Corrupt government punishments are higher for honest officials, i.e. $D_o + A_o > 0$. The dissatisfaction γ is lower than α so that $\gamma - \alpha < 0$. Furthermore, the re-election power of a non-corrupt government is higher, meaning that $D_g < 0$ and $B_g + D_< 0$, but now the efficiency of a non-corrupt government is low, meaning that $E_g + D_g > 0$ and $E_g + B_g + D_g > 0$.

Since the respective incident edges are attractor, we conclude from the theorem in appendix 2.B that there are two stable pure attractors: the "good equilibrium" $(n_c^1, g_c^1, c_c^1) = (0, 0, 0)$ and the bad equilibrium $(n_c^8, g_c^8, c_c^8) = (1, 1, 1)$. Dynamically, this also implies the existence of a completely mixed equilibrium in the interior of the cube, which is, as we discussed previously, a saddle equilibrium point. We plot some trajectories of the system in this case in figure 2.10.

Let us now consider a case inspired from the two-dimensional case discussed in chapter 2.1 where we obtained cycles of corruption. Let us then consider that the following inequalities hold: the fines from a non-corrupt government are inappropriate, meaning that for a non-corrupt government, officials will prefer to be corrupt. This is $D_o > 0$. Corrupt government punishments are higher for corrupt officials, i.e. $D_o + A_o < 0$. The dissatisfaction γ is lower than α so that $\gamma - \alpha < 0$. But consider now that the re-election power of a non-corrupt government is higher, meaning that $D_g < 0$ and $B_g + D_{<}0$, but the efficiency of a non-corrupt government is low, meaning that $E_g + D_g > 0$ and $E_g + B_g + D_g > 0$. This makes he dynamics turn around the cube at the edges at the top of the cube and at the bottom of the cube. The rationale in the two cases is similar: officials have an incentive to be corrupt because of low fines, to which the government responds with increasing corruption because of its inefficiency. The officials corruption decreases because punishments for corrupt officials are high, and finally, high re-election power forces the government to be honest. This gives rise to two-dimensional, clockwise cycles in the top and bottom faces of the unit cube. Indeed, in this case, the interior trajectories approach these two cycles as shown in figure 2.11. Dynamically we also have that there must be a completely mixed equilibrium in the interior of the unit cube, which is this case is also a spiralling saddle. The long run behaviour of trajectories inside the unit cube is shown in figure 2.11.

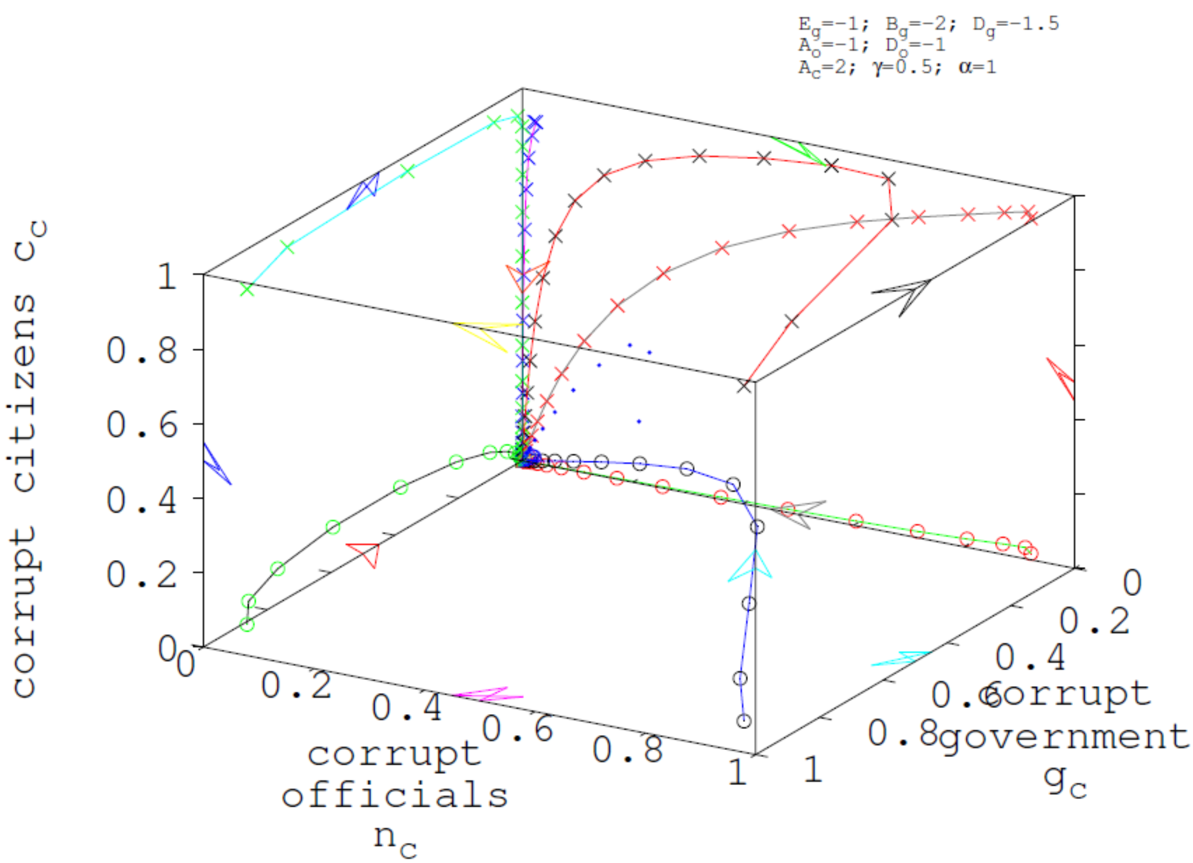

Figure 2.9: Some trajectories of the system when the “good” equilibrium is the unique global attractor.

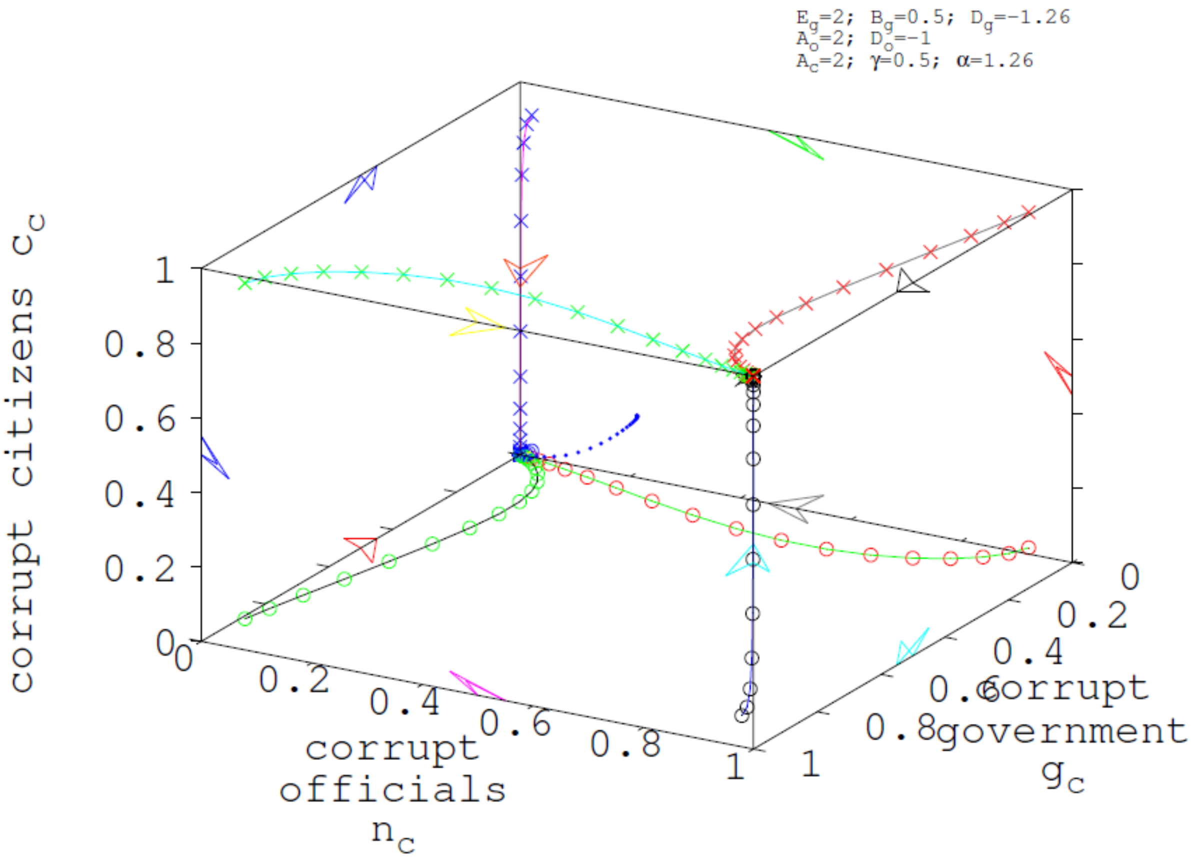

Figure 2.10: Some trajectories of the system when both the "good" equilibrium and the 'bad" equilibrium are attractors.

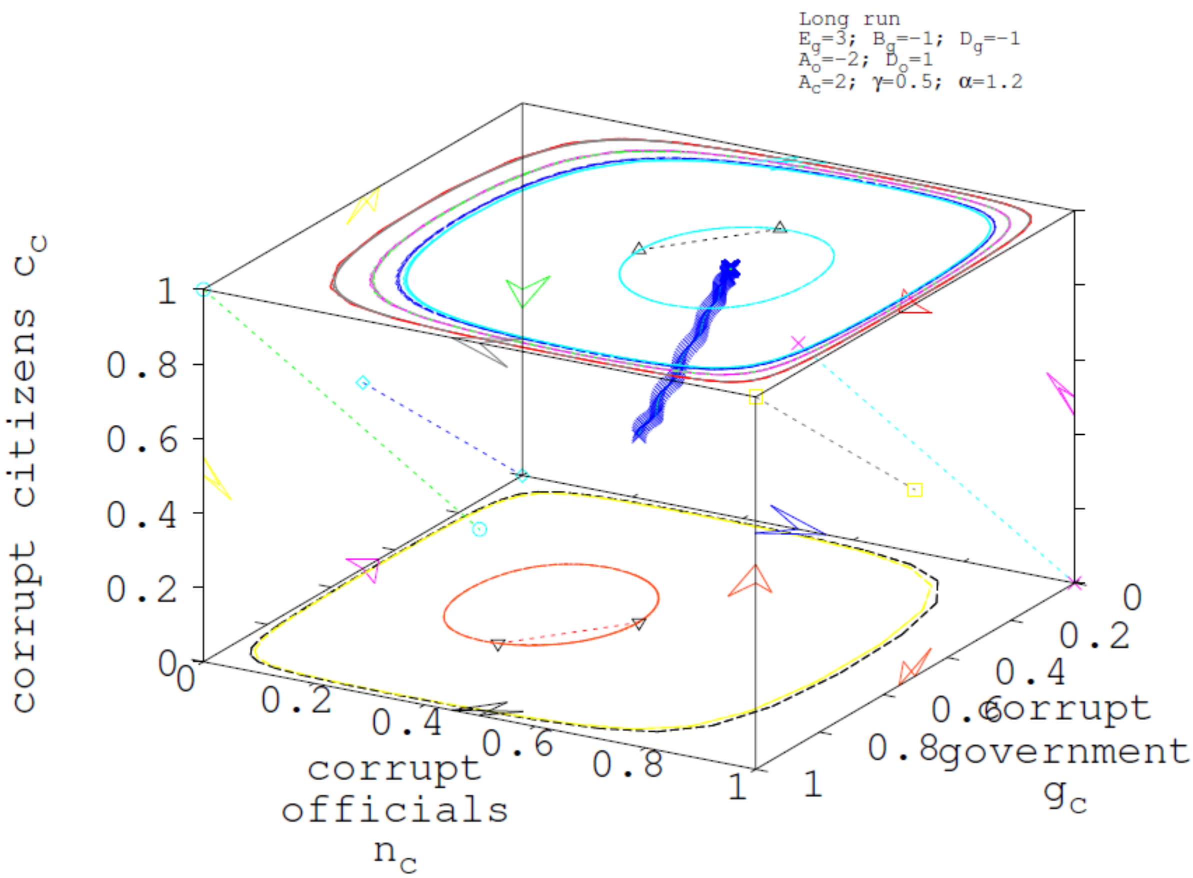

Figure 2.11: The long run behaviour of the trajectories towards cyles of governmental and officials corruption.

2.3 Conclusions

As it is well known, many politicians and ruling elites across the world and from the whole of the political spectrum are involved in processes of corruption. Is it possible to deter this process? To give an answer to this question is the main concern of this paper. In order to do this, we considered an evolutionary model, where the political agents (considered as players of a game in normal-form) compare their respective expected payoffs, and they choose their strategies according to their average performances and the most profitable behaviours end by prevailing. Hence the replicator dynamics appear as a natural mathematical tool to describe the evolution of corruption inside a society.

We first obtained that corruption corrupts, so that corruption is a self-reinforcing mechanism (see Proposition (2.1). When the degree of intolerance is relatively low and the political elite in the government has good prospects of being re-elected and a large interest in gain immediate benefits, the country can be in a corruption social trap, i.e., a self-reinforcing mechanism where corruption generates more corruption.

We constructed an evolutionary version of the game by means of the replicator dynamics. The dynamics has five equilibria, four of them corresponding to pure strategies: i) where both government and officials are corrupt; ii) where both are non-corrupt; iii) where only government is corrupt and officials are not; iv) where government is honest and officials are corrupt. The other equilibrium is a mixed strategy, that is always a Nash equilibrium of the game, where government and officials choose to be corrupt with a given probability. We then did the stability analysis of these equilibria according to the characteristics of government and officials. These characteristics are: a) the fines practised by the non-corrupt government on corrupt officials, that may be appropriate or not; b) the re-election power of government; c) the relative welfare that a corrupt government takes from honest and dishonest officials; d) the efficiency of government in fighting corruption.

We obtained three outcomes: i) the mixed interior equilibrium is a saddle point; ii) the mixed interior equilibrium is a focus point around which solutions of the differential equation oscillate, corresponding to cycles of corruption over time; iii) there is no mixed interior equilibrium and there is a unique pure Nash equilibrium.

When there is a mixed equilibrium that is a saddle point, it is such that its stable manifold separates two different outcomes which correspond to pure Nash equilibria that are stable in the dynamics. In one case, it separates initial conditions that will lead to a general increase in corruption levels and those that will lead to a general decrease in corruption. In the other case, it separates between conditions that will lead to an increase in government corruption and a decrease of corruption by officials and those that will lead to a decrease in government corruption and an increase in officials' corruption. These two asymmetrical situations where corruption increases in one of the populations of the game and decreases in the other, and that correspond to equilibria of the game may be interpreted as situations usually observed in dictatorial regimes. The first one generally corresponds to a political elite that benefits from corruption that is practically confined to the elite itself, together with an efficient censorship system and control of its employees. This corresponds in our model to the punishment of officials for being corrupt and breaking the rules. The other generally corresponds to a situation where there is a *de facto* government of the officials sustained by a power above the law, with a complacent *de jure* government.

The second case occurs when the mixed interior is a focus point of periodic orbits. Corruption cycles arise when government is cost efficient and practices high fines but re-election power is low and penalties of a corrupt government are higher for honest officials. They also arise when re-election power is high but fines and efficiency are low, and penalties are higher for corrupt officials. When corruption cycles occur, the focus point is the only Nash equilibrium since there are no pure Nash equilibria in those cases.

In the third case there is no mixed interior equilibrium. In this case, there is a unique pure Nash equilibrium for the game, and it is stable in the dynamics. The characteristics of the population determine which of the pure strategies of the game is a Nash equilibrium (and stable).

We have that not all situations are compatible with each other. When the non-corruption equilibrium or the corruption equilibrium are stable then the asymmetrical equilibria described above can not be stable, and vice-versa. More precisely, when there are appropriate fines practised by the government that has high re-election power, then only the non-corruption equilibrium and the corruption equilibrium may be stable. If the costs to catch corrupt officials are high, then both are stable and the mixed equilibrium separates initial conditions leading to them. So for some initial conditions, the society may be in a social trap that increases the general levels of corruption and maintains such levels illustrating once again that corruption may be a self-reinforcing mechanism. On the other hand, when re-election power is low and fines are inappropriate, then the non-corruption equilibrium is not stable. Furthermore, in this case, when there are high penalties for corrupt officials by a corrupt government (see remark 2.3), the equilibrium where government is corrupt and officials honest is stable. When penalties are higher for corrupt officials, then either officials corruption equilibrium is stable or the corruption equilibrium is stable, depending on whether there is cost efficiency or inefficiency, respectively. The situation of the corruption equilibrium being stable illustrates a society with high levels of corruption at both the officials' and government's level.

We also constructed a variation of the first game as well as a dynamics for the evolution of corruption where, apart from the evolution governmental *grand* corruption and corruption at the scale of officials, we also track the evolution of citizen compliance towards corruption. We state the characterization of social and political features that have an influence on the dynamics of such game, now also including parameters relative to citizen's relation to corrupt governments and officials. We explain some feature of this new dynamics, particularly the behaviour that occurs at the edges of the unit cube and represents the best responses of the players of the game and its relation to the listed social and political features of the game. We explain the stability of the vertexes, or pure strategy profiles of the game and we provide and interpret some interesting graphical examples of the dynamics.

For the case where the country may be in a social trap with increasing levels of corruption, an external event may be a necessary condition for the country to leave this trap. This social trap may escape any self-monitoring mechanism, and then there is no way to control the controller. However, if the ruling elite has some interest in the re-election this self-reinforcing mechanism can be weakened or broken by a high enough degree of intolerance to corruption by the citizens. The degree of intolerance to corruption plays an important role to make the government fulfil the role that society has assigned it, even when some of its members are attracted by the individual benefits that corruption offers. Even in situations where corruption tends to expand, if the intolerance index has an abrupt change, the regressive process can be reversed[6]. Moreover, we also have seen that cyclical processes may appear in which periods of diminishing corruption are followed by periods of increasing corruption. However, if the corruption index is not high enough, corruption can be self-sustaining by means of cycles of corruption with alternation in power by

[6]Recent events in South Korea, where citizens reacted to the corrupt practices of Prime Minister Park Geun-hye suggest that, if the Index of Intolerance of citizens to Corruption is high enough, it is possible to exert political pressure that can maybe result in stopping the growth of corruption. See http://www.abc.net.au/news/2016-11-15/south-korea-park-geun-hye-hopes-political-crisis-be-contained/8024978.

the elites, or situations where the long-term behaviour approaches equilibria where corruption exists, whether by the officials, the government or both. Furthermore, as we have said when we interpreted those equilibria, they can correspond to situations of democratic deficit due to corrupted institutions and their inability to fulfil their roles, and more extreme situations as dictatorships.

Thus, how to maintain a high index of intolerance to corruption is a fundamental question. Responding correctly to this question may be the key to avoid increasing corruption and cyclical corruption processes. The index of intolerance to corruption depends on the citizen's perception of corruption and decreases when the number of corrupt acts decreases. Since this index reflects a perception of the citizens it can be exploited by a ruling elite to gain personal benefits from corrupt acts. Furthermore, it can depend on other political and economical variables not included in the index of intolerance we introduced, and not addressed in this paper. Another issue is the media coverage and pressure about corruption and the very relevant aspect of how corruption perceived in society. We plan to address some of these in future works.

Confirmation of results by empirical testing and statistical methods is surely a very important issue. However, this was not our goal in this work. We focused instead on a game theoretical model of conflict between several levels of controllers, with the objective of obtaining some insight and a game theoretic and evolutionary reasoning for the question posed in the title of this work, and to capture some essential features of this conflict between over the question of controlling honest vs. corrupt behaviours in employees of the central government (that may be lured into dishonest behaviour). There have been some previous approaches to corruption via game theory, as for instance in [6], focused on the problem of environment protection. As far as we know, our game theoretic approach to corruption through the problem of who will be controlling the controller is new. It should be noted that this work points towards a game theoretical and dynamical approach and methodology that drew some fundamental features of the conflict between agents, so that it might require some changes to the model to proceed to some kind of empirical testing. We will try to address this in future works. It should also be noted that there are statistical studies and case studies in specific countries yielding results that go in different directions. The observance or not of some effects of corruption depend greatly on the underlying social, political and economic structures of the countries. For example, see the introduction of [93] and the references therein about corruption and economic growth. In [72] the author has shown the negative effects of corruption on investment and economic growth, while in [36] the authors show the positive correlation of corruption and direct foreign investment. Some studies conclude that corruption is a cyclical phenomena, associated with alternation in power of the ruling elites (see [93] for the case of this alternation and its relation to corruption in Mexico). Others, in which citizen participation reduces general corruption levels (see [84], where South-Eastern Asian countries are analysed). Thus our model gives a game theoretical background to these different situations.

We can say that we have given at least a partial answer to the question that motivated this paper and the answer is quasi-optimistic, because it seems possible to control the controller. The citizens are the main protagonists in this control process, although, certainly, for citizen participation to be effective, a high index of intolerance to corruption is required.

Summarizing, in this paper we gave a first step to recognize the possibility of fight with success against the corruption and in the knowledge of possible trajectories of the evolution of a corruption process. However, as we said in section 2.1.6 to obtain the analytical solution of the dynamical system is generically not possible, but we can obtain some approximation using numerical methods and do stability and qualitative analysis of the trajectories of the system.

The model can be improved by further studying on the characteristics ofithout the parameters considered. For instance, the degree of intolerance of citizens, that may be modelled as depending on other political, social and economic variables. It will be necessary to consider also cross terms and the corresponding non-linear effects, and to study other types of dynamics, for instance, the role of the imitative behaviour (see [3]). Other accesses to information, like the press and modern media, that can exert relevant influence in the the index of intolerance should also be incorporated in future developments of the model.

2.A Appendix: Citizens Choice

As stated above, citizens are not decision makers in the game we considered in section 2.1, but their choice has a political influence in the the outcome of the game by means of the probabilities of re-election of a corrupt and a non-corrupt government, and so have decisive influence in the decisions of government and officials.

The ruler elite must undergo an electoral process on which it will seek to influence in various ways. We show a possible way to endogenize the choice of citizens, i.e., the probabilities of re-election for a corrupt and a non-corrupt government using the utility functions of the citizens.

For instance in case of a corrupt government buying votes. Note that, even in this case, not necessarily all citizens will receive some payment from a corrupt government. The ruling elite chooses who and how much to pay. To simplify we will consider that some citizens will receive an amount equal to $P > 0$ monetary units and others receive 0.

We have that, fixed $p \in \{0, P\}$ each citizen prefers an honest government to a non-corrupt one, i.e.

$$\text{if } \bar{n}_c < n'_c \text{ then } u_i(\bar{n}_c, p) > u_i(n'_c, p) \ .$$

Citizens measure the level of corruption by the percentage of corrupt officials acting in the public administration. This is a natural assumption, because for a citizen, corruption means having to pay for a service that must be provided free of charge.

Let $\mathcal{K}$ be the subset of citizens receiving the amount P. The subset $\mathcal{K}$ and the value of P are fixed by the ruler elite and depend on the estimates that this elite does.

- Not necessarily all citizen will receive some payment made by a corrupt government. The ruling elite chooses who to pay. Some citizens will receive an amount equal to $P > 0$ monetary units, and others will receive 0.

- The ruler elite of a corrupt government will offer an amount $P = P(n_c)$ to each citizen in a subgroup $\mathcal{K} \subset \mathcal{H}$ in the sense that $u_i(n_c, P) > u_i(n'_c, 0)$ for all $i \in \mathcal{K}$ and for all $n'_c < n_c$. In other words, a citizen in $\mathcal{K}$ prefers to receive the amount of money P and bear the amount n_c of corrupt officers more than a lower amount n'_c of corrupt officials and not receive money from the government.

- The value of $P(n_c)$ may be assumed to increase with n_c. We also assume that the utility function $u_i(n_c, P(n_c))$ is decreasing in n_c. In other words, the amount P is enough to convince the citizen to vote when the number of corrupt officials is higher since it yields a higher utility than lower corruption without any payment, but still, in this situation, the citizen places greater importance in the amount of corrupt officials, and this is why $u_i(n_c, P(n_c))$ is decreasing.

- Let $\bar{P}$ be the maximum amount that a corrupt government can offer to citizens in exchange for their votes. Note that for all $i \in \mathcal{K}$ we have that $u_i(1, 0) \leq u_i(n_c, P(n_c)) \leq u_i(0, \bar{P})$ and for all $i \notin \mathcal{K}$ we have that $u_i(1, 0) \leq u_i(n_c, 0) \leq u_i(0, 0)$.

So the choice of citizens is randomized because the information they have is not complete. We consider that the probability of re-electing the government or not depends directly on their relationship with the officials, and this probability decreases as the number of corrupt officers increases, so that this may be distorted by personal experience. In this case, even high values of P, will not give the result sought by the ruling elite and the citizens will vote for not re-electing a corrupt government.

Let us introduce the following boundary values:

$$m_K = \min_{i\in\mathcal{K}} u_i(1,0) \text{ and } M_K = \max_{i\in\mathcal{K}} u_i(0,\bar{P})$$
$$m_{nK} = \min_{i\notin\mathcal{K}} u_i(1,0) \text{ and } M_{nK} = \max_{i\notin\mathcal{K}} u_i(0,0)\ .$$

We consider that the probability that a citizen votes for the re-election of a government is proportional to the utility function. Then, for a fixed n_c and $P(n_c)$ we can make the following:

- The probability that a citizen votes for the re-election of a corrupt government is given by:

$$\tfrac{m_K}{M_K} \le q^i_{G_c} = \tfrac{\alpha_i}{M_K} u_i(n_c, P(n_c))\ ,\ \forall i\in\mathcal{K} \text{ and}$$
$$\tfrac{m_{nK}}{M_{nK}} \le q^i_{G_c} = \tfrac{\alpha_i}{M_{nK}} u_i(n_c, 0)\ ,\ \forall i\notin\mathcal{K}$$

- The probability that a citizen votes for the re-election of a non-corrupt government is given by:

$$\tfrac{m}{M} \le q^i_{G_{nK}} = \tfrac{\alpha_i}{M} u_i(n_c, 0)$$
$$\text{where } m = \min_{i\in\mathcal{H}} u_i(1,0),\ M = \max_{i\in\mathcal{H}} u_i(0,0)\ .$$

And so

$$\begin{aligned} q_{G_c} &= \Pi_{i\in\mathcal{H}} q^i_{G_c} = \Pi_{i\in\mathcal{K}} \tfrac{\alpha_i}{M_K} u_i(n_c, P(n_c)) \Pi_{i\notin\mathcal{K}} \tfrac{\alpha_i}{M_K} u_i(n_c, 0) \\ q_{G_{nc}} &= \Pi_{i\in\mathcal{H}} q^i_{G_{nc}} = \Pi_{i\in\mathcal{H}} \tfrac{\alpha_i}{M} u_i(n_c, 0)\ . \end{aligned} \tag{2.27}$$

2.B Appendix: Dynamics on the edges of the cube

Consider a dynamics of the form:

$$\begin{cases} \dot{x}_1 = x_1(1-x_1)F_1(x_1,\ldots,x_n), \\ \ldots \\ \dot{x}_n = x_n(1-x_n)F_n(x_1,\ldots,x_n), \end{cases} \tag{2.28}$$

where for all $j \in \{1,\ldots,n\}$, $F_j(x_1,\ldots,x_n)$ are Lipschitz continuous functions.

We will use the following nomenclature:

- Equilibria of the form $(\bar{x}_1,\ldots,\bar{x}_n)$ where for all $i \in \{1,\ldots,n\}$, $\bar{x}_i \in \{0,1\}$ will be called pure equilibria. These points correspond to the vertexes of the $n-$cube $\mathcal{C} = [0,1]^n$.

- The segments $(\bar{x}_1,\ldots,x_i,\ldots,\bar{x}_n)$ where $0 \leq x_i \leq 1$ and $\bar{x}_j \in \{0,1\}$, $j \neq i$ are the edges of the $n-$cube. Observe that each vertex has n incident edges and that for each edge we have that all except one of its coordinates is free. We will call this the free coordinate of the edge. The other $n-1$ coordinates will be called the fixed coordinates of the edge. We also observe that we consider the vertexes to be part of the edges that are incident on it.

- The sets of points of the form $(x_1,\ldots,\bar{x}_i,\ldots,x_n)$ where for some i we have $\bar{x}_i \in \{0,1\}$, and the other coordinates are free is a face of the cube. Observe that a face of a n-*dimensional* cube is a $(n-1)$-*dimensional* cube.

Remark. *In many situations, the setting described above becomes simpler. For instance, in the replicator dynamics, the function F_i is the difference between the expected values of the two strategies available to population i. So the functions F_i do not depend on the distribution/strategy of population i. In other words, we have that $F_i(x_i,x_{-i}) = F_i(x_{-i})$. In this case the value of F_i is constant along each edge of the unit cube whose free coordinate is the i-th coordinate.*

We have that the unit cube $\mathcal{C}$ is invariant for the dynamics, i.e., if the initial condition lies in the unit cube $\mathcal{C}$ then the solution of the system will remain in the cube forever. Analogously, the edges of the unit cube are also invariant since if x_i equals 0 or 1 then $\dot{x}_i = 0$, so that x_i remains constant equal to 0 or 1. Similarly, the faces are also invariant by the same reasons.

According to the notation previously introduced we will sometimes write $F(x) = F_i(x_1,...,x_n) = F_i(x_i,x_{-i})$. The following theorem is straightforward.

Theorem. *Consider a pure equilibrium $\bar{x} = (\bar{x}_1,...,\bar{x}_n)$, i.e., a vertex of the cube. Assume that, as in remark 2.B that the functions F_i are constant along edges, i.e., $F_i(\bar{x}_i,\bar{x}_{-i}) = F_i(\bar{x}_{-i})$. If for some $i \in \{1,...,n\}$ such that $\bar{x}_i = 0$ we have that $F_i(\bar{x}_{-i}) > 0$ then $\bar{x}$ is unstable in the Lyapunov sense. Analogously, if for some $i \in \{1,...,n\}$ such that $\bar{x}_i = 1$ we have that $F_i(\bar{x}_{-i}) < 0$ then $\bar{x}$ is unstable in the Lyapunov sense.*

Proof. It is clear that if $\bar{x}_i = 0$ and $F_i(\bar{x}_{-i}) > 0$ then along the edge $(x_i,\bar{x}_{-i})$ where the free coordinate is x_i the trajectory 'goes away' from the pure equilibrium. So $\bar{x}_i$ cannot be stable. With a similar argument we prove that when $\bar{x}_i = 1$ and $F_i(\bar{x}_{-i}) < 0$ the vertex $\bar{x}$ is unstable. □

The previous theorem states that for a vertex to be stable is necessary that restricted to edges be an attractor. The characterization of whether along the edges a solution of the dynamical system converges or "go away" from a vertex depend on the functions sign of the functions F_i in a neighbourhood of this vertex.

Indeed, by a continuity argument we also have the following theorem, which is a converse to the previous result.

Theorem. *Assume there are no equilibria in the edges apart from the pure equilibria, i.e., apart from the vertexes of the cube. Consider a pure equilibrium point* $\bar{x} = (\bar{x}_i, \bar{x}_{-i})$. *Assume that*

1. *if* $\bar{x}_i = 0$ *then* $F_i(0, \bar{x}_{-i}) < 0$ *and that*
2. *if* $\bar{x}_i = 1$ *then* $F_i(1, \bar{x}_{-i}) > 0$.

Then the pure equilibrium $\bar{x}$ *is a local attractor relative to the unit cube, i.e., it is an attractor for initial conditions in a neighbourhood in the unit cube.*

Proof. Let $\bar{x} = (\bar{x}_i, \bar{x}_{-i})$ be a pure equilibrium for the system (2.28). Since for all $i \in \{1, ..., n\}$, F_i are continuous functions we have that:

- when $\bar{x}_i = 0$, we have that $F_i(0, \bar{x}_{-i}) < 0$, so there is a ball B_i with center at $\bar{x} = (0, \bar{x}_{-i})$ and radius $\epsilon_i > 0$ such that $F_i(x_i, \bar{x}_{-i}) < 0$ for all $x \in \mathcal{B}_i$;
- when $\bar{x}_i = 1$ we have that $F_i(1, \bar{x}_{-i}) > 0$ so there is a ball B_i with center at $\bar{x} = (1, \bar{x}_{-i})$ and radius $\epsilon_i > 0$ such that $F_i(x_i, \bar{x}_{-i}) > 0$ for all $x \in \mathcal{B}_i$.

