

The Emergence and Evolution of Social Norms:
Rational Cooperation in a Finitely Repeated
Prisoners' Dilemma

by

Matthew T. Burke
Class of 2007

A thesis submitted to the
faculty of Wesleyan University
in partial fulfillment of the requirements for the
Degree of Bachelor of Arts
with Departmental Honors in Economics

ABSTRACT

The Emergence and Evolution of Social Norms:
Rational Cooperation in a Finitely Repeated Prisoners' Dilemma

By

Matthew T. Burke

In reaction to the contractarian approach to morals of David Gauthier (1986) that assumes individuals rationally restrict their own self-interest in order to facilitate cooperation and the associated benefits, this paper investigates cooperation among rational players in a finitely repeated Prisoners' Dilemma through game theoretic analysis. This paper attempts to show that rational cooperation is possible in a finitely repeated Prisoners' Dilemma when two-sided information asymmetry is incorporated regarding player types, and that this analysis is consistent with an evolutionarily stable equilibrium.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support of many caring people.

My friends and family – For giving me counsel and encouragement during the tough thesis times.

The Wesleyan University Economics Department – My academic experience at Wesleyan has been focused on achieving the best possible understanding of economics as an undergraduate. This would not have been possible without the devotion of each and every professor to the art and profession of teaching and to every student in the classroom.

And particularly,

Prof. Cameron Shelton and Prof. Richard Adelstein – Thank you both for taking the time, even when I was not your advisee, to meet with me personally and give helpful advice and a new perspective on my project.

And most importantly,

Prof. Gil Skillman – I would like to thank you for showing me to light of game theory which unfortunately came as late as it did in my college career. You're direction and guidance throughout the whole process was invaluable, and I greatly appreciated your constant curiosity and academic vigor. Regardless of my future aspirations as a game theorist, the thought processes and analytical rigor you taught me will be priceless for anything I choose to pursue.

TABLE OF CONTENTS

Introduction	1
Thesis Outline	3
Chapter I: The Prisoners' Dilemma	4
1.1 Introduction	4
1.2 The Prisoners' Dilemma Framework	4
1.3 Intellectual History of the Prisoners' Dilemma	6
1.4 Critique of the Literature	10
1.5 Rational Choice Game Theoretic Approach	11
1.6 Evolutionary Game Theoretic Approach	16
1.7 Conclusion	20
Chapter II: 2 Period Model	22
2.1 Introduction	22
2.2 Two Period Rational Choice Game	24
3.3 Evolutionary Model with Two Period Payoffs	33
Chapter III: 3 Period Model	37
3.1 Introduction	37
3.2 Three Period Rational Choice Game	37
3.3 Evolutionary Model with Three Period Payoffs	45
Chapter IV: Conclusion and Suggestions for Further Research	51
Conclusion	51
Suggestions for Further Research	51

Introduction

The purpose of this thesis is to address cooperation as a social phenomenon. From the convenience of having a door held for you when your hands are full to the vast material wealth provided through specialization in a functioning economy, the benefits of cooperation are almost too great to completely appreciate. Yet, cooperative behavior seems to be diametrically opposed to an important assumption of neoclassical economics, that individuals are rational actors seeking to maximize their own personal utility. The conflict between cooperation and self-interest is inherent in the general strategic interaction known as the Prisoners' Dilemma. The Prisoners' Dilemma arises in a group when the individuals that benefit the most from an interaction are those that do not cooperate when others do. As a whole, the group would benefit most if all members cooperate, but obtaining the optimal level is challenging because individuals, behaving in their self-interest, will tend not to cooperate as the individual that does not cooperate, benefits the greatest from the interaction. In the Prisoners' Dilemma, a group will tend towards a sub-optimal level where individuals expect others not to cooperate unless some other solution is reached.

David Gauthier (1986) suggests that the Prisoners' Dilemma is resolved in society and a cooperative outcome reached through a social contract, and that this social contract is followed by the members of society because individuals recognize that restraining their own self-interest is, in fact, rational. However, by suggesting that cooperation is rational, Gauthier fails to properly address the strategic problem central to the Prisoners' Dilemma. Cooperation may be beneficial to society, but cheating provides a greater immediate payoff.

While Gauthier's fails to properly address the Prisoner's Dilemma, the solution to the dilemma is not self-evident. In rational choice game theory, the cooperative outcome is achievable when Prisoners' Dilemma interactions repeat indefinitely over time, but cooperation is not extendable finitely repeated Prisoners' Dilemma games. In light of experimental evidence, particularly Axelrod (1984), individuals do seem to cooperate more than theory suggests, and further analysis of finitely repeated games is necessary to explain the dichotomy between theory and experiment.

In the study of evolutionary games, cooperation is possible because evolutionary games do not require that individuals always behave as rational actors. Players behave according to pre-programmed behavior adopted laterally or through generations based on expected return from the future use of the behavior. These pre-programmed strategies operate as social norms, and they facilitate cooperation by suggesting certain behavior in specific social dilemmas that, without some exogenous influence, would result in a suboptimal outcome.

If all individuals operated strictly as preprogrammed strategies, then universal cooperation would be achievable if the cooperative strategy were employed by all members of the population. However, this thesis rests on the belief that society is neither composed of all pre-programmed individuals or of all rational individuals. As long as some portion of the population is recognized as rational, analyzing cooperation in the finitely repeated Prisoners' Dilemma is necessary.

Thesis Outline

This thesis will begin with a chapter discussing in greater detail the Prisoners' Dilemma and its inherent strategic problem. The chapter will then look at the ways in which the Prisoners' Dilemma is overcome in rational choice game theory, through indefinitely repeated games, and it will readdress the criticism of Gauthier in the context of game theoretic analysis. In the final section of the first chapter, the Prisoners' Dilemma will be explained through the lens of evolutionary analysis.

In the second and third chapters of this thesis I develop a model proving the possible outcomes for a sequential game with multiple periods of the Prisoners' Dilemma. In an attempt to reveal a situation where cooperation is rational, I create a scenario where rational players and players using pre-programmed strategies interact with two-sided information asymmetry regarding player types. I then imbed the rational choice model, using the respective expected payoffs of each player, into an evolutionary framework to evaluate if the two types of players interacting in the rational choice game would persist in an evolutionarily stable system. The model in the second chapter is of a game with two periods of the Prisoners' Dilemma, and the third chapter analyzes a three period repeated PD with the additional requirement of Bayesian Updating in the middle period.

The use of information asymmetry to explain rational cooperation, is motivated by Kreps et al (1982), and this thesis contributes original material by extending the Kreps et al. model to a game with two sided information asymmetry, and by assessing the validity of the assumption of rational players that the opposing player is irrational with some positive probability. The last chapter concludes the findings of the two models, and it also provides suggestions for further research.

CHAPTER I: THE PRISONERS' DILEMMA

1.1 Introduction

In light of the evidence suggesting individuals cooperate more than might be predicted on the basis of immediate strategic consideration, further consideration of the Prisoners' Dilemma is necessary to reconcile game theoretic theory with the experimental evidence. The Prisoners' Dilemma, has been recognized, in one form or another, throughout history by intellectuals such as Plato, Hobbes, Hume, and Gauthier, although analysis of the dilemma did not enter the realm of economics until the advent of game theoretic analysis in the mid 20th century. Like Gauthier's, most suggested solutions to the dilemma have argued that when individuals in the dilemma recognize that cooperation is in their own self-interest as the benefits from long-run cooperation are much greater than the immediate non-cooperative outcome, they will cooperate. Game theory, in both its rational-choice and evolutionary forms, can be used to assess the coherence and scope of this claim.

1.2. The Prisoners' Dilemma Framework

The story of the Prisoners' Dilemma (PD) goes as follows: two prisoners are interrogated separately about a crime they committed, but not enough evidence exists to convict either of them outright. They both have the option of confessing to the crime or keeping silent, and they both know the other prisoner has this option as well. If both prisoners admit wrong-doing, then each will serve a sentence for five years. If neither confess, enough evidence is available to convict them of a lesser crime and both will be sent to prison for one year. However, if only one confesses, in exchange for testifying, he

is granted immunity and is free to walk while the other must serve a hefty ten year sentence. The following matrix shows the payoffs for this dilemma associated with possible strategies of each player. In common form, the first number in each cell is either the sentence or payoff for Player A and the following number is for Player B.

		B	
		Not Confess	Confess
A	Not Confess	1, 1	-2, 2
	Confess	2, -2	-1, -1

Matrix 2.

The “dilemma” is that each person realizes he will gain more by confessing regardless of whether the other confesses or not, and he will behave accordingly. However, when both confess, each would prefer the situation where neither confessed. Confess is a *Nash equilibrium* as it is an equilibrium where each strategy is the best strategy to play against itself, but the condition for equilibrium is even stronger. Confess is the *dominant strategy* as it returns a greater payoff given either possible strategy of the other player. This results in a Pareto-inferior equilibrium where the *Pareto-dominant* outcome, the outcome that provides the greatest social payoff, is for neither to confess. The Prisoners’ Dilemma can be generalized to a larger game with more players, but for any Prisoner’s Dilemma, players always choose between two strategies, to either *cooperate* or *collude*, or to choose the *cheating* or *defecting* strategy.

1.3 Intellectual History of the Prisoners' Dilemma

As far back as Plato (360 BC), Prisoners' Dilemmas have been recognized. Glaucon, in *The Republic*, presents the following story to Socrates: the ancestor to Gyges, king of Lydia, was a shepherd and who found a magic ring in a cavern. When the setting of the ring was turned, the bearer became invisible, and when the setting was turned back again towards the bearer, the bearer became visible once again. The shepherd realized the power of the ring, and he lost no time in wielding the ring to his advantage. The shepherd disguised himself as a messenger to the king, and upon arrival, he seduced the king's wife, killed the king, and took power over Lydia. Glaucon goes on to suggest to Socrates that no one would have the will not to defect against society when invisibility and therefore freedom from punishment is guaranteed.¹ Socrates goes on to rebut this story by explaining that it is in an individual's self-interest to obey the laws of society regardless of the possible consequences of breaking the law.

