On the Foundations of Basic Property Rights, Part I: A Model of the State-of-Nature with Two Players

by

Abhinay Muthoo

Department of Economics
University of Essex
Colchester CO4 3SQ
England, UK.

Email: `muthoo@essex.ac.uk`
URL: `http://privatewww.essex.ac.uk/ ~ muthoo`

July 2000

Abstract. This paper is concerned with the fundamental issue of the origins (or emergence) of basic property rights such as the right of an individual to the fruits of his labour. I develop a dynamic model of the strategic interaction between two players in the *state-of-nature*, which is an environment characterized by the absence of any rules, regulations, laws and institutions (including property rights and the state). My objective is to address in a rigorous manner the questions of *why, when* and *how* basic property rights can emerge and be made secure. In particular, I explore the roles of the players' fighting and productive skills (or capabilities) on the emergence of secure property rights.

Keywords: Property rights, state-of-nature, natural equilibrium, self-enforcement, incentive-compatibility, inter-player transfers of output.

JEL Classification Numbers: C72, C73, C78, D23, D62, D74, F02, K11, K12, K14, K33, K42, O12

Acknowledgements: I acknowledge, with gratitude, the financial support of the UK Economic and Social Research Council (ESRC) through a research grant (No: R000222995). I have had stimulating discussions on various issues related to the topic of this paper with several friends and colleagues including Roy Bailey, V. Bhaskar, Michihiro Kandoori, Chris Minns, Todd Sandler, Stergios Skaperdas, David Austen-Smith, Hugh Ward, Albert Weale, and Ethan Zorick. For helpful comments by seminar participants at the University of Essex and by the participants of the joint World Bank and Princeton University's Center for International Studies workshop on *The Economics of Political Violence*, I am grateful.

"... there be no Propreity, no Dominion, no Mine and Thine distinct: but only that to be every mans that he can get: and for so long, as he can keep it." THOMAS HOBBES, *Leviathan, 1651.*

1 Introduction

1.1 Motivation

Property rights are important for a variety of reasons. In the absence of well-defined and secure property rights, mutually beneficial transactions may fail to occur, and value-enhancing investments may fail to be undertaken. If, for example, my right over the fruits of my labour are not secure (perhaps because they are vulnerable to theft), then my incentive to work would be adversely affected. The fundamental importance of secure property rights to the economic, social and political development of the poorer parts of the world has been recently emphasized by the World Bank in their 1997 World Development Report *The State in a Changing World*. In many nation-states today, property rights are insecure or non-existent, which is a fundamental reason for these nation-states' failure to develop economically, socially and politically.[1]

There is a large literature that studies the role of the *distribution* of property rights on economic outcomes. An early key contribution was Coase (1960), who argued that in a "frictionless" environment, if property rights are well-defined and secure, then economic efficiency will typically be attained. In particular, the distribution of property rights has no affect on economic efficiency, although economic distribution may be affected by who has what property rights. Subsequently, many authors have explored the role of the distribution of property rights in environments with various kinds of frictions. For example, Grossman and Hart (1986) have argued that in an environment with incomplete contracting, the distribution of property rights will affect economic efficiency. There is now a large literature that builds on Grossman and Hart's 1986 contribution; a main focus of that literature is on the issue of the optimal distribution of property rights.[2]

[1] There are many empirical studies that show the adverse effects of weak or insecure property rights. Let me mention some of them. Besley (1995) finds a significant link between property rights and investment in Ghana. Mauro (1995) and Svensson (1998) show in the context of cross-country studies that weak property rights adversely affects aggregate growth. Demirguc-Kunt and Maksimovic (1998) find that firms invest more in countries with secure property rights. For five transition economies in Eastern Europe and the former Soviet Union, Johnson, McMillan and Woodruff (1999) find that weak property rights limit the reinvestment of profits in startup manufacturing firms.

[2] For a recent survey of that literature, see Hart (1995).

Almost all of the analyses in the economics literature on the role of property rights on economic outcomes *assumes* — sometimes explicitly, but more often only implicitly — that any specified distribution of property rights can be almost *costlessly enforced* (by, for example, a third party such as the courts/state).[3] The broad objective of this literature is to study the impact that different distributions of property rights have on economic outcomes. Much valuable insights have been provided by this literature.

This paper, on the other hand, is concerned with the fundamental issue of the *origins* (or emergence) of secure property rights. The starting point of this paper is the *state-of-nature*, which is an environment characterized by the absence of any property rights (and any other institution such as the state). I shall construct a dynamic, repeated interaction model of the state-of-nature in order to rigorously analyze the origins of basic property rights. The analysis of my model will provide novel insights into the conditions under which basic property rights can and cannot emerge. Furthermore, and equally importantly, I shall explore the thorny issue of how such property rights are enforced and made secure.

Mainstream economists have largely ignored the issue of the origins of secure property rights by taking them as exogenously given. This might partly be because they see this topic lying outside the scope of the discipline of economics. There is, in contrast, a large literature in political and moral philosophy that is concerned with the origins of the state and conceptions of a just society. That literature does indirectly (if not directly) address the issue of the origins of basic property rights. Early notable contributions to this literature were made by the great political philosophers such as Thomas Hobbes, John Locke, David Hume and Jean-Jacques Rousseau in respectively Hobbes (1651), Locke (1690), Hume (1739) and Rousseau (1762). Many important contributions to this topic were made in the twentieth century by John Rawls, Robert Nozick and James Buchanan in respectively Rawls (1972), Nozick (1974), and Buchanan (1975). The strengths of much of the more recent work lie in the formalization of some of the ideas and arguments of the early political and moral philosophers — see, for example, Gauthier (1986), Sugden (1986), Taylor (1987), and Binmore (1994 and 1998).