Now consider the intersection of such balls for every $i \in \{1, \ldots, n\}$, i.e., $\mathcal{V} = \cap_{i=1}^{n} B_i$. This set is open so there exists a ball $\mathcal{B} \subset \mathcal{V}$ such that the sign of each F_i in $\mathcal{B}$ remains the same as in $\mathcal{B}_i$. So in the intersection of this ball with the unit cube, i.e. in $\mathcal{B} \cap \mathcal{C}$, which contains the vertex $\bar{x}$, the sign of the functions F_i remains the same as in the hypothesis of the theorem. So if $\bar{x}_i = 0$ then $F_i(x) < 0$ for all $x \in \mathcal{B} \cap \mathcal{C}$ and if $\bar{x}_1 = 1$ then $F_i(x) > 0$ for all $x \in \mathcal{B} \cap \mathcal{C}$. Then

(I) if $\bar{x}_i = 0$, $\dot{x}_i < 0$ for all $x \in \mathcal{B} \cap \mathcal{C}$ such that $x_i \neq 0$ and

(II) if $\bar{x}_i = 1$, $\dot{x}_i > 0$ for all $x \in \mathcal{B} \cap \mathcal{C}$ such that $x_i \neq 1$.

When $x \in \mathcal{B} \cap \mathcal{C}$ is such that its i coordinate is $x_i = 0$ or $x_i = 1$ then $\dot{x}_i = 0$, and that coordinate remains fixed for all time and $x_i = \bar{x}_i$. So we need only to look at those points in $\mathcal{B} \cap \mathcal{C}$ with $x_i \notin \{0, 1\}$.

So if we denote by $\xi(t, t_0, x_0)$ the solution of the system (2.28) with initial condition $x(t_0) = x_0$, it follows from (I) and (II) that $\xi(t, t_0, x_0) \to \bar{x}$ for all $x_0 \in \mathcal{B} \cap \mathcal{C}$. □

By putting these two theorems together we have that the behaviour along the incident edges is necessary and sufficient to characterize whether the vertex is an attractor or not in the case where there are no other equilibria in the edges apart from the vertexes, for dynamics where the vector field is constant along each edge, for instance, for the replicator dynamics.

Chapter 3

Nash and social tariffs impact in international trade

The first part of this chapter is based on the joint work:

F. Martins, A. A. Pinto, and J. P. Zubelli. Nash and social welfare impact in an international trade model. *Journal of Dynamics and Games*, 4(2):149–173, April 2017.

In this chapter we study a classic international trade model consisting of a strategic game in the tariffs of the governments. The model is a two-stage game where, at the first stage, governments of each country use their welfare functions to choose their tariffs either (i) competitively (Nash equilibrium) or (ii) cooperatively (social optimum). In the second stage, firms choose competitively (Nash) their home and export quantities. We compare the competitive (Nash) tariffs with the cooperative (social) tariffs and we classify the game type according to the coincidence or not of these equilibria as a social equilibrium (when they coincide), a prisoner's dilemma (when they do not coincide and the competitive outcome is dominated by the social) or a lose-win dilemma (when they do not coincide but one of the countries is damaged in terms of welfare in the social optimum). We do this comparison for several relevant utility functions that are economic relevant quantities for the two countries such as the custom revenue of the countries, the consumer surplus of each country, the profit of the firms and the welfare of the countries. The lack of coincidence of these equilibria for the welfare of the governments is a main difficulty in international trade that can be partially dealt with the use of trade agreements that impose the social tariffs and rule the distribution of the corresponding externalities among the two countries. We consider a welfare balanced trade agreement that has the feature of maintaining the welfare shares of the two countries when the social tariffs are enforced. We analyse some possible externalities that may be caused by such trade agreement at the level of relevant quantities such as profits and consumer surplus by also analysing shares of these quantities between the two countries and we analyse the gain obtained by the countries by using such a trade agreement. We conclude that the enforcing of a trade agreement may be a difficult issue because of some powerful externalities that might arise.

This chapter is structured in the following way. In section 3.1 we present the fundamental concepts of Nash and Social tariffs and the comparison between these two equilibria and the type of game that is obtained according to this comparison. In section 3.2 we present the international duopoly model and the most relevant economic quantities of the model. In section 3.3 we solve the second stage game between the firms. In section 3.4 we compute the Nash and Social tariffs for each relevant economic quantity of the duopoly model, considering each one as the utility function of the countries. We compare such tariffs and classify the games according to their types. We summarize the results in tables (3.1), (3.2) and (3.3). In section 3.5 we focus on the case of the welfare of the two countries, showing that the outcome is either of lose-win type (**LW**) or prisoner's dilemma type (**PD**). We present a full characterization of the game outcomes in terms of the tax-free home production indexes (see figure (3.1) and table (3.4)). In section 3.6, instead of the absolute comparison of economic quantities of the two countries at the Nash and social optimum we compute the relative share for these economic quantities and the difference between the Nash and social shares to measure the harm or benefit that the enforcing of the social tariffs can present to each country (see figures (3.2), (3.3) and (3.4)). In section 3.7 we propose a welfare balanced trade agreement, in the sense that it maintains the same welfare shares that were observed at the Nash tariffs. We analyse some possible externalities that arise in such a trade agreement using the shares computed at the previous section. We present some conclusions in section 3.8.

3.1 Strategic tariffs

In this section, we introduce the most relevant game theoretical concepts that we will use in the other sections to understand the strategic behaviour of firms, consumers and governments of the countries. Several classical books in game theory are available. For the interested reader not familiar with the fundamentals of game theory we refer to one of these classics, for instance [40].

We will present some fundamental concepts of game theory such as that of best response, Nash equilibrium and social optimum. We present them in the framework of government's choice of tariffs, although the definitions for other contexts are the same, for instance in terms of the choice of produced quantities by firms in competition, which we will see in sections 3.2 and 3.3.

Let $u_i(t_i, t_j)$ and $u_j(t_i, t_j)$ be two relevant economic quantities of the countries X_i and X_j depending only upon the tariffs t_i and t_j imposed by the governments of the two countries. For instance, for every pair of tariffs (t_i, t_j), the functions $u_i(t_i, t_j)$ and $u_j(t_i, t_j)$ can be the profit of the firms or the consumer surplus at the competitive Nash equilibrium for the quantities produced by the firms.

We are going to interpret $u_i(t_i, t_j)$ and $u_j(t_i, t_j)$ as the utilities of a game where the players are the governments of the countries and their actions are the tariffs (t_i, t_j).

The quantity $t_i^{BR}(t_j) \equiv t_i^{BR}(t_j; u)$ is a *best response* of the country X_i for the utility u_i, if for all tariffs t_i,

$$u_i(t_i^{BR}(t_j), t_j) \geq u_i(t_i, t_j) \ .$$

A pair of tariffs $(t_i^N, t_j^N) \equiv (t_i^N(u), t_j^N(u))$ is a *Nash equilibrium* or a *global strategic optimum*, if for all tariffs t_i

$$u_i(t_i^N, t_j^N) \geq u_i(t_i, t_j^N) \ ,$$

and for all tariffs t_j

$$u_j(t_i^N, t_j^N) \geq u_i(t_i^N, t_j) \ .$$

In other words, a pair of tariffs (t_i^N, t_j^N) is a Nash equilibrium, if

$$t_i^N = t_i^{BR}(t_j^N) \qquad \text{and} \qquad t_j^N = t_j^{BR}(t_i^N) \ .$$

A pair of tariffs $(t_i^P, t_j^P) \equiv (t_i^P(u), t_j^P(u))$ is a *Pareto optimum*, if there is no pair (t_i, t_j) of tariffs such that

$$u_i(t_i, t_j) \geq u_i(t_i^P, t_j^P) \qquad \text{for all} \;\; i, j \in \{1, 2\},$$

and at least one utility $u_i, \; i \in \{1, 2\}$ gets a better payoff with (t_i, t_j) than with (t_i^P, t_j^P), i.e.

$$u_i(t_i, t_j) > u_i(t_i^P, t_j^P) \; .$$

The *social utility* (or *total utility*) u_T is

$$u_T(t_i, t_j) = u_i(t_i, t_j) + u_j(t_i, t_j) \; .$$

The quantity $t_i^{SR}(t_j) \equiv t_i^{SR}(t_j; u)$ is a *social best response*, if for all tariffs t_i

$$u_T(t_i^{SR}(t_j), t_j) \geq u_T(t_i, t_j) \; .$$

A pair of tariffs $(t_i^S, t_j^S) \equiv (t_i^S(u), t_j^S(u))$ is a *social optimum*, if for all tariffs t_i

$$u_T(t_i^S, t_j^S) \geq u_T(t_i, t_j^S) \; ,$$

and for all tariffs t_j

$$u_T(t_i^S, t_j^S) \geq u_T(t_i^S, t_j) \; .$$

In other words, a pair of tariffs (t_i^S, t_j^S) is a social optimum, if

$$t_i^S = t_i^{SR}(t_j^S) \qquad \text{and} \qquad t_j^S = t_j^{SR}(t_i^S) \; .$$

We observe that a social optimum is a Pareto optimum. For games with a unique Nash equilibrium, we describe the three typical games outcomes when we compare the social optimum with the Nash equilibrium.

(**SE**) **Social equilibrium**: When the social optimum coincides with the Nash equilibrium

$$(t_i^S, t_j^S) = (t_i^N, t_j^N)$$

and the social optimum is the only Pareto optimum. In this case, the individualist Nash choice of the tariffs by the governments leads to a social equilibrium. Hence, a priori there is no need of a trade agreement between the two governments of the two countries.

(**PD**) **Prisoner's dilemma**: When the social optimum (t_i^S, t_j^S) is different from the Nash equilibrium

$$t_i^S \neq t_i^N \qquad \text{or} \qquad t_j^S \neq t_j^N$$

and both utilities are bigger in the social optimum than in the Nash equilibrium,

$$u_i(t_i^S, t_j^S) > u_i(t_i^N, t_j^N) \qquad \text{and} \qquad u_j(t_i^S, t_j^S) > u_j(t_i^N, t_j^N) \; .$$

In this case, the game is like the Prisoner's dilemma, where the Nash strategy leads to a lower outcome for both countries than if they would agree among therein (through a trade agreement) in opting for the social optimum. When we obtain

$$u_i(t_i^S, t_j^S) = u_i(t_i^N, t_j^N) \qquad \text{and} \qquad u_j(t_i^S, t_j^S) > u_j(t_i^N, t_j^N) \ .$$

or

$$u_i(t_i^S, t_j^S) > u_i(t_i^N, t_j^N) \qquad \text{and} \qquad u_j(t_i^S, t_j^S) = u_j(t_i^N, t_j^N) \ .$$

we say the game is a (weak) Prisoner's dilemma.

(**LW**) **Lose-win social dilemma**: When the social optimum (t_i^S, t_j^S) is different from the Nash equilibrium

$$t_i^S \neq t_i^N \qquad \text{or} \qquad t_j^S \neq t_j^N$$

and one of the utilities is bigger in the social optimum and the other utility is bigger in the Nash equilibrium, i.e.,

$$u_i(t_i^S, t_j^S) < u_i(t_i^N, t_j^N) \qquad \text{and} \qquad u_j(t_i^S, t_j^S) > u_j(t_i^N, t_j^N) \ ,$$

or

$$u_i(t_i^S, t_j^S) > u_i(t_i^N, t_j^N) \qquad \text{and} \qquad u_j(t_i^S, t_j^S) < u_j(t_i^N, t_j^N) \ .$$

When the game is of lose-win type there are two possible outcomes as described above. We will denote such outcomes respectively by $\mathbf{L}_i\mathbf{W}_j$ and $\mathbf{L}_j\mathbf{W}_i$. The first indicates that the country X_i has a utility loss and country X_j has a utility gain while enforcing the social optimum, and the second indicates the opposite situation.

In this case, the governments can implement an external mechanism (trade agreement) that will make them to opt for the social optimum in such a way that the country that gets an advantage in its utility compensates the loss in the utility of the other country and can also give some extra benefit in order to persuade the other country to implement the social equilibrium.

3.2 International duopoly model

In this section, we introduce the relevant economic quantities of the international duopoly model.

The international duopoly model is a game with two stages (sub-games). In the first stage, both governments choose simultaneously their Nash or social tariffs for a utility given by a relevant economic quantity; and, in the second stage, the firms choose simultaneously their home and export quantities to maximize competitively their profits.

The *home consumption* h_i is the quantity produced by the firm F_i and consumed in its own country X_i. The *export* e_i is the quantity produced by the firm F_i and consumed in the country X_j of the other firm F_j, where $i, j \in \{1, 2\}$ with $i \neq j$. The *tariff rate* t_i is determined by the government of country X_i on the import quantity e_j. The *total quantity* q_i produced by firm F_i is

$$q_i \equiv q_i(h_i, e_i) = h_i + e_i \ .$$

The *aggregate quantity* Q_i sold on the market in the country X_i is

$$Q_i \equiv Q_i(h_i, e_j) = h_i + e_j.$$

The *inverse demand* p_i in the country X_i is

$$p_i \equiv p_i(h_i, e_j) = \alpha_i - Q_i \ ,$$

where α_i is the *demand intercept* of country X_i.

The *payoff* π_i of firm F_i is

$$\pi_i \equiv \pi_i(h_i, e_i, h_j, e_j; t_i, t_j) = (p_i - c_i)h_i + (p_j - c_i)e_i - t_j e_i \ ,$$

where $c_i \geq 0$ is the firm F_i's *unitary production cost* such that $\alpha_i - c_i > 0$, and $t_j \geq 0$ is the tariff fixed by the government of country X_j.

The *custom revenue* CR_i of the country X_i is given by

$$CR_i \equiv CR_i(e_j; t_i) = t_i e_j \ .$$

The *consumer surplus* CS_i in the country X_i is given by

$$CS_i \equiv CS_i(h_i, e_j) = \frac{1}{2} Q_i^2 \ .$$

The *welfare* W_i of the country X_i is

$$W_i \equiv W_i(h_i, e_i, h_j, e_j; t_i, t_j) = CR_i + CS_i + \pi_i \ .$$

3.3 Second stage Nash equilibrium

In this section, we give a presentation of the well-known Nash equilibrium of the second sub-game, i.e., firms choose the home and export quantities that competitively maximize their profits, in the case of complete information, i.e. when both firms have full information on their and others utility functions.

Let $i, j \in \{1, 2\}$ with $i \neq j$. Define

$$T_i \equiv T_i(c_i, c_j) = (\alpha_i + c_i - 2c_j)/2 \ ,$$

$$T_j \equiv T_j(c_i, c_j) = (\alpha_j + c_j - 2c_i)/2 \ .$$

We also define

$$T_i^* \equiv T_i^*(c_i, c_j) = (\alpha_i + c_j - 2c_i)/2 \ ,$$

$$T_j^* \equiv T_j^*(c_i, c_j) = (\alpha_j + c_i - 2c_j)/2 \ .$$

Denoting $\Delta\alpha := \alpha_i - \alpha_j$, we have that

$$T_i^* = T_j + \Delta\alpha/2 \ ,$$
$$T_j^* = T_i - \Delta\alpha/2 \ .$$

This yields

$$T_i + T_j = T_i^* + T_j^* \ . \tag{3.1}$$

We also have that

$$\alpha_i - c_i = \frac{2(T_i + 2T_i^*)}{3} \ ,$$
$$\alpha_j - c_j = \frac{2(T_j + 2T_j^*)}{3} \ .$$

Denoting $\Delta c := c_i - c_j$ we have

$$\Delta c = \frac{2(T_j^* - T_j)}{3} = \frac{2(T_i - T_i^*)}{3} \ .$$

Assumption (A). *$T_i > 0$, $T_j > 0$, $T_i^* > 0$ and $T_j^* > 0$.*

We observe that under assumption (A), $\alpha_i - c_i > 0$ and $\alpha_j - c_j > 0$.

The *best response* $(h_i^{BR}(e_j), e_i^{BR}(h_j; t_j))$ of the firm F_i is the solution of

$$(h_i^{BR}(e_j), e_i^{BR}(h_j; t_j)) = \arg\max_{(h_i, e_i)} \pi_i(h_i, e_i, h_j, e_j; t_i, t_j).$$

Hence

$$\begin{cases} h_i^{BR}(e_j) = \frac{\alpha_i - e_j - c_i}{2} \\ e_i^{BR}(h_j; t_j) = \frac{\alpha_j - h_j - c_i - t_j}{2} \end{cases} .$$

The *Nash equilibrium* $(h_i(t_i), e_i(t_j); h_j(t_j), e_j(t_i))$ is the solution of

$$\begin{cases} (h_i(t_i), e_i(t_j)) = (h_i^{BR}(e_j(t_i)), e_i^{BR}(h_j(t_j); t_j)) \\ \\ (h_j(t_j), e_j(t_i)) = (h_j^{BR}(e_i(t_j)), e_j^{BR}(h_i(t_i); t_i)) \ . \end{cases}$$

So, for every $t_i \in [0, T_i]$ and every $t_j \in [0, T_j]$, the home $h_i(t_i)$ and export $e_i(t_j)$ quantities for the firms at the Nash equilibrium (see [40]) are

$$\begin{aligned} h_i(t_i) &\equiv h_i(c_i, c_j; t_i) = \frac{2T_i^* + t_i}{3} \\ e_i(t_j) &\equiv e_i(c_i, c_j; t_j) = \frac{2(T_j - t_j)}{3} \\ h_j(t_j) &\equiv h_i(c_i, c_j; t_i) = \frac{2T_j^* + t_j}{3} \\ e_j(t_i) &\equiv e_i(c_i, c_j; t_j) = \frac{2(T_i - t_i)}{3} \end{aligned}$$

We note that an increase in the quantities T_i^* and T_j^* generates an increase in the home quantities produced by the firms at a given tariff level. The quantities T_i^* and T_j^* have a clear economic interpretation. An increase in quantity T_i^* is good for firm F_i since this can occur due to three different possibilities, all of whom favour firm F_i: an increase in the home market size α_i of country X_i; an increase in the production costs of the opponent firm c_j; a decrease in firm i own production cost c_i. So any of these three situations generates an increase in home quantities, albeit at different rates: at a given tariff level, a reduction in the own costs increases home production at twice the speed as an increase in the home market size or an increase in the opponent's production costs.

Since the export quantities are non-negative, T_i and T_j are the *maximal tariffs*. When country X_i is tax-free, then firm F_j exports 2/3 of the admissible maximal tariff T_i. The maximal tariff T_i can increase due to several reasons: an increase in the market size α_i of country X_i; an increase in costs c_i of firm F_i; a decrease in the costs c_j of firm F_j. We further observe that at a fixed tariff rate, a decrease in the production cost c_j of firm F_j increases its export quantity faster, as the rate of change in the maximal tariff T_i due to a decrease in c_j is twice as big as the rate of change due to an increase in the market size of country X_i or a competitive loss in firm F_i because of increased production costs.

We also observe that the decrease in one country's export quantity due to an increase in the tariff that is practised by the other country is twice the increment that is provoked in the home quantity of the other country by the same increase in the tariff. In other words, an increase in tariffs lowers export quantities more rapidly than it enhances the home quantities.

Let

$$R_i = \frac{T_i^*}{T_i} \text{ and } R_j = \frac{T_j^*}{T_j} \ .$$

We observe that under assumption (A), $R_i > 0$ and $R_j > 0$. Using equality (3.1) we obtain

$$(1 - R_i)T_i = (R_j - 1)T_j \ .$$

Hence, the ratios R_i and R_j satisfy the relation

$$R_i < 1 \Leftrightarrow R_j > 1 \ .$$

Furthermore, we have that $R_i = 1$ if and only if $R_j = 1$, meaning that for every T_i and T_j, $T_i^* = T_i$ and $T_j^* = T_j$. When $R_i \neq 1$

$$\frac{T_i}{T_j} = \frac{(R_j - 1)}{(1 - R_i)} \ .$$

The *tax-free home production index* is

$$H_i = \frac{h_i^N(0)}{h_i^N(T_i)} = \frac{2T_i^*}{T_i + 2T_i^*} = \frac{2R_i}{1 + 2R_i} \ ,$$

where $h_i^N(0)$ corresponds to the home production of country i when there are tax-free exports from country j to country i, and $h_i^N(T_i)$ to the monopoly home production of country i when j does not export. We have

$$R_i = \frac{H_i}{2(1 - H_i)} \ .$$

Hence, the indexes H_i and H_j satisfy $0 < H_i < 1$, $0 < H_j < 1$ and the relation

$$0 < H_i < 2/3 \Leftrightarrow 2/3 < H_j < 1 \ .$$

Furthermore, we have that $H_i = 2/3$ if and only if $H_j = 2/3$, meaning that for every T_i and T_j, $T_i^* = T_i$ and $T_j^* = T_j$ (and $R_i = 1 = R_j$). When $H_i \neq 2/3$,

$$\frac{T_i}{T_j} = \frac{(H_i - 1)(3H_j - 2)}{(H_j - 1)(2 - 3H_i)} \ .$$

We observe that the country whose tax-free home production index is closer to 1 is the country that faces a lower decrease in their home production quantities while changing from a monopoly situation to a tax-free situation where the other country exports freely.

3.4 Strategic games

In this section, we will analyse the advantages and disadvantages of the use of tariffs for the firms, the consumers and the governments of the countries. To do it, we will use the relevant economic quantities as utilities $u_i(t_i, t_j)$ and $u_j(t_i, t_j)$ of a game where the players are the governments of the countries and their actions are the tariffs (t_i, t_j). For each pair of utilities that we will consider, we will find which of the three typical games occurs: social equilibrium (**SE**), prisoner's dilemma (**PD**), or lose-win social dilemma (**LW**).

3.4.1 Tariff effects in produced quantities and prices

We first consider the case where the utilities are the home quantities, i.e., $u_i = h_i$. The home quantity $h_i(t_i)$ increases with the tariff t_i and the home quantity $h_j(t_j)$ increases with the tariff t_j, and so

$$t_i^{BR}(t_j; h) = T_i \ , \ t_j^{BR}(t_i; h) = T_j \qquad \text{and} \qquad (t_i^N(h), t_j^N(h)) = (T_i, T_j) \ .$$

The social utility $h_T(t_i, t_j)$ is

$$h_T(t_i, t_j) = \frac{2(T_i^* + T_j^*) + t_i + t_j}{3}$$

and so the social utility also increases with both tariffs t_i and t_j, so

$$t_i^{SR}(t_j; h) = T_i \ , \ t_j^{SR}(t_i; h) = T_j \qquad \text{and} \qquad (t_i^S(h), t_j^S(h)) = (T_i, T_j) \ .$$

Hence, there is a unique social optimum (that is the unique Pareto optimum) and coincides with the Nash equilibrium

$$t_i^S(h) = t_i^N(h) = T_i \ .$$

Therefore, the game with utility $u_i = h_i$, is of the type **SE**.

When we consider the utilities to be the export quantities, i.e. $u_i = e_i$, we see that the export quantity $e_i(t_j)$ decreases with the tariff t_j, but does not depend upon tariff t_i. The same occurs with the export quantity $e_j(t_i)$ that decreases with tariff t_i, but does not depend upon tariff t_j. Hence, every tariff

t_i is a best response to any tariff t_j, and vice-versa, and so every pair of tariffs is a Nash equilibrium:

$$t_i^{BR}(t_j;e) \in [0,T_i]\ ,\ t_j^{BR}(t_i;e) \in [0,T_j] \qquad \text{and} \qquad (t_i^N(e), t_j^N(e)) \in [0,T_i] \times [0,T_j]\ .$$

The social utility $e_T(t_i,t_j)$ is

$$e_T(t_i,t_j) = \frac{2(T_i+T_j) - 2(t_i+t_j)}{3}$$

and so

$$t_i^{SR}(t_j;e) = 0\ ,\ t_j^{SR}(t_i;e) = 0 \qquad \text{and} \qquad (t_i^S(e), t_j^S(e)) = (0,0)\ .$$

Hence, there is a unique social optimum, that is the unique Pareto optimum

$$t_i^S(e) = 0\ .$$

Therefore, for the game with utility $u_i = e_i$, we have:

1. For the Nash tariff $(t_i^N, t_j^N) = (0,0)$, the game is of **SE** type;
2. For all other Nash tariffs the game is of **PD** type.

We now consider the case where the utilities are the total quantities produced by the firms, i.e., $u_i = q_i$. The total quantity $q_i(t_i,t_j)$ produced by firm F_i is given by

$$q_i(t_i,t_j) \equiv q_i(c_i,c_j;t_i,t_j) = \frac{1}{3}(2T_i^* + 2T_j + t_i - 2t_j)$$

and so the total quantity $q_i^N(t_i,t_j)$ increases with t_i and decreases with t_j. For firm F_j we have

$$q_j(t_i,t_j) \equiv q_j(c_i,c_j;t_i,t_j) = \frac{1}{3}(2T_j^* + 2T_i + t_j - 2t_i)$$

so the total quantity increases with t_j and decreases with t_i. Thus, there is a unique Nash equilibrium

$$t_i^{BR}(t_j;q) = T_i\ ,\ t_j^{BR}(t_i;q) = T_j \qquad \text{and} \qquad (t_i^N(q), t_j^N(q)) = (T_i, T_j)\ .$$

The social utility $q_T(t_i,t_j)$ is

$$q_T(t_i,t_j) = \frac{4(T_i+T_j) - (t_i+t_j)}{3}$$

and so

$$t_i^{SR}(t_j;q) = 0\ ,\ t_j^{SR}(t_i;q) = 0 \qquad \text{and} \qquad (t_i^S(q), t_j^S(q)) = (0,0)\ .$$

Hence, there is a unique social optimum but it does not coincide with the Nash equilibrium

$$t_i^S(q) \neq t_i^N(q)\ .$$

We have that

$$q_i(t_i^N, t_j^N) < q_i(t_i^S, t_j^S) \text{ if and only if } T_i < 2T_j\ .$$

$$q_j(t_i^N, t_j^N) < q_j(t_i^S, t_j^S) \text{ if and only if } T_j < 2T_i\ .$$

Hence, we have two possible cases:

Case I. $T_i/2 \leq T_j \leq 2T_i$. Then we have

$$q_i(t_i^N, t_j^N) \leq q_i(t_i^S, t_j^S) \qquad \text{and} \qquad q_j(t_i^N, t_j^N) \leq q_j(t_i^S, t_j^S) \ .$$

Therefore, the game is of the type **PD**.

Case II. $T_j < T_i/2$. (The case $2T_i < T_j$ is similar.) Then we have

$$q_i(t_i^N, t_j^N) > q_i(t_i^S, t_j^S) \qquad \text{and} \qquad q_j(t_i^N, t_j^N) < q_j(t_i^S, t_j^S) \ .$$

Therefore, the game is of the type **LW**. More precisely, the outcome is $\mathbf{L}_i\mathbf{W}_j$.

We now consider the utilities to be the aggregate quantities in each country, i.e., $u_i = Q_i$. The aggregate quantity $Q_i^N(t_i)$ in the market of country X_i is

$$Q_i(t_i) \equiv Q_i(c_i, c_j; t_i) = \frac{2(T_i + T_i^*) - t_i}{3} \ ,$$

and the aggregate quantity $Q_j(t_j)$ in the market of country X_j is

$$Q_j(t_j) \equiv Q_j(c_i, c_j; t_j) = \frac{2(T_j + T_j^*) - t_j}{3} \ ,$$

and so the aggregate quantities decrease with the respective tariffs. So, we have

$$t_i^{BR}(t_j; Q) = 0 \quad , \quad t_j^{BR}(t_i; Q) = 0 \qquad \text{and} \qquad (t_i^N(Q), t_j^N(Q)) = (0, 0) \ .$$

The social utility $Q_T(t_i, t_j)$ is

$$Q_T(t_i, t_j) = q_T(t_i, t_j)$$

and so

$$(t_i^S(Q), t_j^S(Q)) = (0, 0) \ .$$

Hence, there is a unique social optimum (that is the unique Pareto optimum) and coincides with the Nash equilibrium

$$t_i^S(Q) = t_i^N(Q) = 0 \ .$$

Therefore, the game with utility $u_i = Q_i$, is of the type **SE**.

We now consider the utility of the countries to be the symmetric of the prices, i.e., $u_i = p_i$. The inverse demand function $p_i(t_i)$ in the country X_i is

$$p_i(t_i) \equiv p_i(c_i, c_j; t_i) = \alpha_i - \frac{2(T_i^* + T_i) - t_i}{3}$$

and the inverse demand function $p_j(t_j)$ in the country X_j is

$$p_j(t_i) \equiv p_j(c_i, c_j; t_j) = \alpha_j - \frac{2(T_j^* + T_j) - t_j}{3} \ ,$$

and so the inverse demand functions of the two countries increase with the respective tariffs. So we have

$$t_i^{BR}(t_j;p) = T_i \ , \ t_j^{BR}(t_i;p) = T_j \qquad \text{and} \qquad (t_i^N(p), t_j^N(p)) = (T_i, T_j) \ .$$

The social utility $p_T(t_i, t_j)$ is

$$p_T(t_i, t_j) = \alpha_i + \alpha_j - \frac{2(T_i + T_j + T_i^* + T_j^*) - (t_i + t_j)}{3}$$

and so

$$t_i^{SR}(t_j;p) = T_i \ , \ t_j^{SR}(t_i;p) = T_j \qquad \text{and} \qquad (t_i^S(p), t_j^S(p)) = (T_i, T_j) \ .$$

Hence, there is a unique social optimum (that is the unique Pareto optimum) and coincides with the Nash equilibrium

$$t_i^S(p) = t_i^N(p) = T_i \ .$$

Therefore, the game with utility $u_i = p_i$, is of the type **SE**. When the utility is $u_i = -p_i$ then the solution is the same as the aggregate quantity in the market of country X_i, yielding a **SE** type game with equilibrium $(0, 0)$.

3.4.2 Governments direct gains from using tariffs

We now analyse the government's direct gains from using tariffs, i.e., the case where the utilities are given by the custom revenues, $u_i = CR_i$.

The custom revenue $CR_i(t_i)$ of country X_i is given by

$$CR_i(t_i) \equiv CR_i(c_i, c_j; t_i) = \frac{2t_i(T_i - t_i)}{3} \ ,$$

and the custom revenue $CR_j(t_j)$ of country X_j is given by

$$CR_j(t_j) \equiv CR_j(c_i, c_j; t_j) = \frac{2t_j(T_j - t_j)}{3} \ .$$

We have that $CR_i(t_i) \geq 0$ and $CR_j(t_j) \geq 0$. In both cases, first-order conditions yield the critical points $T_i/2$ and $T_j/2$, which are indeed maximum points. The custom revenue increases with the tariff $t_i \in [0, T_i/2]$, and it decreases with the tariff $t_i \in [T_i/2, T_i]$,

$$0 = CR_i(0) = CR_i(T_i) \leq CR_i(t_i) \leq CR_i\left(\frac{T_i}{2}\right) = \frac{T_i^2}{6} \ ,$$

and so

$$t_i^{BR}(t_j, CR) = \frac{T_i}{2} \ , \ t_j^{BR}(t_i, CR) = \frac{T_j}{2} \qquad \text{and} \qquad (t_i^N(CR), t_j^N(CR)) = \left(\frac{T_i}{2}, \frac{T_j}{2}\right) \ .$$

The social utility $CR_T(t_i, t_j)$ is

$$CR_T(t_i, t_j) = \frac{2t_i(T_i - t_i)}{3} + \frac{2t_j(T_j - t_j)}{3}$$

and so

$$t_i^{SR}(t_j;CR) = \frac{T_i}{2} \ , \ t_j^{SR}(t_i;CR) = \frac{T_j}{2} \qquad \text{and} \qquad \big(t_i^S(CR), t_j^S(CR)\big) = \left(\frac{T_i}{2}, \frac{T_j}{2}\right) \ .$$

Hence, there is a unique social optimum (that is the unique Pareto optimum) and coincides with the Nash equilibrium

$$t_i^S(CR) = t_i^N(CR) = \frac{T_i}{2} \ .$$

Therefore, the game with utility $u_i = CR_i$, is of the type **SE**.

3.4.3 Consumers savings effects from the use of tariffs

We now analyse the consumer's savings, i.e., the case where the utilities are given by the consumers surplus, $u_i = CS_i$.