Thomas Hobbes (1651), an English philosopher of the 17th century, provides in *Leviathan* a foundation and justification for governments and societies. In Hobbes' state of nature, the hypothetical initial condition of the human race before government, society cannot overcome the Prisoners' Dilemma. Here, the cooperative outcome is never achieved, and society is in a constant "state of war" where individuals fight and cheat incessantly. When society is not in the state of nature, it is at peace, and all individuals enjoy the respective benefits of cooperation. The problem in the state of nature is that no one individual has the incentive to attempt to reach the peaceful level because by doing so, the individual subjects himself to the non-cooperative treatment of others. Hobbes' suggests that society is able to reach the peaceful level through a social contract. A social

contract is an agreement made by free and rational individuals to cooperate with one another, and Hobbes proposes the peaceful level will be obtained if all the members of the society submit to a sovereign power. The Foole in *Leviathan* does not submit to the social contract because he only recognizes the personal benefits of taking advantage of others, assuming that others will not quit cooperating with him, without understanding that he would be better off in the long run if he cooperated and assuming.²

Rousseau (1754), an Enlightenment philosopher, uses an example of a stag hunt in his *Second Discourse* to describe a Prisoners' Dilemma. The story goes as follows: to kill a stag, a group of hunters must surround the stag, requiring the contribution of the entire party. However, if any one hunter defects from the group to kill a rabbit, the hunter that kills the rabbit benefits more from leaving the group than from staying and contributing to the kill of the stag.³ The dilemma here is that while it is in the self-interest of a hunter to defect from the group and kill the rabbit, it is in the best interest of the group for all of the hunters to stay on board and kill the stag. Rousseau does not detail how a group of hunters might overcome this dilemma; rather he uses it as an example of the problems created by social interaction.

David Hume (1739) of the 18th century attends to the Prisoners' Dilemma in an attempt to explain the emergence and maintenance of morals. Hume presents a thought experiment where two neighbors commonly own a meadow, and both neighbors would benefit if the meadow was drained. Hume suggests that the meadow will most likely be drained because the two neighbors know each other and each neighbor can easily see the consequences of not cooperating. However, in a much larger group, draining the meadow would be much more difficult. A plan allowing for an equal contribution from

each individual would be challenging, and it would be even more difficult to execute the plan because each individual would attempt to cheat on the work of the other members of the group.⁴

Approaching the specific concern of the present work, Hume seeks to explain morals in the context of the Prisoners' Dilemma. First he recognizes that, after the invention of property, individuals are able to benefit from simultaneously exchanging tangible commodities. However, he suggests that people would benefit even more if they could be relied on to cooperate over a longer time period. He provides a case where one party provides a good or service in the present, and the other party completes the exchange at some point in the future. This exchange is not possible without the ability of the initial provider to secure a guarantee that the other party will fulfill his end of the exchange. According to Hume, the benefit from promising to fulfill an exchange in the future and fulfilling that promise can easily be recognized, and once it is recognized, individuals will act morally in expectation of future cooperation.⁵

Moving to the 20th century, David Gauthier (1986) justification for the existence of morals is currently one of the most influential theories on the subject. Similar to Hume, Gauthier suggests that people often act morally and overcome the Prisoners' Dilemma because they recognize that obeying morals is in their self-interest. Gauthier defines morals as the "rational constraint on the pursuit of individual interest."⁶ He suggests that by resisting immoral action when others also act morally, individuals reap a greater payoff. Following morals is rational because the expected future benefit is much greater is much greater than following immediate self-interest. To explain the initial emergence of morals, Gauthier takes a contractarian approach. His state of nature is

similar to Locke's state of nature, and he describes the initial situation in terms of what he calls the *Lockean proviso*. The Lockean proviso is an assumption that in the state of nature, individuals are unable to improve their situation by hurting others, and as a result of this assumption an implicit structure of property rights is created. The Lockean proviso is a necessary pre-requisite for any moral contract.

To escape the state of nature, individuals form a contract with each other agreeing to restrain self-interest in the expectation of greater wealth from cooperating. Gauthier equates the initial setting of the contract to a bargaining table. When bargaining, each individual attempts to receive the greatest payoff from the contract, and an agreement is reached through *maximin relative concession*. Each individual tries to minimize the maximum concession he makes to other individuals when bargaining, and individuals are concerned with their relative maximum concession, the concession compared to what other individuals are conceding. Maximin relative concession is favored by the bargainers over the *zero-point*, the result when bargaining fails and the individual does not enjoy the benefits from cooperating, when forming a contract allows for a greater expected payoff than at the zero-point. It is necessary that individuals start at the Lockean proviso because it is the only way to obtain the voluntary compliance of all individuals. Respecting the Lockean proviso is rational because it is a necessary condition for bargaining, and if it is not respected, individuals will not be able to appreciate the benefits of cooperation.

Once Gauthier explains how a moral contract will be reached, he seeks to explain why the contract will be followed *ex post*. Here Gauthier introduces a new concept, what he calls *constrained maximizers*, individuals that have formed the moral contract and

behave accordingly. It is rational for constrained maximizers to act morally when they expect enough of the other members of the group to cooperate so that the payoff from future cooperation is greater than not cooperating. Those individuals that do not comply are irrational, following from Gauthier's initial definition of morals.⁷

1.4 Critique of the Literature

It seems common practice to claim that the Prisoners' Dilemma is overcome when individuals recognize that cooperating is in their best interest, but this makes the assumption that the interests of the individual and the interests of the group as a whole are aligned. Gauthier's formulation ignores the strategic problem of the Prisoners' Dilemma. Although commitment to a "moral" code could improve social outcomes it is not in individuals' unilateral interest to commit to such a code. Mancur Olson (1965) in *The Logic of Collective Action* refuted the above assumption of the alignment of interests. While it is in the common interest of the individuals in the group to reap the benefits of cooperation, it is not in the common interest of the members of the group to abstain from the benefits of not cooperating, a cost necessary to achieve the objective of cooperation. Individuals will only work to fulfill the objective of the group if they are coerced to do so. Olson never specifies the form of the collective action problem, whether it takes the form of a Prisoners' Dilemma or a coordination problem, but much of what Olson suggests about collective action problems seems to apply directly to the Prisoners' Dilemma. Olson applies his work to the analysis of labor unions, Marxist theory, and pressure groups. However, it is Olson's basic attack on the assumption of aligned interests that is most profound and directly applicable to the argument in this thesis.

Olson's critique is an applicable response to all of the previous arguments for rational cooperation. A critique of contract theory and Gauthier's development of rational choice theory will now be given special attention. A contractarian approach to overcoming the Prisoners' Dilemma, used by Hobbes and Gauthier, must explain why individuals are willing to act as though they signed an externally enforceable contract, even though no such contract actually exists. This suggests that unless the contract is binding, individuals will continue to behave in the same manner after the contract as they did before it was created.

Gauthier's work is a special case because his argument involves a modification to conventional theory of strategic behavior. Unlike the other authors that explain the solution to Prisoners' Dilemmas as the recognition by individuals of the long-run benefits of cooperation, Gauthier redefines rationality so that the point of recognition is internalized in rational behavior. Individuals that cooperate are rational and individuals that do not are not. Gauthier justifies this alteration of rational choice by stating that "the theory of rational choice is an ongoing enterprise."⁸ However, to redefine rationality as he does is to assume away the Prisoners' Dilemma. By assuming that immediate defection is irrational, he does not have to address the weighty issue of overcoming cheating in the Prisoners' Dilemma.

1.5 Rational Choice Game Theoretic Approach

Recall the PD situation described earlier. The dominant strategy and resulting unique equilibrium, is for both players to defect, but this equilibrium is not Pareto optimal. The goal of realizing an adequate solution to Prisoners' Dilemma has yet to be

realized, but game theory provides some additional insights and conclusions for certain situations. A commonly suggested solution is to alter the payoffs of the PD so as to provide incentive for the players to choose the cooperative strategy.

In the *Emergence of Norms*, Ullman-Margalit (1977) describes an example which she refers to as the Mortarmen's Dilemma that details the affects of changing the payoffs on the moves players make. The initial Mortarmen's Dilemma takes the form of the Prisoners' Dilemma, and the story goes as follows: two mortarmen are at separate but near by outposts, and they are about to face an enemy attack. Each mortarman may either leave his outpost to escape the attack or stay and fight. The story has three possible outcomes, if both soldiers stay and fight, the enemy is repelled, if both soldiers flee, they are over-run by the enemy and taken as prisoners, and if one stays and the other flees, the soldier that stays resists the enemy long enough so that the other soldier is able to escape but is eventually killed. The dilemma takes the form of the following matrix:

		B	
		Remain	Desert
A	Remain	1, 1	-2, 2
	Desert	2, -2	-1, -1

Matrix 3.

Each soldier realizes that regardless of the action of the other soldier, he is better off fleeing the scene. This logic results in the Pareto-inferior equilibrium of both soldiers deserting even when both would prefer both stayed. The Mortamen's Dilemma can be solved by altering the payoff structure to make remaining at one's post more desirable

than fleeing. The first alteration Ullman-Margalit considers is minelaying where mines are laid surrounding both soldiers' posts. The payoff remains the same for remaining at the post, given either move by the other soldier, but if a soldier tries to flee, he is killed by a mine. This results in the following matrix:

		B	
		Remain	Desert
A	Remain	1, 1	-2, -2
	Desert	-2, -2	-2, -2

Matrix 4.

Now, both soldiers will remain at their post because if they desert they are committing to death, and the only way to escape the situation alive is by remaining at the post. While laying mines seems like a brutal enforcement policy, it is preferred by both players because it guarantees a better outcome than when players are free to desert without cost.

Ullman-Margalit goes on to describe a situation where the soldiers are part of a society where honor is internalized and soldiers prefer to die honorably in battle than be taken as a captive and live in disgrace as a deserter. The matrix is given as follows:

		B	
		Remain	Desert
Remain		1, 1	-1, 0

A		
Desert	0, -1	-2, -2

Matrix 5.

To remain is the best strategy of a soldier given whether the other soldier remains or deserts, and it should be noted that honor functions not only as a punishment for those who desert but also as an award for those who die honorably. By changing the payoff structures, an incentive to cooperate is created, and an inspection of matrices 4 and 5 will reveal that these no longer look like the original PD matrix. While this solves the PD, it is through exogenous action, and it does not seem to follow the hypothesis that cooperation can be achieved even without the presence of enforcement.