Although the political and moral philosophy literature contains a wealth of ideas and informal arguments, and also some formal analyses, in this literature there is no *model*

[3]For example, as is discussed at great length in Buchanan (1975), this assumption lies at the heart of the many contributions (such as Cheung, 1963, and Demsetz, 1967) on the economics of property rights made in the 1960s and early 1970s. Indeed, those contributions do not concentrate on explaining the emergence of secure property rights. Instead, their focus is on explaining changes in property rights in a society with an established government.

that is used consistently, in a unified and systematic manner, to address the issues that arise in the study of the origins of basic property rights and other institutions. In order to probe more deeply (and in a rigorous and consistent manner) into the ideas and informal arguments put forward in this literature, and explore their range of validity, and in order to obtain new insights, we need to construct an appropriately detailed, rigorous and plausible model of the state-of-nature. This is precisely what I shall do in this paper. The power of my model will be illustrated by the precisely defined insights that it provides about the origins of basic property rights.

Some authors (such as Taylor, 1987) have informally argued that the state-of-nature can be modelled as a standard repeated Prisoners' Dilemma game. This is overly simplistic and somewhat inaccurate, partly because it leads to restrictive conclusions. The model that I develop in this paper will be a richer, more accurate and general representation of the strategic interaction in the state-of-nature. Equally importantly, it will provide novel insights about the costs and benefits of basic property rights. In particular, I will show that they depend on the distribution of the initially endowed characteristics amongst the players. For example, whether or not basic property rights can potentially emerge, and the related issue of whether or not they can be enforced, depends crucially on the distribution of the initially endowed productive and fighting skills amongst the players. The crucial role played by the *heterogeneity* in such skills on the origins of such property rights will be rigorously developed in this paper.

In developing my model I adopt Hobbes's viewpoint that in the state-of-nature, the only "natural" right of man is the right to the use of his own physical power. The emergence of secure property rights, and the nature of any transfers of output amongst the various players, depends crucially on the use, and potential use of, violence and force. Indeed, this perspective would appear to be vindicated by international events in today's world; for example, it is all too clear that the United States of America's influence and power is derived as much from its productive skills and technologies as from its fighting skills and military technologies.

James Buchanan in Buchanan (1975) presents an informal framework and analysis of, in particular, the emergence of property rights that is also based on the Hobbessian perspective. Although he does not conduct a formal and rigorous analysis, and does not, in particular, provide an analysis of the conditions under which property rights can be made secure (which is a crucial issue in understanding the origins of basic property rights), his analysis provided me with much food for thought. John Umbeck in Umbeck (1981) has developed a relatively more formal analysis of the origins of basic property rights based also upon the Hobbessian perspective. But, like Buchanan (1975), the crucial issue of the conditions under which (and how) the basic property rights can be made secure is not properly addressed. In the absence of any third party in a state-of-nature with two players, the issue of how any property rights are made secure becomes thorny. Indeed, property rights have to be *self-enforcing* in order to be secure. But that requires adopting a dynamic perspective, and constructing a *dynamic*, repeated interaction model of the state-of-nature. I should emphasize that most scholars do, however, recognize that the "enforcement" issue is of fundamental importance in understanding the origins of institutions such as property rights. See, in particular, North (1990), which is a classic treatise that discusses this and other related issues.

Given the absence of a global (or supranational) state, the world, at the international level, is essentially in a state-of-nature. As such, my model may also be interpreted in terms of the issue of the emergence of secure, "international-level" property rights. My analysis and results may, therefore, contribute to the understanding of the circumstances under which such rights can and cannot emerge, and, furthermore, the cirumstances under which nation-states would and would not engage in war.[4]

As a final motivating factor, let me suggest that my model may also be useful in terms of providing some insights into civil wars. A basic cause of many kinds of civil wars is the absence of secure property rights. In particular, the state is so weak that it is unable to enforce them. For example, the large-scale civil violence in the former Soviet Union is linked to the end of the Cold War and the collapse of the Soviet state. The war in the former Yugoslavia may be as well. Given the state-of-nature type of environment within which many civil wars take place, my results may help towards understanding the conditions that are required to hold for such conflicts to be peacefully negotiated.[5]

1.2 Overview

In this paper I study environments with exactly two players. This allows me to develop an indepth understanding of several key aspects of the topic under consideration without having to be concerned with the thorny issue of coalition formation that rears its head when there are three or more players around. However, it may be noted that the model of the two-player environment developed in this paper provides the back-

[4]There is, of course, a large literature in *International Relations Theory* that is concerned with such issues. A good starting point is Powell (1999), who, unlike much of this literature, emphasizes the role of formal, game-theoretic models in addressing the issues.

[5]There is a large literature on the economics of civil wars. For a good introduction to this literature, see Collier and Hoeffler (1998), and, Fearon and Laitin (1999).

bone for the model of the three-player environment developed in my companion paper, Muthoo (In Preparation), in which the issue of who forms a coalition with whom and why is a focus of the analysis.