The consumer surplus $CS_i(t_i)$ of country X_i is

$$CS_i(t_i) \equiv CS_i(c_i, c_j; t_i) = \frac{\big(2(T_i + T_i^*) - t_i\big)^2}{18} \ ,$$

and the consumer surplus $CS_j(t_j)$ of country X_j is

$$CS_j(t_j) \equiv CS_j(c_i, c_j; t_j) = \frac{\big(2(T_j + T_j^*) - t_j\big)^2}{18} \ .$$

The first-order conditions to minimize these quantities are

$$t_i = 2(T_i + T_i^*) \text{ and } t_j = 2(T_j + T_j^*) \ .$$

which are respectively bigger than T_i and T_j. So we have that

$$t_i^{BR}(t_j;CS) = 0 \ , \ t_j^{BR}(t_i;CS) = 0 \qquad \text{and} \qquad (t_i^N(CS), t_j^N(CS)) = (0,0) \ .$$

The social utility $CS_T(t_i, t_j)$ is

$$CS_T(t_i, t_j) = \frac{\big(2(T_i + T_i^*) - t_i\big)^2}{18} + \frac{\big(2(T_j + T_j^*) - t_j\big)^2}{18}$$

and so

$$t_i^{SR}(t_j;CS) = 0 \ , \ t_j^{SR}(t_i;CS) = 0 \qquad \text{and} \qquad (t_i^S(CS), t_j^S(CS)) = (0,0) \ .$$

Hence, there is a unique social optimum (that is the unique Pareto optimum) and coincides with the Nash equilibrium

$$t_i^S(CS) = t_i^N(CS) = 0 \ .$$

Therefore, the game with utility $u_i = CS_i$, is of the type **SE**.

3.4.4 Firms profits effects from the use of tariffs

We now consider the case where the utilities are the profits of the firms, i.e., $u_i = \pi_i$.

The profit $\pi_i(t_i, t_j)$ of the firm F_i is

$$\pi_i(t_i, t_j) \equiv \pi_i(c_i, c_j; t_i, t_j) = \frac{1}{9}[(2T_i^* + t_i)^2 + 4(T_j - t_j)^2] .$$

and the profit of the firm F_j is

$$\pi_j(t_i, t_j) \equiv \pi_j(c_i, c_j; t_i, t_j) = \frac{1}{9}[(2T_j^* + t_j)^2 + 4(T_i - t_i)^2] .$$

Thus, the profit $\pi_i(t_i, t_j)$ increases with t_i and decreases with t_j, and vice-versa for the profit $\pi_j(t_i, t_j)$. So

$$t_i^{BR}(t_j, \pi) = T_i \ , \ t_j^{BR}(t_i, \pi) = T_j \qquad \text{and} \qquad (t_i^N(\pi), t_j^N(\pi)) = (T_i, T_j) .$$

The social utility $\pi_T(t_i, t_j)$ is

$$\pi_T(t_i, t_j) = \frac{1}{9}[(2T_i^* + t_i)^2 + (2T_j^* + t_j)^2 + 4(T_i - t_i)^2 + 4(T_j - t_j)^2] .$$

Hence,

$$\frac{\partial \pi_T}{\partial t_i} = \frac{4(T_i^* - 2T_i) + 10t_i}{9} .$$

Noting that

$$\frac{\partial^2 \pi_T}{\partial t_i^2} = \frac{10}{9} > 0 ,$$

we obtain that the local maxima of π_T is attained at the boundary points of the admissible tariffs

$$t_i^{SR}(t_j; \pi) \in \{0, T_i\} .$$

Similarly,

$$t_j^{SR}(t_i; \pi) \in \{0, T_j\} .$$

We have that

$$\pi_T(T_i, t_j) - \pi_T(0, t_j) = \frac{T_i}{9}(4T_i^* - 3T_i) .$$

and

$$\pi_T(t_i, T_j) - \pi_T(t_i, 0) = \frac{T_j}{9}(4T_j^* - 3T_j) .$$

Hence, a priori, there are four possibilities for the social optimum. However, the tariff pair $(0, 0)$ cannot be achieved as a social optimum because the conditions $4T_i^* < 3T_i$ and $4T_j^* < 3T_j$ are incompatible. So we are left with three possible cases:

Case I. $4T_i^* > 3T_i$ and $4T_j^* > 3T_j$. Equivalently, $R_i > 3/4$ and $R_j > 3/4$ or $H_i > 3/5$ and $H_j > 3/5$. We have

$$t_i^{SR}(t_j; \pi) = T_i \qquad \text{and} \qquad t_j^{SR}(t_i; \pi) = T_j .$$

Thus,

$$(t_i^S(\pi), t_j^S(\pi)) = (T_i, T_j) .$$

Hence, there is a unique social optimum (that is the unique Pareto optimum) and coincides with the

Nash equilibrium

$$t_i^S(\pi) = t_i^N(\pi) = T_i \ .$$

Therefore, the game is of the type **SE**.

Case II. $4T_i^* < 3T_i$. Equivalently, $R_i < 3/4$ or $H_i < 3/5$. The case $4T_j^* < 3T_j$, or equivalently, $R_j > 3/4$ or $H_i > 3/5$ is similar. We have

$$t_i^{SR}(t_j;\pi) = 0 \qquad \text{and} \qquad t_j^{SR}(t_i;\pi) = T_j \ .$$

Therefore,

$$(t_i^S(\pi), t_j^S(\pi)) = (0, T_j) \ .$$

Hence, there is a unique social optimum but it does not coincide with the Nash equilibrium

$$t_i^N(\pi) \neq t_i^S(\pi) \qquad \text{and} \qquad t_j^N(\pi) = t_j^S(\pi) = T_j \ .$$

Furthermore,

$$\pi_i(t_i^N, t_j^N) > \pi_i(t_i^S, t_j^S) \qquad \text{and} \qquad \pi_j(t_i^N, t_j^N) < \pi_j(t_i^S, t_j^S) \ .$$

Therefore, the game is of the type **LW**. More precisely, the outcome is $\mathbf{L}_i\mathbf{W}_j$.

Case III. $4T_i^* = 3T_i$. or equivalently $R_i = 3/4$ or $H_i = 3/5$. The case $4T_j^* = 3T_j$ or equivalently, $R_j = 3/4$ or $H_j = 3/5$ is similar. In this case

$$\pi_T(T_i, t_j) = \pi_T(0, t_j) \ .$$

So that

$$t_i^{SR}(t_j;\pi) = \{0, T_i\} \qquad \text{and} \qquad t_j^{SR}(t_i;\pi) = T_j \ .$$

Therefore, there are two social optima.

$$(t_i^S(\pi), t_j^S(\pi)) = (0, T_j) \qquad \text{and} \qquad (t_i^S(\pi), t_j^S(\pi)) = (T_i, T_j) \ .$$

One of them $(t_i^S(\pi), t_j^S(\pi)) = (T_i, T_j)$ coincides with the Nash equilibrium, in which case the game is of **SE** type.

In the other social optimum, $(t_i^S(\pi), t_j^S(\pi)) = (0, T_j)$, by definition of Nash equilibrium we have that

$$\pi_i(t_i^N, t_j^N) > \pi_i(t_i^S, t_j^S) \ .$$

So, also by definition of social optimum we have

$$\pi_j(t_i^N, t_j^N) < \pi_j(t_i^S, t_j^S) \ ,$$

and hence the game is of **LW** type. More precisely, the outcome is $\mathbf{L}_i\mathbf{W}_j$.

3.4.5 Game outcomes

In this section we further discuss the game outcomes obtained for different utilities given by relevant economic quantities.

For every pair of tariffs (t_i, t_j), we found the Nash equilibrium for the second sub-game, i.e. the home and export quantities such that firms competitively maximize their profits. Then, using the Nash equilibrium for the home and export quantities we found the tariffs that lead to a Nash equilibria or to a social equilibria for different utilities.

We observed that for the home quantities the Nash equilibria and the social optimal tariffs are the same and equal to the maximal tariffs. So, countries decide to block exports, both when in competition and in cooperation. For the export quantities all tariffs lead to a Nash equilibrium but only the $(0,0)$ tariffs is a social optimum, with all the others yielding a prisoner's dilemma game. For the aggregate quantities in the market of each country, prices, custom revenues and consumer surpluses we found that the Nash tariffs coincide with the social tariffs, thus, the games with these utilities are of Social Equilibrium (**SE**) type. For the aggregate quantities and the consumer surpluses the tariffs are zero, corresponding to free export; for the custom revenues they are half of the maximal tariffs; and for prices they are the maximal tariffs. Hence, some of the difficulties of imposing tariffs arise from these social equilibria having different tariffs. We summarize these results in table (3.1).

	SE game					
Economic quantity	h	e	Q	p	CR	CS
Nash (Social) tariff of country X_i	T_i	0	0	T_i	$T_i/2$	0
Nash (Social) tariff of country X_j	T_j	0	0	T_j	$T_j/2$	0

Table 3.1: The Nash (Social) tariffs for the home quantity, export quantity, total quantity in the market, inverse demand, custom revenue and consumer surplus, resulting in a social equilibrium. h - Home quantities; e - Export quantities; Q - Aggregate quantity in each country; p - Inverse demand (price); CR - Custom revenue; CS - Consumer surplus.

For the total quantities produced by the firms we found that the Nash tariffs are the maximal tariffs, and the social tariffs are the zero tariffs. The game can be either of Prisoner's Dilemma (**PD**) type or of Lose-Win (**LW**) type, depending on the maximal tariffs. When the maximal tariffs of the two countries are sufficiently similar, then the game is of **PD** type. The game is **LW** for the bounds presented in table (3.2). When the game is of **LW** type then the country that is in a losing position is the one that has the highest maximal tariff. The reason for this is that the social tariffs are the zero tariffs, so that both countries become tax-free and so export more, but produce less for the home market. When the maximal tariffs are relatively similar, the effect of the increase in exports is more relevant since as we observed exports change more than the home quantity when maximal tariffs change, and so both countries would be better at the social optimum hence yielding **PD**. When one of the maximal tariffs is sufficiently larger then the one with the largest maximal tariff has a large decrease in home production which puts it in a losing position in terms of total output of the firms.

For the profits of the firms we found that the Nash tariffs are the maximal tariffs and the social tariffs can be either the maximal tariffs, or one of the two symmetrical cases where one country chooses the maximal tariff and the other chooses the zero tariff. Which one of the cases occurs depends on the

Total quantities (q_i, q_j) produced by the firms			
Condition	Nash tariffs	Social tariffs	Game type
If $2T_j < T_i$	(T_i, T_j)	$(0, 0)$	$\mathbf{L}_i\mathbf{W}_j$
If $T_i/2 \leq T_j \leq 2T_i$	(T_i, T_j)	$(0, 0)$	**PD**
If $2T_i < T_j$	(T_i, T_j)	$(0, 0)$	$\mathbf{L}_j\mathbf{W}_i$

Table 3.2: Comparing total quantities (q_i, q_j) of the two countries with Nash tariffs and social tariffs with different cost similarities and concluding the game type.

tax-free home production indexes. When a firm has a tax-free home production index lower than the threshold 3/5, the game is of **LW** type, and this firm has a profit loss. In these cases, the country with the higher tax-free home production index earns more profit at the social optimum. The social optimum tariff for that country with the tax-free home production index below the threshold is to become tax-free, and he has a loss in profit, and for the country with the highest index is to block exports from the other country. For the country that has a lower tax-free home production index this means it has greater loss while becoming tax-free, meaning that the firm is less competitive when his home market is shared between the two countries. Hence, its firm produces its tax-free home production quantity and does not export, while the firm of the other country has the monopoly in its own market and exports for the tax-free country. Even if the country with the lowest index produces much less for its home market in a tax-free situation, it is socially more profitable to allow the more competitive foreign country to enter he market and not allow the weaker country to export. So, the less competitive firm has interest in becoming tax-free if he gets some compensation by the other firm. Possibilities that may be discussed between the two countries might include merging of the two firms, R&D exchange, or allowing representation of the more competitive firm to be settled in the less competitive country. When both firms have a high tax-free home production index the game is of **SE** type. In this case both firms produce their monopoly quantities. We summarize these results in table (3.3).

Profits (π_i, π_j) of the firms			
Condition	Nash tariffs	Social tariffs	Game type
If $H_i < 3/5$	(T_i, T_j)	$(0, T_j)$	$\mathbf{L}_i\mathbf{W}_j$
If $H_i > 3/5$ and $H_j > 3/5$	(T_i, T_j)	(T_i, T_j)	**SE**
If $H_j < 3/5$	(T_i, T_j)	$(T_i, 0)$	$\mathbf{L}_j\mathbf{W}_i$

Table 3.3: Comparing profits of the firms of the two countries with Nash tariffs and social tariffs, where H_i and H_j are the tax-free home production indexes.

3.5 Nash and social welfares

In this section, we consider the utility of the governments to be the welfare of the country, i.e. $u_i = W_i$. We will compute the Nash equilibrium tariffs and the social tariffs and analyse the game type obtained according to the tax-free home production index.

3.5.1 Computation of the equilibria

The welfare $W_i(t_i, t_j)$ of the country X_i is given by

$$\begin{aligned} W_i(t_i, t_j) &= \tfrac{1}{9}\Big[(2T_i^* + t_i)^2 + 4(T_j - t_j)^2\Big] + \tfrac{2}{3}t_i(T_i - t_i) \\ &+ \tfrac{1}{18}(2(T_i + T_i^*) - t_i)^2 \ . \end{aligned}$$

The welfare $W_j(t_i, t_j)$ of the country X_j is given by

$$\begin{aligned} W_j(t_i, t_j) &= \tfrac{1}{9}\Big[(2T_j^* + t_j)^2 + 4(T_i - t_i)^2\Big] + \tfrac{2}{3}t_j(T_j - t_j) \\ &+ \tfrac{1}{18}(2(T_j + T_j^*) - t_j)^2 \ . \end{aligned}$$

We have that

$$\frac{\partial W_i}{\partial t_i} = \frac{4T_i + 2T_i^*}{9} - t_i \ , \qquad \frac{\partial W_j}{\partial t_j} = \frac{4T_j + 2T_j^*}{9} - t_j \ ,$$

and

$$\frac{\partial^2 W_i}{\partial t_i^2} = -1 \ , \qquad \frac{\partial^2 W_j}{\partial t_j^2} = -1 \ .$$

Therefore, the maximum points of the polynomials $W_i(t_i, t_j)$ and $W_j(t_i, t_j)$ in t_i and t_j are, respectively

$$A_{W,i} = \frac{2(T_i^* + 2T_i)}{9} > 0 \qquad \text{and} \qquad A_{W,j} = \frac{2(T_j^* + 2T_j)}{9} > 0 \ .$$

Noting that $A_{W,i} < T_i$ is equivalent to $2T_i^* < 5T_i$, and that $A_{W,j} < T_j$ is equivalent to $2T_j^* < 5T_j$, we get that the best responses are

$$t_i^{BR}(t_j, W) = \begin{cases} A_{W,i} \ , & \text{if } T_i^* < \frac{5T_i}{2} \ ; \\ \\ T_i \ , & \text{otherwise} \ . \end{cases} \qquad t_j^{BR}(t_i, W) = \begin{cases} A_{W,j} \ , & \text{if } T_j^* < \frac{5T_j}{2} \ ; \\ \\ T_j \ , & \text{otherwise} \ . \end{cases}$$

The social utility $W_T(t_i, t_j)$ is

$$W_T(t_i, t_j) = W_i(t_i, t_j) + W_j(t_i, t_j) \ .$$

Hence, we have that

$$\frac{\partial W_T}{\partial t_i} = \frac{2T_i^* - 4T_i - t_i}{9} \ .$$

and

$$\frac{\partial^2 W_T}{\partial t_i^2} = -\frac{1}{9} \ .$$

Let

$$B_{W_S,i} = 2(T_i^* - 2T_i) \ ,$$

and analogously,

$$B_{W_S,j} = 2(T_j^* - 2T_j) \ .$$

Noting that $0 < B_{W_S,i} < T_i$ is equivalent to $2T_i < T_i^* < 5T_i/2$, we get that the social best responses are

$$t_i^{SR}(t_j;W) = \begin{cases} 0 \ , & \text{if } T_i^* \leq 2T_i \ ; \\ \\ B_{W_S,i} \ , & \text{if } 2T_i < T_i^* < \frac{5T_i}{2} \ ; \\ \\ T_i \ , & \text{if } T_i^* \geq \frac{5T_i}{2} \ . \end{cases}$$

Similarly, we get

$$t_j^{SR}(t_i;W) = \begin{cases} 0 \ , & \text{if } T_j^* \leq 2T_j \ ; \\ \\ B_{W_S,j} \ , & \text{if } 2T_j < T_j^* < \frac{5T_j}{2} \ ; \\ \\ T_j \ , & \text{if } T_j^* \geq \frac{5T_j}{2} \ . \end{cases}$$

Hence, there are several possible cases when comparing the Nash and social tariffs. Since $0 \leq H_i \leq 2/3 \leq H_j \leq 1$ or $0 \leq H_j \leq 2/3 \leq H_i \leq 1$ some possibilities are incompatible with each other. We will make use of the notation ΔW to denote the difference between the welfare of a country when applying the Nash tariffs and when applying the social tariffs. We are left with five cases:

Case I. $T_i^* \leq 2T_i$ and $T_j^* \leq 2T_j$. Equivalently, $R_i \leq 2$ and $R_j \leq 2$ or $H_i \leq 4/5$ and $H_j \leq 4/5$. The Nash equilibrium is

$$(t_i^N(W), t_j^N(W)) = (A_{W,i}, A_{W,j}) \ .$$

The social optimum is

$$(t_i^S(W), t_j^S(W)) = (0, 0) \ .$$

The social optimum does not coincide with the Nash equilibrium.

The welfare at the Nash equilibrium is

$$W_j(A_{W,i}, A_{W,j}) = CR_j(A_{W,j}) + CS_j(A_{W,j}) + \pi_j(A_{W,i}, A_{W,j}) \ ,$$

where

$$CR_j(A_{W,j}) = -\frac{4(T_j^* + 2T_j)(2T_j^* - 5T_j)}{243} \ ,$$

$$CS_j(A_{W,j}) = \frac{2(8T_j^* + 7T_j)^2}{729}$$

and

$$\pi_j(A_{W,i}, A_{W,j}) = \frac{4\left(4(5T_j^* + T_j)^2 + (2T_i^* - 5T_i)^2\right)}{729} \ .$$

The welfare at the social optimum is

$$W_j(0,0) = CS_j(0) + \pi_j(0,0) \ ,$$

where

$$CS_j(0) = \frac{2(T_j^* + T_j)^2}{9}$$

and

$$\pi_j(0,0) = \frac{4}{9}\left((T_j^*)^2 + T_j^2\right) \ .$$

Letting

$$\Delta W_{1,j} = W_j(A_{W,i}, A_{W,j}) - W_j(0,0) \ ,$$

we have

$$\Delta W_{1,j} = \frac{2\left(9(T_j^* + 2T_j)^2 - 8(2T_i + T_i^*)(7T_i - T_i^*)\right)}{729} \ .$$

Case Ia). When $R_i \neq 1$, and so $R_j \neq 1$,

$$\Delta W_{1,j} = \frac{2}{729}\left(\frac{9(R_i - 1)^2(R_j + 2)^2}{(R_j - 1)^2} + 8(R_i + 2)(R_i - 7)\right) \ .$$

Hence, depending on the ratios R_i and R_j, or the tax-free home production indexed H_i and $_j$, the game has three outcomes (see figure (3.1)): (1) If $R_i = 1 - \beta$ and $R_j = 1 + \beta$, with β close to 0, the game is of Prisoner's dilemma (**PD**) type; (2) If R_i is closer to 1 than R_j, the outcome is $(\mathbf{L}_i\mathbf{W}_j)$; (3) If R_j is closer to 1 than R_i the outcome is $(\mathbf{L}_j\mathbf{W}_i)$.

Case Ib). If $R_i = 1$ (equivalently $R_j = 1$), or $H_i = 2/3$ and $H_j = 2/3$, we have that $T_i = T_i^*$ and $T_j = T_j^*$. This means that the two countries have the same production costs $c_i = c_j = c$, and $T_i = \frac{\alpha_i - c}{2}$ and $T_j = \frac{\alpha_j - c}{2}$. As before, we compute the welfare deltas

$$\begin{aligned}
\Delta W_{1,i} &= \frac{2(9T_i^2 - 16T_j^2)}{81} \ , \\
\Delta W_{1,j} &= \frac{2(9T_j^2 - 16T_i^2)}{81} \ .
\end{aligned}$$

So, we have that

$$\Delta W_{1,i} > 0 \text{ iff } 3/4T_i > T_j \ ,$$

$$\Delta W_{1,j} > 0 \text{ iff } T_j > 4/3T_i \ ,$$

Hence, depending on T_i and T_j, the game has three outcomes: (1) if T_i is close to T_j, the game is of Prisoner's dilemma (**PD**) type; (2) when T_i is sufficiently larger than T_j, the outcome is ($\mathbf{L}_i\mathbf{W}_j$); (3) when T_j is sufficiently larger than T_i, the outcome is ($\mathbf{L}_j\mathbf{W}_i$).

Recall that in this case countries have the same production costs, so their difference arises solely from the demand intercepts. So equivalently, we have the following three cases: (1) if the two demand intercepts of the two countries are similar, then the game is of Prisoner's dilemma (**PD**) type; (2) when the demand intercept α_i of country X_i is sufficiently larger than the demand intercept α_j of country X_j, the outcome is ($\mathbf{L}_i\mathbf{W}_j$); (3) when the demand intercept α_j of country X_j is sufficiently larger than the demand intercept α_i of country X_i, the outcome is ($\mathbf{L}_j\mathbf{W}_i$). The country with the greatest demand intercept has greater market size since it has greater demand than the other country. So if countries have similar demand intercepts and consequently similar market sizes then they are in a Prisoner's dilemma situation, otherwise the country with the greatest market size (i.e. greatest demand) will be harmed by the enforcing of the social tariffs and the country will be benefited by such enforcement.

Case II. $2T_j < T_j^* < 5T_j/2$. Equivalently, $2 < R_j < 5/2$ or $4/5 < H_j < 5/6$. In this case we also have that $T_i^* \leq 2T_i$, equivalently, $R_i \leq 2$ or $H_i \leq 4/5$. The Nash equilibrium is

$$(t_i^N(W), t_j^N(W)) = (A_{W,i}, A_{W,j}) \ .$$

Hence, the welfare at the Nash equilibrium is the same as in case ia). The social optimum is

$$(t_i^S(W), t_j^S(W)) = (0, B_{W_S,j}) \ .$$

Since

$$B_{W_S,j} < A_{W,j}$$

is equivalent to

$$T_j^* < 5/2T_j$$

which is verified in this case, then

$$t_i^S(W) \neq t_i^N(W) \qquad \text{and} \qquad t_j^S(W) \neq t_j^N(W) \ .$$

The welfare at the social optimum is

$$W_j(0, B_{W_S,j}) = CR_j(B_{W_S,j}) + CS_j(B_{W_S,j}) + \pi_j(0, B_{W_S,j}) \ ,$$

where

$$CR_j(B_{W_S,j}) = -\frac{4(T_j^* - 2T_j)(2T_j^* - 5T_j)}{3} \ ,$$

$$CS_j(B_{W_S,j}) = 2T_j^2$$

and

$$\pi_j(0, B_{W_S,j}) = \frac{4\left(4(T_j^* - T_j)^2 + T_i^2\right)}{9} \ .$$

Letting

$$\Delta W_{2,i} = W_i(A_{W,i}, A_{W,j}) - W_i(0, B_{W_S,j})$$

and

$$\Delta W_{2,j} = W_j(A_{W,i}, A_{W,j}) - W_j(0, B_{W_S,j}) \ .$$

We have

$$\begin{aligned}
\Delta W_{2,i} &= \frac{2\left(9(2T_i+T_i^*)^2 - 160(2T_j^* - 5T_j)^2\right)}{729} \\
\Delta W_{2,j} &= \frac{16\left(18(2T_j^* - 5T_j)^2 - (2T_i+T_i^*)(7T_i - T_i^*)\right)}{729} \ .
\end{aligned}$$

Hence,

$$\begin{aligned}
\Delta W_{2,i} &= \frac{2}{729}\left(\frac{9(R_j-1)^2(R_i+2)^2}{(R_i-1)^2} - 160(2R_j-5)^2\right) \\
\Delta W_{2,j} &= \frac{16}{729}\left(\frac{18(R_i-1)^2(2R_j-5)^2}{(R_j-1)^2} - (R_i+2)(7-R_i)\right) \ .
\end{aligned}$$

Thus, depending on the ratios R_i and R_j, the game has three outcomes (see figure (3.1)): (i) For instance, if R_i is close to 0 and R_j close to 2 the outcome is $\mathbf{L}_j\mathbf{W}_i$; (ii) For instance, if R_i is close to 0, but not too close, and R_j is close to 2, but not too close, the game is of Prisoner's dilemma (**PD**) type; (iii) For the majority of values of R_i and R_j the outcome is $\mathbf{L}_i\mathbf{W}_j$.

Case III. $0 < 5T_j/2 \leq T_j^*$. Equivalently, $R_j > 5/2$ or $5/6 < H_j < 1$. In this case we also have that $T_i^* \leq 2T_i$, equivalently, $R_i \leq 2$ or $H_i \leq 4/5$. The Nash equilibrium is

$$(t_i^N(W), t_j^N(W)) = (A_{W,i}, T_j) \ .$$

The social optimum is

$$(t_i^S(W), t_j^S(W)) = (0, T_j) \ .$$

The Welfare at the Nash equilibrium is

$$W_j(A_{W,i}, T_j) = CS_j(T_j) + \pi_j(A_{W,i}, T_j) \ ,$$

where

$$CS_j(T_j) = \frac{(2T_j^* + T_j)^2}{18}$$

and

$$\pi_j(A_{W,i}, T_j) = \frac{81(2T_j^* + T_j)^2 + 4(2T_i^* - 5T_i)^2}{729} \ .$$

The welfare at the social optimum is

$$W_j(0,T_j) = CS_j(T_j) + \pi_j(0,T_j) \ ,$$

where

$$\pi_j(0,T_j) = \frac{(2T_j^* + T_j)^2 + 4T_i^2}{9} \ .$$

Furthermore, by definition of Nash equilibrium, clearly we have that

$$W_i(A_{W,i},T_j) > W_i(0,T_j) \ .$$

The fact that the social optimum is $(0,T_j)$ together with this last inequality yields

$$W_j(A_{W,i},T_j) < W_j(0,T_j) \ .$$

Hence, in this case there is a unique social optimum that does not coincide with the Nash equilibrium. Furthermore, the game is of the type **LW**. More precisely, the outcome is $\mathbf{L}_i\mathbf{W}_j$.

Case IV. $2T_i < T_i^* < 5T_i/2$. This case is dual to case II.

Case V. $0 < 5T_i/2 \leq T_i^*$. This case is dual to case III.

3.5.2 Welfare game outcomes

Here, we present the regions where the outcomes are of type $\mathbf{L}_i\mathbf{W}_j$, **PD** and $\mathbf{L}_j\mathbf{W}_i$. We will analyse the case $0 < H_i < 2/3 < H_j < 1$, the other case is dual to this one. Observe that in the case we study, since H_i is lower than H_j, country X_i has a higher decrease in home quantities when changing from maximal tariffs to a tax-free situation. The corner point $H_i = 2/3$ and $H_j = 2/3$ was analysed separately before (case Ib)). We summarize the game type obtained in table (3.4).

Welfares (W_i, W_j) of the countries			
Condition	Nash tariffs	Social tariffs	Game type
$H_j \geq 5/6$	$(A_{W,i}, T_j)$	$(0, T_j)$	$\mathbf{L}_i\mathbf{W}_j$
$4/5 < H_j < 5/6$	$(A_{W,i}, A_{W,j})$	$(0, B_{W_S,j})$	**LW** or **PD**
$H_j \leq 4/5$	$(A_{W,i}, A_{W,j})$	$(0, 0)$	**LW** or **PD**

Table 3.4: Comparing welfares of the two countries with Nash tariffs and social tariffs where H_i and H_j are the tax-free home production indexes satisfying $0 < H_i < 2/3 < H_j < 1$.

For the welfare of the countries we found two thresholds for the tax-free home production index H_j: the *social monopoly-tax threshold* 5/6 and the *social tax-free threshold* 4/5 (see figure (3.1)). For all values of the tax-free home production indexes H_i and H_j, the firm F_i chooses the Nash tariff $A_{W,i}$ and vanishes its tariff at the social equilibrium. In case III, when $H_j \geq 5/6$, the game is of $\mathbf{L}_i\mathbf{W}_j$ type and the country X_j has a welfare gain. The country X_j applies the maximal tariff both at the Nash and Social equilibria. When $H_j < 5/6$, the game has three outcomes: prisoner's dilemma **PD**, and lose-win $\mathbf{L}_i\mathbf{W}_j$ and $\mathbf{L}_j\mathbf{W}_i$. In this case, firm F_j chooses the Nash tariff $A_{W,j}$. Its social tariff is $B_{W_S,j}$

if $4/5 < H_j < 5/6$, and it vanishes its tariff at the social equilibrium if $2/3 < H_j \leq 4/5$ (see table (3.4)). We observe from figure (3.1) that in case II the majority of the parameter yields either a **PD** or a $\mathbf{L}_i\mathbf{W}_j$ type game, meaning that country X_j has a welfare gain, except for a very small parameter region where country X_j has a welfare loss. In case I, the three game types may occur. When H_j gets lower and closer to 2/3 (meaning that the tax-free home production of X_j gets lower in comparison to the monopoly home quantity), it gets more likely that the game type is $\mathbf{L}_j\mathbf{W}_i$, with country X_j losing welfare. For lower values of H_i, H_j doesn't need to be so lower in order to have a $\mathbf{L}_j\mathbf{W}_i$ type game, and there is a threshold in H_i (approximately 0.2) such that the game is always of this type in case I.

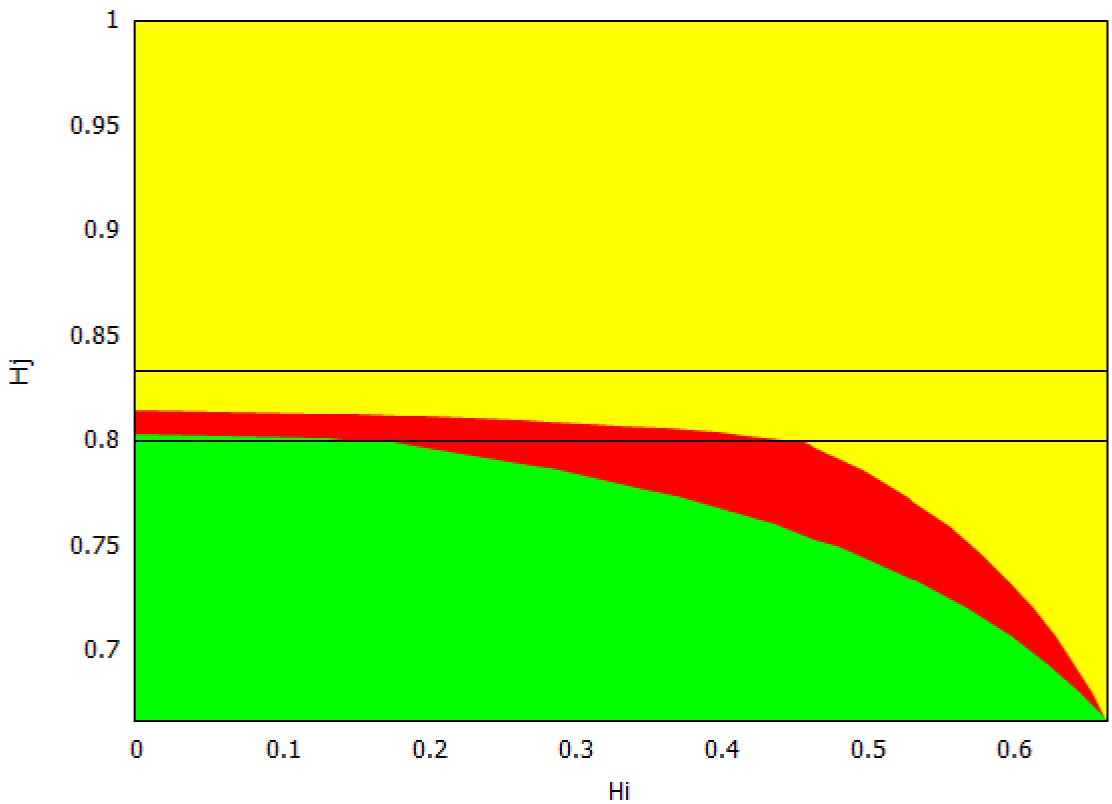

Figure 3.1: The Welfare Game Type: Green - $\mathbf{L}_j\mathbf{W}_i$; Red - **PD**; Yellow - $\mathbf{L}_i\mathbf{W}_j$.