Now, consider a scenario where the Prisoners' Dilemma repeats. In contrast to the single-shot PD, collusion is possible between rational players when the PD repeats. However, when the PD is finitely repeated, both players know when the last period of the game will occur, and collusion is not a possible equilibrium strategy because it is not a *subgame perfect equilibrium* (SPE). A SPE is defined as a strategy pair that constitutes a Nash equilibrium for each proper subgame as well as for the entire game. Backward induction is used to determine the rational strategy for each period, and the logic is as follows: in the last period of a finitely repeated PD, defect is clearly the dominant strategy since there is no future outcome to influence. In the second to last period, defect is the dominant strategy because it will return a higher payoff given either of the possible strategies of the other player and given that rational players will defect in the last period. This logic continues backward through the game until it is deduced that rational players will defect in every period of the repeated period game. The outcome suggested by

backward induction is drastically different from the outcome Gauthier assumes, that players will cooperate. The logic of backward induction cannot be overlooked in defining equilibrium play, and further analysis is necessary to determine when individuals would cooperate consistently with subgame perfect equilibrium.

In the more general case of repetition, when repetition is infinite, collusion is achievable because backward induction, from a known final period, is no longer the basis for determining equilibrium play. *Contingent strategies*, strategies based on the play of the opposing player in the previous period are now strategically viable. For example, a tit-for-tat strategy (TFT) is a specific type of contingent strategy where those using it start the game with collude, and after the first period, collude if the other player colluded in the previous period and defect if the other player defected in the previous period. A player employing this strategy will thus collude as long as the other player colludes and defect as long as the other player defects.

Contingent strategies make collusion possible, and they foster an outcome much more efficient than a strategy of only defect. When rational players determine if the TFT strategy is part of a Nash equilibrium pair, they will generally be interested in comparing the present value of future collusion with the value of immediate defection. In games where the Prisoners' Dilemma repeats indefinitely, contingent strategies are also strategically viable. The indefinitely repeated Prisoners' Dilemma is a more realistic account of repeated PD interaction, but in addition to the considerations in the infinitely repeated, players, when determining the rational strategy, now must also consider the probability that the game will end in the next period.

In games with an infinitely repeating or indefinitely repeating Prisoners' Dilemmas, collusion is part of a viable strategy. However, the only rational strategy in a finitely repeated PD is universal defection thwarting the attempt to explain collusion through rational game theory. The next section will address the PD in an evolutionary framework to explore if greater results can be obtained.

1.6 Evolutionary Game Theoretic Approach

In evolutionary game theory, players are no longer strict utility-maximizers, rather they behave as a *type*. Each move a type plays in a game is dictated according to a "preprogrammed" strategy specifying how the player should behave given the specifics of the social interaction. In any strategic interaction, a rational player assesses the possible outcomes given his possible strategies, but when a type interacts, he behaves strictly according to his strategy without considering costs or benefits of his strategy or any other strategy. In biology, the genetic makeup of an individual, the *genotype*, dictates the behavior of the organism, the *phenotype*. Organisms that exhibit a phenotype that is better suited to the environment will have a greater probability of reproductive success, defined quantitatively as *fitness*. Through the process of natural selection, organisms with a favorable genotype will tend to survive and increase in number, possibly reaching a stable state where the dynamic ceases. Random *mutations* of the genotype may alter the phenotype and thus the dynamic between organisms. If the mutation is not well suited for its surroundings, it will die out as natural selection favors fitter genotypes. However, if it is better suited to the environment than the other

genotypes of the population, it will *invade* the existing population and come to exist in the population at a relatively higher frequency.

It was John Maynard Smith who first suggested the applicability of the study of biological evolution to social phenomenon.⁹ In social evolutionary theory, strategies, like phenotypes, are determined before any interaction, and they survive and propagate based on their fitness. However, strategies reproduce, not through the genotype, but through the adoption by other individuals. Individuals may adopt behavior for a variety of reasons, but the important assumption is that types that are more successful will tend to be adopted in greater frequency. *Equilibrium* in the population is reached when types no longer change frequency over time, and a type is an *evolutionary-stable strategy* (ESS) if it cannot be invaded by other types.

Two forms of evolutionary equilibria are possible in a game with two types, *monomorphic* and *dimorphic*. A monomorphic equilibrium is an equilibrium in which only one type prevails, and a dimorphic equilibrium is an equilibrium where two types persist. A dimorphic equilibrium is defined as the point where the expected payoffs of each type are equal for some population distribution, and the equilibrium will be either *stable* or *unstable*. A dimorphic equilibrium is stable if types have a higher expected payoff when their frequency is lower than at equilibrium. This causes the population distribution to return to the equilibrium when perturbed. If the types are in an unstable equilibrium, each type has a lower expected payoff when at a lower frequency than equilibrium. When this occurs, the population distribution will not return to equilibrium if perturbed. The unstable equilibrium gives a balancing point where, away from this point, the replication dynamic pushes the population towards a distribution of all one type.

Similar to the rational choice game, a PD can also be modeled using evolutionary game theory. The game is now played, not between two rational players, but between two types. Assume that the population is composed of two types, a type that always confessing (C), and a type that never confesses (NC). The payoffs are the same as in the initial Prisoners' Dilemma description and repeated here:

		Not Confess	Confess
Not Confess	1, 1	-2, 2	
Confess	2, -2	-1, -1	

Matrix 6.

Regardless of which type a C player matches against, the C player will always do better than an NC type, and thus C has a greater expected payoff than NC for any population distribution. To explain more formally, let q be the proportion of the population that is NC, and $(1-q)$ the proportion that is C, and allow for a large group of players where players interact randomly in the form of the Prisoners' Dilemma. The expected payoff to a C in any one interaction is $q(2)+(1-q)(-1)$, and the expected payoff to an NC is $q(1)+(1-q)(-2)$. For $q \in (0,1)$, it is easy to see that $q(2)+(1-q)(-1) > q(1)+(1-q)(-2)$, and the expected payoff to a C is always greater than the expected payoff to a NC. A population of all NC is unstable because if an NC mutates into a C, the C will have a greater expected payoff and will replicate with a greater frequency than the NC types of the population. When this occurs, C players will

continue to replicate in greater frequency for $q \in (0,1)$ until the population is composed of all C types. This is a monomorphic equilibrium, and it is stable because if at $q = 0$, a C mutates into an NC, the expected payoff to C players continues to be greater for the small q , and the replication process will return the population distribution to $q = 0$. Thus, C is an ESS.

Now, supposed when two types meet, they interact for two periods of the Prisoners' Dilemma. In a game with multiple periods, a strategy must give a complete account of the move a player will use in each period based on the moves of the other player in prior periods. For this game, C types will confess in both periods. Now, instead of NC types, the C types will interact with TFT types where TFTs collude in the first period, and in the second period collude if he sees collude in the first period and defect if he sees defect. The payoffs take the following form:

		Not	
		Confess	Confess
Not Confess	Not Confess	2, 2	-3, 1
	Confess	1, -3	-2, -2

Matrix 7.

Let p represent the probability a player is TFT, and let $1-p$ represent the probability a player is a C type. The expected payoff of a TFT player is $p(2) + (1-p)(-3)$, and a C type is $p(1) + (1-p)(-2)$. Notice that C types do not always dominate. Rather, when the population is mostly C types, C types have a greater expected payoff, and when

the population is mostly TFT players, TFT players have a greater expected payoff. Therefore, neither strategy can be invaded by the other strategy at their respective monomorphic equilibria, and therefore, both types are evolutionary stable strategies. When the population is initially at a mixture between the two types, an equilibrium is possible. The equilibrium occurs when the payoffs between the types are equal so that $3p - 2 = 5p - 3$, and solving for p , the population will be at equilibrium when $p = 1/2$. If $p > 1/2$, TFT types will have a greater expected payoff, and the dynamic will push p to the stable equilibrium at $p = 1$. When $p < 1/2$, the dynamic will move in the opposite direction towards $p = 0$.

The results of the one-shot evolutionary PD game are the same as the results in a one-shot rational choice game. Comparing the single-shot PD in the rational choice game with the single-shot PD as an evolutionary game, “confess” is a Nash equilibrium in the rational choice game and an ESS in the evolutionary model. If rational choice players were to play the evolutionary model with two periods of the PD as a single period, two Nash equilibria would exist, “confess” and “TFT.” The pattern is not a random one, evolutionary stable strategies must also be Nash equilibria when the game is played by rational individuals. However, while the TFT strategy in the two period game is a Nash equilibria, it is not a subgame-perfect equilibrium because rational players will use backward induction to determine that “confess” is the dominant strategy.

1.7 Conclusion

Gauthier’s assumption of rational cooperation has been proven through rational choice game theory to be valid under conditions of infinitely and indefinitely repeated

games. However, the only rational strategy in finitely repeated games is universal defect, and this, as of now, defeats Gauthier's assumption of rational cooperation. In the evolutionary model, a TFT type has proven to resist invasion by a type that always defects, but the evolutionary model cannot account for a stable evolutionary equilibrium where both TFT and cheating types coexist. Thus, only populations of all "confessing" or all TFT types are possible, neither of which are acceptable conclusions as this would deny the intuition that rational and some form of pre-programmed types exist in society. The next chapter, through the right kind of information asymmetry, will attempt to prove that cooperation is possible in a finitely repeated PD game.

CHAPTER II: TWO PERIOD MODEL

2.1 Introduction

The purpose of this chapter is to determine if “social norms,” interpreted as the willingness of players to cooperate in a setting in which “cheating” or “taking advantage” might otherwise be expected, can arise in the context of a two period rational choice PD game. Kreps et al.¹⁰ explain the possibility of collusion in finitely repeated game through the incorporation of one-sided information asymmetry. Starting with this approach, this chapter will extend Kreps et al. to incorporate two sided information asymmetry, and it will investigate whether rational players’ beliefs regarding player types is consistent with a stable evolutionary equilibrium.