In section 2 below, I develop the basic framework that will be built upon in subsequent sections to address the main issues of concern. Specifically, in subsection 2.1, I lay down the *base* model of the state-of-nature, which constitutes the infinite repetition of a two-stage game. This model captures the essential, basic elements of the strategic interaction between two players in an environment characterized by the absence of property rights. In particular, two important constraints (one of which I relax in section 5) built into the structure of the base model are that no explicit communication between the players is allowed, and that the players cannot make any (potentially desirable) investments in their respective productive and fighting skills. Then, in subsections 2.2 and 2.3, respectively, I characterize and study the properties of the *natural equilibrium* — which, formally speaking, is the unique *stationary* subgame perfect equilibrium of the base model. The "natural equilibrium" terminology, which is taken from Buchanan (1975), aptly denotes the (equilibrium) outcome in the absence of secure property rights. A key insight obtained here is that the absence of secure property rights means that the *fear* of war is not absent, notwithstanding the possibility that in the natural equilibrium no war actually takes place; and moreover, it is the fear of war that determine's each player's *ex-ante* incentives to work and produce output. Another interesting insight obtained here is that an improvement in a player's fighting skill makes him worse-off in the natural equilibrium *if* he is militarily strong; but better-off otherwise.

Section 3 derives and studies the cost and benefit to *each* player of establishing the property rights under consideration. This involves comparing each player's payoff in the natural equilibrium with his payoff in, what I shall call, the *property rights equilibrium*, which is the unique subgame perfect equilibrium of the base model *on the assumption* that secure property rights do exist. This exercise of exploring the players' *private* incentives to establish the property rights (and how such incentives depend on the players' potentially differential fighting and productive skills) provides some key insights about the conditions under which secure property rights can and cannot emerge. For example, it will be shown that there exists configurations of the players' fighting and productive skills under which the players' private incentives to establish the property rights are in *conflict*, in the sense that while one player prefers the property rights equilibrium over the natural equilibrium, the other player has the opposite preference. This is the case when, for example, one player is quite unproductive but very strong, while the other player is quite productive but very weak.

As is intuitive, in this case the former player will prefer the natural equilibrium over the property rights equilibrium, while the opposite is the case for the latter player. Another insight obtained is that in order to promote incentives to establish the property rights, improvements in the players' productive skills (or economic prosperity) should go hand-in-hand with improvements in their fighting skills (or military technologies).

A key aspect of the analysis of the emergence of secure property rights, which is conducted in section 4, concerns the study of the appropriate incentive-compatibility conditions that are required to hold for the emergence of *self-enforcing* property rights. Formally, this analysis concerns the conditions under which the *outcome* associated with the property rights equilibrium can be sustained as the outcome of some (necessarily *non-stationary*) subgame perfect equilibrium of the base model. It will be shown that for *some* configurations of the players' fighting and productive skills, and, discount rates, secure property rights *can* emerge in the state-of-nature. However, it will be shown that there exists configurations of values of such parameters under which secure property rights can *never* emerge in the state-of-nature. For example, unless the players do not discount future payoffs, secure property rights are unlikely to emerge when their private incentives are in conflict. Several key insights into the roles of the players' fighting and productive skills on the emergence or otherwise of secure property rights are derived. For example, it will be shown that improvements in the military technology of a militarily strong player (such as the USA) can enhance the likelihood of the emergence of secure property rights. On the other hand, improvements in the military technology of a militarily weak player may increase the likelihood that existing secure property rights become insecure. All the results and insights obtained in this section are based on the assumption that players cannot engage in *inter-player transfers of output*, an assumption that is relaxed in the next section.

In Section 5, I extend the base model by giving the players the option to engage in explicit communication. The main purpose of such communication is to allow the players the opportunity to bargain over (potentially beneficial) transfers of output between them. Such bargained agreements, which are required to be self-enforcing, enhance the likelihood of the emergence of secure property rights. In particular, it will be shown that even if the players do discount future payoffs (to some extent) and their private incentives are in conflict, then the mechanism of inter-player transfers of output can provide the basis for the emergence of secure property rights. The results and insights obtained in this section are partly based on the fact that the players always have a *collective* incentive to establish the property rights, which can be potentially exploited through the mechanism of inter-player transfers of output.[6]

[6]The notion that there exists a "collective" incentive will be formally defined later on. Suffice it

I conclude in Section 6 with a summary of the main contribution of this paper, and, with a discussion of the main limitations of my model and analyses. The latter will suggest potentially fruitful extensions of my model that could be developed and studied in future research.

1.3 Other Models of the State-of-Nature

My model — and the focus of my analysis — is quite different and unrelated to other models of the state-of-nature. However, some of these other models do share some (general) features with my model. I now highlight some of the main similarities and differences between them and my model.[7]

Bush and Mayer (1974) were perhaps the first to study a formal model of the state-of-nature. A basic weakness of their model is that it is static, and thus, they could not address the key issue of the enforcement of property rights, an issue, however, that they recognise to be an important one.

It is worth stressing upfront that an important and fundamental difference between my (dynamic, repeated interaction) model and all the other models of the state-of-nature is that my model of the interaction between the two players in each period — which, in the game theory terminology, denotes my *stage game* — is quite different to the static model of Bush and Mayer (1974) which (or some variant of it) forms the basis of the other models of the state-of-nature.