We have observed that $B_{W_S,j} < A_{W,j}$. So, we have that for both countries the social tariffs are lower than the Nash tariffs, and they are lower or both countries except for the situation where one of the countries plays its maximal tariff. So, any trade agreement that enforces the social tariffs will therefore yield lower tariffs than the tariffs used at a competitive (Nash) equilibrium for at least one of the countries. By direct inspection of the custom revenue, consumer surplus and profit functions, we find that if the two countries decide to impose the social tariffs the following holds: a) the custom revenue of country X_i vanishes, since his tariff vanishes at the social optimum, and in some situations (cases I and III), it also vanishes for country X_j, since country X_j's social tariff is, respectively, 0 and the maximal tariff. When it doesn't vanish, which occurs only for values of H_j between the two thresholds, then the its custom revenue may increase or decrease; b) the consumer surplus of the countries always increases; and c) the profits may increase or decrease.

3.6 Nash and social welfare shares

In absolute terms, if the countries decide to impose the social tariffs the following hold: a) the custom revenue of country X_i vanishes, since his tariff vanishes at the social optimum, and in some situations it also vanishes for country X_j. When it doesn't vanish for country X_j, which occurs only for values of H_j between the two thresholds, then it may increase or decrease; b) the consumer surplus of the countries increases; and c) the profits may increase or decrease. However, perhaps more important than the absolute value of gains or losses in these economic quantities are the relative shares of these quantities. For instance, even if the consumer surplus always increases, the relative share of the consumer surplus may change with the application of the social tariffs, thus resulting in a change in the balance between the two countries, since a dominant country in one aspect may cease to be dominant with the social tariffs. Because of this, in this section we will focus on the shares of the economic quantities considered above.

In order to simplify the notation in the following, we will make the following convention: a superscript N will denote evaluation at the Nash equilibrium tariffs, while a superscript S wil denote evaluation at the social optimum tariffs. For instance, in the case of the welfare we will denote by $W_i^N = W_i(t_i^N, t_j^N)$ the Nash equilibrium welfare, i.e., the welfare at the Nash equilibrium tariffs, and by $W_i^S = W_i(t_i^S, t_j^S)$ the social optimum welfare, i.e., the welfare at the social optimum tariffs that maximize the joint welfare of the two countries. Analogously, we will use the same notation to the other quantities, such as profits, consumer surplus and custom revenue.

Using the previous notation, the *joint Nash welfare* is $W_T^N = W_i^N + W_j^N$ and the *joint social welfare* is $W_T^S = W_i^S + W_j^S$. The *Nash welfare share* and the *social welfare share* are

$$ShW_j^N = \frac{W_j^N}{W_T^N} \qquad \text{and} \qquad ShW_j^S = \frac{W_j^S}{W_T^S} \ .$$

So $ShW_i^N + ShW_j^N = 1$ and $ShW_i^S + ShW_j^S = 1$. The *Nash-social welfare share difference* is

$$\Delta ShW_j^S = ShW_j^N - ShW_j^S \ ,$$

and so $ShW_j^N = ShW_j^S + \Delta ShW_j^S$, and $\Delta ShW_j = -\Delta ShW_i$. The Nash-social welfare share difference is very relevant to compare the relative advantage of one country over the other country between the Nash and Social equilibria.

Similarly we define the *Nash consumer surplus share* and the *Social Consumer surplus share*

$$ShCS_j^N = CS_j^N / CS_T^N \qquad \text{and} \qquad ShCS_j^S = CS_j^S / CS_T^S \ .$$

Hence, the *Nash-social consumer surplus share difference* is

$$\Delta ShCS_j^S = ShCS_j^N - ShCS_j^S \ .$$

We also define the *Nash profit share* and the *Social profit share*

$$Sh\pi_j^N = \pi_j^N / \pi_T^N \qquad \text{and} \qquad Sh\pi_j^S = \pi_j^S / \pi_T^S \ .$$

Hence, the *Nash-social profit share difference* is

$$\Delta Sh\pi_j^S = Sh\pi_j^N - Sh\pi_j^S \ .$$

The *Nash custom revenue share* is defined by

$$ShCR_j^N = CR_j^N / CR_N^S \ .$$

We will not compute the *Social custom revenue share* since in the social tariffs this share is either undefined, because both countries have zero custom revenue, or it is equal to 0 for one country and 1 for the other country in the case when the tax-free home production index is between the two thresholds we described above.

The Nash-social consumer surplus share difference and Nash-social profit share difference are also very relevant to compare the relative advantage of one country over the other country between Nash and Social equilibrium from the perspective of the consumer and of the firm. Therefore, we will present in several figures the shares and the share differences of these relevant economical quantities to exhibit their properties in terms of the tax free home production indices (H_i, H_j), for $0 \leq H_i \leq 2/3 \leq H_j \leq 1$. We will consider the case where $H_i = 2/3$ and $H_j = 2/3$ separately since the computations simplify significantly and we are able to obtain workable and simple expressions for the shares.

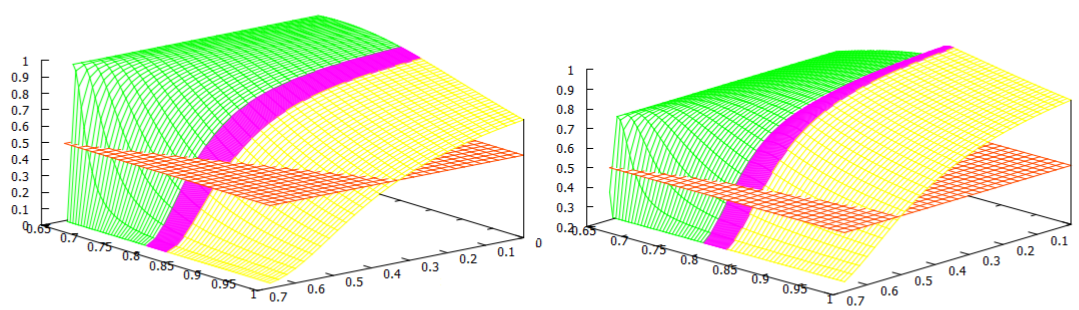

Figure 3.2: Left: The Nash welfare share ShW_j^N of country X_j. Right: The Social welfare share ShW_j^S of country X_j.

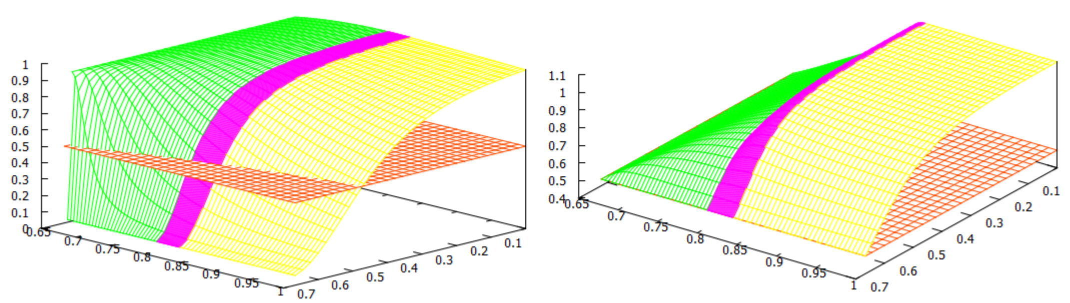

Figure 3.3: Left: The Nash profit share $Sh\pi_j^N$ of country X_j. Right: The Social profit share $Sh\pi_j^S$ of country X_j.

SPECIAL CASE: If $H_i = 2/3$ and $H_j = 2/3$ depending on T_i and T_j, then as we have seen previously the game has three outcomes: (1) If T_i is close to T_j, the game is of Prisoner's dilemma (**PD**) type; (2) when T_i is sufficiently larger than T_j, the outcome is $(\mathbf{L}_i\mathbf{W}_j)$; (3) when T_j is sufficiently larger than T_i, the outcome is $(\mathbf{L}_j\mathbf{W}_i)$.

The Nash and social Welfare shares are

$$ShW_j^N = \frac{63T_j^2 + 2T_i^2}{65(T_j^2 + T_i^2)} \quad \text{and} \quad ShW_j^S = \frac{3T_j^2 + T_i^2}{4(T_j^2 + T_i^2)} \ .$$

The Nash-social welfare share difference is

$$\Delta ShW_j = \frac{57(T_j - T_i)(T_j + T_i)}{260(T_j^2 + T_i^2)} \ .$$

The Nash and social profit shares are

$$Sh\pi_j^N = \frac{16T_j^2 + T_i^2}{17(T_j^2 + T_i^2)} \quad \text{and} \quad Sh\pi_j^S = \frac{1}{2} \ .$$

The Nash-social profit share difference is

$$\Delta Sh\pi_j = \frac{15(T_j - T_i)(T_j + T_i)}{34(T_j^2 + T_i^2)} \ .$$

The Nash and social consumer surplus shares are

$$ShCS_j^N = ShCS_j^S = \frac{T_j^2}{T_j^2 + T_i^2} \ .$$

The Nash custom revenue share is

$$ShCR_j^N = \frac{T_j^2}{T_j^2 + T_i^2} \ .$$

The Nash and social welfare shares exhibit similar behaviours with both shares being higher for bigger values of T_j. However, for the the Nash welfare share, its maximum value is 1, while for the social welfare share, the maximum is 0.75. The Nash profit share is higher for bigger values of T_j and its maximum value is 16/17. The social profit share is equal 1/2, meaning that at the social tariffs profit is evenly split between the two firms. The Nash-social welfare share difference and the Nash-social profit share difference are respectively between $-50/260$ and $50/260$ (approximately 19%), and between $-15/34$ and $15/34$ (approximately 44%). They are 0 along the line $T_i = T_j$, and when $T_i > T_j$, country X_i has a positive share difference, while for $T_i < T_j$ country X_j has a positive share difference. The Nash and social consumer surplus shares are the same, and are equal to the Nash custom revenue share. These shares are higher for bigger values of T_j and its maximum is 1.

In figures (3.2),(3.3),(3.4), we exhibit low values of H_i with the following properties: a) the Nash and social welfares, profits, and consumer surpluses shares of country X_j are higher than 1/2; b) country X_j is the looser in game type for low values of H_j and country X_i is the looser in game type for high values of H_j. We exhibit high values of H_i with the following properties: a) the Nash and social welfares, profits, and consumer surpluses shares of country X_i are higher than 1/2; b) country X_i is the loser.

The Nash and social welfare, profit and consumer surplus shares show similar qualitative features

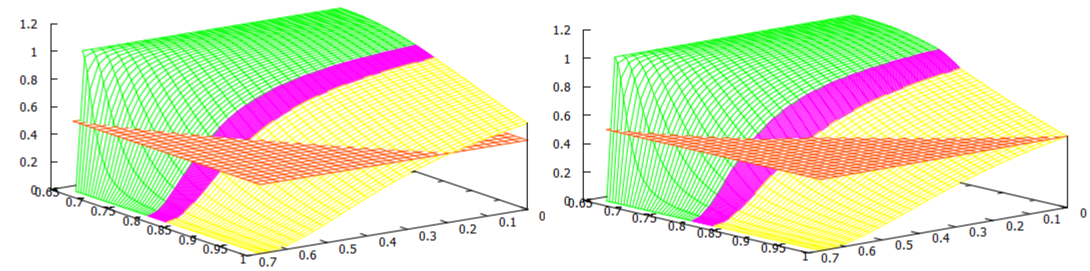

Figure 3.4: Left: The Nash Consumer Surplus share $ShCS_j^N$ of country X_j. Right: The Social consumer surplus share $ShCS_j^S$ of country X_j.

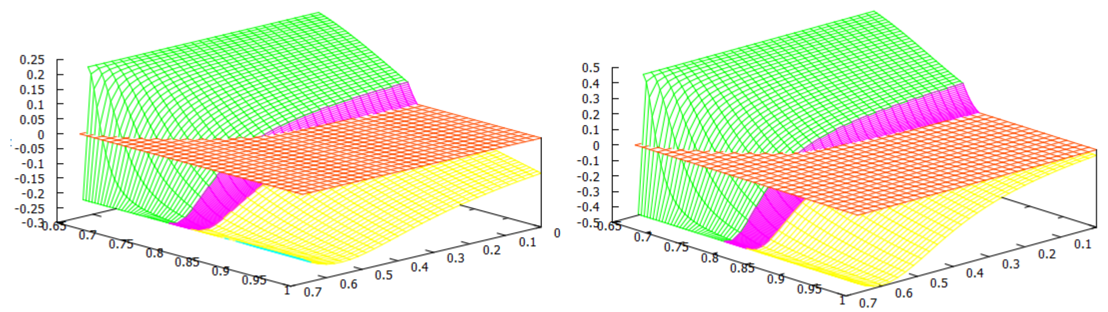

Figure 3.5: Left: The Nash-social welfare share difference ΔShW_j of country X_j. Right: The Nash-social profit share difference $\Delta Sh\pi_j$ of country X_j.

but significantly different quantitative properties (see figures (3.2),(3.3),(3.4). For all these shares: a) for low values of H_i the shares are higher for country X_j; b) for high values of H_i and high values of H_j, the shares are higher for country X_i; and c) for high values of H_i and low values of H_j, both cases occur. The isocurves of equal share (0.5) are close to segment lines starting at point $(2/3, 2/3)$ but finishing at different points: for the Nash welfare, it finishes close to the point $(2/3, 1)$; for the social welfare, it finishes close to the point $(0.5, 1)$; for the Nash consumer surplus, it finishes close to the point $(0.2, 1)$; for the social consumer surplus, it finishes close to the point $(0, 1)$; for the Nash profits, it finishes close to the point $(0.5, 1)$. For the custom revenue the situation is different, as the isocurve is not close to being a straight line and it finishes at point $(0, 7/9)$. The social profit share of country X_j is always at least 0.5, attaining its minimum 0.5 at the segment lines $H_i = 2/3$ and $H_j = 2/3$, and attaining its maximum 1 at the segment line with endpoints $(0, 5/6)$ and $(0, 1)$.

The Nash-social welfare share difference and the Nash-social profit share difference show similar qualitative features but significantly different quantitative properties (see figure (3.5)). All the isocurves (0) corresponding to equal Nash and social shares are in the prisoner's dilemma region or close to it and start at the point $(2/3, 2/3)$. However, the Nash-social welfare share difference isocurve (0) finishes close to $(0, 0.81)$, more precisely, between 0.81 and 0.815, and the Nash-social profit share difference isocurve (0) finishes close to $(0, 5/6)$. The Nash-social welfare share difference varies approximately between $-0,25$ and $0,25$, and the Nash-social profit share difference varies approximately between $-0,5$ and $0,5$. The Nash-social consumer surplus share difference is qualitatively different from the previous two share differences since it is always positive. The Nash-social consumer surplus share difference for country X_j

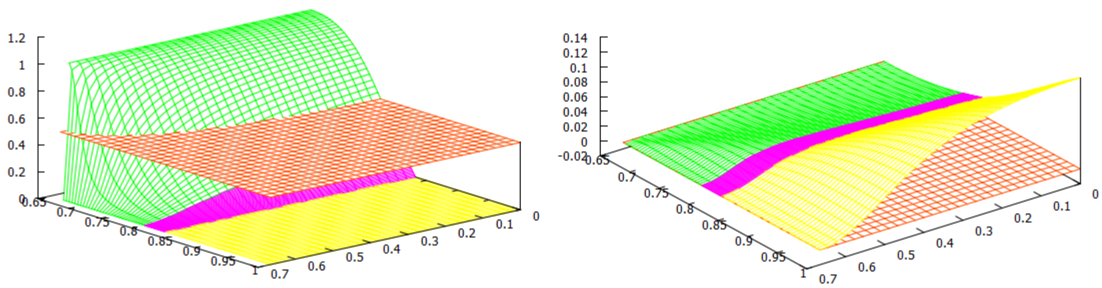

Figure 3.6: Left: The Nash custom revenue share $ShCR_j^N$ of country X_j. Right: The Nash-social consumer surplus share difference $\Delta ShCS_j$ of country X_j.

varies approximately between 0 and $0,15$. The $\mathbf{L}_j\mathbf{W}_i$ and the prisoner's dilemma **PD** regions can be noticed as corresponding to lower values of the Nash-social consumer surplus share difference for country X_j (see figure (3.6)).

3.7 Welfare balanced international trade agreements

In this section we indicate some of the positive and negative externalities of a welfare balanced international trade agreement between the two countries.

Let ΔW_T be the difference between the joint welfare computed at the social equilibrium and the joint welfare at the Nash equilibrium

$$\Delta W_T = W_T^S - W_T^N \ .$$

A γ-trade agreement determines the following γ-payoffs V_i and V_j for the countries X_i and X_j:

$$\begin{aligned} V_i &= W_i^N + \gamma \Delta W_T \\ V_j &= W_j^N + (1-\gamma)\Delta W_T \ . \end{aligned}$$

where γ is the *countries' bargaining power index* and ΔW_T is the *trade agreement welfare* gain.
Let the *trade agreement index* be

$$g = g(H_i, H_j) = \frac{W_T^S}{W_T^N} \ .$$

For the *welfare balanced bargaining power index*

$$\gamma^N = \frac{W_i^N}{W_T^N} \ ,$$

the γ^N-payoffs of the *welfare balanced* trade agreement are

$$\begin{aligned} V_i^N &= g W_i^N \\ V_j^N &= g W_j^N \ . \end{aligned}$$

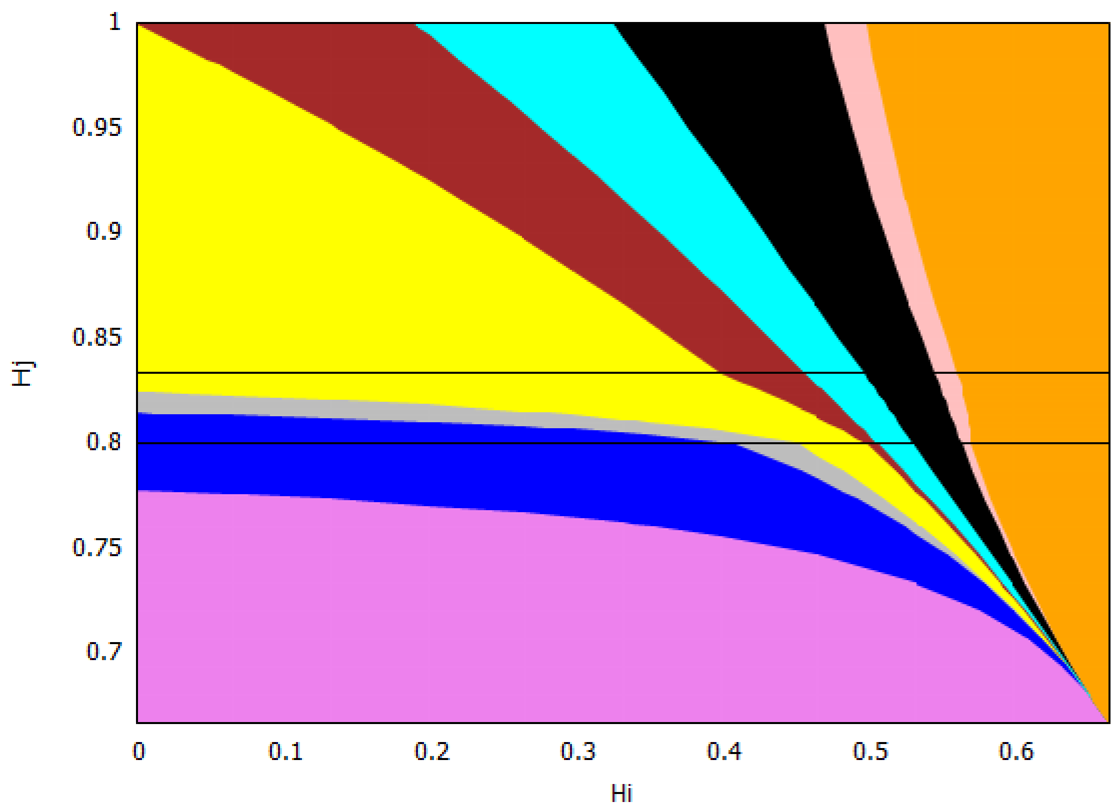

Figure 3.7: The Nash share isocurves and the social share isocurves (0.5) and the share difference isocurves (1) for welfare, profit, consumer surplus and custom revenue.

The total welfare of the trade agreement is the total social welfare:

$$V_i^N + V_j^N = W_T^S \ .$$

The *compensations* of the trade agreement in units of the welfare of the trade agreement, i.e., in total social welfare units, are the share differences

$$\frac{V_i^N - W_i^S}{W_T^S} = \Delta ShW_i \quad , \quad \frac{V_j^N - W_j^S}{W_T^S} = \Delta ShW_j \ .$$

The welfare balanced trade agreement has two sides: a) the two countries impose the welfare social tariffs; b) the two countries attain the joint welfare at the social optimum that is $g > 1$ times higher than the joint Nash equilibrium welfare, and the joint social welfare is split in a way such that both countries keep the same welfare shares that they had at the Nash equilibrium. Thus, the country with positive Nash-social welfare share difference must be indemnified by the other country and the amount of this compensation in units of the total welfare of the trade agreement (i.e., in units of joint social welfare) is determined exactly by the Nash-social welfare share difference.

The two countries enforce the social tariffs, so each country obtains the social profits, consumer surplus and custom revenue. This may cause some collateral effects in the economy of the countries. For instance, its produced quantities may change so that one of the countries is no longer the dominant force in terms of output. This influences other relevant economic quantities such as profits, consumer surplus and custom revenue. These effects, that we may call the externalities of the trade agreement, are decisive to whether or not the country signs the trade agreement, in the sense that the country may consider that the welfare compensation stated above is not sufficient to outweigh such collateral effects. Possible effects might include, for instance, unemployment raise in the country due to a decrease in production, and a subsequent wave of migration from that country to the other, or the effect of a decrease in the profits of firms, that may cause firms to invest abroad, also possibly triggering the problem of unemployment in the home country. Other consequences may be a fall of the consumer's surplus of one country relative to the other country, and a fall in revenues from using tariffs. Countries may try to mitigate and overcome these difficulties by including other features (that we do not explore in this work) in the trade agreement apart from the welfare compensation. For example, these features may be financing to the industry of the impaired country; R&D exchange between countries; compensation and investment in other sectors of the economy of the impaired country; *etc.*.

In light of this, we identify and analyse parameter regions where these difficulties occur. More precisely, these are the regions where the relevant economics we have considered above, such as profits, consumer surplus and custom revenue are such that the social shares are higher for one country while the Nash shares are higher for the other country. When this occurs, then there is a change in the dominant force with respect to that economic quantity, and these externalities are a disincentive for that country to sign the trade agreement, and might cause that country to not sign it, even in situations where he is compensated in welfare by the other country as ruled by the welfare balanced trade agreement. Other situations that may occur is that one of the countries may have a benefit in some economic quantity when the social tariffs are put into practice, but simultaneously have to to indemnify the other country. Thus, the country must perform a thorough analysis of the consequences to know if that benefit is worthy while compared to the obligation of compensating the other country.

In figure (3.7) we plot all the isocurves for the share differences (0), the Nash shares and the social shares (0.5). The violet region is delimited by the isocurve of the Nash custom revenue share. The blue region is delimited from the grey region by the isocurve of the Nash-social welfare share difference, and the grey region is delimited from the yellow region by the isocurve of the Nash-social profit share difference. From the yellow region we have the brown, cyan, black, pink and orange regions which are respectively delimited by: the isocurve of the social consumer surplus share; the isocurve of the Nash consumer surplus share; the isocurve of the Nash welfare share; the isocurve of the Nash profit share; the isocurve of the social welfare share.

The relation of the isocurves with the frontiers of the **LW** and **PD** regions presented in figure (3.1) is relatively complex. The violet and blue regions are completely contained in the union of the $\mathbf{L}_i\mathbf{W}_j$ and **PD** regions. The grey region intersects the three game type regions. The **PD** region is not contained in the grey region and below. In figure (3.8), we represent the decomposition of the **PD** region according to the violet, blue and grey regions, and in red is the portion of the **PD** that lies outside of the union of the violet, blue and grey regions. The upper frontier of the **PD** region enters through the grey region and finishes just above the frontier between the blue and grey regions. It always remains above the blue region since the blue region never intercepts the $\mathbf{L}_i\mathbf{W}_j$ region. Also, the violet region is almost totally inside the $\mathbf{L}_j\mathbf{W}_i$ region, except for a small portion that is of **PD** type, as represented in the left figure. Regarding the other regions, we first observe that the brown, cyan, black, pink and orange regions do not intercept the $\mathbf{L}_i\mathbf{W}_j$ regions. This means that in these regions, country X_i is always a winner in terms of absolute welfare. They do intercept the **PD** regions, although the parameter region for which

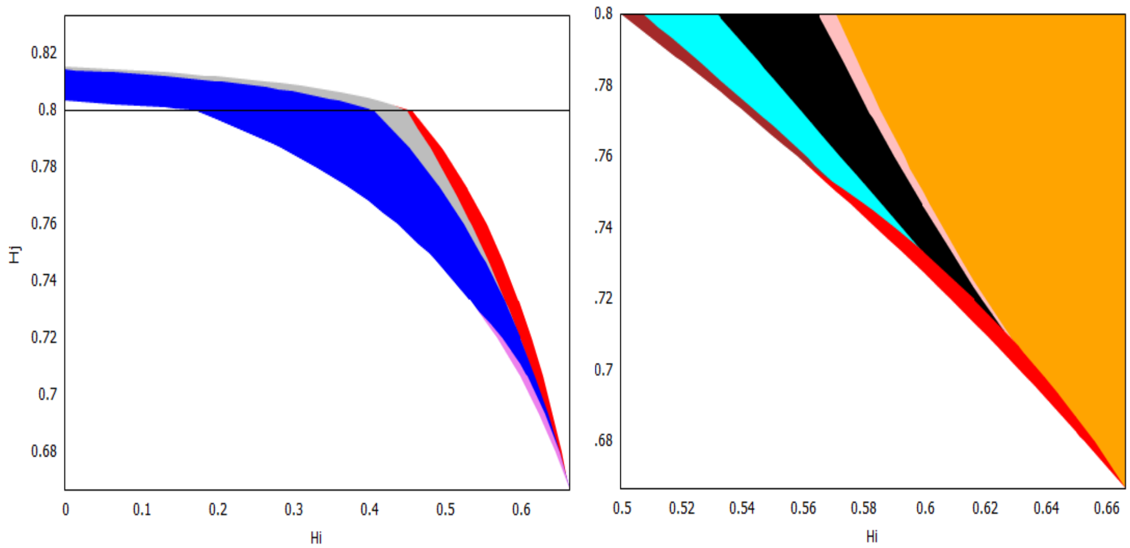

Figure 3.8: Left: The decomposition of the **PD** region according to the violet, blue and grey regions; Right: In red, the portion of the brown, cyan, black, pink and orange regions that intersects the **PD** region.

this occurs is very small, as we show in figure (3.8).

Since the trade agreement is based on the welfare of the two countries, we start by analysing the welfare. The Nash welfare share of country X_j is higher than that of X_i in and below the cyan region. The social welfare share is higher for country X_j except in the orange region. The Nash-social welfare share difference of country X_j is negative above the union of the violet and blue regions. Hence, in this region, by the welfare balanced trade agreement, country X_j must compensate X_i, but since the welfare balanced trade agreement maintains the Nash welfare share, he still has higher welfare share up to the cyan region. In the union of the black, pink and orange regions country X_i is compensated and retains its higher welfare share. The advantage of the welfare balanced trade agreement is that country X_i remains with a higher welfare share not only in the orange region but also in the pink and black regions where he had a higher Nash welfare share, but a lower social welfare share. In and below the blue region, country X_j has a positive Nash-social profit share difference and so is compensated by X_i and he has a higher welfare share. When the game type is $\mathbf{L}_i\mathbf{W}_j$ or $\mathbf{L}_j\mathbf{W}_i$ the country that has a loss in absolute welfare also has a loss in share and so is compensated by the other country in the welfare balanced trade agreement. When the game type is **PD**, so that both countries have an absolute gain, then the Nash-social share difference can be positive or negative, so country X_j may be asked to indemnify X_i or may be indemnified by X_i.

Regarding the custom revenue isocurves, we have that in the violet region country X_j has a higher Nash custom revenue share than country X_i. At the social optimum, country X_i always gives up its custom revenue in favour of going tax-free, which improves its consumer surplus in absolute terms, and in terms of share as we will see below. Country X_j does not apply tariffs in the region below the social tax-free threshold, but prefers to apply the maximal tax in the region above the social monopoly-tax threshold not allowing X_i to export both at the Nash and at the social optimum. It makes its tariff move from 0 to the maximal tariff in between these two regions, with this being the only region where there is a positive custom revenue at the social optimum for one of the countries, in the case, country X_j (and his share is 1). As a result, for values of H_j between these two thresholds, county X_j has some positive revenue, which may be important for his decision of signing or not the trade agreement, since it may be seen as an advantage against, for instance, the compensation that he must do to other country, which occurs outside the blue region, or the loss in share profit that occurs in the grey region, or in the brown region where his Nash consumer surplus share is higher than X_i's, but his social share is not. As noted before, the violet region is contained in the union of the $\mathbf{L}_j\mathbf{W}_i$ and **PD** regions, so, in absolute terms regarding welfare, country X_i is a winner, or equivalently, in the $\mathbf{L}_i\mathbf{W}_j$ region, country X_j has a lower Nash custom revenue share. We observe that it is possible for country X_j to have a lower share within the $\mathbf{L}_j\mathbf{W}_i$ region.

The Nash consumer surplus share of country X_j is higher than that of country X_i in and below the brown region. The Social consumer surplus share of country X_j is higher than that of X_i in and below the yellow region. We have observed previously that the consumer surplus increases for both countries when the social tariffs are applied. In spite of this increase in absolute terms, the Nash-social consumer surplus share difference of country X_j is always positive (see figure (3.6)), meaning that his share of consumer surplus at the Nash equilibrium is bigger than at the social optimum. This loss is, at most, approximately 0.14. This occurs mainly because country X_j exports more, since country X_i is tax-free at the social optimum, increasing the consumer surplus of X_i. Country X_i also exports more to X_j since the tariffs are lowered, but X_j only goes tax free for H_j below the tax-free threshold, so his consumer consumer surplus also increases, but his share diminishes.

For regions above the brown, country X_j already has a lower share than X_i at the Nash equilibrium, and will get an even lower share with the trade agreement. In the brown region, X_j has a higher share at the Nash equilibrium, but no longer has advantage at the social optimum. In the yellow regions and below, country X_j has a bigger share at the Nash equilibrium, and in spite of the loss in share, he still has a higher social share. In the regions where X_j has a lower social share, that is, in and above the brown regions, the trade agreement rules that he indemnifies X_i. This might be a difficulty for country X_j, that might have to use his custom revenue (that is positive for H_j between the two thresholds as we noted above), or somehow use the increase in the profits share in these regions (which occurs since they lie above the grey region), possibly through taxation, as well the fact that its firm always has the most part of the joint profits of the two countries at the social optimum.

Regarding this last observation, we now analyse the profits shares. The Nash profit share of the country X_j is higher than that of country X_i in and below the black region and it is lower in the union of the pink and orange regions. The social profit share of country X_j is always bigger for country X_j (see figure (3.3)). The Nash-social profit share difference of country X_j is always negative above the grey region, so the profit share increases with the use of social tariffs. For this region the game type is $\mathbf{L}_i\mathbf{W}_j$ except for the small red region in figure (3.8) where the game type is **PD**, so X_j always has a gain in absolute welfare. In these regions the trade agreement presents an advantage for the firm of country X_j, since it increases its profit share. In the yellow, brown, cyan and black regions, country X_j has a higher profit share and reinforces its position with the trade agreement, getting a higher share. The advantage for the firm is also evident in the pink and orange regions, where the firm of X_j is not the dominant firm in terms of Nash profit share, but does become dominant with the trade agreement. However, in these regions, country X_j has to compensate country X_i according to the trade agreement. Because of this, country might have to impose taxes on the profits of the firms to fulfil the trade agreement.

In the yellow, brown and cyan regions, after the compensation, X_j has a higher welfare share. In the black, pink and orange regions, he has no that advantage in welfare share. In and below the grey region, country X_j has a decrease in profit share, but always has a higher social share. In the blue and violet regions country X_j is indemnified by country X_i. This compensation may be used by the country to compensate the loss of the profit share of its firm by investing in it. In the grey region, the opposite occurs, and country X_j has to indemnify X_i. In this case he may need to use its custom revenue, or tax the profits to indemnify X_i. This might prove difficult for the country since in the grey region the firm faces a decline in its profit share (although he is still the dominant firm).