Kreps et al. describe a scenario in which it is possible for two types of players to exist, rational and TFT. In each period, rational players choose the strategy that yields the greatest payoff while TFT players are “programmed” to choose a strategy that is based on the strategy of the other player in the previous period. Specifically, TFT players respond to collude in the previous period with collude and defect with defect. In Kreps, et al. when two rational players meet, Player A and Player B, they interact in a repeated PD where the end period is common knowledge. Player A is fully informed of Player B’s type, and Player B assesses that Player A is a TFT player with some positive probability. If Player A’s type were to become common knowledge, the game becomes a finitely repeated PD game where universal defection is the only subgame-perfect equilibrium. Player B will collude if his expected payoff from colluding in the current and future periods is greater than the expected payoff from defecting in the current period. Given this, it will be in Player A’s best interest, even though he knows that Player B is a

rational player, to collude when his expected payoff from collusion is greater than his expected payoff from defecting in the current period. Thus, under these circumstances, Player A deceives Player B into thinking that he is a TFT player because it yields a greater payoff, and Player B colludes because it improves his payoff as well. As long as Player A's type does not become common knowledge, collusion is possible in a finitely repeated PD.¹¹ By extending the Kreps et al. story to two-sided asymmetry regarding player types, a more complete account of collusion in a finitely repeated game is available. It can now be assumed that rational players come to the game without any prior knowledge of the opposing player where each rational player's belief about the opposing player's type is based solely on the prior probability the player is TFT and on the equilibrium play of the player.

The assumption of Kreps et al. that a rational player believes the opposing player is irrational with some probability is more consistent with evolutionary game theoretic analysis. These irrational players, using a tit-for-tat strategy, do not choose the strategy that offers the highest expected payoff, instead playing their pre-programmed strategy regardless of the situation. While these TFT players appear irrational in a rational choice game, an evolutionary story suggests that TFT behavior is learned based expectation that it will provide a high payoff relative to other possible strategies. The replication process of the TFT type can only be modeled in an evolutionary game, and it is through these means that whether or not TFT types will persist in a population with rational players can be determined.

2.2 Two-Period Rational Choice Game

The Model

The game of concern is a repeated two-person noncooperative game. Each period ($n = 1, 2$) consists of a Prisoners' Dilemma (PD) where $a > x > 0 > b$ and $a + b < 2x$ ¹² with the following normal form:

		Player B	
		Collude	Defect
Player A	Collude	x, x	b, a
	Defect	a, b	$0, 0$

Players move simultaneously in each period, and payoffs for the supergame are the undiscounted sum of single-period payoffs. Players are fully informed of the number of remaining periods and of the previous moves of both players, but each player's type is not common knowledge. Two types exist, rational and tit-for tat (TFT), where the prior probability that a given player is TFT is represented by $\delta \in (0,1)$, and the prior probability that a player is rational is $1 - \delta$. Rational players use strategies that are part of a subgame perfect equilibrium strategy pair where it is required that the strategy be rational for each period of the game as well as for the whole game. Since the subgame perfect equilibria for the game are unknown at this point, backward induction must be employed, starting from the last period as in the following lemma, to determine which equilibrium path will result.

Subgame perfect equilibrium in the 2-period PD supergame

LEMMA 1: Rational players always defect in the last period, regardless of information about their opponent's type.

PROOF: In the last period, defect is the dominant strategy because it returns a higher payoff than collude against defect as $0 > b$ and a higher payoff against collude as $a > x$. Therefore, a rational player will defect whether the opponent is rational or TFT.

Q.E.D.

PROPOSITION: In a two period rational choice PD game where each period is in the form of a PD, i) if $\delta > \max\{V(a), V(b)\}$, $[(c,d)(c,d)]$ is a unique SPE, ii) if $x > a + b$ and $\delta \in [V(a), V(b)]$, $[(c,d)(c,d)]$ and $[(d,d)(d,d)]$ are both SPEs, and a mixed strategy equilibrium is possible where rational players use (c,d) with a probability of σ_i^ given by $\sigma_i^* = [\delta(b-x) - b] / [(1-\delta)(x - [a+b])]$, iii) if $\delta \in (V(a), V(b))$ both $[(c,d)(d,d)]$ and $[(d,d)(c,d)]$ are SPEs, and if rational players mix strategies they will again use (c,d) with a probability of σ_i^* given by $\sigma_i^* = [\delta(b-x) - b] / [(1-\delta)(x - [a+b])]$, and iv) if $\delta < \min\{V(a), V(b)\}$, $[(d,d)(d,d)]$ is a unique SPE.*

PROOF: Following from lemma 1, in $n = 2$, both rational players will always defect, and the question becomes whether or not a rational player should collude in $n = 1$. The strategy (c,d) is an equilibrium response to (c,d) if it maximizes a rational player's

payoff. Now, suppose two rational players, Player A and Player B, meet, and Player B chooses (c, d) . Player A will receive a payoff of 0 in $n = 2$ as both players defect. If Player A colludes in $n = 1$, he will receive an immediate payoff of x , and if he defects, he will receive an immediate payoff of a . If Player A meets a TFT player, in $n = 1$, Player A will receive an immediate payoff of a if he defects and x if he colludes as a TFT player always colludes in the first period. Given that Player A will defect in the last period and that the TFT player will respond to collude with collude and defect with defect, Player A will receive a payoff of a if he colluded in $n = 1$ and 0 if he defected. Player A assesses the probability that he will meet a TFT player as δ and $1 - \delta$ as the probability he will meet a rational player. The expected payoff for Player A using the strategy (c, d) is thus $\delta a + x$, and if he uses (d, d) his expected payoff is a . Player A will prefer playing (c, d) over (d, d) when $a \leq \delta a + x$, and therefore, (c, d) is an SPE strategy pair when $\delta \geq V(a)$.

Again, suppose two players meet, however this time, let Player B use the strategy (d, d) . When Players A and B meet, Player A will receive a payoff of b if he colludes in $n = 1$, and 0 if he defects. In $n = 2$, Player A will receive a payoff of 0. When Player A meets a TFT player, the payoffs for the game are the same as above, $x + a$ if he uses (c, d) and a if he uses (d, d) . Again, Player A assess the probability that the other player is a TFT player as δ , and the expected game payoff for using (c, d) is $\delta(x + a) + (1 - \delta)b$ and for using (d, d) is δa . When $\delta a \geq \delta(x + a) + (1 - \delta)b$ (d, d) is an SPE strategy pair, and solving for δ , when $\delta \leq V(b)$. When $\delta > \max\{V(a), V(b)\}$, $[(c, d)(c, d)]$ is a unique SPE

as $[(d,d)(d,d)]$ is not an SPE, and when $\delta < \min\{V(a), V(b)\}$, $[(d,d)(d,d)]$ is unique as $[(c,d)(c,d)]$ is no longer an SPE.

If $V(b) < \delta < V(a)$ and $x < a+b$ neither $[(c,d)(c,d)]$ and $[(d,d)(d,d)]$ are SPEs.

Now, assume that when this occurs, two rational players meet, Player A and Player B, and Player B uses (d,d) . If Player A uses (c,d) , he will have an expected payoff of $\delta(x+a) + (1-\delta)b$ and if he uses (d,d) , he will have an expected payoff of δa . Player A will play (c,d) over (d,d) when $\delta < V(b)$. Now assume that Player A uses (c,d) . If Player B plays (c,d) his expected payoff will be $x + \delta a$, and if he plays (d,d) his expected payoff will be a . So Player B will use (c,d) when $\delta < V(a)$. Additionally, for these strategy pairs to be SPEs, the opposite must be true: if A plays (d,d) , B must want to play (c,d) , and when A plays (c,d) , B must want to play (d,d) .

The payoffs are symmetrical, and therefore it follows that B will play (c,d) when $\delta > V(b)$, and A will play (d,d) when $\delta < V(a)$. Thus, $[(c,d)(d,d)]$ and $[(d,d)(c,d)]$ are SPEs when $V(b) < \delta < V(a)$.

Mixed strategy equilibria are also possible equilibria in this rational choice game. Assume that when two rational players meet, Player A and Player B, Player B uses (c,d) with a probability of σ_B and (d,d) with a probability of $1-\sigma_B$. The expected payoff for Player A, represented as $E(\pi_A)$, when Player A uses (c,d) is given by

$E(\pi_A) = \sigma_B(x + \delta a) + (1-\sigma_B)[\delta(x+a) + (1-\delta)b]$, and the expected payoff for Player A

when he uses (d,d) is given by $E(\pi_A) = \sigma_B(a) + (1 - \sigma_B)(\delta a)$. Now, assume that Player A uses (c,d) with a probability of σ_A , his expected payoff will then be given by

$$E(\pi_A) = \sigma_A \left(\sigma_B(x + \delta a) + (1 - \sigma_B) [\delta(x + a) + (1 - \delta)b] \right) + (1 - \sigma_A) [\sigma_B(a) + (1 - \sigma_B)(\delta a)].$$

Assuming that a mixed strategy equilibrium exists, such that $\sigma_i^* \in (0,1)$, $i = A,B$, to

determine the point at which Player B is indifferent to Player A's mix of strategies, take the first order condition for this payoff function with respect to σ_A :

$$dE(\pi_A)/d\sigma_A = \sigma_B(x + \delta a) + (1 - \sigma_B) [\delta(x + a) + (1 - \delta)b] - [\sigma_B(a) + (1 - \sigma_B)(\delta a)] = 0.$$

Solving for σ_B^* yields $\sigma_B^* = [\delta(b - x) - b] / [(1 - \delta)(x - [a + b])]$, and this is the optimal mix for Player B, the point at which Player A is indifferent between either strategy Player B chooses. It follows that since the expected payoffs between Players A and B are symmetrical, σ_A^* , the point at which Player B is indifferent between the strategy Player A uses is given by $\sigma_A^* = [\delta(b - x) - b] / [(1 - \delta)(x - [a + b])]$, the same value for σ_B^* . Thus all rational players, when mixed strategies are available, use (c,d) with a probability of σ^* where $\sigma^* = [\delta(b - x) - b] / [(1 - \delta)(x - [a + b])]$. *Q.E.D.*

So, to no surprise, (d,d) is an SPE for the two-period game, but interestingly enough, (c,d) is an SPE for a certain range of δ as well. The belief of rational players that $\delta > 0$ needs to be confirmed in the evolutionary game, but first a brief analysis of the situation where both (d,d) and (c,d) are equilibria will be addressed.