Houba and Weikard (1995) study a repeated interaction, bargaining model in which the two parties negotiate over how much time each party should spend in each period on productive and predatory activities. Since their analysis focuses on the case in which the parties (effectively) do not discount future payoffs, *any* negotiated agreement can be self-enforcing for *any* parameter values.[8] Their main result is that there exists bargaining equilibria in which the parties fail to reach any agreement (and thus, in each period, the parties chooses their actions "non-cooperatively"). With the exception of their model, my model is the only other model of the state-of-nature with repeated interaction and in which the parties have the option to engage in negotiations over

to say here that it captures the notion that the players, when treated as "jointly", prefer the property rights equilibrium over the natural equilibrium.

[7]By restricting this short review to those models which share some features with my model, I am sure that I will fail to cite some good models of the state-of-nature. I apologise to authors of such models for this omission.

[8]This implication is an immediate consequence of the well-known "Folk Theorem" result from the *Theory of Repeated Games*.

various relevant issues. But, beyond these two common, general features, my model and the focus of my analysis is quite different. For example, their assumption that the parties' discount factors are arbitrarily close to one effectively solves the enforcement problem. In contrast, my analysis is specifically concerned with the (more plausible) case in which the parties do discount future payoffs. This is because my objective is to develop a thorough understanding of the roles of the key parameters (namely, the parties' productive skills, fighting skills, and, discount factors) on the emergence or otherwise of secure property rights. Furthermore, in my model the parties negotiate over whether or not to establish the property rights and over the level of ouput that one player will transfer to the other player. It seems rather persuasive that in the state-of-nature, the parties would engage in such negotiations. As such, my aim is to explore the extent to which the mechanism of *inter-player transfers of output* can enhance the likelihood of the emergence of secure property rights.

There is a large and growing literature that looks at the problem of resource allocation in a world where productive resources can be used to appropriate wealth as well as to create it. Some notable examples of such models of conflict include Skaperdas (1992), Grossman and Kim (1995) and Hirshleifer (1995). For a nice introduction to this literature, see Garfinkel and Skaperdas (2000), which contains references to the many other papers in this literature. A few basic ideas from that literature are related to those in this paper such as the absence of the state and secure property rights. However, this literature is not explicitly concerned with the issue of the emergence of secure property rights and/or cannot address this issue since almost all of these models are static. Their focus is on, amongst other issues, studying how various parameters affect the equilibrium allocation of resources amongst productive and predatory activities. A key, common feature (or assumption) that underlies most of the models in this literature is that the players' outputs are *aggregated into a common pool* that is at risk of being redistributed by force. As such, these models may be interpreted as models of the right of access to *common* property.[9] In contrast, my model is concerned with *private* property rights. Furthermore, while a key concept in this literature is the technology of conflict (with the so-called *Constant Success Function* the oft-used one), this plays a minor role in my model, partly since the focus of my analysis lies elsewhere.

Bates, Greif and Singh (2000) study a repeated interaction model that has some similar features to the model that I study below. However, their focus is on the conditions under which an unproductive third-party (a government) acts as the enforcer of prop-

[9]This point has, in fact, been recognised by — and forms the basis of the studies in — Grossman and Kim (1995), and, Neary (1996).

erty rights in return for some output (taxes) from the other two productive players. As such, this contrasts with my analysis which focuses on the issue of the emergence of secure property rights *without* recourse to any third party.

2 The Basic Framework

2.1 The Base Model

Time is divided into an infinite number of periods, $1, 2, 3, \ldots,$ where each period consists of $T > 0$ units of time. There are two players, A and B. The decisions that each player has to take in each period, and the structure of the interaction between them is defined in the following two-stage game, which, for future reference, I denote by $\mathcal{G}$.

Stage 1: [How much to work?]. At the beginning of each period the two players simultaneously choose the quantities of time that they respectively will work. If player i ($i = A, B$) works for L_i units of time, where $0 \leq L_i \leq T$, then he produces $f_i(L_i)$ units of output. The production function f_i is twice differentiable, increasing ($f_i' > 0$) and concave ($f_i'' \leq 0$), and furthermore, $f_i(0) = 0$.

Stage 2: [To fight or not to fight?]. At the end of each period both players observe the quantities of output produced by each player, and then they simultaneously decide whether or not to fight. If both players choose not to fight, then player i's levels of consumption and leisure in this period are respectively $f_i(L_i)$ and $T - L_i$. On the other hand, if at least one player decides to fight, then a fight takes place. There are three possible (randomly determined) outcomes of a fight, namely:

• With probability p_i player i wins the fight and steals all of player j's ($j \neq i$) output, where $p_A > 0$, $p_B > 0$ and $p_A + p_B \leq 1$. In this case player i's levels of consumption and leisure in this period are respectively $f_A(L_A) + f_B(L_B)$ and $T - L_i$, while player j's levels of consumption and leisure in this period are respectively 0 and $T - L_j$.[10]

• With probability $1 - p_A - p_B$ no one wins the fight, the players retreat and no one steals anything. In this case player i's ($i = A, B$) levels of consumption and leisure in this period are respectively $f_i(L_i)$ and $T - L_i$.

[10]It is implicitly being assumed that no output can be consumed until after the outcome of a fight. This modelling assumption is a simple way to capture the role of a fight on each player's incentives to work.

The (von Neumann-Morgenstern) utility to player i in each period is $U_i(c, l)$, where c and l are respectively his levels of consumption and leisure in that period. I shall assume that U_i takes the following (quasi-linear) form: $U_i(c, l) = c + v_i(l)$, where v_i is twice differentiable, increasing ($v_i' > 0$) and strictly concave ($v_i'' < 0$), and furthermore, $v_i(0) = 0$.[11] Each player's objective is to maximize the present discounted value of his expected utility, where $\delta_i \in [0, 1)$ denotes player i's (per-period) discount factor.