We observe that in the $\mathbf{L}_j\mathbf{W}_i$ region, meaning that country X_j is a loser in absolute terms, he has a higher Nash welfare, profit and consumer surplus shares. In fact, this occurs for a wider region, since when country X_j is indemnified by X_i (which occurs below the blue region), he has an advantage in Nash shares in all these quantities: welfare, profits and consumer surplus. In these regions, country X_j also has a higher social profit and consumer surplus shares, so, with the use of the social tariffs, he also is the dominating country in consumer surplus and has the dominating firm. If the trade agreement is not balanced and just applies the social tariffs, he would remain as the dominating country in terms of welfare, since his social welfare share is higher than X_i's in this region. Hence, country X_i (which is the country with greatest decline in home production in a tax-free situation, since H_i is lower than X_j) can't simultaneously be a winner and have higher shares. This advantage of X_j in all shares occurs up to the frontier of yellow and brown regions. But when the game type is $\mathbf{L}_i\mathbf{W}_j$ then the Nash and social welfare, consumer surplus shares, and the Nash profit share of country X_j may be higher or lower than the ones of X_i. These shares are higher for X_i for higher values of H_i and lower when H_i is lower. When H_j increases, then H_i does not need to be so high in order to assure higher shares for country X_i. For the Nash custom revenue share the situation is different, since in the $\mathbf{L}_j\mathbf{W}_i$ and $\mathbf{PD}$ regions country X_i can have a higher share than country X_j. However, in the $\mathbf{L}_i\mathbf{W}_j$ region, he always has the lower Nash custom revenue share.

If the Nash-social welfare share difference or the Nash-social consumer surplus share difference or the Nash-social profit share difference is large then the countries have to be very careful in making a trade agreement because small differences in the trade agreement can mean significant social and economic changes for the countries. For values of (H_i, H_j) in the prisoner's dilemma region or close to the prisoner's dilemma region, the Nash-social welfare share difference, the Nash-social consumer surplus share difference and the Nash-social profit share difference have lower values than away from the prisoner's dilemma region (see figures (3.5), (3.6)). When H_j gets closer to $2/3$ the welfare compensation that X_i has to give to X_j gets higher, making the agreement more risky for X_j. We note that one of the countries might have to pay up to 25% of the joint social welfare to the other government, making the agreement very relevant and difficult to establish.

By plotting the values of the trade agreement index g we see that it attains lower values for values of (H_i, H_j) in the prisoner's dilemma region or close to the prisoner's dilemma region and it attains higher values away from the prisoner's dilemma region (see figure (3.9)). The trade agreement index attains its maximum (approximately 1.22) near the point $(0.37, 1)$. Hence, in the region where the game is of prisoner's dilemma type, or where it is close to prisoner's dilemma, albeit all difficulties that may arise to achieve a welfare balanced trade agreement, the gain will be lower than in other regions away from the prisoner's dilemma region. Furthermore, in most part of the union of the violet and blue regions (i.e., the region where country X_j is indemnified in its welfare by country X_i) country X_j is also the

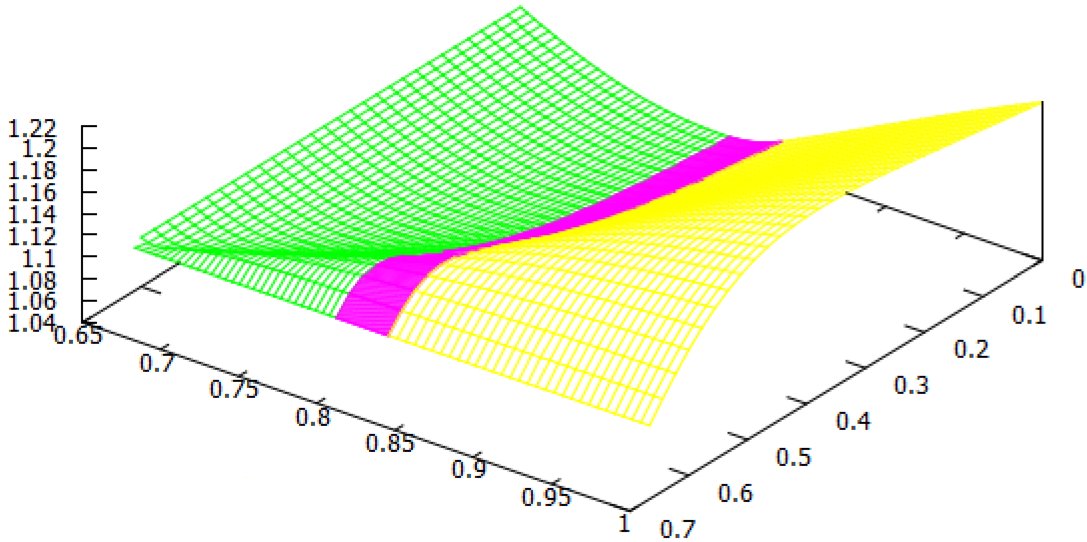

Figure 3.9: The Trade Agreement Index g.

loser, so that he might get greedy (and even more greedy in the violet region because of its higher Nash custom revenue share) and prefer a bargaining power index that is more favourable to him, i.e., a γ smaller than g, rendering the welfare balanced trade agreement unstable. However, when H_j gets closer to 2/3, the value of g can increase slightly, thus improving the advantage of the agreement and perhaps work as an incentive to its enforcement.

3.8 Conclusions

We have considered an international trade model with two countries, where the governments of each country choose whether or not to impose tariffs in the import. We have considered a two-stage game, where in the first stage, governments choose their tariffs and in the second stage, firms in each country competitively choose their home and export quantities. For every pair of tariffs, we found the Nash equilibria for the second sub-game. For the first sub-game the governments can choose competitive (Nash) tariffs or social (cooperative) tariffs. We considered different economic quantities as the utility of the governments, namely, total quantities in the market of a country, total quantities produced by the firms, prices, profits of the firms, consumer's savings, custom revenues of the countries and the welfare of the countries. For each utility we have classified the game according to the social and Nash tariffs of the governments. For the welfare of the countries we proved that the outcome of the game is either a prisoner's dilemma or a lose-win dilemma. This classification for the different utilities and for the welfare suggests where the difficulties in establishing trade agreements may appear, since in a loose-win dilemma the losing country has to be somehow compensated by means of clauses in trade agreement in order to accept the enforcement of the social tariffs. Furthermore, a trade agreement might present some externalities that need to be considered. We considered a welfare balanced trade agreement between the two countries, where each countries maintains its competitive (Nash) shares when the social tariffs are enforced. The gain of each country is proportional to the trade agreement index that we have explicitly computed and analysed. Even with a welfare balanced trade agreement, some important negative externalities arise, related to the productive capacity of the country, whether in profits or quantities produced, or in the consumer surplus. We discussed some of these negative externalities that present major difficulties for the establishment of the welfare balanced trade agreement by both parties, and the parameter regions where they occur. Regarding the externalities, the following questions, among others, can be raised about the trade agreement: a) what additional measures should be part of the trade agreement to mitigate the negative externalities of the country that has a decrease in his production and/or the profits of his firm, his custom revenue or the surplus of his consumers. These measures may include: R&D swap between both countries, financing to industry and financing in other economical sectors, among others. We note that countries that can make balanced trade agreements in more than one sector such that the total compensation of the agreements is not relevant might be in a better position to negotiate; b) what additional measures should be part of the trade agreement to force both countries to agree to set the social tariffs in such a way that the agreement is theoretically durable and sustainable in time, preferably rendering it self-enforcing.

Future work can consist for instance in introducing some of the features mentioned above, such as studying conditions for the self-enforcing the agreement, ideally rendering it durable and stable in time, to study the effects of the swap of R&D between the two firms to decrease their production costs, or to include the effects of subsidies, fines, price dumping and merging and shut-down of firms. The inclusion of one or more of these features into the trade agreement may be a way to overcome some of the externality effects that may arise.

Chapter 4

The fundamental bifurcation for evolutionary matrix models with multiple traits

This chapter is based on the joint work:

J. M. Cushing, A. A. Pinto F. Martins, and A. Veprauskas. A bifurcation theorem for evolutionary matrix models with multiple traits. *The Journal of Mathematical Biology*, 75(2):491–520, August 2017.

In this chapter we consider an evolutionary game theoretic version of a general nonlinear matrix model that includes the dynamics of a vector of mean phenotypic traits subject to natural selection. We prove a fundamental bifurcation theorem for this evolutionary model when the projection matrix is primitive by showing the existence of a continuum of positive equilibria that bifurcates from the extinction equilibrium as the inherent population growth rate passes through 1, in which case the extinction equilibrium loses stability. We also study the stability of the bifurcating equilibria by relating it to the direction of the bifurcation, which is forward (backward) if, near the bifurcation point, the positive equilibria exist for inherent growth rates greater (less) than 1. We obtain that forward bifurcations are stable whereas backward bifurcations are unstable. We apply the results to an evolutionary version of a modified Ricker model with an added Allee component. This application illustrates the theoretical results and, in addition, several other interesting dynamic phenomena, such as period-doubling bifurcations, and backward bifurcation induced strong Allee effects, i.e. coexistence of a stable extinction equilibrium with a stable positive (survival) equilibria attractor or with positive (survival) non-equilibrium attractors such as cycles or more complicated attractors.

This chapter is structured in the following way. We start by discussing the problem in the non-evolutionary setting in section 4.1. We provide the results regarding the fundamental bifurcation of the model in this setting. In section 4.2 we describe the evolutionary model. We then study the stability of the extinction equilibrium in section 4.3 and in Section 4.4 we determine the nature of the fundamental bifurcation, i.e. the bifurcation that occurs when extinction stability is lost. In Section 4.5 an application is made to an evolutionary version of a modified Ricker model with an added Allee component (low density positive feedback effect). We finish with some concluding remarks in section 4.6.

4.1 A bifurcation theorem for the non-evolutionary matrix model.

We consider a discrete time model

$$\hat{x}(t+1) = P(\hat{x}(t))\hat{x}(t), \quad t \in \mathbb{N}_0 = \{0, 1, 2, \dots\} \tag{4.1}$$

for the dynamics of a biological population whose individuals are classified into a finite number of discrete classes. Here $\hat{x} : \mathbb{N}_0 \to \overline{\mathbb{R}}_+^m$ is a sequence of m-dimensional column vectors consisting of class specific population densities, where $\overline{\mathbb{R}}_+^m$ is the closure of the positive cone $\mathbb{R}_+^m$ in m-dimensional Euclidean space $\mathbb{R}^m$. Recursive formulas (4.1), called matrix models, are widely utilized to describe the dynamics of populations in which individuals are classified according to age, size, life cycle stage, spatial location, genetic composition, *etc.*, indeed virtually any classification scheme of interest [19, 23].

The entries $p_{ij}(\hat{x})$ of the projection matrix $P(\hat{x})$ are chosen by a modeller to describe class-specific, per capita (individual) birth and survival rates and to account for transitions of individuals from one class to another. As indicated, these entries can be dependent on the densities in the demographic vector $\hat{x}$, dependencies that make the dynamic model nonlinear. Classic examples of matrix models for structured population dynamics include the age, size, and stage structured models of Leslie and Lewis [60, 61, 63], Usher [98], and Lefkovitch [59].

Of fundamental importance to a biological population is its avoidance of extinction. We refer to the equilibrium $\hat{x} = \hat{0}$ solution of (4.1) as the *extinction equilibrium*. If the extinction equilibrium is an attractor, then the population is threatened with extinction. This leads to the study of the stability properties (local and global) of the extinction equilibrium. An equilibrium $\hat{x}$ is *locally stable* if given any $\varepsilon > 0$ there exists a $\delta > 0$ such that for any initial condition satisfying $|\hat{x}(0) - \hat{x}| < \delta$ it follows that the solution satisfies $|\hat{x}(t) - \hat{x}| < \varepsilon$ for all $t \in N_0$. An equilibrium is a *local attractor* if there exists a $\delta_0 > 0$ such that $|\hat{x}(0) - \hat{x}| < \delta$ implies $\lim_{t\to+\infty} \hat{x}(t) = \hat{x}$. An equilibrium is *locally asymptotically stable* if it is both locally stable and a local attractor. Throughout this chapter, stable (or stability) means local asymptotically stable (or local asymptotic stability). Unstable means not locally asymptotically stable. The linearisation principle [37] leads one to consider the eigenvalues of the Jacobian obtained from (4.1) evaluated at the extinction equilibrium, which is the *inherent projection matrix* $P(\hat{0})$ (inherent means density free). If all eigenvalues of $P(\hat{0})$ lie in the complex unit circle, then the extinction equilibrium is locally asymptotically stable, which threatens the model population with (asymptotic) extinction. If at least one eigenvalue is outside the complex unit circle, then the extinction equilibrium is unstable, which opens the possibility of population persistence. The nature of the bifurcation that occurs when the extinction equilibrium loses stability forms a fundamental bifurcation theorem in population dynamics. We describe this theorem below (Theorem 4.1). For a good introduction to bifurcation theory with applications, and also a description of some usual methods used in the theory, such as the Lyapunov-Schmidt method that we will make use of in our main results, we refer the interested reader to [52].

We make the following assumptions on the entries $p_{ij}(\hat{x})$ in the projection matrix $P(\hat{x})$. Let $\Omega \subseteq \mathbb{R}^m$ denote an open neighbourhood of $\hat{0} \in \mathbb{R}^m$ and $C^2(\Omega \to \overline{\mathbb{R}}_+)$ denote the set of twice continuously differentiable functions that map Ω to $\overline{\mathbb{R}}_+$.

H1: *$P(\hat{x}) = [p_{ij}(\hat{x})]$ is primitive for all $\hat{x} \in \Omega$ and $p_{ij} \in C^2(\Omega \to \overline{\mathbb{R}}_+)$.*

Recall that a nonnegative matrix (i.e. one all of whose entries are nonnegative) is primitive if it is irreducible and has a *strictly* dominant eigenvalue. Another common definition of primitivity is that there is a power of the matrix such that all its entries are positive. Perron-Frobenius theory implies that the spectral radius $\rho[A]$ of a primitive matrix A is a strictly dominant, positive and simple eigenvalue which possesses a positive eigenvector in $\mathbb{R}^m_+$. Moreover, no other eigenvalue has a nonnegative eigenvector, i.e. an eigenvector in $\overline{\mathbb{R}}^m_+$. See [12]. We denote the strictly dominant eigenvalue of $P(\hat{x})$ by

$$r(\hat{x}) := \rho\left[P\left(\hat{x}\right)\right].$$

Observe that $r(\cdot) \in C^2(\Omega \to \mathbb{R}^1_+)$. The number $r(\hat{0})$ is the *inherent growth rate* of the population (the growth rate in the absence of density effects). For notational simplicity we denote this number by

$$r_0 := r(\hat{0}).$$

For our purposes, we normalize the entries of P in a way that

$$P(\hat{x}) = r_0 Q(\hat{x})$$

where the normalized matrix $Q(\hat{x}) = [q_{ij}(\hat{x})]$ satisfies H1 and

$$\rho[Q(\hat{0})] = 1.$$

Then the matrix equation (4.1) becomes

$$\hat{x}(t+1) = r_0 Q(\hat{x}(t))\hat{x}(t), \quad t \in \mathbb{N}_0. \tag{4.2}$$

We denote the entries of the matrix $Q(\hat{x})$ by $q_{ij}(\hat{x})$.

Definition 4.1. *We say that a pair $(r_0, \hat{x}) \in R \times \Omega$ is an equilibrium pair of (4.2) (or equivalently of (4.1)) if $\hat{x} = r_0 Q(\hat{x})\hat{x}$. Observe that $(r_0, \hat{0})$ is an equilibrium pair for every $r_0 \in R$; we call $(r_0, \hat{0})$ an extinction equilibrium pair. An equilibrium pair $(r_0, \hat{x})$ is a positive equilibrium pair if $\hat{x} \in R^m_+$ and it is stable if $\hat{x}$ is a locally asymptotically stable equilibrium of (4.2) (equivalently (4.1)).*

We need the quantity

$$\kappa := -\hat{w}_L^T\left[\nabla^0_{\hat{x}} q^T_{ij} \hat{w}_R\right]\hat{w}_R$$

where T denotes transposition, the gradient $\nabla_{\hat{x}}$ *of* $q_{ij}(\hat{x})$ with respect to $\hat{x}$ is a column m-vector, and $\nabla^0_{\hat{x}} q^T_{ij}$ denotes the transpose of the gradient *evaluated at the bifurcation point* $(r_0, \hat{x}) = (1, \hat{0})$. With this superscript notational convention, we can equivalently write

$$\kappa = -\hat{w}_L^T\left[\nabla^0_{\hat{x}} p^T_{ij} \hat{w}_R\right]\hat{w}_R. \tag{4.3}$$

Here the vectors $\hat{w}_L^T$ and $\hat{w}_R$ are the *(positive) left and right eigenvectors* of $Q(\hat{0})$ (equivalently of $P(\hat{0})$ when $r_0 = 1$) associated with eigenvalue 1, normalized so that

$$\hat{w}_L^T \hat{w}_R = 1.$$

Note that $[\nabla^0_{\hat{x}} p^T_{ij} \hat{w}_R]$ is an $m \times m$ matrix. The derivative $\partial^0_{x_k} p_{ij}$ measures the effect that an increase in the density of class k has on the entry p_{ij} of the population projection matrix P (at low population density). The number $\hat{w}_L^T \left[\nabla_{\hat{x}}^0 p_{ij}^T \hat{w}_R\right] \hat{w}_R$ is a weighted sum (with positive coefficients) of all density effects on all entries p_{ij}. This number therefore represents a summary measure of the effects that (low level) class densities has on the population (as does κ, the minus sign being introduced only for notational convenience in Theorem 4.1.)

From the linearisation principle and from Theorems 1.2.4 and 1.2.5 in [23] we have the following result.

Theorem 4.1. *Assume the matrix $P(\hat{x})$ in (4.1) satisfies H1.*

(a) The extinction equilibrium $(r_0, \hat{0})$ is stable for $r_0 < 1$ and is unstable for $r_0 > 1$.

(b) There exists a continuum $\mathcal{C}$ of positive equilibrium pairs $(r_0, \hat{x}) \in \mathbb{R}_+ \times \mathbb{R}_+^m$ of the matrix equation (4.1) which bifurcates from $(1, \hat{0})$ (i.e. contains the extinction pair $(1, \hat{0})$ in its closure). Near the bifurcation point, the positive equilibrium pairs on $\mathcal{C}$ have the parameterization

$$\hat{x}(\varepsilon) = \hat{w}_R \varepsilon + \mathcal{O}(\varepsilon^2)$$
$$r_0^*(\varepsilon) = 1 + \kappa\varepsilon + \mathcal{O}(\varepsilon^2)$$

for $\varepsilon \gtrapprox 0$.

(c) We say the bifurcation of positive equilibria is forward (respectively, backward) if, in a neighbourhood of $(1, \hat{0})$, the positive equilibrium pairs on $\mathcal{C}$ are such that $r_0 > 1$ (respectively, $r_0 < 1$). If $\kappa > 0$ then the bifurcation of $\mathcal{C}$ at $(1, \hat{0})$ is forward and the equilibrium pairs on $\mathcal{C}$ in a neighbourhood of $(1, \hat{0})$ are (locally asymptotically) stable. If $\kappa < 0$ then the bifurcation is backward and the equilibrium pairs on $\mathcal{C}$ in a neighbourhood of $(1, \hat{0})$ are unstable.

Note how, in this Theorem, the direction of the bifurcation determines the stability of the bifurcating equilibria. A forward bifurcation, occurring when the extinction equilibrium loses its stability as r_0 increases through 1 (removing the threat of extinction), creates stable positive (non-extinction) equilibrium states.

Theorem 4.1 asserts stability or instability of the bifurcating positive equilibria $\mathcal{C}$ locally only, i.e. for equilibrium pairs on $\mathcal{C}$ near the extinction equilibrium $(1, \hat{0})$ only. However, the continuum $\mathcal{C}$ is known to exist globally in the sense that it connects to the boundary of the set on which the matrix model is defined, i.e., it connects to the set $\{+\infty\} \times (\partial\Omega \cap \mathbb{R}_+^m)$, where $\partial\Omega$ denotes the boundary of Ω . In most applications, Ω includes the closure $\overline{\mathbb{R}}_+^m$ of the positive cone, which implies that either the component r_0 is unbounded or the norm $|\hat{x}|$ is unbounded in $\mathbb{R}_+$ (or both). When r_0 is unbounded we have that there is at least one non-extinction equilibrium for each $r_0 > 1$ [23, 28].

A derivative $\partial_{x_k} p_{ij}$ is often negative in population models because of an assumption that an increase in density x_k will have a deleterious effect on some vital rate (birth rate, survival probability, growth rate, metabolic rates, and so on). These kinds of negative feedback phenomena are common in population models that describe density regulation mechanisms for population growth. If all the derivatives $\partial_{x_k}^0 p_{ij}$ are negative (or zero), that is to say, if all density effects in a model are negative feedback effects, then clearly $\kappa > 0$ and the bifurcation of the continuum $\mathcal{C}$ is forward and hence stable.

A positive derivative $\partial_{x_k}^0 p_{ij}$ is called a *component Allee effect* [22]. Clearly, the existence of a component Allee effect is necessary for a backward bifurcation (i.e. for $\kappa < 0$). If all component Allee effects are sufficiently large so that $\kappa < 0$, then the bifurcation of positive equilibria at $(1, \hat{0})$ is backward and hence unstable. A common occurrence in this case is the creation of a *strong Allee effect*, i.e. the presence of two attractors, one of which is an extinction equilibrium and the other of which is positive.

Thus, population survival is initial condition dependent. This scenario can only occur when $r_0 < 1$ and the extinction equilibrium is stable. A backward bifurcation does not create a stable positive equilibrium, however. A strong Allee effect usually arises in models with backward bifurcations. This is because it is usually assumed that negative feedback effects predominate at high densities (even if they do not at low densities) which has the consequence of "turning" the continuum $\mathcal{C}$ around at a critical (saddle-node bifurcation) value of $r_0 < 1$ with a concomitant stabilization of the positive equilibria. We will not pursue this phenomena here, which occurs outside a neighbourhood of the bifurcation point. See [26].

In the matrix model (4.1), the vital rates and transitions modelled by the entries $p_{ij}(\hat{x})$ of the projection matrix $P(\hat{x})$ change temporally only due to changes in the demographic vector $\hat{x} = \hat{x}(t)$. There are, of course, numerous other reasons why these vital rates and transitions might change in time, for example, they might fluctuate randomly due to demographic or environmental stochasticity or periodically due to regular environmental oscillations (seasonal, monthly or daily fluctuations). Another reason these vital rates and transitions might change in time is that they are subject to selective pressures from Darwinian evolution. Our goal in this chapter is to investigate an extension of the fundamental bifurcation theorem for the non-evolutionary model (4.1), as given in theorem 4.1 above, to an evolutionary game theoretic version of (4.1). modelling the dynamics of natural selection from Darwinian evolution.

More precisely we extend the fundamental bifurcation Theorem 4.1 to an evolutionary version of the matrix model (4.2) under the assumption that the projection matrix depends on a suite of phenotypic traits subject to natural selection. This generalizes the results in [24] where models with only a single trait are considered.

4.2 Darwinian dynamics with multiple evolutionary traits

We consider an evolutionary version of the matrix model (4.1) developed in [100]. In that modeling methodology a (focal) individual's vital rates, as described by the entries of the projection matrix, are influenced by a collection of scalar traits $\hat{v} = (v_1, \ldots, v_n)^T$ and the population means of these traits $\hat{u} = (u_1, \ldots, u_n)^T$. By this assumption, an individual's fitness depends on both its own suite of traits $\hat{v}$ and the traits possessed by other individuals in the population $\hat{u}$ (frequency dependence). We indicate this by the notation $P(\hat{x}, \hat{u}, \hat{v})$, which in turn implies that the spectral radius of $P(\hat{x}, \hat{u}, \hat{v})$ is also dependent on $\hat{v}$ and $\hat{u}$:

$$r(\hat{x}, \hat{u}, \hat{v}) := \rho[P(\hat{x}, \hat{u}, \hat{v})].$$

Darwinian dynamics track the dynamics of the structured population $\hat{x}(t)$ and the vector of population mean traits $\hat{u}\,(t) = (u_1(t), \ldots, u_n(t))^T$, the latter by means of the assumption that changes in the mean trait are proportional to the fitness gradient of the focal individual [1, 31, 2, 57, 58, 73, 100]. We extend the resulting evolutionary matrix model, as found in [100], to include a vector of traits $\hat{v} = (v_1, \ldots, v_n)^T$. Different fitness functions can be found throughout the literature, but the most common choice is the exponential growth rate $\ln r$ [83].

Another choice used by some researchers is the *net reproduction number* $R_0(\hat{x}, \hat{u}, \hat{v})$, which is widely used in epidemiological models. In many applications it is quite common that the net reproduction number is easier to compute than the inherent growth rate and that there are tractable analytical formulas for R_0 in terms of the entries of the projection matrix, which is not always the case for r. However, for the net reproduction number to be well defined some extra condition must be imposed in the model. In this work we will use the inherent growth rate $r(\hat{x}, \hat{u}, \hat{v})$, but by the results in [25], our results remain unchanged if, when the net reproduction number is well-defined, $r(\hat{x}, \hat{u}, \hat{v})$ is replaced by $R_0(\hat{x}, \hat{u}, \hat{v})$.

The model equations for the coupled population and trait dynamics provided by evolutionary game theory are [73, 100]

$$\hat{x}(t+1) = P(\hat{x}, \hat{u}, \hat{v})|_{(\hat{x},\hat{u},\hat{v})=(\hat{x}(t),\hat{u}(t),\hat{u}(t))}\, \hat{x}(t) \tag{4.4}$$

$$\hat{u}(t+1) = \hat{u}(t) + M\ \nabla_{\hat{v}} \ln r(\hat{x}, \hat{u}, \hat{v})|_{(\hat{x},\hat{u},\hat{v})=(\hat{x}(t),\hat{u}(t),\hat{u}(t))} \tag{4.5}$$

where $M = (\sigma_{ij})$ is a symmetric $n \times n$ variance-covariance matrix for trait evolution and *the gradient* $\nabla_{\hat{v}} \ln r(\hat{x}, \hat{u}, \hat{v})$ *is a column n-vector*, whose i^{th} entry is

$$\partial_{v_i} \ln r(\hat{x}, \hat{u}, \hat{v}) := \frac{\partial \ln r(\hat{x}, \hat{u}, \hat{v})}{\partial v_i}\ .$$

The entry σ_{ij} of M, $i \neq j$, is the covariance of the i^{th} phenotypic trait and the j^{th} phenotypic trait. The diagonal entries

$$\sigma_i^2 := \sigma_{ii} \geq 0$$

are the variances of the i^{th} trait (from its mean u_i) occurring in the population at each time t (which are assumed constant). We assume the usual conditions for a covariance matrix, namely that M is positive semi-definite and symmetric. Recall that M is a positive semi-definite matrix if $\hat{v}^T M \hat{v} \geq 0$ for all $\hat{v} \in \mathbb{R}^m$, and that M is symmetric if $\sigma_{ij} = \sigma_{ji}$ for all $1 \leq i, j \leq m$. If the matrix M is the null matrix, then no evolution occurs and $\hat{u}(t)$ remains constant for all t. In this case Theorem 4.1 holds when applied to (4.4) with the mean trait $\hat{u}(t) \equiv \hat{u}(0)$ held fixed.

We write (4.4) and (4.5) as

$$\hat{x}(t+1) = P(\hat{x}(t), \hat{u}(t), \hat{u}(t))\hat{x}(t) \tag{4.6a}$$

$$\hat{u}(t+1) = \hat{u}(t) + M\ \nabla_{\hat{v}} \ln r(\hat{x}\,(t)\,, \hat{u}\,(t)\,, \hat{u}\,(t)) \tag{4.6b}$$

where we use the simplifying notation

$$P(\hat{x}(t), \hat{u}(t), \hat{u}(t)) := P(\hat{x}, \hat{u}, \hat{v})|_{(\hat{x},\hat{u},\hat{v})=(\hat{x}(t),\hat{u}(t),\hat{u}(t))}$$

$$\nabla_{\hat{v}} \ln r(\hat{x}(t), \hat{u}(t), \hat{u}(t)) := [\nabla_{\hat{v}} \ln r(\hat{x}, \hat{u}, \hat{v})]|_{(\hat{x},\hat{u},\hat{v})=(\hat{x}(t),\hat{u}(t),\hat{u}(t))}\ .$$

Remark 4.1. *We will need to differentiate functions of the three variables $(\hat{x}, \hat{u}, \hat{v})$ after letting $\hat{v} = \hat{u}$ with respect to the components u_i of $\hat{u}$ and from them construct gradients and Jacobians with respect to $\hat{u}$. Such a derivative is the sum of the partial derivatives with respect to u_i and v_i. For example, the derivative of $r(\hat{x}, \hat{u}, \hat{u}) := r(\hat{x}, \hat{u}, \hat{v})|_{\hat{v}=\hat{u}}$ with respect to u_i is*

$$\frac{\partial}{\partial u_i}\left[r(\hat{x}, \hat{u}, \hat{v})|_{\hat{v}=\hat{u}}\right] + \frac{d}{dv_i}\left[r(\hat{x}, \hat{u}, \hat{v})|_{\hat{v}=\hat{u}}\right]$$

which we write as

$$\frac{\partial r(\hat{x}, \hat{u}, \hat{u})}{\partial u_i} + \frac{\partial r(\hat{x}, \hat{u}, \hat{u})}{\partial v_i}.$$

With this notation, the gradient of $r(\hat{x}, \hat{u}, \hat{u})$ with respect to the components u_i of $\hat{u}$ constructed from *these partial derivatives is*

$$\nabla_{\hat{u}} r(\hat{x}, \hat{u}, \hat{u}) + \nabla_{\hat{v}} r(\hat{x}, \hat{u}, \hat{u}).$$

Let V be an open connected set in $\mathbb{R}^n$ and let $\Omega \subseteq \mathbb{R}^m$ be an open set containing the origin $\hat{0} \in R^m$. We assume the following about the projection matrix $P(\hat{x}, \hat{u}, \hat{v})$ and the variance-covariance matrix M.