The coordination game, as an aside

A Coordination game is a game with multiple Nash equilibriums where players are not competing for the highest payoff, but rather both players prefer the same equilibrium. If $x > a + b$ and $\delta \in [V(a), V(b)]$, a coordination game ensues as both $[(c, d)(c, d)]$ and $[(d, d)(d, d)]$ are SPEs. While, for the sake of this model, rational players will be assumed to converge at the Pareto-dominant equilibrium of $[(c, d)(c, d)]$, in reality equilibrium selection is more complex. A coordination failure may arise in a game with two equilibria if players are unable to select the same equilibrium strategy. The equilibria of the coordination game may be Pareto-ranked meaning that a particular equilibrium will provide both players with a greater payoff than if the players were at a different equilibrium. The Pareto-dominant equilibrium is the natural focal point in the game as it is the equilibrium that both players would unequivocally prefer. Even though the Pareto-dominant equilibrium is the natural focal point, it does not guarantee that players choose the Pareto dominant strategy. Experimentally, players are often show selecting strategies other than the Pareto-dominant strategy, and Cooper provides an extensive survey of this experimental literature.¹³

To guarantee a Pareto-dominant strategy is chosen, a player needs to know, with complete certainty, that the other player will choose the Pareto-dominant strategy. Players may use cheap talk, make strategic moves, or follow social norms in an attempt to achieve the Pareto-dominant equilibrium. Cheap talk is an opportunity for players to send costless messages to the other player signaling that a specific strategy will be chosen. Strategic moves, similar to cheap talk, are actions taken by players used to signal a specific strategy. Social norms pertaining to coordination games are somewhat different

than PD norms as the interests of the players are not opposed. However, they operate similarly and may guide strategy toward the optimal group outcome. These tools for determining and signaling a strategy are non-binding, and while they may make it more likely that a specific strategy is chosen, they make no guarantee. Some chance continues to exist that a player will not choose the Pareto-dominant equilibrium.

In an attempt to explain the evidence that coordination problems are not solved naturally in noncooperative games, Harsanyi and Selten argue that players consider risk dominance when choosing a strategy.¹⁴ In a coordination game, playing the Pareto-dominant strategy is riskier than the alternate strategy, and because of this, a player must make a trade-off between risk and return when choosing a strategy. The decision to play the Pareto-dominant strategy depends on the probability that the other player will play the Pareto-dominant strategy. The higher the probability that the other player will choose the Pareto-dominant strategy, the more likely a rational player will choose the Pareto-dominant strategy. A strategy is risk dominant when the strategy provides the greatest expected payoff given the probability that the opposing player will use that strategy. The risk dominant equilibrium may support the Pareto-dominant equilibrium, but this is not necessarily so.

Given some expectation of the probability a player will choose a strategy, risk dominance explains equilibrium selection in coordination games. However, Harsanyi and Selten's solution does not provide an explanation for why a rational player should expect another player to behave in this manner. A rational player may assign some probability to another player's strategy selection based on the existence of incomplete information, or in an evolutionary model, based on the possibility of mutations. Papers

by Carlsson and van Damme and Kandori, et al. incorporate these ideas in a game theoretic framework to give credibility to the expectations of a rational player. Both of these papers achieve results that are consistent with the risk dominance theory of Harsanyi and Selten.

In Carlsson and van Damme's paper, incomplete information is used to explain equilibrium selection in coordination games.¹⁵ Carlsson and van Damme create a coordination game where the payoff of an equilibrium, equilibrium A for example, is uniformly distributed around the Pareto-dominant payoff allowing for the possibility that equilibrium A is not the Pareto-dominant equilibrium. Players do not know the payoff of equilibrium A, but they observe a signal that is correlated with this payoff. After receiving the signal, players do not know the exact payoff they will receive based on the strategy they choose. Players understand a possibility exists that the other player received a signal revealing equilibrium A as Pareto-inferior. Each player chooses their strategy given the available information, and the strategy chosen is dependent upon the payoff range of equilibrium A and of the strength of the signal.

Risk dominance may also be justified by the existence of mutations in an evolutionary framework. Mutations cause strategies to be chosen that are not optimal, and this creates additional noise in the equilibrium system. A rational player will use the mutation rate to determine the probability another player will choose the Pareto-inferior equilibrium. In Kandori et al., mutations reflect players' inability observe and interpret the information necessary to calculate an optimal strategy, and also players' inability to act upon the information they receive.¹⁶ Kandori et al. provide a more detailed analysis,

and they are able to show that the long run equilibrium tends towards the risk dominance equilibrium of Harsanyi and Selten.

Returning to the two period rational choice game, if $x > a + b$ and $\delta \in [V(a), V(b)]$ and the coordination game occurs, rational players will use the strategy (c, d) , the Pareto-dominant strategy, when it is the risk dominant strategy. To determine when the Pareto-dominant strategy is also the risk dominant strategy, let $V(s)$ be the net gain a rational player will receive from choosing (c, d) over (d, d) when the opposing player uses (c, d) with a probability of s and (d, d) with a probability of $1 - s$. The expected payoff to the rational player using the strategy of (c, d) is $\delta(x + a) + (1 - \delta)[sx + (1 - s)b]$, and the expected payoff when using (d, d) is $\delta a + (1 - \delta)sa$. Thus, the net gain for a rational player choosing (c, d) is given by $V(s) = \delta(x + a) + (1 - \delta)[sx + (1 - s)b] - [\delta a + (1 - \delta)sa]$. To evaluate s , we need to look at the point where the rational player is indifferent between the two strategies. This occurs when $V(s) = 0$, and when $V(s) = 0$,

$$s^* = \frac{\delta(b - x) - b}{\delta(b - x + a) + x - b - a}$$

It follows that if rational players expect to face an opponent that uses (c, d) with a probability of s , then when $s \geq s^*$, the Pareto-dominant equilibrium equals the risk dominant equilibrium and rational players will use the strategy (c, d) . When $s < s^*$, the Pareto-inferior equilibrium is risk dominant and rational players use (d, d) . For the remainder of the paper it will be assumed that rational players choose the Pareto-optimal equilibrium strategy when a coordination game occurs, but the coordination game adds an extra layer of analysis and possibly a topic for further research.

3.3 Evolutionary Model with Two Period Payoffs

The belief of rational players that $\delta > 0$ in the previous model is robust with respect to evolutionary processes if rational and TFT players are present in an equilibrium of an evolutionary game with the relevant frequencies. To study rational players in the context of an evolutionary game, rational players must also be considered types, here called maximizing types, and they are assumed to use a payoff maximizing strategy consistent with the previous rational choice game. The results of the rational choice game are abbreviated for the evolutionary model to simplify the results. The mixed strategy equilibrium as well as the two equilibria, $[(c,d)(d,d)]$ and $[(d,d)(c,d)]$, for their respective range of δ . This simplification should not dilute the results.

The Model

The following model is used to describe the evolutionary process and equilibrium conditions for a group of maximizing and TFT types, and it builds on the analysis of cultural equilibrium and type replication of Bowles and Gintis (1998):¹⁷ suppose a population exists with a large number of individuals of both maximizing (m) and TFT (t) types. The members of the population interact randomly in pairs in the form of the two-period rational choice game of 3.2, and attention will continue to be restricted to the subgame-perfect equilibrium strategies of maximizing types for the ranges of δ previously described. At the end of a game, each player compares his expected payoff to the expected payoff of a random player. If the “other” player has a higher payoff, and if the “other” player is a different type, the player will tend to replace himself with the other type. If the “other” player has a lower payoff, the player will do nothing.

The expected future payoffs for each type are calculated using the probability of meeting each type of player (δ for TFT types and $(1-\delta)$ for maximizing types), and the payoffs for playing each strategy against a TFT or maximizing strategy. The expected payoff for a maximizing type is given by $b_m(\delta) = \delta\pi(m,t) + (1-\delta)\pi(m,m)$, and the expected payoff for a TFT type is given by $b_t(\delta) = \delta\pi(t,t) + (1-\delta)\pi(t,m)$ where $\pi(i,j)$ represents the payoff to a player of type i in the supergame with a player of the type j . The expected payoffs for maximizing and TFT types are represented by $b_m(\delta)$ and $b_t(\delta)$ respectively. The growth rate of δ over time is represented by r_t , and this is given by $r_t = b_t(\delta) - \bar{b}(\delta)$ where $\bar{b}(\delta)$ is the average payoff within the group and is represented by $\delta b_t(\delta) + (1-\delta)b_m(\delta)$. This expression simplifies so that $r_t = (1-\delta)(b_t(\delta) - b_m(\delta))$. Types are in equilibrium when evolution, defined as the process where the frequency of types changes over time, no longer occurs.

From $r_t = (1-\delta)(b_t(\delta) - b_m(\delta))$, it follows that the population is in equilibrium when $r_t \leq 0$. If $\delta^* = 1$, $r_t = 0$, and the population is at a monomorphic equilibrium of all TFT types. This is a *stable* equilibrium if, for a value of δ slightly below 1, $b_t(\delta) > b_m(\delta)$. At $\delta^* = 0$, maximizing types are in a monomorphic equilibrium if $b_t(\delta) \leq b_m(\delta)$ so that $r_t \leq 0$, and if $b_t(\delta) < b_m(\delta)$, the equilibrium is stable. A dimorphic equilibrium is also possible, and this will occur if $b_t(\delta) = b_m(\delta)$. The equilibrium is *stable* at some value $\delta^* \in [0,1]$ if a small increase or decrease in δ is naturally corrected back to δ^* so that $dr_t/d\delta < 0$. This means that

$\pi(m,t) - \pi(m,m) - \pi(t,t) + \pi(t,m) > 0$ ¹⁸ must be true for all stable dimorphic equilibria, and if this requirement does not hold, the types are in a unstable equilibrium at δ^* .

Evolutionary Equilibria

PROPOSITION: *Assuming maximizing types use only pure strategies, and that when multiple subgame perfect equilibria exist maximizing players use the Pareto-dominant strategy, when $\delta < 1$, a monomorphic equilibrium of maximizing types is the only equilibrium, and it is evolutionarily stable. At $\delta = 1$, TFT types are in an unstable monomorphic equilibrium.*

PROOF: Given the conditions for equilibrium,

- I. in an evolutionary game, TFT and maximizing types, using the strategy (d,d) , are in equilibrium when their expected payoffs are equal so that $\delta a = \delta 2x + (1 - \delta)b$. This equality will be true for some δ , given the payoffs of the PD, when $a + b < 2x$, and this must be true for the initial conditions specified. Solving for δ at the dimorphic equilibrium, $\delta^* = b / (a + b - 2x)$. For this equilibrium to exist, (d,d) must be an SPE at δ^* , and this is true when $b / (a + b - 2x)$ is less than $b / (b - x)$. However, this is never true so that at δ^* , maximizing types use (c,d) , and no equilibrium between maximizing types using (d,d) and TFT types is possible. When maximizing types use (d,d) , they have a greater expected payoff than TFT types, and the replication process will cause δ to decrease until an equilibrium at $\delta = 0$ is reached. This equilibrium is stable because for δ close to 0, $b_t(\delta) < b_m(\delta)$.