Notice, therefore, that the base model constitutes the infinite repetition of the two-stage game $\mathcal{G}$. It defines, in particular, the players' basic strategic interaction in an environment in which neither player has property rights over the fruits of his labour — that is, over the output that he produces by using as inputs his labour and his productive skills. Hence, the relevance (in each period) of stage 2 when each player considers stealing the output of the other player. As mentioned earlier, my objective in studying this model is to address the questions of *why, when and how* such property rights can emerge and be made secure. In particular, I shall study how the players' fighting skills (as captured by the probabilities p_A and p_B) and productive skills (as captured by the production functions f_A and f_B) impact upon the emergence or otherwise of secure property rights.

A natural interpretation of the base model is implicit in its formal description: it represents the interaction between two human beings (of the same gender) in the state-of-nature. The following alternative interpretation is also potentially applicable: the model represents the interaction between two organizations (such as two nation-states or two mafias) in the state-of-nature. While the former interpretation is perhaps more useful from a theoretical perspective, the latter has much relevance to the world in which we currently live.

It may be noted that in the base model a player can invest neither in his productive skills nor in his fighting skills; each player has to live forever with what Mother Nature endowed him with at the beginning of period 1. Furthermore, another implicit assumption underlying the base model is that a player's endowments of time and skills cannot be stolen. This means, in particular, that slavery is not allowed. I shall return to these issues in section 6, where I shall also discuss the issue of relaxing various other assumptions such as the implicit assumptions that for each $i = A, B$, p_i and δ_i

[11] I adopt this particular utility function partly to simplify the analyses (the additive separability feature), and partly to capture the assumption that each player has risk-neutral preferences over consumption. As is intuitive, under such an assumption the emergence of secure property rights is that much harder.

are exogenous and history-independent.[12]

I shall use the subgame perfect equilibrium concept (SPE, for short) to analyze the base model. Furthermore, in order to somewhat simplify the analysis, but without any significant loss of generality, I shall rule out SPE in which a player uses a weakly dominated strategy.

2.2 The Natural Equilibrium

In this subsection I shall characterize the unique *stationary* SPE of the base model. As I mentioned in section 1, following Buchanan (1975) I call this the natural equilibrium, partly because in this equilibrium secure property rights do not exist. As is well-known, in a stationary SPE each player's *equilibrium* strategy is stationary: that is, each player's equilibrium actions in each period do not depend on the actions taken by the players in any previous period. Formally, a stationary strategy for player i ($i = A, B$) is defined by a number L_i and a function $\phi_i : [0, T]^2 \to \{f, nf\}$, with the following interpretation. In *each* period player i chooses to work for L_i units of time, and, $\phi_i(L_A, L_B)$ indicates whether player i fights (f) or does not fight (nf) in each period at stage 2 conditional on the choices L_A and L_B made by the players at stage 1.

A stationary SPE of the base model is the repeated play of a SPE of the two-stage game $\mathcal{G}$. Hence, in order to derive the stationary SPE of the base model, I now derive the SPE of the two-stage game $\mathcal{G}$. Consider, therefore, the two-stage game $\mathcal{G}$, and fix an arbitrary pair (L_A, L_B) chosen at stage 1. If the players end up in a fight at stage 2, then player i's *expected* payoff is

$$E_i^f = p_i U_i(\widehat{c}, l_i) + p_j U_i(0, l_i) + (1 - p_i - p_j) U_i(c_i, l_i),$$

where $j \neq i$, $\widehat{c} = f_A(L_A) + f_B(L_B)$, $c_i = f_i(L_i)$ and $l_i = T - L_i$. Letting E_i^{nf} denote player i's payoff if the players do not end up in a fight, where $E_i^{nf} = U_i(c_i, l_i)$, it follows

[12] An alternative interpretation of δ_i is that $1 - \delta_i$ is the probability with which player i dies in each period. With this interpretation, it is natural to then argue that the probability with which player i dies in any period depends on, amongst other variables, the outcome of the fight in that period (if a fight occurs) and on the level of his consumption in that period. Furthermore, player i's probability p_i of winning a fight might depend on how healthy he is, which should depend on the levels of his past consumption.

(after substituting for the assumed quasi-linear form of player i's utility function) that

$$E_i^f \gtreqless E_i^{nf} \Longleftrightarrow p_i f_j(L_j) \gtreqless p_j f_i(L_i). \tag{1}$$

The left-hand side of the second inequality in (1) is player i's expected net gain from the fight: since the fight brings in an additional quantity of output (equal to $f_j(L_j)$) with probability p_i. On the other hand, the right-hand side of the second inequality in (1) is his expected net loss from the fight: since in the fight he would lose all his output with probability p_j. In order to simplify the analysis, I assume that when indifferent between having a fight and not having a fight, each player prefers the latter. It thus follows that at stage 2 each player has a weakly dominant action (WDA): if $E_i^f \leq E_i^{nf}$ then player i's WDA is not to fight, while if $E_i^f > E_i^{nf}$ then player i's WDA is to fight. For future reference, I state this result in the following lemma:

Lemma 1. *Consider the two-stage game $\mathcal{G}$, and fix any pair (L_A, L_B) chosen at stage 1. Then, at stage 2, player i's ($i = A, B$) weakly dominant action (WDA) is not to fight if and only if $p_i f_j(L_j) \leq p_j f_i(L_i)$.*

Lemma 1 implies that (in any SPE of the two-stage game $\mathcal{G}$) a fight will not take place at stage 2 if and only if the pair (L_A, L_B) chosen at stage 1 satisfies the following equation:

$$p_B f_A(L_A) = p_A f_B(L_B). \tag{2}$$

Notice that equation 2 implies, in particular, that each player's expected net loss from a fight equals his expected net gain from a fight. It is instructive to note that a fight takes place only if exactly one player's WDA is to fight. To put it differently, it is never the case that both players' WDA is to fight.