H2. *$P(\hat{x}, \hat{u}, \hat{v})$ is primitive for $(\hat{x}, \hat{u}, \hat{v}) \in \Omega \times V \times V$, $p_{ij} \in C^2(\Omega \times V \times V \to \overline{\mathbb{R}}_+)$, $p_{ij}(\hat{x}, \hat{u}, \hat{v}) = \tilde{p}_{ij}(\hat{v})\bar{p}_{ij}(\hat{x}, \hat{u}, \hat{v})$ such that $\bar{p}_{ij}(\hat{0}, \hat{u}, \hat{v}) \equiv 1$, and M is invertible.*

Remark 4.2. *The assumption on p_{ij} in H2 implies that trait frequency dependence has no effect in the absence of density effects.* ***Specifically,*** *$p_{ij}(\hat{0}, \hat{u}, \hat{v}) = \tilde{p}_{ij}(\hat{v})$. A mathematical implication of this assumption is that all derivatives of $p_{ij}(\hat{0}, \hat{u}, \hat{v})$ with respect to components u_i of $\hat{u}$ are identically equal to 0 for all $\hat{v}$:*

$$\nabla_{\hat{u}} p_{ij}(\hat{0}, \hat{u}, \hat{v}) \equiv 0_n. \tag{4.7}$$

This means the inherent projection matrix $P(\hat{0}, \hat{u}, \hat{v})$ is independent of $\hat{u}$ and hence so is its dominant eigenvalue $r(\hat{0}, \hat{u}, \hat{v})$. Thus $\nabla_{\hat{u}} r(\hat{0}, \hat{u}, \hat{v}) \equiv 0_n$ for all $\hat{v}$ hence $\nabla_{\hat{u}} \left[r(\hat{0}, \hat{u}, \hat{v})\big|_{\hat{v}=\hat{u}} \right] \equiv 0_n$. Using the notation convention in Remark 4.1 we have

$$\nabla_{\hat{u}} r(\hat{0}, \hat{u}, \hat{u}) \equiv 0_n. \tag{4.8}$$

Remark 4.3. *The assumption on M in H2, that it is invertible, is for example satisfied if traits are not strongly correlated.*

Our approach is to consider the bifurcation of equilibria from an *extinction equilibrium* which, by definition, is an equilibrium $(\hat{x}, \hat{u}) = (\hat{0}, \hat{u})$ of (4.6). From (4.6b) we find that $(\hat{0}, \hat{u}^*)$ is an equilibrium if and only if

$$\nabla_{\hat{v}} r(\hat{0}, \hat{u}^*, \hat{u}^*) = \hat{0}_n$$

(where $\hat{0}_n$ is the origin in $\mathbb{R}^n$), in which case we say $\hat{u}^*$ is a *critical trait.* As a bifurcation parameter we use the dominant eigenvalue of $P(\hat{0}, \hat{u}^*, \hat{u}^*)$, which we denote by

$$r_0^* := \rho[P(\hat{0}, \hat{u}^*, \hat{u}^*)].$$

This is the inherent growth rate of the population when the trait is held fixed at the critical trait $\hat{u} = \hat{u}^*$. As in the non-evolutionary case, we normalize the entries of the projection matrix so that

$$P(\hat{x}, \hat{u}, \hat{v}) = r_0^* Q(\hat{x}, \hat{u}, \hat{v})$$

where Q satisfies H2 and

$$\rho[Q(\hat{0}, \hat{u}^*, \hat{u}^*)] = 1.$$

Letting

$$\bar{r}(\hat{x}, \hat{u}, \hat{v}) := \rho[Q(\hat{x}, \hat{u}, \hat{v})]$$

we have

$$r(\hat{x}, \hat{u}, \hat{v}) = r_0^* \, \bar{r}(\hat{x}, \hat{u}, \hat{v}), \quad \bar{r}(\hat{0}, \hat{u}^*, \hat{u}^*) = 1. \tag{4.9}$$

The Darwinian equations (4.6) are now

$$\hat{x}(t+1) = r_0^* Q(\hat{x}(t), \hat{u}(t), \hat{u}(t))\hat{x}(t) \tag{4.10a}$$

$$\hat{u}(t+1) = \hat{u}(t) + \frac{1}{\bar{r}(\hat{x}, \hat{u}, \hat{u})} M \ \nabla_{\hat{v}} \bar{r}(\hat{x}, \hat{u}, \hat{u}). \tag{4.10b}$$

Note that the bifurcation parameter r_0^* does not appear in the trait equation (4.10b). This is because in the trait equation (4.6b) we have

$$\nabla_{\hat{v}} \ln r(\hat{x}(t), \hat{u}(t), \hat{u}(t)) = \frac{1}{r(\hat{x}, \hat{u}, \hat{u})} \nabla_{\hat{v}} r(\hat{x}, \hat{u}, \hat{u}) = \frac{1}{r_0^* \bar{r}(\hat{x}, \hat{u}, \hat{u})} r_0^* \nabla_{\hat{v}} \bar{r}(\hat{x}, \hat{u}, \hat{u})$$

in which r_0^* cancels.

We say that a pair $(r_0^*, (\hat{x}, \hat{u})) \in \mathbb{R} \times (\Omega \times V)$ is an *equilibrium pair* if

$$\hat{x} = r_0^* Q(\hat{x}, \hat{u}, \hat{u})\hat{x} \tag{4.11a}$$

$$\hat{0}_n = \nabla_{\hat{v}} \bar{r}(\hat{x}, \hat{u}, \hat{u}). \tag{4.11b}$$

Note that

$$\nabla_{\hat{v}} \bar{r}(\hat{x}, \hat{u}, \hat{u}) = \hat{0}_n \text{ if and only if } \nabla_{\hat{v}} r(\hat{x}, \hat{u}, \hat{u}) = \hat{0}_n.$$

Definition 4.2. *We say an equilibrium pair $(r_0^*, (\hat{x}, \hat{u}))$ is a positive equilibrium if $\hat{x} \in \mathbb{R}_+^m$. An extinction equilibrium pair is an equilibrium pair of the form $(r_0^*, (\hat{0}, \hat{u}))$.*

Observe that $(r_0^*, (\hat{0}, \hat{u}))$ is an *extinction equilibrium pair* if and only if $\hat{u} = \hat{u}^*$ is a *critical trait* and, conversely, if $\hat{u} = \hat{u}^*$ is a critical trait, then $(r_0^*, (\hat{0}, \hat{u}^*))$ is an extinction equilibrium pair for all values of r_0^*.

4.3 Stability of extinction equilibria

We want to analyze the stability properties of an extinction equilibrium pair.

Definition 4.3. *We say that an equilibrium pair $(r_0^*, (\hat{x}, \hat{u}))$ is stable if $(\hat{x}, \hat{u})$ is (locally asymptotically) stable as an equilibrium of the Darwinian dynamics (4.6).*

To use the Linearisation Principle, we compute the Jacobian matrix for the system (4.10a)-(4.10b)

$$\mathcal{J}(r_0^*, \hat{x}, \hat{u}) = \begin{pmatrix} r_0^* J(\hat{x}, \hat{u}, \hat{u}) & r_0^* \Psi(\hat{x}, \hat{u}, \hat{u}) \\ \Upsilon(\hat{x}, \hat{u}, \hat{u}) & \Phi(\hat{x}, \hat{u}, \hat{u}) \end{pmatrix}$$

where $J(\hat{x}, \hat{u}, \hat{u})$ is the $m \times m$ Jacobian matrix of $Q(\hat{x}, \hat{u}, \hat{u})\hat{x}$ with respect to $\hat{x}$ and $\Psi(\hat{x}, \hat{u}, \hat{u})$ is the $m \times n$ matrix whose n columns are

$$\partial_{u_i} Q(\hat{x}, \hat{u}, \hat{u})\hat{x} + \partial_{v_i} Q(\hat{x}, \hat{u}, \hat{u})\hat{x}, \quad i = 1, 2, \cdots, n. \tag{4.12}$$

The dynamics of the i^{th} mean trait u_i are given by

$$u_i(t+1) = u_i(t) + \sum_{k=1}^{n} \sigma_{ik} \partial_{v_k} \ln \bar{r}(\hat{x}, \hat{u}, \hat{u}).$$

and therefore $\Upsilon(\hat{x}, \hat{u}, \hat{u})$ is the $n \times m$ matrix whose i^{th} row $(i = 1, 2, \cdots, n)$ is the transpose of the gradient

$$\nabla_{\hat{x}} \sum_{k=1}^{n} \sigma_{ik} \partial_{v_k} \ln \bar{r}(\hat{x}, \hat{u}, \hat{u}) = \sum_{k=1}^{n} \sigma_{ik} \nabla_{\hat{x}} \partial_{v_k} \ln \bar{r}(\hat{x}, \hat{u}, \hat{v}) \Big|_{\hat{v}=\hat{u}}$$

and

$$\Phi(\hat{x}, \hat{u}, \hat{u}) = I_{n\times n} + MH(\hat{x}, \hat{u}, \hat{u}) \tag{4.13}$$

where $H(\hat{x}, \hat{u}, \hat{u})$ is a $n \times n$ matrix whose kj^{th} entry is the u_k derivative of $\partial_{v_j} \ln \bar{r}(\hat{x}, \hat{u}, \hat{u})$, i.e.

$$H(\hat{x}, \hat{u}, \hat{u}) := \left[\partial_{u_k v_j} \ln \bar{r}(\hat{x}, \hat{u}, \hat{v}) \big|_{\hat{v}=\hat{u}} + \partial_{v_k v_j} \ln \bar{r}(\hat{x}, \hat{u}, \hat{v}) \big|_{\hat{v}=\hat{u}} \right]$$

By assumption H2, the projection matrix $P(\hat{0}, \hat{u}, \hat{v}) = [\tilde{p}_{ij}\,(\hat{v})]$ is independent of $\hat{u}$ and as a result

$$\partial_{u_k v_j} \ln \bar{r}(\hat{x}, \hat{u}, \hat{v}) \big|_{(\hat{x}, \hat{u}, \hat{v})=(\hat{0}, \hat{u}, \hat{u})} \equiv 0$$

for all $\hat{u}$. It follows that

$$H(\hat{0}, \hat{u}^*, \hat{u}^*) = \left[\partial^0_{v_k v_j} \ln \bar{r} \right] \tag{4.14}$$

where we have adopted the superscript notation

$$\partial^0_{v_k v_j} \ln \bar{r} := \partial_{v_k v_j} \ln \bar{r}(\hat{x}, \hat{u}, \hat{v}) \big|_{(\hat{x}, \hat{u}, \hat{v})=(\hat{0}, \hat{u}^*, \hat{u}^*)}.$$

Thus, the matrix $H(\hat{0}, \hat{u}^*, \hat{u}^*)$ is the *Hessian* of $\ln \bar{r}(\hat{x}, \hat{u}, \hat{v})$ with respect to $\hat{v}$ evaluated at $(\hat{x}, \hat{u}, \hat{v}) = (\hat{0}, \hat{u}^*, \hat{u}^*)$.

The Jacobian $\mathcal{J}$ evaluated at an extinction equilibrium pair $(r_0^*, (\hat{x}, \hat{u})) = (r_0^*, (\hat{0}, \hat{u}^*))$ is

$$\mathcal{J}(r_0^*, \hat{0}, \hat{u}^*) = \begin{pmatrix} r_0^* J(\hat{0}, \hat{u}^*, \hat{u}^*) & 0_{m\times n} \\ \Upsilon(\hat{0}, \hat{u}^*, \hat{u}^*) & \Phi(\hat{0}, \hat{u}^*, \hat{u}^*) \end{pmatrix} = \begin{pmatrix} r_0^* Q(\hat{0}, \hat{u}^*, \hat{u}^*) & 0_{m\times n} \\ \Upsilon(\hat{0}, \hat{u}^*, \hat{u}^*) & \Phi(\hat{0}, \hat{u}^*, \hat{u}^*) \end{pmatrix} \tag{4.15}$$

where $0_{m\times n}$ denotes the null matrix with dimension $m \times n$. We note that $J(\hat{0}, \hat{u}^*, \hat{u}^*)$ is the Jacobian with respect to $\hat{x}$ of (4.10a) when the trait is held fixed at $\hat{u}^*$. The eigenvalues of (4.15) are the m eigenvalues of $r_0^* Q(\hat{0}, \hat{u}^*, \hat{u}^*)$ and the n eigenvalues of $\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)$.

Recall that r_0^* is the (strictly) dominant eigenvalue of $r_0^* Q(\hat{0}, \hat{u}^*, \hat{u}^*)$. Thus, if $r_0^* > 1$ the extinction equilibrium $(\hat{0}, \hat{u}^*)$ is unstable. On the other hand, if $r_0^* < 1$ then stability (by linearisation) is determined by the n eigenvalues of $\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)$. Using the linearisation principle for discrete dynamical systems [37], we obtain the following result, which is an extension, for the evolutionary case with multiple traits, of the first statement in Theorem 4.1.

Theorem 4.2. *Assume H2 holds and that $\hat{u}^* \in V$ is a critical trait.*

(a) If $r_0^* < 1$ and $\rho[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)] < 1$, then the extinction equilibrium pair $(r_0^*, (\hat{0}, \hat{u}^*))$ is stable.

(b) If $r_0^* < 1$ and $\rho[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)] > 1$, then the extinction equilibrium pair $(r_0^*, (\hat{0}, \hat{u}^*))$ is unstable.

(c) If $r_0^* > 1$, then the extinction equilibrium pair $(r_0^*, (\hat{0}, \hat{u}^*))$ is unstable.

To investigate the spectral radius $\rho[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)]$, which appears in Theorem 4.2, we make further assumptions on the matrix M.

H3. *The variance-covariance matrix M is diagonally dominant: $\sigma_i^2 \geq \sum_{j \neq i} |\sigma_{ij}|$.*

In [99] it is shown, under assumption H3, that $\rho[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)] < 1$ if $H(\hat{0}, \hat{u}^*, \hat{u}^*)$ is negative definite and that $\rho[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)] > 1$ if $H(\hat{0}, \hat{u}^*, \hat{u}^*)$ is positive semi-definite or indefinite provided the variances σ_i^2 are small.

Corollary 4.1. *Assume H2 and H3 hold and that $\hat{u}^* \in V$ is a critical trait. If the variances σ_i^2 are small, then the extinction equilibrium pair $(r_0^*, (\hat{0}, \hat{u}^*))$ is*

(a) stable if $r_0^* < 1$ and the Hessian (4.14) is negative definite;

(b) unstable if $r_0^* > 1$ or if the Hessian (4.14) is either indefinite or positive semi-definite.

With regard to the variances, the assumption $\sigma_i^2 < 1/\rho\left[H(\hat{0}, \hat{u}^*, \hat{u}^*)\right]$ is sufficient in Corollary 4.1**.** We are particularly interested in the case when the extinction equilibrium $(r_0^*, (\hat{0}, \hat{u}^*))$ loses stability as r_0^* increases through 1. This occurs in Theorem 4.2 when $\rho\left[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)\right] < 1$. It also occurs in Corollary 4.1, when $H(\hat{0}, \hat{u}^*, \hat{u}^*)$ is negative definite. This suggests the possibility of a (transcritical) bifurcation occurring at $r_0^* = 1$ that can result in a branch of stable positive (non-extinction) equilibria. We address this question in Section 4.4.

As an example consider the case when there is no covariant evolution of the traits (i.e. that the off diagonal terms in M are all equal to 0 and the diagonal terms σ_i^2 are positive) and when

$$\partial^0_{v_i v_j} \bar{r} = 0 \text{ for } i \neq j, \tag{4.16}$$

so that traits evolve nearly independently in a neighbourhood of $(\hat{x}, \hat{u}) = (\hat{0}, \hat{u}^*)$. With these assumptions the matrix $\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)$ is diagonal and its eigenvalues are $1 + \sigma_i^2 \partial^0_{v_i v_i} \bar{r}$. From Theorem 4.2 we obtain the following corollary.

Corollary 4.2. *Assume H2 holds and that $\hat{u}^* \in V$ is a critical trait. Further assume $\sigma_{ij} = 0$ and (4.16) for all $i \neq j$. The extinction equilibrium pair $(r_0^*, (\hat{0}, \hat{u}^*))$ is*

(a) stable if $r_0^* < 1$ and $\left|1 + \sigma_i^2 \partial^0_{v_i v_i} \bar{r}\right| < 1$ for all i;

(b) unstable if $r_0^* > 1$;

(c) unstable for any $r_0^* > 0$ if $\left|1 + \sigma_i^2 \partial^0_{v_i v_i} \bar{r}\right| > 1$ for at least one i.

Note. In Corollary 4.2(c) the extinction equilibrium pair $(r_0^*, (\hat{0}, \hat{u}^*))$ is unstable, for any value of r_0^*, if $\partial^0_{v_i v_i} \bar{r} > 0$ for at least one i. On the other hand, if $\partial^0_{v_i v_i} \bar{r} < 0$ for all values of i, then extinction equilibrium pair $(r_0^*, (\hat{0}, \hat{u}^*))$ is stable for $r_0^* < 1$ and small variances σ_i^2.

4.4 A bifurcation theorem for the evolutionary model

The loss of stability by the extinction equilibrium pair when r_0^* increases through 1 suggests the possibility of a (transcritical) bifurcation at the value $r_0^* = 1$. This is due to the fact that an eigenvalue of the Jacobian leaves the complex unit circle as r_0^* increases through 1. In this section we establish a bifurcation theorem for the evolutionary model (4.10).

We begin by assuming that $\hat{u}$ can be expressed as a function of $\hat{x}$ by means of the equilibrium equation (4.11b).

H4. *Let $\hat{u}^* \in V$ be a critical trait. Assume there exists a function $\hat{\xi} \in C^2(N \to V)$, where N is a open neighbourhood of $\hat{0}$ in $\mathbb{R}^m$, such that $\hat{\xi}(\hat{0}) = \hat{u}^*$ and $\nabla_{\hat{v}}\bar{r}(\hat{x}, \hat{\xi}(\hat{x}), \hat{\xi}(\hat{x})) = \hat{0}_n$ for $\hat{x} \in N$.*

Let $J^0_{\hat{u}}(\nabla_{\hat{v}}\bar{r})$ and $J^0_{\hat{x}}(\nabla_{\hat{v}}\bar{r})$ denote the *Jacobian matrices of the gradient* $\nabla_{\hat{v}}\bar{r}(\hat{x}, \hat{u}, \hat{u})$ with respect to $\hat{u}$ and $\hat{x}$ respectively evaluated at $(\hat{x}, \hat{u}, \hat{u}) = (\hat{0}, \hat{u}^*, \hat{u}^*)$. The following assumption and the Implicit Function Theorem guarantee that H4 holds.

H5. *Let $u^* \in V$ be a critical trait for which $J^0_{\hat{u}}(\nabla_{\hat{v}}\bar{r})$ is a non-singular matrix.*

Remark 4.4. *The product rule applied to*

$$\nabla_{\hat{v}} \ln \bar{r}(\hat{x}, \hat{u}, \hat{v})|_{\hat{v}=\hat{u}} = \frac{1}{\bar{r}} \nabla_{\hat{v}}\bar{r}(\hat{x}, \hat{u}, \hat{v})|_{\hat{v}=\hat{u}}$$

evaluated at $\hat{x} = \hat{0}$ implies, together with $\bar{r}^0 = 1$ and $\nabla^0_{\hat{v}}\bar{r} = \hat{0}$ (by the definition of a critical trait), that $J^0_{\hat{u}}\nabla_{\hat{v}} \ln \bar{r} = J^0_{\hat{u}}\nabla_{\hat{v}}\bar{r}$. Now we have that

$$J^0_{\hat{u}}\nabla_{\hat{v}} \ln \bar{r} = \left[\partial^0_{v_k v_j} \ln \bar{r}\right].$$

Thus, under assumption H2 we see (from (4.14)) that in H5

$$J^0_{\hat{u}}(\nabla_{\hat{v}}\bar{r}) = H(\hat{0}, \hat{u}^*, \hat{u}^*).$$

We now recall a basic alternative result from linear algebra, known as the Fredholm Alternative that we shall use the proof of our main theorem. The Fredholm Alternative holds in infinite dimensional spaces but this formulation is enough for our needs.

Theorem 4.3 (Fredholm Alternative)**.** *Consider a linear map in $\mathbb{R}^n$ represented by a matrix A. Given a column vector b, **exactly one** of the following holds:*

1. *$Ax = b$ has a solution x;*

2. *$A^T y = 0$ has a solution y such that $y^T b \neq 0$.*

   Analogously, the linear equation $Ax = b$ is solvable if and only if for every solution y of the equation $A^T y = 0$ one has that $y^T b = 0$, i.e., b is orthogonal to the kernel of A^T.

Proof. If $Ax = b$ is solvable and y is such that $A^T y = 0$ then $y^T b = (y^T A)x = 0$ so that 2 doesn't hold. Conversely if 2 holds then if 1 is solvable then $0 = x^T(A^T y) = y^T Ax = y^T b \neq 0$ which is a contradiction. □

Remark 4.5. *If condition* 1 *in the theorem holds, then the equation is solvable so that A is surjective. Since A is a linear map in $\mathbb{R}^n$ then A is also injective, so given b the solution x of the equation is unique. So, for each b the orthogonality in condition* 2 *guarantees unique solvability of the equation in condition 1.*

We now state and prove the main theorem of this work, which characterizes the fundamental bifurcation of equilibria and its stability.

Theorem 4.4. *Assume $\hat{u}^* \in V$ is a critical trait. Assume H2 and H4 hold and that $\kappa^* \neq 0$.*

(1) There exists a continuum $\mathcal{C}^*$ of positive equilibrium pairs $(r_0^*, (\hat{x}, \hat{u})) \in \mathbb{R}_+ \times (\mathbb{R}_+^m \times V)$ of (4.10) that bifurcates from the extinction pair $(1, (\hat{0}, \hat{u}^*))$ (i.e. that contains the extinction pair in its closure).

(2) Assume H2 and H5 hold. In a neighbourhood of $(1, (\hat{0}, \hat{u}^*))$, the positive equilibrium pairs have the parametric representation

$$\hat{x}(\varepsilon) = \hat{w}_R \varepsilon + \mathcal{O}(\varepsilon^2) \tag{4.17a}$$

$$\hat{u}(\varepsilon) = \hat{u}^* + \hat{u}_1 \varepsilon + \mathcal{O}(\varepsilon^2) \tag{4.17b}$$

$$r_0^*(\varepsilon) = 1 + \kappa^* \varepsilon + \mathcal{O}(\varepsilon^2) \tag{4.17c}$$

for small $\varepsilon \gtrapprox 0$ where $\hat{w}_L$ and $\hat{w}_R$ are, respectively, a positive left and right eigenvector of $Q(\hat{0}, \hat{u}^*, \hat{u}^*)$ associated with eigenvalue 1 (equivalently of $P(\hat{0}, \hat{u}^*, \hat{u}^*)$ when $r_0^* = 1$) normalized so that $\hat{w}_L^T \hat{w}_R = 1$ and

$$\hat{u}_1 := -\left[J_{\hat{u}}^0 (\nabla_{\hat{v}} \bar{r})\right]^{-1} J_{\hat{x}}^0 (\nabla_{\hat{v}} \bar{r}) \hat{w}_R. \tag{4.18}$$

$$\kappa^* := -\hat{w}_L^T \left(\left[\nabla_{\hat{x}}^0 q_{ij}^T \hat{w}_R \right] \right) \hat{w}_R \ . \tag{4.19}$$

Furthermore, we have the following alternatives.

(a) If $\rho[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)] < 1$ and $\kappa^* > 0$, then the bifurcation of $\mathcal{C}^*$ is forward and the positive equilibrium pairs on $\mathcal{C}^*$ are stable.

(b) If $\rho[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)] < 1$ and $\kappa^* < 0$, then the bifurcation is backward and the positive equilibrium pairs on $\mathcal{C}^*$ are unstable.

(c) If $\rho[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)] > 1$, then positive equilibrium pairs in the continuum $\mathcal{C}^*$ are unstable regardless of the direction of bifurcation.

Remark 4.6. *Because κ^* is calculated from evaluations at the bifurcation point $(r_0^*, (\hat{x}, \hat{u})) = (1, (\hat{0}, \hat{u}^*))$ and because only the sign of κ^* is involved in determining the direction of bifurcation and stability, Theorem 4.4, parts (a) and (b), remains valid if in the formula (4.19) and in H4 and H5 we replace q_{ij} by p_{ij} and $\bar{r}$ by r.*

Proof. (1) Under H4, the equilibrium equations (4.11) reduce to the single equation

$$\hat{x} = r_0^* Q(\hat{x}, \hat{\xi}(\hat{x}), \hat{\xi}(\hat{x})) \hat{x} \tag{4.20}$$

for $\hat{x} \in N$. Theorem 4.1 applies to this equation with matrix $Q(\hat{x}, \hat{\xi}(\hat{x}), \hat{\xi}(\hat{x}))$ in place of $Q(\hat{x})$ (and N in place of Ω). This results in the existence of a continuum $\mathcal{C}$ of positive solution pairs $(r_0^*, \hat{x})$ of (4.20) that bifurcates from $(1, \hat{0})$. The continuum $\mathcal{C}$ in turn gives rise to a continuum

$$\mathcal{C}^* := \{(r_0^*, (\hat{x}, \hat{u})) \mid (r_0^*, \hat{x}) \in \mathcal{C}, \ \hat{u} = \hat{\xi}(\hat{x})\}$$

of equilibrium pairs $(r_0^*, (\hat{x}, \hat{u}))$ of (4.10) that bifurcates from the extinction equilibrium $(1, (\hat{0}, \hat{u}^*))$ at $r_0^* = 1$

(2) The parametrisation of $\mathcal{C}$ Theorem 4.1 implies that, near the bifurcation point, the positive equilibrium pairs on the continuum $\mathcal{C}^*$ have a parametrisation (4.17). The coefficient κ^* is given by

the formula (4.3) for κ but with $p_{ij}(\hat{x})$ replaced by $p_{ij}(\hat{x}, \hat{\xi}(\hat{x}), \hat{\xi}(\hat{x}))$. The gradient of $p_{ij}(\hat{x}, \hat{\xi}(\hat{x}), \hat{\xi}(\hat{x}))$ relative to $\hat{x}$ is

$$\nabla^0_{\hat{x}} p^T_{ij} + \nabla^0_{\hat{u}} p^T_{ij} \nabla^0_{\hat{x}} \hat{\xi} + \nabla^0_{\hat{v}} p^T_{ij} \nabla^0_{\hat{x}} \hat{\xi} \ .$$

The second term is zero by remark 4.2. To make this calculation we start by noting that the coefficient $\hat{u}_1 = \nabla^0_{\hat{x}} \hat{\xi} \hat{w}_R$ can be calculated by an implicit differentiation of the equation $\nabla_{\hat{v}} \bar{r}(\hat{x}(\varepsilon), \hat{\xi}(\hat{x}(\varepsilon)), \hat{\xi}(\hat{x}(\varepsilon))) = \hat{0}_n$ with respect to ε and a subsequent evaluation at $\varepsilon = 0$. From this calculation we obtain the formula (4.18) for $\hat{u}_1$. So we obtain

$$\kappa^* = -\hat{w}^T_L \left[\nabla^0_{\hat{x}} p^T_{ij} \hat{w}_R \right] \hat{w}_R - \hat{w}^T_L \left[\nabla^0_{\hat{v}} p^T_{ij} \hat{u}_1 \right] \hat{w}_R \ .$$

The second term in this formula is zero as we shall prove in Lemma 4.1. So we obtain formula (4.19) for κ^*.

To investigate the stability of bifurcating positive equilibrium pairs we make use of the parametrisation (4.17) which allows us to parametrise by ε the Jacobian $J(r, \hat{x}, \hat{u})$ when it is evaluated at the equilibria (4.17) and, subsequently, to parametrise this Jacobian's eigenvalues by ε. From this parametrisation we can approximate the eigenvalues of the Jacobian for $\varepsilon \gtrapprox 0$.

At $\varepsilon = 0$ the eigenvalues of the Jacobian $J(1, \hat{0}, \hat{u}^*)$ are the eigenvalues of $J(\hat{0}, \hat{u}^*, \hat{u}^*)$ and the eigenvalues of $\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)$. The spectrum of the Jacobian $J(r^*_0(\varepsilon), \hat{x}(\varepsilon), \hat{u}(\varepsilon))$ approaches, by continuity, the spectrum of $J(1, \hat{0}, \hat{u}^*)$ as ε tends to 0. Therefore, if $\rho[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)] > 1$ then for $\varepsilon \gtrapprox 0$ the Jacobian $J(r^*_0(\varepsilon), \hat{x}(\varepsilon), u(\varepsilon))$ also has spectral radius greater than 1. Consequently, the positive equilibria are unstable near the bifurcation point. This establishes 2(c).

Suppose $\rho[\Phi(\hat{0}, \hat{u}^*, \hat{u}^*)] < 1$. Since the strictly dominant eigenvalue of $J(\hat{0}, \hat{u}^*, \hat{u}^*)$ is 1, it follows that the dominant eigenvalue of $\mathcal{J}(1, \hat{0}, u^*)$ is 1. To determine stability of the bifurcating positive equilibria, by means of the linearisation principle, we must investigate if the dominant eigenvalue of $\mathcal{J}(r^*_0(\varepsilon), \hat{x}(\varepsilon), \hat{u}(\varepsilon))$, which equals 1 when $\varepsilon = 0$, is greater or less than 1 for $\varepsilon \gtrapprox 0$. Using the aforementioned parametrisations we can obtain a parametrisation of the strictly dominant eigenvalue for small perturbations ϵ by means of the Lyapunov-Schmidt method (also called Lyapunov-Schmidt reduction). Using this reduction technique we are able to reduce the problem of studying the dominant eigenvalue of of $\mathcal{J}(r^*_0(\varepsilon), \hat{x}(\varepsilon), \hat{u}(\varepsilon))$ and its associated eigenvector to a lower dimensional problem involving the eigenvector of matrix J, whose eigenvector structure we know. This reduction involves the derivative $\hat{u}_1$ of the implicit function $\hat{\xi}$ that we have computed above.

Let

$$\mu(\varepsilon) = 1 + \mu_1 \varepsilon + \mathcal{O}(\varepsilon^2)$$

denote the dominant eigenvalue of $\mathcal{J}(r^*_0(\varepsilon), \hat{x}(\varepsilon), \hat{u}(\varepsilon))$. Whether $\mu(\varepsilon)$ is less than or greater than 1 for $\varepsilon \gtrapprox 0$, and hence whether the bifurcating positive equilibria are stable or unstable near the bifurcation point, can be determined by the sign of μ_1: for $\varepsilon \gtrapprox 0$, the bifurcating positive equilibria are stable if $\mu_1 < 0$ and unstable if $\mu_1 > 0$.

To calculate a formula for μ_1 we begin by letting

$$\hat{W}_R(\varepsilon) = \hat{W}_{R0} + \hat{W}_{R1} \varepsilon + \mathcal{O}(\varepsilon^2)$$

denote a right eigenvector of the Jacobian $\mathcal{J}$ associated with the dominant eigenvalue $\mu(\varepsilon)$, so that

$$\mathcal{J}(r_0^*(\varepsilon), \hat{x}(\varepsilon), \hat{u}(\varepsilon))\hat{W}_R(\varepsilon) = \mu(\varepsilon)\hat{W}_R(\varepsilon). \tag{4.21}$$

Setting $\varepsilon = 0$ in (4.21) we obtain

$$\mathcal{J}(1, \hat{0}, \hat{u}^*)\hat{W}_{R0} = \hat{W}_{R0}.$$

We can write

$$\hat{W}_{R0} = \begin{pmatrix} \hat{w}_{R0}^m \\ \\ \hat{w}_{R0}^n \end{pmatrix}$$

where $\hat{w}_{R0}^m$ and $\hat{w}_{R0}^n$, are column vectors with m and n entries respectively. The vector $\hat{w}_{R0}^m$ is a right eigenvector of $J(\hat{0}, \hat{u}^*, \hat{u}^*) = Q(\hat{0}, \hat{u}^*, \hat{u}^*)$ associated with the eigenvalue 1 and consequently $\hat{w}_{R0}^m = \hat{w}_R$. A calculation shows

$$\hat{W}_{R0} = \begin{pmatrix} \hat{w}_R \\ \\ \hat{u}_1 \end{pmatrix}.$$

The vector $\hat{W}_{L0}^T$ where

$$\hat{W}_{L0} = \begin{pmatrix} \hat{w}_L \\ \\ \hat{0}_n \end{pmatrix}$$

is a left eigenvector of $\mathcal{J}(1, \hat{0}, \hat{u}^*)$ (where $\hat{w}_L^T$ is a left eigenvector of $J(\hat{0}, \hat{u}^*, \hat{u}^*)$). Note that $\hat{W}_{L0}^T\hat{W}_{R0} = \hat{w}_L^T\hat{w}_R = 1$.