- II. When a maximizing type uses the strategy (c,d) , his expected payoff is $\delta a + x$, and the expected payoff of a TFT type is $x(\delta + 1) + b(1 - \delta)$. At the dimorphic equilibrium, $\delta a + x = x(\delta + 1) + b(1 - \delta)$, but this statement is unambiguously false as $\delta a + x$ will always be greater than $x(\delta + 1) + b(1 - \delta)$. Therefore, no dimorphic equilibrium exists between the two.
- a. At $\delta = 1$, TFT types are in equilibrium as the frequency of types is unchanging. This equilibrium cannot be stable because $b_i(\delta) < b_m(\delta)$ for a δ close to 1, and when maximizing types invade a population of TFT types, the replication dynamic will cause δ to decrease.
 - b. When $\delta < 1$, maximizing types using (c,d) have a greater expected payoff than TFT types, and δ decreases until $\delta < V(a)$ at which point (c,d) is no longer an SPE, and maximizing types use (d,d) as a strategy. *Q.E.D.*

The only stable evolutionary equilibrium is one with all rational players using the strategy of (d,d) . This is a powerful result because, while Kreps et al. provide a general case for cooperation and do not specify the necessary number of periods for rational players to cooperate, their results are not applicable to a two period game with two-sided asymmetry. It seems intuitive that as the number of periods increases the expected payoff for cooperative players will increase. Thus, the next step is to perform a similar analysis for a game with three periods of the Prisoners' Dilemma.

CHAPTER III: THREE PERIOD MODEL

3.1 Introduction

The purpose of this chapter is to determine the implications of extending the two-period model with two-sided information asymmetry about player types developed previously to a scenario with two repetitions of the Prisoners' Dilemma ($n = 1, 2, 3$). Now, strategy choices of rational players affect future payoffs, in addition to current payoffs through the effect of strategies on the updating beliefs of the opposing player. In this model, rational players update their beliefs about player types after each period of the game using Bayes' Rule based on equilibrium play. Moves in the current period are determined in accordance with beliefs about player types deduced from the prior periods of the game. A complete strategy now must in general specify a rational player's move in each period given the possible moves of the other player in prior periods. As in the two period model, the plausibility of the Kreps et al. story regarding cooperation among rational players in a finitely repeated PD will continue to be evaluated in an evolutionary context. For the story to hold, $\delta > 0$ must be consistent with a stable evolutionary equilibrium.

3.2 Three Period Rational Choice Game

According to the subgame perfect equilibrium requirement, in a sequential move game, at any point in the game, a rational player will choose the strategy that optimizes his payoff in the current period and over the remainder of the game. Given uncertainty about player types, rational players need also consider the available information about the opposing player's type when determining a strategy. In addition to the SPE requirement,

a Bayesian perfect equilibrium (BPE) requires that, at any point in the game, players assess the probabilities of other players' types using Bayes' rule and assuming equilibrium play up to the current period and in future periods. BPE also requires that players choose the optimal strategy in light of this information. At any point in the game, a rational player, Player i , uses Bayes' law to assess the probability that the other player, a rational Player j , is TFT where z_n^i represents the posterior probability assessed by Player i in period n that Player j is TFT. Likewise, in period n , Player j assesses the probability that Player i is TFT, and this is represented by z_n^j .

Now consider the three-period PD supergame introduced earlier. Since we do not know yet what the Bayesian Perfect equilibria are for this game, to determine which equilibrium path will result for a range of z_n^i , we must build up to it by backward induction. Therefore, consider the situation in $n = 3$, the last period. For any value of z_3^i , Player i will always defect in the last period as it is the dominant strategy for the one-shot PD game. The TFT player, on the other hand, will use a strategy dependent on the opponent's move in the previous period.

Given the outcomes in $n = 3$, now consider the strategic setting in period $n = 2$. We know from the two period model that in the second period, defection is consistent with SPE if $z_2^i \leq V(b)$ and is uniquely so if $z_2^i < \min\{V(a), V(b)\}$. Collusion is consistent with SPE in the second period if $z_2^i \geq V(a)$ and is uniquely so if $z_2^i > \max\{V(a), V(b)\}$. Player i derives z_2^i based on the opponent's move in $n = 1$ in accordance with Bayes' rule and with the assumption that the other player is using an equilibrium strategy, in so far as possible. If Player i sees defect in $n = 1$ when he expected both players to collude, Player

i will update in a similar fashion as the rational players of Kreps et al. Player i will conclude that since defection is a possible move in the first period of rational players but not TFT players, then the opposing player must be rational. The equilibrium move of a Player i in $n = 2$ depends both on z_2^i , the updated belief of Player i , and his own move in $n = 1$. In the first period, z_1^i is equal to the prior probability δ that a player is TFT. In the following lemma, the derivation of the posterior probability a player is TFT is considered. The following model will not pursue mixed equilibria strategy or the non-symmetrical strategies for the simplicity of the model.

LEMMA 1: If a rational opponent, Player B, is expected to collude in the first period, a rational player, Player A, will infer that $z_2^A = \delta$ if the opponent colludes and $z_2^A = 0$ if the opponent defects. If a rational opponent is expected to defect in the first period, Player A will infer $z_2^A = 0$ if he sees defection in $n = 1$, and will infer $z_2^A = 1$ if he sees collusion in $n = 1$.

PROOF: Given that a rational opponent is expected to collude in $n = 1$, if a rational player sees collude in $n = 1$, $z_2^A = \delta$ because he expects both types of players to collude with a probability of 1 in $n = 1$. The probability a player is TFT given the move used by the other player in the previous period is equal to

$$\left[\Pr(D_1)\Pr(C:D_1) \right] / \left[\Pr(D_1)\Pr(C:D_1) + \Pr(D_2)\Pr(C:D_2) \right]$$
 where C is the observed play of the other player, D_1 is a TFT player, and D_2 is a rational player. When Player A expects both players to collude, C , with a probability of 1, z_2^A is determined as follows

where $\Pr(C:D_1)=1$ and $\Pr(C:D_2)=1$ so $\Pr(D_1:C) = \frac{[(\delta)(1)]}{[(1-\delta)(1) + (\delta)(1)]}$, and therefore, $\Pr(D_1:C) = z_2^A = \delta$.

If the other player instead chose to defect in $n = 1$ when Player A expected that his opponent would collude if rational in $n = 1$, then Player A determines that since it is only possible for rational players to defect in $n = 1$, the probability a TFT type will defect, C, is given by $\Pr(C:D_1)=0$, and it follows that $\Pr(D_1:C) = \frac{[(\delta)(0)]}{[(1-\delta)(1) + (\delta)(0)]}$ and $\Pr(D_1:C) = z_2^A = 0$.

Alternatively, if a rational opponent is expected to defect in $n = 1$, then $z_2^A = 0$ if defection is observed in $n = 1$ since TFT types are always expected to collude in the first period. Let defect be given by C so that $\Pr(C:D_1)=0$ and $\Pr(C:D_2)=1$. It now follows that $\Pr(D_1:C) = \frac{[(\delta)(0)]}{[(1-\delta)(1) + (\delta)(0)]}$ and $\Pr(D_1:C) = z_2^A = 0$. By the same reasoning, $z_2^A = 1$ if the opponent colludes in $n = 1$, since the rational player is expected to defect, and this is given by $\Pr(C:D_1)=1$ and $\Pr(C:D_2)=0$ where C represents collude so that $\Pr(D_1:C) = \frac{[(\delta)(1)]}{[(1-\delta)(0) + (\delta)(1)]}$, and therefore, $\Pr(D_1:C) = z_2^A = 1$.
Q.E.D.

LEMMA 2: After $n = 1$, four cases are possible that dictate the optimal strategy of a rational player in $n = 2$. If a rational player colludes in $n = 1$ and observes collusion, (c,d) is a possible equilibrium strategy for the remainder of the game if $\delta \geq V(a)$, and (d,d) is a possible equilibrium strategy for the remainder of the game if $\delta \leq V(b)$. If a rational player defects and observes collusion, (c,d) is a possible equilibrium strategy for

the remainder of the game if $\delta \geq b/-a$, and (d,d) is a possible equilibrium strategy for the remainder of the game if $\delta < b/-a$. If a rational player colludes or defects in $n = 1$ and observes defection, (d,d) is the only possible strategy for the remainder of the game.

PROOF: Assume two rational players meet, Player A and Player B. If both players collude in $n = 1$, $z_2^A = \delta$ and $z_2^B = \delta$. The SPE requirement from the two period game, as discussed in the previous chapter, implies that (c,d) is an equilibrium strategy for the remainder of the game if $\delta \geq V(a)$. If Player A defects in $n = 1$ and Player B colludes, $z_2^A = \delta$ and $z_2^B = 0$. If the opposing player is Player B, Player B will use (d,d) as $z_2^B < \min\{V(a), V(b)\}$, and if the opposing player is a TFT player, the TFT player will defect in $n = 2$ and either collude or defect in $n = 3$ depending on the move Player A makes in $n = 2$. If Player A uses (c,d) his expected payoff is $\delta(b+a) + (1-\delta)b$, and if he uses (d,d) his expected payoff is 0. Using (c,d) is an equilibrium strategy in the continuation game for Player A if it yields a payoff greater than or equal to the payoff from (d,d) , and this occurs if $\delta \geq b/-a$ and (d,d) is an equilibrium strategy in the continuation of the game if $\delta \leq b/-a$. If Player A colludes or defects in $n = 1$ and Player B defects, $z_2^A = 0$, and Player A will defect for the remainder of the game because $z_2^A < \min\{V(a), V(b)\}$. *Q.E.D.*

COROLLARY 1: Collusion is possible in $n = 2$ when $\delta \geq V(a)$ if a rational player sees collusion in $n = 1$ and if the rational player colluded in $n = 1$. Collusion is also possible when $\delta \geq b/-a$ if the opposing player colluded in $n = 1$, and the rational

player defected. Defection is possible in $n = 2$ if a rational player sees defection in $n = 1$, or if $\delta \leq b/-a$ and the rational player defected in $n = 1$.