It is straightforward to show that for any pair (L_A, L_B) chosen at stage 1, player i's (equilibrium) expected payoff is $\Pi_i(L_i, L_j)$, where[13]

$$\Pi_i(L_i, L_j) = (1 - p_j) f_i(L_i) + p_i f_j(L_j) + v_i(T - L_i), \tag{3}$$

which may be interpreted as follows. In the case of a fight, player i's total expected

[13] This is established as follows. First, consider a pair (L_A, L_B) that satisfies equation 2. In that case the expected payoff to player i is $f_i(L_i) + v_i(T - L_i)$, which, however, equals $\Pi_i(L_i, L_j)$ since (by equation 2) $p_i f_j(L_j) = p_j f_i(L_i)$. Now consider a pair (L_A, L_B) that does not satisfy equation 2. In that case the expected payoff to player i is (after simplifying) $\Pi_i(L_i, L_j)$.

consumption is the sum of the first two terms: with probability $1 - p_j$ he consumes all of his output, and with probability p_i he consumes all of player j's output. In the case of no fight (when $p_i f_j(L_j) = p_j f_i(L_i)$), he consumes all of his output and none of player j's output.

Since the partial derivative of Π_i with respect to L_i is independent of L_j, and since (given that $f_i'' \leq 0$ and $v_i'' < 0$) Π_i is strictly concave in L_i, it follows that at stage 1 player i has a unique strictly dominant action (SDA), which is to work for L_i^N units of time, where L_i^N is the unique solution of the following maximization problem:

$$\max_{0 \leq L_i \leq T} (1 - p_j) f_i(L_i) + v_i(T - L_i).$$

In order to simplify the exposition, but without any significant loss of generality, I shall adopt the assumptions on the parameters stated below in Assumption 1 which are sufficient for the maximization problem stated above to have an *interior* solution (i.e., $0 < L_i^N < T$).[14]

Assumption 1. $(1 - p_j) f_i'(0) > v_i'(T)$ and $f_i'(T) < v_i'(0)$ (where $i, j = A, B$ with $i \neq j$).

Given Assumption 1, it follows that L_i^N is the unique solution of the following (first-order) condition:

$$(1 - p_j) f_i'(L_i) = v_i'(T - L_i). \tag{4}$$

The left-hand of (4) is player i's marginal benefit from working, while the right-hand side is his marginal cost from doing so. I have thus established the following proposition:

Proposition 1 (The Natural Equilibrium (NE)). *The base model has a unique stationary SPE — which will be called the natural equilibrium (NE, for short). Player i's ($i = A, B$) equilibrium strategy is $\langle L_i^N, \phi_i^N \rangle$, where L_i^N is the unique solution of (4), and*

$$\phi_i^N(.) = \begin{cases} f & \text{if } (L_A, L_B) \in \Sigma / \Omega_i \\ nf & \text{if } (L_A, L_B) \in \Omega_i, \end{cases}$$

[14] While the first inequality in Assumption 1 is also necessary, the second is not necessary. However, it is convenient to adopt it here since it will prove to be necessary for the existence of an *interior* solution in another maximization problem stated in section 3.

where $\Sigma = \{(L_A, L_B) : 0 \leq L_A, L_B \leq T\}$ *and* $\Omega_i = \{(L_A, L_B) \in \Sigma : p_i f_j(L_j) \leq p_j f_i(L_i)\}$ *with* $j \neq i$. *In the natural equilibrium, player* i*'s payoff in each period is*

$$V_i^N = \Pi_i(L_i, L_j) = (1 - p_j) f_i(L_i^N) + p_i f_j(L_j^N) + v_i(T - L_i^N).$$

Notice (from equation 4) that L_i^N does not depend on the probability p_i with which player i steals player j's output. This is because that probability has no affect on his marginal benefit (or marginal cost) from working, although it will affect his NE payoff. On the other hand, L_i^N is influenced by the probability p_j with which player j steals player i's output. For example, an increase in p_j — by decreasing player i's marginal benefit from working — decreases L_i^N. Interestingly, however, for some parameter values the NE satisfies equation 2, which means that no fighting takes place in equilibrium.[15] However, even then the NE working levels are influenced by the players' fighting skills. The intuition behind this is that even when no fighting occurs in the NE, player i's ($i = A, B$) equilibrium marginal benefit from working is equal to the left-hand side of equation 4. This is because if player i unilaterally deviates and chooses $L_i \neq L_i^N$, then a fight will occur — since such a deviation will imply that equation 2 is no longer satisfied. The following (Hobbesian) interpretation of this result is instructive. The absence of property rights over one's output means that the *fear* of war is not absent, notwithstanding that in the NE no war takes place; it is the *fear* of war that ultimately determines each player's incentives to work.