If we differentiate (4.21) with respect to ε and set $\varepsilon = 0$ in the result, we obtain

$$\left(\mathcal{J}(1, \hat{0}, \hat{u}^*) - I_{m+n}\right)\hat{W}_{R1} = \mu_1\hat{W}_{R0} - \frac{d}{d\varepsilon}\mathcal{J}(r_0^*(\varepsilon), \hat{x}(\varepsilon), \hat{u}(\varepsilon))\bigg|_{\varepsilon=0}\hat{W}_{R0}$$

where I_{m+n} denotes the identity matrix of size $m+n$. According to Fredholm alternative, this equation is solvable for $\hat{W}_{R1}$ if and only if the right hand side is orthogonal to the kernel of $(\mathcal{J}(1, \hat{0}, \hat{u}^*) - I_{m+n})^T$. This kernel is spanned by the left eigenvectors of $\mathcal{J}(1, \hat{0}, \hat{u}^*)$ associated with the eigenvalue 1. Therefore, the right hand side must be orthogonal to $\hat{W}_{L0}$. This implies

$$\begin{aligned}\mu_1 &= \hat{W}_{L0}^T\frac{d}{d\varepsilon}\mathcal{J}(r_0^*(\varepsilon), \hat{x}(\varepsilon), \hat{u}(\varepsilon))\bigg|_{\varepsilon=0}\hat{W}_{R0} \\ &= \begin{pmatrix} \hat{w}_L \\ \\ \hat{0}_n \end{pmatrix}^T \frac{d}{d\varepsilon}\begin{pmatrix} r_0^*(\varepsilon)\, J(\hat{x}(\varepsilon), \hat{u}(\varepsilon), \hat{u}(\varepsilon)) & r_0^*(\varepsilon)\,\Psi(\hat{x}(\varepsilon), \hat{u}(\varepsilon), \hat{u}(\varepsilon)) \\ \\ \Upsilon(\hat{x}(\varepsilon), \hat{u}(\varepsilon), \hat{u}(\varepsilon)) & \Phi(\hat{x}(\varepsilon), \hat{u}(\varepsilon), \hat{u}(\varepsilon)) \end{pmatrix}\Bigg|_{\varepsilon=0}\begin{pmatrix} \hat{w}_R \\ \\ \hat{u}_1 \end{pmatrix}\end{aligned}$$

or

$$\mu_1 = \hat{w}_L^T\frac{d}{d\varepsilon}\left[r_0^*(\varepsilon)\, J(\hat{x}(\varepsilon), \hat{u}(\varepsilon), \hat{u}(\varepsilon))\right]\bigg|_{\varepsilon=0}\hat{w}_R + \hat{w}_L^T\frac{d}{d\varepsilon}\left[r_0^*(\varepsilon)\,\Psi(\hat{x}(\varepsilon), \hat{u}(\varepsilon), \hat{u}(\varepsilon))\right]\bigg|_{\varepsilon=0}\hat{u}_1.$$

We consider the two terms in this sum separately. With regard to the first term in μ_1, the product and chain rules imply

$$\begin{aligned}
&\hat{w}_L^T \frac{d}{d\varepsilon}\left[r_0^*(\varepsilon)\, J(\hat{x}(\varepsilon),\hat{u}(\varepsilon),\hat{u}(\varepsilon))\right]\Big|_{\varepsilon=0} \hat{w}_R \\
&= \kappa^* \hat{w}_L^T Q(\hat{0},\hat{u}^*,\hat{u}^*)\hat{w}_R + \hat{w}_L^T \frac{d}{d\varepsilon} J(\hat{x}(\varepsilon),\hat{u}(\varepsilon),\hat{u}(\varepsilon))\Big|_{\varepsilon=0} \hat{w}_R \\
&= \kappa^* + \hat{w}_L^T \left(\left[\nabla_{\hat{x}}^0 q_{ij}^T \hat{w}_R\right] + \left[\nabla_{\hat{u}}^0 q_{ij}^T \hat{u}_1\right] + \left[\nabla_{\hat{v}}^0 q_{ij}^T \hat{u}_1\right] + \left[\partial_{x_i}^0 q_j \hat{w}_R\right]\right)\hat{w}_R
\end{aligned}$$

where we have defined the row vector

$$\partial_{x_i}^0 q_j := \left[\begin{array}{cccc} \partial_{x_i}^0 q_{j1} & \partial_{x_i}^0 q_{j2} & \cdots & \partial_{x_i}^0 q_{jm} \end{array}\right].$$

With regard to the second term in μ_1, we recall that $\Psi(\hat{x},\hat{u},\hat{u})$ is the $m \times n$ matrix whose n columns are (4.12) and as a result, upon evaluation at the bifurcation point, the only contribution to the derivative in the second term arises from the derivatives of $\Psi(\hat{x},\hat{u},\hat{u}) = [\psi_{ij}(\hat{x},\hat{u},\hat{u})]$ with respect to the components of $\hat{x}$. Therefore the second term in μ_1 is

$$\begin{aligned}
\hat{w}_L^T \frac{d}{d\varepsilon}\left[r_0^*(\varepsilon)\,\Psi(\hat{x}(\varepsilon),\hat{u}(\varepsilon),\hat{u}(\varepsilon))\right]\Big|_{\varepsilon=0} \hat{u}_1 &= \hat{w}_L^T \left[\nabla_{\hat{x}}^0 \psi_{ij}^T \hat{w}_R\right]\hat{u}_1 \\
&= \hat{w}_L^T \left(\left[\partial_{u_i}^0 q_j \hat{w}_R\right] + \left[\partial_{v_i}^0 q_j \hat{w}_R\right]\right)\hat{u}_1
\end{aligned}$$

where we have defined the row vectors

$$\partial_{u_i}^0 q_j := \left[\begin{array}{cccc} \partial_{u_i}^0 q_{j1} & \partial_{u_i}^0 q_{j2} & \cdots & \partial_{u_i}^0 q_{jm} \end{array}\right], \quad \partial_{v_i}^0 q_j := \left[\begin{array}{cccc} \partial_{v_i}^0 q_{j1} & \partial_{v_i}^0 q_{j2} & \cdots & \partial_{v_i}^0 q_{jm} \end{array}\right].$$

H2 implies $\partial_{u_k}^0 q_{ij} = 0$ for all u_k and all i, j, and so

$$\mu_1 = \kappa^* + \hat{w}_L^T\left[\nabla_{\hat{x}}^0 q_{ij}^T \hat{w}_R\right]\hat{w}_R + \hat{w}_L^T\left[\nabla_{\hat{v}}^0 q_{ij}^T \hat{u}_1\right]\hat{w}_R + \hat{w}_L^T\left[\partial_{x_i}^0 q_j \hat{w}_R\right]\hat{w}_R + \hat{w}_L^T\left[\partial_{v_i}^0 q_j \hat{w}_R\right]\hat{u}_1.$$

Noting that

$$\hat{w}_L^T\left[\nabla_{\hat{x}}^0 q_{ij}^T \hat{w}_R\right]\hat{w}_R = \hat{w}_L^T\left[\partial_{x_i}^0 q_j \hat{w}_R\right]\hat{w}_R, \quad \hat{w}_L^T\left[\nabla_{\hat{v}}^0 q_{ij}^T \hat{u}_1\right]\hat{w}_R = \hat{w}_L^T\left[\partial_{v_i}^0 q_j \hat{w}_R\right]\hat{u}_1$$

we have

$$\mu_1 = \kappa^* + 2\left(\hat{w}_L^T\left[\nabla_{\hat{x}}^0 q_{ij}^T \hat{w}_R\right]\hat{w}_R + \hat{w}_L^T\left[\nabla_{\hat{v}}^0 q_{ij}^T \hat{u}_1\right]\hat{w}_R\right).$$

By Lemma 4.1 the last term in this formula is zero so we get

$$\mu_1 = \kappa^* + 2\hat{w}_L^T\left[\nabla_{\hat{x}}^0 q_{ij}^T \hat{w}_R\right]\hat{w}_R$$

which, by (4.19), implies $\mu_1 = -\kappa^*$. As a result, $\kappa^* > 0$ implies both that the bifurcation is forward and that the bifurcating positive equilibria are stable for $\varepsilon \gtrapprox 0$. On the other hand, $\kappa^* < 0$ implies that the bifurcation is backward and that the bifurcating positive equilibria are unstable for $\varepsilon \gtrapprox 0$. □

The only piece missing in the proof of the theorem is the formula that we prove in the following lemma.

Lemma 4.1. *Assume H2 and H5 hold. Then* $\hat{w}_L^T[\nabla_{\hat{v}}^0 q_{ij}^T \hat{u}_1]\hat{w}_R = 0$.

Proof. Consider the equality

$$P(\hat{0}, \hat{u}, \hat{u})\hat{w}_R(\hat{0}, \hat{u}) = r(\hat{0}, \hat{u}, \hat{u})\hat{w}_R(\hat{0}, \hat{u}). \tag{4.22}$$

which holds by the definition of $r(\hat{0}, \hat{u}, \hat{u})$ as an eigenvalue with a positive right eigenvector $\hat{w}_R(\hat{0}, \hat{u})$. Let $\hat{p}_i = \hat{p}_i(\hat{0}, \hat{u}, \hat{u})$ denote the i-th column of $P = P(\hat{0}, \hat{u}, \hat{u})$. We want to take the Jacobian of both sides of equation (4.22) with respect to $\hat{u}$. To do this we let $J_{\hat{y}}[\hat{\omega}(\hat{y})]$ denote the Jacobian of a vector valued function $\hat{\omega}(\hat{y})$ of a vector $\hat{y}$.

The right side of (4.22) is a vector valued function of the form $\tau(\hat{y})\hat{\omega}(\hat{y})$ for a scalar valued function $\tau(\hat{y})$. Applying the general formula

$$J_{\hat{y}}[\tau(\hat{y})\hat{\omega}(\hat{y})] = \hat{\omega}(\hat{y})\nabla_{\hat{y}}\tau(\hat{y})^T + \tau(\hat{y})J_{\hat{y}}[\hat{\omega}(\hat{y})] \tag{4.23}$$

and recalling (4.8) in Remark 4.2, we find that the Jacobian of the right side of (4.22) with respect to $\hat{u}$ is

$$\hat{w}_R(\hat{0}, \hat{u})\left(\nabla_{\hat{u}} r^T + \nabla_{\hat{v}} r^T\right) + rJ_{\hat{u}}[\hat{w}_R(\hat{0}, \hat{u})] = \hat{w}_R(\hat{0}, \hat{u})\nabla_{\hat{v}} r^T + rJ_{\hat{u}}[\hat{w}_R(\hat{0}, \hat{u})].$$

To calculate the Jacobian of the left-hand side of (4.22), we write

$$P\hat{w}_R(\hat{0}, \hat{u}) = \sum_{i=1}^{m} w_i^R(\hat{0}, \hat{u})\hat{p}_i$$

where $w_i^R(\hat{0}, \hat{u})$ are the components of the vector $\hat{w}_R(\hat{0}, \hat{u})$ and apply the product rule (4.23) to each term. Noting (4.7) in Remark 4.2 we get

$$PJ_{\hat{u}}[\hat{w}_R(\hat{0}, \hat{u})] + \sum_{i=1}^{m} w_i^R(\hat{0}, \hat{u})J_{\hat{v}}[\hat{p}_i].$$

Equating the Jacobians of the left and right sides of (4.22) we have

$$PJ_{\hat{u}}[\hat{w}_R(\hat{0}, \hat{u})] + \sum_{i=1}^{m} w_i^R(\hat{0}, \hat{u})J_{\hat{v}}[\hat{p}_i] = \hat{w}_R(\hat{0}, \hat{u})\nabla_{\hat{v}} r^T + rJ_{\hat{u}}[\hat{w}_R(\hat{0}, \hat{u})]. \tag{4.24}$$

or

$$(P - rI_m)J_{\hat{u}}[\hat{w}_R(\hat{0}, \hat{u})] = \hat{w}_R(\hat{0}, \hat{u})\nabla_{\hat{v}} r^T - \sum_{i=1}^{m} w_i^R(\hat{0}, \hat{u})J_{\hat{v}}[\hat{p}_i]$$

which in turn can be rewritten as the m equations

$$(P - rI_m)\partial_{u_i}(\hat{w}_R(\hat{0}, \hat{u})) = (\partial_{v_i} rI_m - \partial_{v_i} P)\hat{w}_R(\hat{0}, \hat{u}) \ \text{ for } \ 1 \le i \le n.$$

The matrix $P - rI_m$ is singular and the kernel of its transpose, or equivalently its left kernel is spanned by the left eigenvector $\hat{w}_L^T(\hat{0}, \hat{u})$. By the Fredholm alternative, the solvability of these equations imply the m orthogonality conditions

$$\hat{w}_L^T(\hat{0}, \hat{u})(\partial_{v_i} rI_m - \partial_{v_i} P)\hat{w}_R(\hat{0}, \hat{u}) = 0$$

are satisfied. Solving for $\partial_{v_i} r$ and recalling that the eigenvectors are normalized so that $\hat{w}^L(\hat{0},\hat{u})^T\hat{w}_R(\hat{0},\hat{u}) = 1$, we find

$$\partial_{v_i} r = \hat{w}_L^T(\hat{0},\hat{u})\partial_{v_i} P\hat{w}_R(\hat{0},\hat{u}) \text{ for } 1 \leq i \leq n.$$

Since $\partial^0_{v_i} r = 0$ by definition of a critical trait vector $\hat{u}^*$, when setting $\hat{u} = \hat{u}^*$ and $r_0^* = 1$ in these expressions we get

$$\hat{w}_L^T \partial^0_{v_k} Q\hat{w}_R = 0 \text{ for } 1 \leq k \leq n. \tag{4.25}$$

Let $u_{1,k}$ denote the scalar components of the vector $\hat{u}_1$. Then

$$\nabla^0_{\hat{v}} q_{ij}^T \hat{u}_1 = \sum_{k=1}^n u_{1,k}\partial^0_{v_k} q_{ij}$$

and

$$\left[\nabla^0_{\hat{v}} q_{ij}^T \hat{u}_1\right] = \sum_{k=1}^n u_{1,k}\left[\partial^0_{v_k} q_{ij}\right] = \sum_{k=1}^n u_{1,k}\partial^0_{v_k} Q.$$

From

$$\hat{w}_L^T[\nabla^0_{\hat{v}} q_{ij}\hat{u}_1]\hat{w}_R = \sum_k u_{1,k}\left(\hat{w}^L\partial^0_{v_k} Q\hat{w}_R\right)$$

and (4.25) it follows that $\hat{w}_L^T[\nabla^0_{\hat{v}} q_{ij}\hat{u}_1]\hat{w}_R = 0$. □

From (4.13) we obtain (as in [99]) the following corollary of Theorem 4.4.

Corollary 4.3. *Assume H2, H3 and H5 hold and that $\hat{u}^* \in V$ is a critical trait. Let $\mathcal{C}^*$ be the continuum of positive equilibrium pairs that bifurcates from the extinction pair $(1,(\hat{0},\hat{u}^*))$ guaranteed by Theorem 4.4. If the variances σ_i^2 are small, then in a neighbourhood of the bifurcation point $(1,(\hat{0},\hat{u}^*))$ we have the following alternatives.*

(a) The bifurcation of $\mathcal{C}^*$ is forward and stable if the Hessian (4.14) is negative definite and $\kappa^* > 0$.

(b) The bifurcation of $\mathcal{C}^*$ is backward and unstable if the Hessian (4.14) is negative definite and $\kappa^* < 0$.

(c) The positive equilibrium pairs in the continuum $\mathcal{C}^*$ are unstable if the Hessian (4.14) is positive semi-definite or indefinite (regardless of the direction of bifurcation).

For the case of no trait covariance considered in Corollary 4.2, we obtain the following result from Corollary 4.3.

Corollary 4.4. *Assume H2 and H5 hold and that $\hat{u}^* \in V$ is a critical trait. Further assume $\sigma_{ij} = 0$ and (4.16) for all $i \neq j$. Let $\mathcal{C}^*$ be the continuum of positive equilibrium pairs that bifurcates from the extinction pair $(1,(\hat{0},\hat{u}^*))$ guaranteed by Theorem 4.4. Then in a neighbourhood of the bifurcation point $(1,(\hat{0},\hat{u}^*))$ we have the following alternatives.*

(a) The bifurcation of $\mathcal{C}^*$ is forward and stable if $\kappa^* > 0$ and $\left|1+\sigma_i^2\partial^0_{v_iv_i}\bar{r}\right| < 1$ for all i.

(b) The bifurcation of $\mathcal{C}^*$ is backward and unstable if $\kappa^* < 0$ and $\left|1+\sigma_i^2\partial^0_{v_iv_i}\bar{r}\right| < 1$ for all i.

(c) The positive equilibrium pairs in the continuum $\mathcal{C}^*$ are unstable if $\left|1+\sigma_i^2\partial^0_{v_iv_i}\bar{r}\right| > 1$ for at least one i (regardless of the direction of bifurcation).

Note. In Corollary 4.4 we see that the positive equilibrium pairs in the continuum $\mathcal{C}^*$ are unstable if $\partial^0_{v_iv_i}\bar{r} > 0$ for at least one i. On the other hand, if $\partial^0_{v_iv_i}\bar{r} < 0$ for all values of i, then the positive equilibrium pairs in the continuum $\mathcal{C}^*$ are stable if $\kappa^* > 0$ and the variances σ_i^2 are small enough so that $\left|1 + \sigma_i^2 \partial^0_{v_i v_i} \bar{r}\right| < 1$, i.e. $\sigma_i^2 < -2/\partial^0_{v_i v_i} \bar{r}$.

4.5 An application

Consider the single difference equation

$$x(t+1) = bx(t) \exp(-cx(t)) \exp\left(-\frac{\alpha}{1+sx(t)}\right) \tag{4.26}$$

with coefficients $b, c > 0$ and $\alpha, s \geq 0$. When $\alpha = 0$ this map is the famous Ricker equation (see [82]) which is one of the most well known equations that incorporates negative effects that population density can have on population growth. Equation (4.26) is studied in [90] as a model equation that incorporates a positive effect of increased population density (a so-called component Allee effect [22]) in the presence of a predator. This is the well-known predator-saturation effect in ecology and is one of the most commonly attributed causes of Allee effects [30], [22].

The factor $\exp(-\alpha/(1+sx))$ in (4.26) is an increasing function of x and represents the probability of escaping predation. We can interpret α as the *intensity of predation* and s a measure of how effective the protection from predation attributed to population density x, which we will refer to as the *predation protection factor*. Re-writing (4.26) as

$$x(t+1) = r_0 \bar{r}(x) x(t) \tag{4.27}$$

with

$$r_0 := b\exp(-\alpha), \quad \bar{r}(x) := \exp\left(-cx(t) + \alpha\frac{sx(t)}{1+sx(t)}\right)$$

we see that r_0 is the inherent (density-free) per capita birth rate, which equals b in the absence of predation $\alpha = 0$. This equation, studied in [90], not surprisingly can exhibit the same kind of period-doubling route-to-chaos as r_0 increases as does the famous Ricker equation when $\alpha = 0$. (The right side of (4.27) defines a unimodal map). The bifurcation that occurs at $r_0 = 1$ where the extinction equilibrium $x = 0$ destabilizes is, according to Theorem 4.1, forward and therefore stable if $\kappa = c - \alpha s > 0$. This inequality holds if the effect of predation, as measured by the product of the predation intensity α and (per capita) predation protection factor s, is small compared to that of the negative density effects measured by c. This occurs, of course, for the Ricker equation when $\alpha = 0$.

On the other, if the reverse is true and the effect of predation αs is large compared to c, then by Theorem 4.1 the bifurcation at $r_0 = 1$ is backward and unstable. In this case, i.e. when $\alpha s > c$, we can also say some things about the bifurcating continuum $\mathcal{C}$ of positive equilibrium pairs (r_0, x) outside the neighbourhood of the bifurcation point $(r_0, x) = (1, 0)$. The equation

$$1 = r_0 \exp\left(-cx + \alpha\frac{sx}{1+sx}\right)$$

satisfied by positive equilibria $x > 0$, when re-written as

$$r_0 = \exp\left(cx - \alpha\frac{sx}{1+sx}\right)$$

describes the continuum $\mathcal{C}$ of positive equilibrium pairs (r_0, x). The graph of r_0 as a function of x contains the point $r_0 = 1$ at $x = 0$, decreases as x increases to a unique critical point $x_{cr} > 0$ at which r_0 attains a global minimum $r_{cr} > 0$, and increases without bound for $x > x_{cr}$. See figure (4.1).

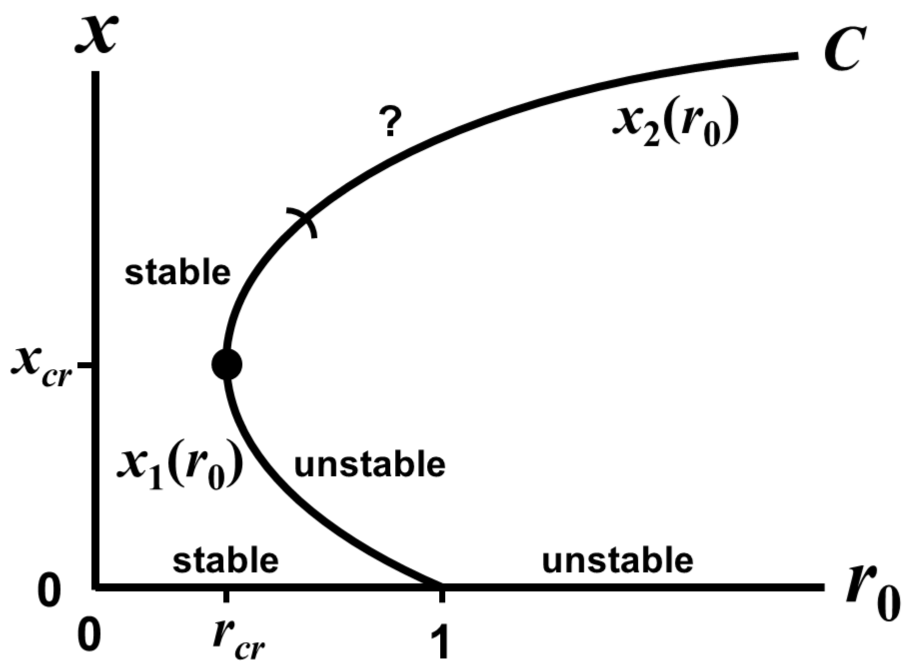

Figure 4.1: Generic plot of the bifurcating continuum $\mathcal{C}$ for equation (4.27) when $\alpha s > c$ and, consequently, a backward (unstable) bifurcation occurs at the point $(r_0, x) = (1, 0)$. The question mark indicates that although the positive equilibria on the upper branch $x_2(r_0)$ are (locally asymptotically) stable near the saddle-node bifurcation point (r_{cr}, x_{cr}), they can, depending on model parameter values, destabilize further along the continuum $\mathcal{C}$.

From the parabola-like shape of this graph, we see that the inverse function, treating x as a function of r_0, has two branches: an upper branch of positive equilibria $x_2(r_0)$ for $r_0 \geq r_{cr}$ and a lower branch of positive equilibria $x_1(r_0) < x_2(r_0)$ for $r^{cr} \leq r_0 < 1$ which satisfies $x_1(1) = 0$. The value $r_0 = r_{cr}$ is a saddle-node (blue-sky) bifurcation (or tipping) point at which the lower branch $x_1(r_0)$ and and upper branch $x_2(r_0)$ coalesce. The following facts follow from general results in [27]: for $r_0 < r_0^{cr}$ the extinction equilibrium is globally asymptotically stable; the equilibria $x_1(r_0)$ are unstable and the equilibria $x_2(r_0)$ are (locally asymptotically) stable for $r_0 \gtrapprox r_{cr}$. The upper branch $x_2(r_0)$ might not remain stable for all $r_0 > r_{cr}$, however, but might undergo a period doubling cascade to chaos. If a destabilization of $x_2(r_0)$ occurs at a point $r_0 \geq 1$, then on the interval $r_{cr} < r_0 < 1$ there are two stable equilibria, the extinction equilibrium and the positive equilibrium $x_2(r_0)$. This scenario is called a *strong Allee effect*. It asserts that survival is possible for some $r_0 < 1$ provided a population's initial condition lies outside the basin of attractor of the extinction equilibrium (the Allee basin). If, on the other hand, $x_2(r_0)$ loses stability at a point in the interval $r_{cr} < r_0 < 1$, then there still occurs a strong Allee effect but one with a non-equilibrium survival attractor (e.g. a periodic cycle or a more complicated attractors).

Sample forward and backward bifurcation diagrams are shown in figure (4.2). That secondary period-doubling bifurcations cascade to complex (presumably chaotic) dynamics in both cases is not unexpected, given that (4.27) is based on the Ricker nonlinearity. The backward bifurcation in figure (4.2B) is an example illustrating a saddle-node bifurcation (at $r_0 \approx 0.4$) that results a multi-stable equilibrium (strong Allee) scenario, as shown in figure (4.1). In this example the positive equilibria destabilize (into a period doubling route to chaos) just outside the Allee interval $0.4 < r_0 < 1$. In other examples, using different parameter values, this destabilization can occur at a value of $r_0 < 1$ so that the multi-attractor scenario of the strong Allee effect involves a stable cycle or even more complicated attractor. For examples and further results concerning the relationship between backward bifurcations and strong Allee effects, see [26]. The complex dynamics that can arise in this model, particularly when positive non-equilibrium attractors are present for $r_0 < 1$ are studied in [90], although not from this bifurcation point-of-view.

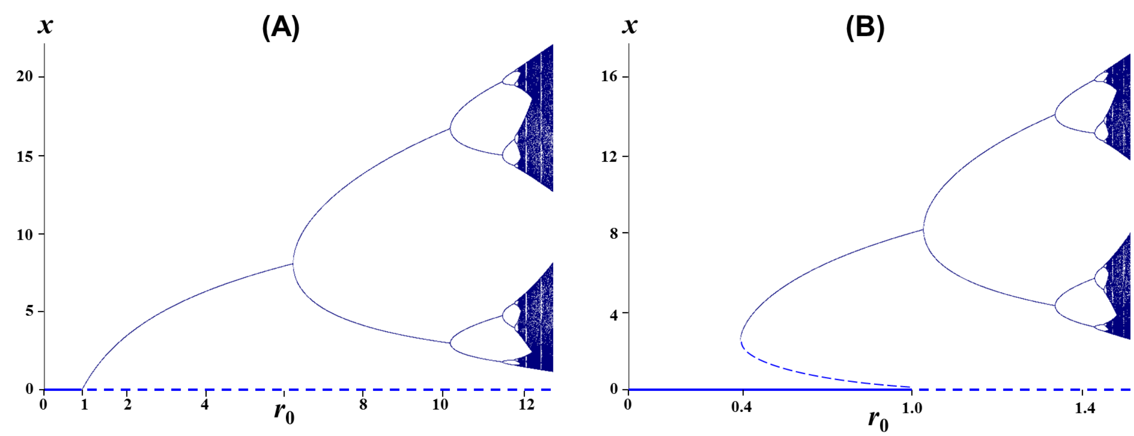

Figure 4.2: Sample bifurcation diagrams for equation (4.27) with $c = 0.3$ and $\alpha = 3$ and different values of the protection from predation parameter s.
A. $s = 0.05$ and $\kappa = c - \alpha s = 0.15 > 0$ so that the bifurcation at $r_0 = \beta e^{-3} = 1$ is forward and stable.
B. $s = 1$ and $\kappa = -2.7 < 0$ so that the bifurcation at $r_0 = \beta e^{-3} = 1$ is backward and unstable (dashed line).

We now consider an evolutionary version of equation (4.27) to which we can apply the results of Sections 4.3 and 4.4. For our application we consider the case when the inherent (density and predation free) per capita birth rate b and the predation protection factor s are subject to evolutionary adaptation. We think of these per capita quantities as characteristics of an individual and that they are determined by a suite of phenotypic traits $\hat{v}$ of the individual. Thus, $b = b(\hat{v})$ and $s = s(\hat{v})$. We assume that there is a trait vector that maximizes b and one that maximizes s, but these optimizing trait vectors are not the same. The idea is that there are trade-offs in the allocation of energy, behavioural activities, and resources towards reproduction and towards the avoidance of predators. For example, traits that promote physiological and behavioural characteristics promote successful herding or flocking or schooling in order to avoid predation are not necessarily traits that make for optimal reproduction.

Since we have set no units or scales for the traits, we assume without loss in generality that b is maximal at $\hat{v} = \hat{0}$ and s is maximal at $\hat{v} = (1, 1, \cdots, 1)^T$. Specifically, we assume (as is often done in evolutionary models [100]) that these coefficients have a multi-variate Gaussian-type distribution about these maximal points:

$$b(\hat{v}) = \beta \exp\left(-\sum_{i=1}^{n} \frac{v_i^2}{2b_i}\right), \quad s(\hat{v}) = s_0 \exp\left(-\sum_{i=1}^{n} \frac{(v_i-1)^2}{2s_i}\right)$$

where b_i and s_i are positive real numbers (variances), $\beta > 0$ is the maximal possible value of $b(\hat{v})$, and $s_0 \geq 0$ is the maximal possible value of $s(\hat{v})$. The resulting 1×1 projection matrix $P(x, \hat{u}, \hat{v})$ for (4.27) is independent of $\hat{u}$ and its single entry $p_{11}(x, \hat{v})$ equals the dominant eigenvalue, i.e. $p_{11}(x, \hat{v}) = r(x, \hat{v})$ where

$$r(x, \hat{v}) = \beta e^{-\alpha} \exp\left(-\sum_{i=1}^{n} \frac{v_i^2}{2b_i}\right) \exp\left(-cx + \alpha \frac{s(\hat{v})x}{1+s(\hat{v})x}\right).$$

The Darwinian equations (4.6) are

$$\begin{aligned} x(t+1) &= r(x(t), \hat{u}(t))\, x(t) \\ \hat{u}(t+1) &= \hat{u}(t) + M\ \nabla_{\hat{v}} \ln r(x(t), \hat{u}(t)) \end{aligned}$$

with

$$\nabla_{\hat{v}} \ln r(x, \hat{v}) = -\begin{pmatrix} \frac{v_1}{b_1} \\ \vdots \\ \frac{v_n}{b_n} \end{pmatrix} - \alpha \frac{s(\hat{v})\, x}{(1+s(\hat{v})\, x)^2} \begin{pmatrix} \frac{v_1-1}{s_1} \\ \vdots \\ \frac{v_n-1}{s_n} \end{pmatrix}.$$

Since

$$\nabla_{\hat{v}} \ln r(0, \hat{v})\big|_{\hat{v}=\hat{u}} = -\begin{pmatrix} \frac{u_1}{b_1} \\ \vdots \\ \frac{u_n}{b_n} \end{pmatrix}$$

we see that the only critical trait is

$$\hat{u}^* = \hat{0}.$$

and hence the only extinction equilibrium is $(x, \hat{u}) = (0, \hat{0})$. Our bifurcation parameter $r_0^* = r(0, \hat{0})$ reduces to

$$r_0^* = \beta e^{-\alpha}.$$

Under the added assumption that the traits are not correlated, so that the variance-covariance matrix $M = diag\left(\sigma_i^2\right)$ is a diagonal matrix, the model equations for our evolutionary version of (4.27) are

$$x(t+1) = r_0^* \exp\left(-\sum_{i=1}^{n} \frac{u_i^2(t)}{2b_i}\right) x(t) \exp\left(-cx(t) + \alpha \frac{s(\hat{u}(t))x(t)}{1+s(\hat{u}(t))x(t)}\right) \tag{4.28a}$$

$$u_i(t+1) = u_i(t) - \sigma_i^2 \left(\frac{u_i(t)}{b_i} + \alpha \frac{u_i(t)-1}{s_i} \frac{s(\hat{u}(t))x(t)}{(1+s(\hat{u}(t))x(t))^2}\right) \tag{4.28b}$$

for $i = 1, \cdots, n$. Our goal is to apply Corollary 4.4 to these difference equations, toward which end we must consider H2 and H5.

H2 holds with $\tilde{p}_{11}(\hat{v}) = r_0(\hat{v})$ and $\bar{p}(x, \hat{u}, \hat{v}) = \bar{r}(x, \hat{v})$, since $\bar{r}(0, \hat{v}) \equiv 1$. A calculation shows that the Hessian $H(0,\hat{u}^*,\hat{u}^*) = [\partial^0_{v_i v_j} \ln \bar{r}]$ is the diagonal matrix

$$H(0,\hat{u}^*,\hat{u}^*) = diag\left[-\frac{1}{b_i}\right] \ .$$

Thus, H5 holds (see Remark 4.4). By Theorem 4.4 the bifurcating continuum of positive equilibrium pairs $(r_0^*, (x,\hat{u}))$, near the bifurcation point $(1,(0,\hat{0}))$, has the parametric representation

$$r_0^*(\varepsilon) = 1 + \kappa^* \varepsilon + O(\varepsilon^2) \tag{4.29a}$$

$$x(\varepsilon) = \varepsilon + O(\varepsilon^2) \tag{4.29b}$$

$$\hat{u}(\varepsilon) = \hat{u}_1 \varepsilon + O(\varepsilon^2) \tag{4.29c}$$

for $\varepsilon \gtrapprox 0$ where, by the formulas (4.18) and (4.19),

$$\kappa^* = c - \alpha s_0 \exp\left(-\sum\nolimits_{i=1}^n \frac{1}{2s_i}\right) \tag{4.30}$$

$$\hat{u}_1 = \alpha s_0 \exp\left(-\sum\nolimits_{i=1}^n \frac{1}{2s_i}\right) \begin{pmatrix} \frac{b_1}{s_1} \\ \vdots \\ \frac{b_n}{s_n} \end{pmatrix} \in \mathbb{R}^n_+.$$

By Corollary 4.4(a,b), the direction of bifurcation determines the stability of the bifurcating positive equilibria provided

$$\sigma_i^2 < 2b_i \text{ for all } i = 1, \cdots, n$$

that is to say, provided the speed of evolution is not too fast. Under this assumption, we have the following conclusions concerning the bifurcation at $r_0^* = 1$ for the Darwinian model (4.28).