PROPOSITION: *Assuming that rational players do not use mixed strategies, the strategy [play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$] is a BPE when $\delta \geq V(a)$, the strategy [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] is a BPE when $\delta \leq V(b)$, and [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] is a BPE when $x > a + b$ and $V(a) \leq \delta \leq V(b)$.*

PROOF: Suppose two rational players, A and B, meet, and all rational players, including B use the strategy [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$]. After $n = 1$, Player A assesses that $z_2^A = 0$, and therefore the only possible equilibrium strategy in the continuation game is defection. This leaves two possible strategies for A to play [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$], and [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$]. The former strategy will return a payoff of δa as a TFT player will collude in the first period and then respond to defect with defect thereafter, and the latter strategy will return a payoff of $\delta(a + x) + (1 - \delta)b$ as a TFT player colludes in $n = 1$ and $n = 2$. Player A will use [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] when $\delta a > \delta(a + x) + (1 - \delta)b$ occurring when $\delta \leq V(b)$.

Suppose two rational players, A and B, meet, and all rational players, including B use the strategy *[play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$]*. Suppose that Player A plays c in $n = 1$. After the first period, Player A assesses that $z_2^A = \delta$. Two strategies are possible for Player A in the continuation game, *[in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$]* and *[in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$]*. If Player A uses the former strategy his expected payoff is $\delta a + 2x$, and if he uses the latter strategy his expected payoff is $x + a$. For the strategy *[play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$]* to be employed, $\delta \geq V(a)$ must be true.

If Player A defects in $n = 1$, again two strategies are possible in the continuation game, *[in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$]* and *[in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$]*. If Player A uses the former strategy his expected payoff will be $a(\delta + 1) + b$, and the strategy *[in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$]* will always have a greater expected payoff as $\delta a + 2x$ will always be greater than $a(\delta + 1) + b$ given the initial condition that $a + b < 2x$. If Player A uses *[in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$]*, his expected payoff is a , and Player A will use *[play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$]* when $\delta a + 2x \geq a$ so that $\delta \geq (a - 2x)/a$. $V(a)$ will always be greater than $(a - 2x)/a$, and therefore $\delta \leq V(a)$ is the bound for *[play c in $n = 1$; in*

$n = 2$, play c if opponent played c and d if opponent played d ; play d in $n = 3$] as a BPE strategy pair.

Now, assume that when Player A and B meet, Player B, if rational uses the strategy [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d ; play d in $n = 3$]. After the first period, Player A concludes that $z_2^A = \delta$. If Player A uses the same strategy as Player B, his expected payoff is $\delta a + x$ for the entire game. If he uses [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d ; play d in $n = 3$], his expected payoff is a as a TFT player will respond to defection in $n = 1$ with defect in $n = 2$, and a rational player will defect in $n = 2$ as is the expected move. If A uses [play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d ; play d in $n = 3$], his expected payoff is $\delta(2x + a) + (1 - \delta)(x + b)$. Comparing A's expected payoff from the strategy [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d ; play d in $n = 3$] and from the strategy [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d ; play d in $n = 3$], A will employ the latter strategy when $\delta a + x \geq a$ or that $\delta \geq V(a)$. Now, comparing A's expected payoff when using the strategy [play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d ; play d in $n = 3$] to using the strategy [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d ; play d in $n = 3$], A will use the latter strategy when $\delta a + x \geq \delta(2x + a) + (1 - \delta)(x + b)$ or that $\delta \leq V(b)$. Given these two conditions for δ , Player A will use [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d ; play d in $n = 3$] when $V(a) \leq \delta \leq V(b)$. If Player A uses [play d in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d ;

play d in $n = 3$], he will have an expected payoff of $a(\delta + 1) + b$, and he will use the same strategy as Player B if $x > a + b$. Since this only occurs when $V(a) \leq \delta \leq V(b)$, [play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$] is a BPE strategy pair when $V(a) \leq \delta \leq V(b)$. Q.E.D.

Collusion in the first and second period are proven to be possible under certain conditions, and the results essentially extended backward from the two period model. This model included a new layer, Bayesian Updating, that requires players to update their beliefs about the opposing player's type after each period. The updating functioned in such a way that if the opposing player behaved as a TFT player would, no more information could be gathered about the player's type at the end of the period. However, if the opposing player took an action that was dissimilar to a move a TFT would make, the opposing player signaled rationality. Again, this story needs to be evaluated in the evolutionary context.

3.3 Evolutionary Model with Three Period Payoffs

In the same manner as in the evolutionary model discussed in the previous chapter, maximizing (m) and TFT (t) types, drawn randomly from a large population, interact randomly in pairs in a large group. The interaction takes the form of the three period rational choice game of section 4.2. The expected payoff for a maximizing type is given by $b_m(\delta) = \delta\pi(m, t) + (1 - \delta)\pi(m, m)$, and the expected payoff for a TFT type is given by $b_t(\delta) = \delta\pi(t, t) + (1 - \delta)\pi(t, m)$. The rate of replication is given by

$r_t = \delta b_t(\delta) + (1 - \delta)b_m(\delta)$. Types are in equilibrium when evolution, defined as the process where the frequency of types changes over time, no longer occurs. If $\delta^* = 1$, $r_t = 0$, and the population is at a monomorphic equilibrium of all TFT types. This is a *stable* equilibrium if, for a value of δ slightly below 1, $b_t(\delta) > b_m(\delta)$. At $\delta^* = 0$, maximizing types are in a monomorphic equilibrium if $b_t(\delta) \leq b_m(\delta)$ so that $r_t \leq 0$, and if $b_t(\delta) < b_m(\delta)$, the equilibrium is stable. A dimorphic equilibrium is also possible, and this will occur if $b_t(\delta) = b_m(\delta)$. The equilibrium is stable when

$$\pi(m, t) - \pi(m, m) - \pi(t, t) + \pi(t, m) > 0.$$

PROPOSITION: *Assuming players use pure strategies, $\delta = 1$ is an unstable equilibrium, and $\delta = 0$ is a stable equilibrium. When $2x > a$ and maximizing types use [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$], maximizing and TFT types in an unstable equilibrium at $\delta^*(d, d, d) = b/(a + b - 3x)$. When $2x < a$, no polymorphic equilibrium exists. When $2x > a$, maximizing types use [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$], and $\delta > \delta^*(d, d, d)$ the population cycles indefinitely around the point where maximizing types switch from [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] to either [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] or [play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$]. When this occurs, $\delta > 0$ is stable, but no equilibrium exists.*

PROOF: Using the payoffs from the rational choice game,

I. Maximizing types using [play d in $n=1$; in $n=2$, play d if opponent played c and d if opponent played d; play d in $n=3$], and TFT types are in equilibrium when their respective expected payoffs are equal where $\delta a = \delta 3x + (1 - \delta)b$. For the expected payoffs to be equal for some value of δ , $a + b$ must be less than $3x$ which is the case given the collusion condition of $a + b < 2x$ presented initially. The equilibrium is given by $\delta^*(d, d, d) = b/(a + b - 3x)$, and it is knife-edge equilibrium because

$$\pi(m, t) - \pi(m, m) - \pi(t, t) + \pi(t, m) > 0 \text{ is not true.}$$

- a. When $\delta < \delta^*(d, d, d)$, maximizing types have a greater expected payoff, and the replication dynamic will push the population to an equilibrium at $\delta = 0$. The equilibrium is stable because $\delta^*(d, d, d)$ is unstable.
- b. When $\delta > \delta^*(d, d, d)$, TFT types will have a greater expected payoff, and δ will increase until $\delta > V(b)$ at which point [play d in $n=1$; in $n=2$, play d if opponent played c and d if opponent played d; play d in $n=3$] is no longer a BPE and maximizing types begin using the strategy [play c in $n=1$; in $n=2$, play c if opponent played c and d if opponent played d; play d in $n=3$].
- c. For $\delta^*(d, d, d)$ to be possible as a knife-edge equilibrium, [play d in $n=1$; in $n=2$, play d if opponent played c and d if opponent played d; play d in $n=3$], must be an BPE when $\delta^*(d, d, d) = b/(a + b - 3x)$. Comparing $b/(a + b - 3x)$ and $V(b)$, $\delta^*(d, d, d)$ is greater than $V(b)$ when $2x < a$, and $\delta^*(d, d, d)$ is less than $V(b)$ when $2x > a$. So when $2x < a$, $\delta^*(d, d, d) > V(b)$, and [play d in $n=1$; in $n=2$, play d if opponent played c and d if opponent played d; play d

in $n = 3$], is not a BPE at $\delta^*(d, d, d)$. No dimorphic equilibrium exists, maximizing types always have a greater expected payoff than TFT types when using [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$], and when maximizing types use [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$], the only equilibrium is a stable equilibrium at $\delta = 0$. When $\delta^*(d, d, d) < V(b)$, a knife-edge equilibrium exists at $\delta^*(d, d, d)$ as long as maximizing types use [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$], when [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] and [play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$] are also Bayesian perfect equilibria which occurs when $V(a) \leq \delta^*(d, d, d)$. When this is true, (a) and (b) explain the replication process and equilibrium selection.

II. Maximizing types using [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] are in equilibrium with TFT types when their payoffs are equal at $\delta a + x = \delta 2x + x + (1 - \delta)b$, so that $\delta^*(c, d, d) = b / (a + b - 2x)$, and this equality is possible when given the original collusion condition, $a + b < 2x$. This equilibrium is a knife-edge as $\pi(m, t) - \pi(m, m) - \pi(t, t) + \pi(t, m) > 0$ is false.

- a. When $\delta < \delta^*(c, d, d)$, maximizing types using $(c; d$ given c or $d; d)$ have a greater expected payoff, and δ will decrease until maximizing types begin using [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$].