2.3 Fighting Skills and Natural Equilibrium Payoffs

I now explore the impact of the players' fighting skills on each player's NE payoff. First, I derive the effect of a marginal change in p_j on V_i^N ($j \neq i$). It is trivial to verify (by, for example, using the Envelope Theorem) that $\partial V_i^N / \partial p_j = -f_i(L_i^N)$. Hence, a marginal increase in player j's fighting skill decreases player i's NE payoff, and *vice-versa*. This is a fairly intuitive result.

I now consider the impact of a marginal change in p_i on player i's NE payoff. It will

[15] For example, equation 2 is satisfied when the players are identical: they have identical preferences, identical productive skills and identical fighting skills — since it would then follow that $L_A^N = L_B^N$. Equation 2 may also be satisfied when the players are not identical.

be shown that it is not monotonic in p_i. First, it is straightforward to verify that

$$\frac{\partial V_i^N}{\partial p_i} = f_j(L_j^N) + p_i f_j'(L_j^N)\frac{\partial L_j^N}{\partial p_i}. \tag{5}$$

A marginal change in p_i has two opposing effects on player i's NE payoff. The first term on the right-hand side of (5), which is strictly positive, may be called the *direct effect* of a marginal change in p_i; it results from the fact that a marginal increase (for example) in player i's fighting skill gives him (in expected terms) more of player j's output. On the other hand, the second term on the right-hand side of (5), which is strictly negative, may be called the *strategic effect* (or *indirect effect*) of a marginal change in p_i; it results from the fact that a marginal increase in player i's fighting skill decreases player j's incentive to work, which, in turn, decreases the quantity of output that player i can potentially steal. In Proposition 2 below, I establish that if p_i is sufficiently large, then a marginal increase in p_i decreases V_i^N; otherwise, the opposite holds. The intuition for this result comes from noting that if player i is sufficiently strong, then the strategic effect dominates the direct effect. This is most transparent in the following, extreme cases: if p_i is close to one then (since L_j^N is close to zero) the direct effect is close to zero, while if p_i is close to zero then the strategic effect is close to zero. More precisely, I establish the following result:[16]

Proposition 2 (Fighting Skills and Natural Equilibrium Payoffs). *Fix $i, j = A, B$ with $i \neq j$.*
(a) Player i's NE payoff V_i^N is strictly decreasing in p_j.
(b) If $f_j''' \leq 0$ and $v_j''' \leq 0$, then there exists $p_i^* \in (0,1)$ — where p_i^* is independent of p_j — such that

$$\frac{\partial V_i^N}{\partial p_i} \gtreqless 0 \quad \text{if} \quad p_i \lesseqgtr p_i^*.$$

Proof. While part (a) has been established above, the formal proof of part (b) is in the Appendix. □

The results stated in Proposition 2 imply that if some player's fighting skill improves, then his opponent's NE payoff decreases, while his own NE payoff decreases or increases depending on whether he is strong or weak. Since it seems intuitive that each player would have a relatively greater incentive to establish the property rights the

[16]For future reference, I also state in this proposition the result established above concerning the effect of p_j on V_i^N.

smaller is his NE payoff, these results suggest that military strength may enhance the likelihood of the emergence of secure property rights. This suggestion — which I fully and properly explore below in subsection 3.4 — is similar to the *Mutual Assured Destruction* doctrine, which states that the prospects for peace are relatively better when nation-states are militarily strong than militarily weak.

Another insight contained in Proposition 2 is that if each player could choose the level of his fighting skill (through appropriate investments), then he would choose to be strong, but not too strong. The intuition for this is that he would want to be strong enough so as to be somewhat successful in fighting, but not too successful; for otherwise his opponent would have no incentive to produce any output. Thus, ignoring the direct costs of investment in fighting skills, the level of such investment must balance, at the margin, the benefit obtained from increasing the likelihood of being able to win a fight against the indirect cost from inducing one's opponent to produce less output.

2.4 Productive Skills and Natural Equilibrium Payoffs

I now explore the impact of the players' productive skills on each player's NE payoff. Suppose that player i becomes more productive, which I formalize as follows. Player i's new production function is $\widehat{f}_i$, where for any $L_i > 0$, $\widehat{f}_i(L_i) > f_i(L_i)$ and $\widehat{f}'_i(L_i) > f'_i(L_i)$. Thus, not only is his total output higher for any level of labour input, but also his marginal product is higher. An example is when $f_i(L_i) = \lambda_i L_i$ and $\widehat{f}_i(L_i) = \widehat{\lambda}_i L_i$, where $\widehat{\lambda}_i > \lambda_i > 0$.

It is straightforward to show (using (4)) that player i would increase the amount of time spent working: that is, $\widehat{L}_i^N > L_i^N$. However, notice that player j would not change the amount of time that he spends working: that is, $\widehat{L}_j^N = L_j^N$. This is because player i's productive skills do not affect player j's marginal benefit or marginal cost from working. Of course, it does affect his NE payoff. Indeed, it is straightforward to verify that both players' NE payoffs increase as player i becomes more productive.

Since, as I mentioned above, it seems intuitive that each player would have a relatively greater incentive to establish the property rights the smaller is his NE payoff, these results suggest that an improvement in a player's productive skills may diminish the likelihood of the emergence of secure property rights. This suggestion — which I fully and properly explore below in subsection 3.5 — is not necessarily correct, since it fails to take into account the positive effect of an increase in player i's productive skill on his own payoff when secure property rights do exist.