1. *(Forward bifurcations)* The bifurcation of the continuum $\mathcal{C}$ of positive equilibria for (4.28b) is forward and consequently stable if

$$\alpha s_0 \exp\left(-\sum\nolimits_{i=1}^n \frac{1}{2s_i}\right) < c. \tag{4.31}$$

This occurs if the negative density effects, as described by the Ricker coefficient c, are large enough to dominate the positive effects from the Allee effect attributed to density protection from predation, as encapsulated by the quantity on the left side of the inequality (4.31). Thus, mechanisms that promote a forward bifurcation are: a low predation intensity α, a low maximum possible predation protection coefficient s_0, and small variances s_i (i.e. the largest predator protection coefficients $s(\hat{v})$ are attained only for trait vectors narrowly distributed around the maximal trait vector $\hat{v} = (1,\cdots,1)^T$).

Note that the entries in $\hat{u}_1$ in (4.29c) are positive if $\alpha s_0 > 0$, i.e. if both predation and predation protection are present. In this case, we see that near the bifurcation point, the trait components $u_i(\varepsilon)$ of the bifurcating positive equilibria are positive. As a result, for $r_0^* \gtrapprox 1$ the stable, positive equilibria have trait components that do not maximize the inherent birth rate. Indeed, an even stronger conclusion follows directly from the trait equilibrium equations (4.28b):

$$\frac{u_i}{b_i} + \alpha \frac{u_i - 1}{s_i} \frac{s(\hat{u})x}{(1 + s(\hat{u})x)^2} = 0 \text{ for all } i = 1, \ldots, n. \tag{4.32}$$

This shows, when $\alpha > 0$ and $s_0 > 0$, that for any positive equilibrium $(x, \hat{u})$ of (4.28), the equilibrium trait components u_i cannot equal 0 or 1. For those equilibrium pairs from the continuum $\mathcal{C}$ the trait components form a continuum of equilibrium trait vectors $\hat{u}$ which must, therefore, have components that lie entirely in the interval $0 < u_i < 1$ (whether the equilibria are stable or not). It follows that for those positive equilibria from $\mathcal{C}$ which are in fact stable (such as those for $r_0^* \gtrapprox 1$), we can say that evolution selects a vector of traits that neither maximizes the inherent birth rate $b(\hat{u})$ (which occurs at $\hat{u} = \hat{0}$) nor the predator protection coefficient $s(\hat{u})$ (which occurs at $\hat{u} = (1, 1, \cdots, 1)^T$). One might say, then, that evolution trades-off a smaller inherent birth rate in favor of some predator protection.

When predation and/or predation protection is absent ($\alpha = 0$ and/or $s_0 = 0$) in the model, then clearly inequality (4.31) holds and the bifurcation at $r_0^* = 1$ is forward and stable. In this case, the equilibrium equation (4.32) for the traits u_i implies $u_i = 0$ for any positive equilibrium pair and, not surprisingly, evolution selects to maximize the inherent bifurcation rate $b(\hat{u})$. □

(2) *(Backward bifurcations)*. The bifurcation of the continuum $\mathcal{C}$ of positive equilibria for (4.28b) is backward and consequently unstable if

$$\alpha s_0 \exp\left(-\sum\nolimits_{i=1}^{n} \frac{1}{2s_i}\right) > c. \tag{4.33}$$

This occurs only if predation is present $\alpha > 0$ and density protection from predation is also present $s_0 > 0$. Inequality (4.33) holds if predation intensity α and/or predator protection s_0 are large (relative to the negative density effects c). Also promoting a backward bifurcation are large variances s_i, that is to say, when a high level of predator protection $s(\hat{v})$ is attained for a wide distribution of trait vectors $\hat{v}$. □

Our general results in Section 4.4 concern equilibrium properties in a neighbourhood of the bifurcation point and do not imply anything about the dynamics outside such a neighbourhood. As in the non-evolutionary model (4.27), we expect it to be true that the positive equilibria on the continuum $\mathcal{C}$ for the evolutionary model (4.28) do not necessarily retain the stability properties that they possess near the bifurcation point. In particular, in the case of a forward/stable bifurcation we would expect that, at least for some model parameter values, the stable positive equilibria will destabilize with increasing r_0^* and even give rise to a sequence of bifurcations that result complicated, chaotic dynamics. In the case of a backward/unstable bifurcation, in addition to this phenomenon, we would also anticipate the potential for strong Allee effects on an interval of r_0^* values less than 1. We will not study these questions about the dynamics of (4.28) in this work where our theory is focussed on the local bifurcation at $r_0^* = 1$.

However, we can provide a few selected numerical simulations that, in addition to illustrating the local bifurcation predicted by our theorems, also illustrate the kinds of secondary bifurcations and strong Allee effects as in the non-evolutionary case (cf. figure (4.2)). Figure (4.3) shows two sample bifurcation diagrams for the evolutionary model (4.28) with two traits, i.e. $n = 2$. We observe that the oscillations of the traits u_i as the bifurcation parameter changes are of small amplitude. The plots in figure (4.3A) are from parameter values for which $\kappa^* > 0$ and, hence, a forward, stable bifurcation occurs at $r_0^* = 1$. As with the non-evolutionary version of the model in figure (4.1), further increases in r_0^* result in the familiar period doubling route to chaotic dynamics. In figure (4.3B) the same parameter values are used except that the predator protection coefficient s_0 is increased to the extent that $\kappa^* < 0$ and, as a result, a backward, unstable bifurcation occurs. The result is a bifurcation diagram that shows a saddle-node bifurcation (at $r_0^* \approx 0.3$) creating an interval of strong Allee effects with both a stable extinction equilibrium $(x, \hat{u}) = (0, \hat{0})$ and a stable positive equilibrium. In this example, one sees from figure (4.3B) that the positive equilibrium loses stability through a period doubling at a value of r_0^* less than 1. This results in an interval of r_0^* values less than 1 for which there is a strong Allee effect that involves a stable positive 2-cycle instead of a positive equilibrium. Indeed, in figure (4.4A) we observe that the sample orbits approach a stable positive equilibrium so that we have a strong Allee effect involving stable positive equilibria whereas in figure (4.4B) we observe that the sample orbits approach a stable 2-cycle so that we witness a strong Allee effect involving stable 2-cycles.

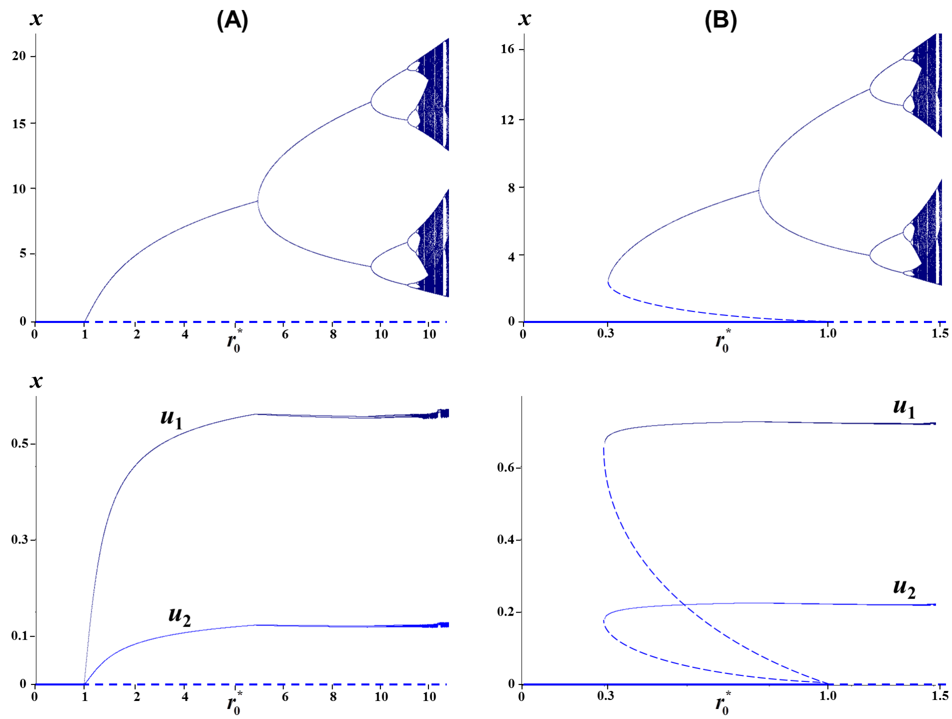

Figure 4.3: Sample bifurcation diagrams for equations (4.28) with $c = 0.3, \alpha = 3, b_1 = 3, b_2 = 1, s_1 = 1$, and $s_2 = 3$ and variances $\sigma_1 = 0.45$ and $\sigma_2 = 0.2$ for different values of the maximal protection from predation parameter s_0 .
A. $s_0 = 0.1$ and $\kappa^* = 0.3 - 0.3e^{-2/3} \approx 1.46 \times 10^{-1} > 0$ so that the bifurcation at $r_0^* = \beta e^{-3} = 1$ is forward and stable.
B. $s_0 = 1$ and $\kappa^* = 0.3 - 3e^{-2/3} \approx -1.24 < 0$ so that the bifurcation at $r_0^* = \beta e^{-3} = 1$ is backward and unstable (dashed lines).

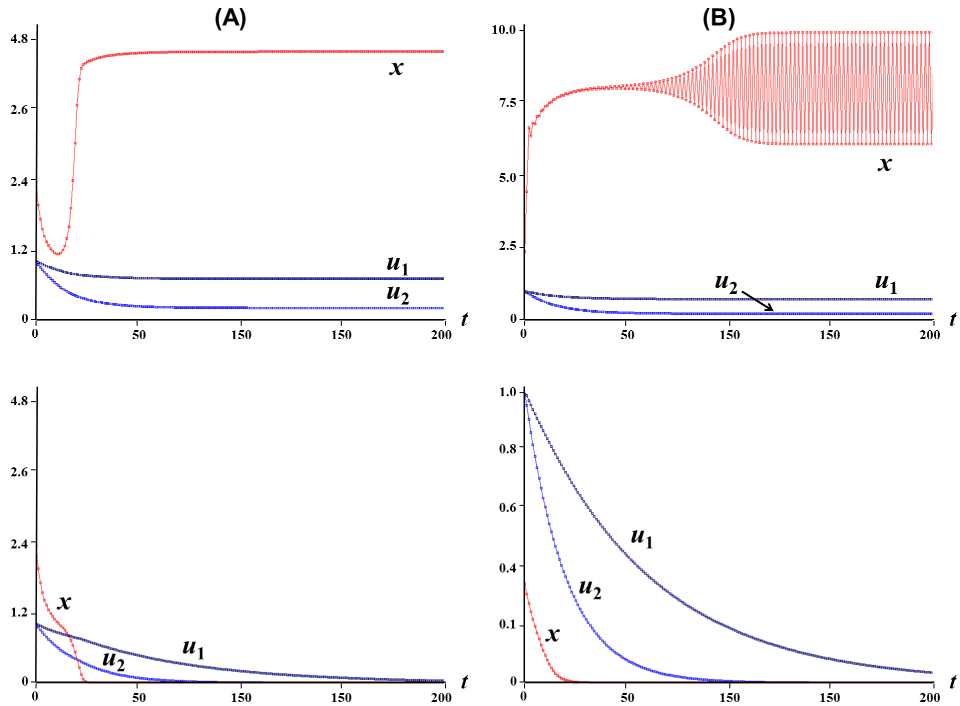

Figure 4.4: Shown are sample orbits for equations (4.28) with the same parameter values used in Figure 3B when a backward bifurcation creates an interval of $r_0^* < 1$ for which there is a strong Allee effect.
A. For $r_0^* = \beta e^{-3} = 0.4$ the bifurcation diagram Figure 3B shows a stable extinction equilibrium and a stable positive equilibrium. The upper graph in column A shows plots of the solution with initial conditions $x = 2.35$, $u_1 = u_2 = 1$ and that tends to the positive equilibrium. The lower graph shows plots of the solution with initial conditions $x = 2.33$, $u_1 = u_2 = 1$ and that tends to the extinction equilibrium $(x, \hat{u}) = (0, \hat{0})$.
B. For $r_0^* = \beta e^{-3} = 0.9$ the bifurcation diagram Figure 3B shows a stable extinction equilibrium and a stable 2-cycle. The upper graph in column B shows plots are of the solution with initial conditions $x = 2.35$, $u_1 = u_2 = 1$ and that tends to the positive 2-cycle. The lower graph in column B shows plots of the solution with initial conditions $x = 0.33$, $u_1 = u_2 = 1$ and that tends to the extinction equilibrium $(x, \hat{u}) = (0, \hat{0})$.

4.6 Concluding remarks

A fundamental property of population dynamic models, when the extinction state destabilizes due to a change in a model parameter, is the occurrence of a bifurcation which results in the presence of positive equilibria. Typically the stability of these bifurcating equilibria depend on the direction of bifurcation (Theorem 4.1). In this chapter we investigate this basic bifurcation phenomenon for an evolutionary version of a general matrix model for the dynamics of a structured population. The model assumes that the entries of the model's projection matrix (i.e. the per capita birth, survival and class transition rates) depend on a vector of phenotypic traits, each of which is subject to Darwinian evolution, and tracks the dynamics of the population and the vector of mean traits [100]. We define the notion of a critical trait vector, which is associated with the existence of an extinction equilibrium in the model, and obtain conditions under which an extinction equilibrium destabilizes (Theorem 4.2) and conditions under which a continuum of positive equilibria bifurcates from the extinction equilibrium, as the inherent population growth rate (at the critical trait) increases through 1 (Theorem 4.4). We further obtain conditions under which stability of the bifurcating equilibria is determined by the direction of bifurcation and conditions under which it is not (Theorem 4.4).

It is shown in [75] that the bifurcating continuum $\mathcal{C}^*$ of positive equilibria in Theorem 4.4 has a global extent in $\mathbb{R}_+ \times (\mathbb{R}_+^m \times V)$ in that it connects to the boundary of this cone (∞ is included in the boundary). In general, however, the stability/instability results in Theorem 4.4 hold only in a neighbourhood of the bifurcation point. This is illustrated in the example studied in Section 4.5 where secondary bifurcations occur outside the neighbourhood of the bifurcation point. Whether or not such bifurcations occur are model dependent (which is true in non-evolutionary matrix models as well).

In non-evolutionary matrix models, backward bifurcations are often associated with strong Allee effects, i.e. multiple attractors for values of $r_0^* < 1$ one of which is extinction and the other which is a survival attractor [26]. While conditions sufficient for the occurrence of a backward bifurcation are given in Section 4.4, its relation to strong Allee effects is not investigated in this work and may present a possibility of future work. A backward induced strong Allee effect is shown to occur, by simulations, in the example studied in Section 4.5. Furthermore, other possibilities of future work in related areas might include the study of other types of bifurcations, such as bifurcation of cycles, which have a very relevant biological meaning, to both non-evolutionary and evolutionary models and study situations where the projection matrix is not primitive but only irreducible. When the projection matrix is primitive, then the direction of the bifurcation determines the stability of the bifurcating equilibria as we have seen. This is due to the fact that at least locally around the bifurcation point only one eigenvalue of the projection matrix leaves the unit circle, that is the strictly dominant eigenvalue provided by Perron-Frobenius theorem. When the matrix is not primitive certainly this situation is a greater challenge and it is not expected that stability would be determined solely by the direction of bifurcation since the bifurcation would have higher co-dimension with more than one eigenvalue leaving the unit circle simultaneously, according to Perron-Frobenius theorem.

Bibliography

[1] P. A. Abrams. Modelling the adaptive dynamics of traits involved in inter- and intraspecific interactions: An assessment of three methods. *Ecology Letters*, 4(2):166–175, 2001.

[2] P. A. Abrams. 'Adaptive dynamics' vs. 'adaptive dynamics'. *Journal of Evolutionary Biology*, 18(5):1162–1165, 2005.

[3] E. Accinelli and E. Carrera. Corruption driven by imitative behaviour. *Economics Letters*, 117(1):84–87, 2012.

[4] E. Accinelli and E. Carrera. Corruption, inequality, and income taxation. In Jean-Pierre Bourguignon, Rolf Jeltsch, Alberto Pinto, and Marcelo Viana, editors, *Dynamics, Games and Science*, volume 1 of *CIM Series in Mathematical Sciences*, chapter 1, pages 1–16. Springer-Verlag, 2015.

[5] E. Accinelli, F. Martins, J. Oviedo, A. A. Pinto, and L. Quintas. Who controls the controller? a dynamical model of corruption. *The Journal of Mathematical Sociology*, 41(4):220–247, 2017.

[6] E. Accinelli, L. Policardo, and E. Carrera. On the dynamics and effects of corruption on environmental protection. In Alberto Pinto and David Zilberman, editors, *Modelling, Dynamics, Optimization and Bioeconomics I*, volume 73 of *Springer Proceedings in Mathematics & Statistics*, chapter 3, pages 23–41. Springer-Verlag, 2014.

[7] H. W. Allen and K. W. Allen. Vote fraud and data validity. In Jerome Clubb, William Flanigan, and Nancy Zingale, editors, *Analyzing Electoral History: A Guide to the Study of American Voter Behavior*, pages 153–195. Sage Publications, 1981.

[8] Kyle Bagwell and Robert W. Staiger. A theory of managed trade. *The American Economic Review*, 80(4):779–795, 1990.

[9] Kyle Bagwell and Robert W. Staiger. Enforcement, private political pressure, and the general agreement on tariffs and trade/world trade organization (GATT/WTO) escape clause. *The Journal of Legal Studies*, 34(2):471–513, 2005.

[10] Nigel Banik, F. A. Ferreira, J. Martins, and A. A. Pinto. An economical model for dumping by dumping in a cournot model. In M. M. Peixoto, A. A. Pinto, and D. A. Rand, editors, *Dynamics, Games and Science II, DYNA 2008, in Honour of Maurício Peixoto and David Rand*, volume 2 of *Springer Proceedings in Mathematics*, pages 141–154. Springer-Verlag Berlin Heidelberg, 2011.

[11] G. Becker. A theory of competition among pressure groups for political influence. *The Quarterly Journal of Economics*, 98(3):371–400, August 1983.

[12] A. Berman and R. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[13] Cristina Bicchieri and John Duffy. Corruption cycles. *Political Studies*, 45(3):477–495.

[14] James A. Brander. Intra-industry trade in identical commodities. *Journal of International Economics*, 11(1):1 – 14, 1981.

[15] James A. Brander and Barbara J. Spencer. Export subsidies and international market share rivalry. *Journal of International Economics*, 18(1):83 – 100, 1985.

[16] Brasil. Brasil corruption report, 2017. URL: http://www.business-anti-corruption.com/country-profiles/brazil.

[17] Jeremy I. Bulow, John D. Geanakoplos, and Paul D. Klemperer. Multimarket oligopoly: Strategic substitutes and complements. *Journal of Political Economy*, 93(3):488–511, 1985.

[18] S. V. Campbell. Perception is not reality: The FCPA, Brazil, and the mismeasurement of corruption. *Minnesota Journal of International Law*, 22(1):247–281, 2013.

[19] H. Caswell. *Matrix Population Models: Construction, Analysis and Interpretation*. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts, second edition, 2001.

[20] M. Choubdar, E. Farias, F. A. Ferreira, and A. A. Pinto. Uncertainty costs on an international duopoly with tariffs. In *Recent Advances in Applied Economics, Proceedings of the 6th International Conference on Applied Economics, Business and Development (AEBD'14)*, pages 13–16, Lisbon, Portugal, 2014.

[21] M. Choubdar, J. P. Zubelli, and A. A. Pinto. Nash and social welfare impact in international trade. In *Recent Advances in Applied Economics, Proceedings of the 6th International Conference on Applied Economics, Business and Development (AEBD'14)*, pages 23–26, Lisbon, Portugal, 2014.

[22] F. Courchamp, L. Berec, and J. Gascoigne. *Allee Effects in Ecology and Conservation*. Oxford University Press, Oxford, Great Britain, 2008.

[23] J. M. Cushing. *An Introduction to Structured Population Dynamics*, volume 71 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998.

[24] J. M. Cushing. A bifurcation theorem for darwinian matrix models. *Nonlinear Studies*, 17(1):1–13, 2010.

[25] J. M. Cushing. On the relationship between r and R_0 and its role in the bifurcation of stable equilibria of darwinian matrix models. *Journal of Biological Dynamics*, 5(3):277–297, 2011.

[26] J. M. Cushing. Backward bifurcations and strong allee effects in matrix models for the dynamics of structured populations. *Journal of Biological Dynamics*, 8(1):57–73, 2014.

[27] J. M. Cushing. One dimensional maps as population and evolutionary dynamic models. In J. M. Cushing, M. Saleem, H. M. Srivastava, M. A. Khan, and M. Merajuddin, editors, *Applied Analysis in Biological and Physical Sciences, ICMBAA, Aligarh, India, June 2015*, volume 186 of *Springer Proceedings in Mathematics & Statistics (PROMS)*, pages 41–62. Springer India, New Delhi, India, 2016.

[28] J. M. Cushing. Matrix models and population dynamics. In James P. Keener Mark Lewis, A. J. Chaplain and Philip K. Maini, editors, *Mathematical Biology*, volume 14 of *IAS/Park City Mathematics Series*, pages 47–150. American Mathematical Society, Providence, I,2009.

[29] J. M. Cushing, A. A. Pinto F. Martins, and A. Veprauskas. A bifurcation theorem for evolutionary matrix models with multiple traits. *The Journal of Mathematical Biology*, 75(2):491–520, August 2017.

[30] Brian Dennis. Allee effects: Population growth, critical density, and the chance of extinction. *Natural Resource Modeling*, 3(4):481–538, 1989.

[31] F. Dercole and S. Rinaldi. *Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications*. Princeton University Press, Princeton, New Jersey, 2008.

[32] Avinash Dixit. International trade policy for oligopolistic industries. *The Economic Journal*, 94 supplement:1–16, 1984.

[33] Avinash Dixit. Strategic aspects of trade policy. In T. Bewley, editor, *Advances in Economic Theory*, chapter 9, pages 329–362. Cambridge University Press, 1987.

[34] Avinash Dixit and Gene M. Grossman. Targeted export promotion with several oligopolistic industries. *Journal of International Economics*, 21(3):233 – 249, 1986.

[35] Jonathan Eaton and Gene M. Grossman. Optimal trade and industrial policy under oligopoly. *The Quarterly Journal of Economics*, 101(2):383–406, May 1986.

[36] P. Egger and H. Winner. Evidence on corruption as an incentive for foreign direct investment. *European Journal of Political Economy*, 21(4):932–952, 2005.

[37] S. N. Elaydi. *An Introduction to Difference Equations*. Springer-Verlag, New York, 1996.

[38] European Union (EU). European anti-corruption report, February 2014. URL: https://ec.europa.eu/home-affairs/sites/homeaffairs/files/e-library/documents/policies/organized-crime-and-human-trafficking/corruption/docs/acr_2014_en.pdf.

[39] Eric O'N. Fisher and Charles A. Wilson. Price competition between two international firms facing tariffs. *International Journal of Industrial Organization*, 13(1):67 – 87, 1995.

[40] R. Gibbons. *A Primer in Game Theory*. Pearson Prentice Hall, 1999.

[41] J. Grego. *A History of Parliamentary Elections and Electioneering in the Old Days*. Chatto and Willis, Picadilly, 1886. Re-edited by Nabu Press, November 2013. URL: https://archive.org/details/historyofparliam00greg.

[42] G. M. Grossman. Strategic export promotion: A critique. In Paul Krugman, editor, *Strategic Trade Policy and the New International Economics*, chapter 3, pages 47–68. MIT Press, Cambridge, MA, 2007.

[43] G. M. Grossman and E. Helpman. Protection for sale. *The American Economic Review*, 84(4):833–850, September 1994.

[44] S. Gupta, H. Davoodi, and R. Alonso-Terme. Does corruption affect income inequality and poverty? *Economics and Governance*, 3:23–45, 2002.

[45] E. Helpman. Increasing returns, imperfect markets, and trade theory. In R. W. Jones and P. B. Kenen, editors, *Handbook of International Economics Volume I*, chapter 7, pages 325–365. North Holland Press, 1984.

[46] M. Hirsch, S. Smale, and R. Devanay. *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Elsevier Ltd, Third edition, 2012.

[47] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.

[48] L. W. Huberts. *Expert views on corruption around the world*. Vrij University, Amsterdam, 1996.

[49] L. Hurwicz. But who will guard the guardians?, 2007. Nobel Prize Lecture. URL: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2007/hurwicz_lecture.pdf.

[50] A. K. Jain. Corruption: A review. *Journal of Economics Surveys*, 15(1):1–51, 2001.

[51] Juvenal. *The Satires*. Oxford's World Classics. Clarendon Press, 1991. The Latin original is available in http://www.thelatinlibrary.com/juvenal.html and an English translation is available in https://en.wikisource.org/wiki/The_Satires_of_Juvenal.

[52] H. Keilhöfer. *Bifurcation Theory: An Introduction with Applications to PDE's*, volume 156 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004.

[53] J.-M. M. Kilolo. Country size, trade liberalization and transfers. MPRA Paper, University Library of Munich, Germany, 2013. URL: https://mpra.ub.uni-muenchen.de/47996/1/MPRA_paper_47996.pdf.

[54] I. Kornai. Hidden in an envelope. gratitude payments to medical doctors in hungary. In L. Dahrendorf and Y. Elkana, editors, *The paradoxes of unintended consequences*, pages 195–214. Central European University Press, 2000.

[55] Kala Krishna. Trade restrictions as facilitating practices. *Journal of International Economics*, 26(3):251 – 270, 1989.

[56] A. Kruger. The political economy of the rent seeking society. *American Economic Review*, 64(3):291–303, 1974.

[57] R. Lande. Natural selection and random genetic drift in phenotypic evolution. *Evolution*, 30(2):314–334, 1976.

[58] R. Lande. A quantitative genetic theory of life history evolution. *Ecology*, 63(3):607–615, 1982.

[59] L. P. Lefkovitch. The study of population growth in organisms grouped by stages. *Biometrics*, 21(1):1–18, 1965.

[60] P. H. Leslie. On the use of matrices in certain population mathematics. *Biometrika*, 33(3):183–212, 1945.

[61] P. H. Leslie. Some further notes on the use of matrices in population mathematics. *Biometrika*, 35(3/4):213–245, 1948.

[62] L. Lessig. *Republic, Lost: How Money Corrupts Congress - and a Plan to Stop It.* Twelve, 2011.

[63] E. G. Lewis. On the generation and growth of a population. *Sankhyā: The Indian Journal of Statistics (1933-1960)*, 6(1):93–96, 1942.

[64] P.-C. Liao. Rivalry between exporting countries and an importing country under incomplete information. *Academia Economic Papers*, 32(4):605–630, December 2004.

[65] Nuno Limão and Kamal Saggi. Tariff retaliation versus financial compensation in the enforcement of international trade agreements. *Journal of International Economics*, 76(1):48 – 60, 2008.

[66] Nuno Limão and Kamal Saggi. Size inequality, coordination externalities and international trade agreements. *European Economic Review*, 63:10 – 27, 2013.

[67] Francis T. Lui. An equilibrium queuing model of bribery. *Journal of Political Economy*, 93(4):760–781, 1985.

[68] Francis T. Lui. A dynamic model of corruption deterrence. *Journal of Public Economics*, 31:215–236, November 1986.

[69] F. Martins, A. A. Pinto, and J. P. Zubelli. Nash and social welfare impact in an international trade model. *Journal of Dynamics and Games*, 4(2):149–173, April 2017.

[70] J. Martins, N. Banik, and A. A. Pinto. Optimal strategies for international expansion with and without making dumping. To appear.

[71] J. Martins, N. Banik, and A. A. Pinto. A repeated strategy for dumping. In Lluís Alseda, Jim M. Cushing, Saber Elaydi, and A. A. Pinto, editors, *Discrete Dynamical Systems and Applications, ICDEA 2012*, volume 180 of *Springer Proceedings in Mathematics & Statistics*, pages 145–153. Springer-Verlag Berlin Heidelberg, 2016.

[72] P. Mauro. Corruption and growth. *The Quarterly Journal of Economics*, 110:681–712, August 1995.

[73] Brian J. McGill and Joel S. Brown. Evolutionary game theory and adaptive dynamics of continuous traits. *Annual Review of Ecology, Evolution, and Systematics*, 38(1):403–435, 2007.

[74] J. McMillan. *Game Theory in International Economics.* Harwood Academic Publishers, Chur, Switzerland, 1986.

[75] Emily P. Meissen, Kehinde R. Salau, and Jim M. Cushing. A global bifurcation theorem for darwinian matrix models. *Journal of Difference Equations and Applications*, 22(8):1114–1136, 2016.

[76] A. Mishra. Persistence of corruption: some theoretical perspectives. *World Development*, 34:349–358, 2006.

[77] J. S. Nye. Corruption and political development: a cost-benefit analysis. *The American Political Science Review*, 61:417–427, 1967.

[78] M. Olson. Dictatorship, democracy, and development. *The American Political Science Review*, 87:567–576, 1993.

[79] Platão. *A República*. Europa-América, 1975.

[80] L. Quintas. Computing equilibrium points for Q-cyclic games. *Proceding of the II Latin American Congress of Applied Mathematics, Río de Janeiro, Brazil*, II:576–598, 1983.

[81] L. Quintas. A note on polymatrix games. *International Journal of Game Theory*, 18:261–272, September 1989.

[82] W. E. Ricker. Stock and recruitment. *Journal of the Fisheries Research Board of Canada*, 11(5):559–623, 1954.

[83] D. A. Roff. *The Evolution of Life Histories: Theory and Analysis*. Chapman and Hall, New York, 1992.

[84] H. Root. *Small countries, Big Lessons: Governance and the Rise of East Asia*. Oxford University Press, 1996.

[85] S. Rose-Ackerman. The economics of corruption. *Journal of Public Economics*, 4:187–203, February 1975.

[86] S. Rose-Ackerman. *Corruption. A Study in Political Economy*. Academic Press, 1978.

[87] B. Rothstein. *Social traps and the Problem of Trust*. Cambridge University Press, 2005.

[88] R. Sah. Corruption across countries and regions: Some consequences of local osmosis. *Journal of Economic Dynamics and Control*, 31:2573–2598, August 2007.

[89] F. Schaffer and A. Schedler. What is vote buying? the limits of the market model, 2005. URL: https://polsci.umass.edu/file/904/download?token=wqyQOKoP.

[90] Sebastian J. Schreiber. Allee effects, extinctions, and chaotic transients in simple population models. *Theoretical Population Biology*, 64(2):201 – 209, 2003.

[91] M. Seligson. The impact of corruption on regime legitimacy: A comparative study of four Latin American countries. *Journal of Politics*, 64:408–433, 2002.

[92] A. Shleifer and R. Vishny. Corruption. *The Quarterly Journal of Economics*, 108:599–617, 1993.

[93] I. M. Soto and W. Cortez. The impact of political alternation on corruption in Mexico. *Revista de Ciencia Política*, 35(2):371–392, 2015. URL: http://www.redalyc.org/articulo.oa?id=32442271006.

[94] Barbara J. Spencer and James A. Brander. International R&D rivalry and industrial strategy. *The Review of Economic Studies*, 50(4):707–722, 1983.

[95] R. Staiger. International rules and institutions for trade policy. In R. W. Jones and P. B. Kenen, editors, *Handbook of International Economics Volume III*, chapter 29, pages 1495–1551. North Holland Press, 1984.

[96] D. F. Thompson. *Ethics in Congress: From Individual to Institutional Corruption*. Brookings Institution Press, 1995.

[97] T. Thompson and A. Shah. Transparency international's corruption perceptions index: Whose perceptions are they anyway? discussion draft, 2005. URL: http://siteresources.worldbank.org/PSGLP/Resources/ShahThompsonTransparencyinternationalCPI.pdf.

[98] M. B. Usher. A matrix approach to the management of renewable resources, with special reference to selection forests. *Journal of Applied Ecology*, 3(2):355–367, 1966.

[99] Amy Veprauskas and J. M. Cushing. Evolutionary dynamics of a multi-trait semelparous model. *Discrete & Continuous Dynamical Systems - B*, 21(2):655–676, 2016.

[100] T. L. Vincent and J. S. Brown. *Evolutionary Game theory, Natural selection, and Darwinian dynamics*. Cambridge University Press, Cambridge, UK, 2005.

[101] Gottfried von Haberler. *The Theory of International Trade with its Application to Commercial Policy*. Macmillan, New York, 1937. Translated from the original from 1933 in German by Alfred Stonier and Fredric Benham. This work is freely available in the website of the von Mises institute: https://mises.org/library/international-trade.

[102] J. Von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Brookings Institution Press, 1944. This work is in the public domain. It is freely available in https://archive.org/details/in.ernet.dli.2015.215284.

[103] W. J. Weibull. *Evolutionary Game Theory*. The MIT Press, 1995.

[104] Wikipedia. Corruption in Brazil, 2017. URL: https://en.wikipedia.org/wiki/Corruption_in_Brazil.