- b. When $\delta > \delta^*(c, d, d)$, TFT types have a greater expected payoff, and δ will increase until maximizing types begin using [play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$].
- c. For $\delta^*(c, d, d)$ to be possible as a knife-edge equilibrium, [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] must be an BPE when $\delta^*(c, d, d) = b/(a + b - 2x)$. The pure strategy [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] is a BPE when $\delta \leq V(b)$, and when compared, $\delta^*(c, d, d)$ is always be greater than $V(b)$. Maximizing types will use [play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$] at $\delta^*(c, d, d)$, and when maximizing types do use [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$], they will always have a greater expected payoff than TFT types. Using [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] will cause δ to decrease until maximizing types begin using the strategy [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$].
- III. Maximizing types employing [play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$] will be in equilibrium with TFT types when $\delta a + 2x = \delta x + 2x + (1 - \delta)b$. However, these expected payoffs can never be equal, and the expected payoff to maximizing types will always be greater than the expected payoff to TFT types.

- a. At $\delta = 1$, TFT types are in an unstable monomorphic equilibrium. An invading maximizing type using [play c in $n = 1$; in $n = 2$, play c if opponent played c and d if opponent played d; play d in $n = 3$] will have a greater expected payoff than TFT types, and the system will not return to the monomorphic equilibrium.
- b. When $\delta < 1$, maximizing types have a greater expected payoff, and δ will decrease until maximizing types begin using the strategy [play c in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$] or [play d in $n = 1$; in $n = 2$, play d if opponent played c and d if opponent played d; play d in $n = 3$]. *Q.E.D.*

The results of the evolutionary analysis for the three period game differ significantly from the two period game. Here, under certain circumstances collusion is possible in a stable environment. However, collusion does not occur as part of a stable equilibrium, but rather it exists as part of a cycle between colluding and defecting. While the system is not stable, it confirms the belief of rational players that TFT players can exist in a population with some positive probability. While the results are subtle, they confirm the results of Kreps et al. for the three period case with two-sided asymmetry about player types. The benefit of collusion increased between when moving to a game with greater periods, and this seems to predict the possibility of even greater collusion in games with more than three periods.

CHAPTER IV: CONCLUSION AND SUGGESTIONS FOR FURTHER RESEARCH

Conclusion

The finitely repeated Prisoners' Dilemma is an important area of development for game theoretic analysis because of the apparent discrepancy between experimental and theoretic results. The work in this thesis explains collusion in a specific form of interaction, the finitely repeated PD, that has remained unexplained and assumed by Gauthier as negligible in his social contract theory of morals.

In light of Kreps et al. and my own analysis, cooperation appears possible in a finitely repeated Prisoners' Dilemma with a certain amount of information asymmetry. In both the two period and three period models, collusion is rational under certain conditions of δ . However, TFT types are not stable in an evolutionary system given the expected payoffs from the two period PD. In the three period model, TFT types, while not in equilibrium, are stable for a positive level of δ . The situation is quite unique where no equilibrium is reached because rational players toggle between Nash equilibrium strategies. One of these strategies involves collusion in the first period revealing that not only are the beliefs of rational players that TFT types exist with some probability but they also find it rational to collude in certain circumstances.

Suggestions for Further Research

This paper's analysis has shown that cooperation is possible in a 3 period repeated Prisoners' Dilemma, and that the types in the population are stable for certain conditions in an evolutionary system. The first step in extending this thesis would be to develop a

model where universal collusion is a possible equilibrium. In the three period model the rational strategy of a player cycles between defecting and colluding in the first period of the game. An interesting addition would include the possibility of rational players cycling between different levels of collusion. This may be possible in as early as a 4 period model if rational players cycle between colluding in the first period and colluding in periods including and after the first period.

Another interesting direction in which this paper could be developed is towards a generalized model for n repeated periods of the Prisoners' Dilemma. Between the two and three period models, the benefits of cooperation increase with an increase in the number of periods. This suggests that cooperation will become more and more likely as the number of periods increases.

In addition to or in replacement of information asymmetry as a necessary condition for rational cooperation, other scenarios could be explored that allow for cooperation in finitely repeated Prisoner Dilemmas. This is a very interesting topic that this paper has only lightly touched, and from the nexus of rational choice and evolutionary game theory, the model and the concept can be expanded in either direction.

NOTES

¹ Plato, *The Republic* (New York: Cambridge University Press, 2000), bk. 2 359c-60b.

² Thomas Hobbes, *Leviathan* (New York: Penguin Books, 1985).. For critique of the Foole, see especially Hobbes, *Leviathan*, 205.

³ Rousseau, *The Discourses and Other Early Political Writings* (New York: Cambridge University Press, 1997), 163.. Rousseau does not explicitly state the payoffs from killing a stag and killing a rabbit, nor does he state the likelihood that either animal will be successfully killed given either decision the hunter interested in defecting makes. If it is the case that the members of the group do not conflict in their interests but strictly prefer that the same decision is made by all hunters, then the stag hunt is a coordination problem not a Prisoners' Dilemma.

⁴ David Hume, *A Treatise of Human Nature*, ed. John Cottingham, Oxford Philosophical Texts (New York: Oxford University Press, 2000). cited in Mancur Olson, *The Logic of Collective Action* (Cambridge, MA: Harvard University Press, 1965), 33-34.

⁵ Hume, *A Treatise of Human Nature*, bk. 3, part 2, sect. 8, 345-52.

⁶ David Gauthier, *Morals by Agreement* (Oxford: Clarendon Press, 1986), 2.

⁷ Ibid.

⁸ Ibid., 4.

⁹ John Maynard Smith, *Evolution and the Theory of Games* (New York: Cambridge University Press, 1982).

¹⁰ David M. Kreps et al., "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory* (1982): 247-50.

¹¹ Collusion in a finitely repeated period game with information asymmetry may also be achieved if a rational player imputes some positive probability that the other player simply enjoys collusion. Modeling collusion in such a way is practical because it does not rely on individuals behaving according to pre-programmed strategies, and therefore it can support collusion without relying on an evolutionary scenario. However, this assumes that the utility functions of certain individuals are different where colluding increases utility greater than the payoff for collusion in a PD. This begs the question as to why some players do not get an extra payoff from defecting in response to someone else's collusion. The model used in this paper rests on the assumption that at least some individuals, at least part of the time, behave as types.

¹² $a + b < 2$ is necessary because otherwise the Pareto-dominant strategy is to alternate between defect-collude and collude-defect, and I would like to assume that consistent cooperation is always Pareto optimal.

¹³ Russell W. Cooper, *Coordination Games* (New York: Cambridge University Press, 1999), 1-10.

¹⁴ John C. Harsanyi and Reinhard Selten, *A General Theory of Equilibrium Selection in Games* (Cambridge: MIT Press, 1988).

¹⁵ Hans Carlsson and Eric van Damme, "Global Games and Equilibrium Selection," *Econometrica* 61, no. 5 (1993).

¹⁶ Michihiro Kandori, George J. Mailath, and Rafael Rob, "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica* 61, no. 1 (1993).

¹⁷ To see this equilibrium definition refer to Samuel Bowles and Herbert Gintis, "How Communities Govern: The Structural Basis of Prosocial Norms," in *Economics, Values, and Organization*, ed. Avner Ben-Ner and Louis Putterman (New York: Cambridge University Press, 1998).. While other equilibrium definitions exist, for example Smith, *Evolution and the Theory of Games*., the Bowles and Gintis (1998) definition is most consistent with the explanation for the replication of types used in this paper.

¹⁸ This function is drawn from the derivative of the growth function with respect to δ : $dr_t / d\delta$ where

$$r_t = b_t(\delta) - \bar{b}(\delta) = b_t(\delta) - [\delta b_t(\delta) + (1 - \delta)b_m(\delta)] \text{ so}$$

$$dr_t / d\delta = -(b_t - b_m) + (1 - \delta)(\pi(t, t) - \pi(t, m) - \pi(m, t) + \pi(m, m)) \text{ and since } dr_t / d\delta > 0$$

the condition $\pi(m, t) - \pi(m, m) - \pi(t, t) + \pi(t, m) > 0$ must also be true.

BIBLIOGRAPHY

- Axelrod, Robert. *The Evolution of Cooperation*. New York: Basic Books, 1984.
- Basu, Kaushik. *Prelude to Political Economy*. New York: Oxford University Press, 2000.
- Ben-Ner, Avner, and Louis Putterman. "Values and Institutions in Economic Analysis." In *Economics, Values, and Organization*, edited by Avner Ben-Ner and Louis Putterman, 3-72. New York: Cambridge University Press, 1998.
- Bowles, Samuel, and Herbert Gintis. "How Communities Govern: The Structural Basis of Prosocial Norms." In *Economics, Values, and Organization*, edited by Avner Ben-Ner and Louis Putterman, 206-30. New York: Cambridge University Press, 1998.
- Carlsson, Hans, and Eric van Damme. "Global Games and Equilibrium Selection." *Econometrica* 61, no. 5 (1993): 989-1018.
- Cooper, Russell W. *Coordination Games*. New York: Cambridge University Press, 1999.
- Dixit, Avinash, and Susan Skeath. *Game of Strategy*. New York: W.W. Norton & Company, Inc., 1999.
- Gauthier, David. *Morals by Agreement*. Oxford: Clarendon Press, 1986.
- Harsanyi, John C., and Reinhard Selten. *A General Theory of Equilibrium Selection in Games*. Cambridge: MIT Press, 1988.
- Hobbes, Thomas. *Leviathan*. New York: Penguin Books, 1985.
- Hume, David. *A Treatise of Human Nature*. Edited by John Cottingham, Oxford Philosophical Texts. New York: Oxford University Press, 2000.
- Kandori, Michihiro, George J. Mailath, and Rafael Rob. "Learning, Mutation, and Long Run Equilibria in Games." *Econometrica* 61, no. 1 (1993): 29-56.
- Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson. "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." *Journal of Economic Theory* (1982): 245-52.
- Olson, Mancur. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press, 1965.
- Plato. *The Republic*. New York: Cambridge University Press, 2000.
- Rousseau. *The Discourses and Other Early Political Writings*. New York: Cambridge University Press, 1997.

Smith, John Maynard. *Evolution and the Theory of Games*. New York: Cambridge University Press, 1982.

Ullmann-Margalit, Edna. *The Emergence of Norms*. Oxford: Clarendon Press, 1977.