3 The Costs and Benefits of Basic Property Rights

I now study the players' *private* incentives to establish the property rights under consideration. Such private incentives define the (private) costs and benefits to the players from establishing the property rights. I shall be particularly concerned with the roles of the players' fighting and productive skills on their respective private incentives.

3.1 The Property Rights Equilibrium

In order to derive a player's private incentive to establish the property rights, I compare his NE payoff with his payoff in the *property rights equilibrium*, where the latter denotes the unique SPE of the base model *on the assumption* that secure property rights exist; that is, on the assumption that the players are (irrevocably) committed not to fight at stage 2 in any period. In the unique property rights equilibrium, player i chooses to work for L_i^F units of time, where L_i^F is the unique solution of the following maximisation problem:

$$\max_{0 \leq L_i \leq T} f_i(L_i) + v_i(T - L_i).$$

Given Assumption 1, it follows that L_i^F is an interior solution; it is the unique solution of the following (first-order) condition:

$$f_i'(L_i) = v_i'(T - L_i). \tag{6}$$

For future reference, I state this result in the following lemma:[17]

Lemma 2 (The Unique Property Rights Equilibrium (PRE)). *In the unique property rights equilibrium (PRE, for short), player i ($i = A, B$) works for L_i^F units of time, where L_i^F is the unique solution of equation 6. In the property rights equilibrium, player i's payoff in each period is*

$$V_i^F = f_i(L_i^F) + v_i(T - L_i^F).$$

It immediately follows from Proposition 1 and Lemma 2 that $L_i^N < L_i^F$ ($i = A, B$).

[17] It may be useful to note that the PRE can be stated as the following pair of stationary strategies: for each i ($i = A, B$), $L_i = L_i^F$ and $\phi_i^F(.) = nf$ for any pair $(L_A, L_B) \in \Sigma$. Of course, Proposition 1 implies that the PRE is not an SPE of the base model.

This "under-investment" result may be fruitfully interpreted as arising from a "hold-up" problem: in the absence of secure property rights, player i does not receive the full *marginal* return from his work, and hence, he does not work at his first-best level. Indeed, this interpretation is instructive as it draws attention to the close connection between insecure property rights and hold-up problems. An important reason, for example, for relatively little productive investment in the poorer parts of the world is that the absence of secure property rights leads to hold-up problems, which, in turn, adversely affects *ex-ante* incentives to invest.

3.2 Private Incentives

I now turn to the issue of the players' costs and benefits from establishing secure property rights. Define

$$\Delta_i \equiv V_i^F - V_i^N.$$

If $\Delta_A \geq 0$ and $\Delta_B \geq 0$, then there is no cost from having these property rights, only benefits. However, if for some i and j (where $(i, j) = (A, B)$ or $(i, j) = (B, A)$) $\Delta_i > 0$ and $\Delta_j < 0$, then there is a cost *and* a benefit from having these property rights; the cost is to player j and the benefit to player i. It should be noted that both Δ_A and Δ_B cannot be negative, since — it is trivial to verify that — $\Delta_A + \Delta_B \geq 0$.

A key insight obtained below is that the set of parameter values such that for some i ($i = A$ or $i = B$) $\Delta_i < 0$ is *non-empty.* This insight challenges the not uncommon viewpoint that the Prisoners' Dilemma game captures the (per-period) strategic interaction in the state-of-nature; that is, the viewpoint that each and every player is strictly better-off when property rights are secure. My insight is somewhat transparent in the case, for example, when one player is fairly strong but fairly unproductive, while the opposite is the case with regard to the other player. In that case it is intuitive that the strong player would loose out when property rights are secure (since he would then not be able to steal any output from the other, more productive player).[18] It may thus be noted that an important advantage of my model of the state-of-nature over the Prisoners' Dilemma game based model is that it allows one to study the implications of *heterogeneity* in the players' fighting and productive skills on the emergence of secure property rights.

[18] As I discuss in section 5, since the *sum* of the payoffs in the PRE exceeds the *sum* of the payoffs in the NE, secure property rights might be established if, for example, the more productive player transfers some output (in each period) to the less productive, but stronger player.

It is straightforward to show that

$$\Delta_i \gtreqless 0 \Longleftrightarrow \left[V_i^F - \left[f_i(L_i^N) + v_i(T - L_i^N) \right] \right] \gtreqless \left[p_i f_j(L_j^N) - p_j f_i(L_i^N) \right]. \qquad (7)$$

By definition, the left-hand side of the second inequality in (7) is greater than zero. Hence, if the NE satisfies equation 2 — that is, if there is no fighting in the NE — then the right-hand side of the second inequality in (7) equals zero. This implies the following lemma:

Lemma 3 (No fighting in the NE and Private Incentives). *If the natural equilibrium satisfies equation 2 — that is,* $\phi_i^N(L_A^N, L_B^N) = nf$ *for* $i = A, B$ *— then* $\Delta_A > 0$ *and* $\Delta_B > 0$.

The result contained in this lemma implies that if in the NE there is no fighting, then both players have an incentive to establish the property rights under consideration. The following corollary is an immediate consequence of Lemma 3:

Corollary 1 (Identical Players). *If both players are identical — that is, they have identical preferences* ($v_A(l) = v_B(l)$ *for all* $l \in [0, T]$)*, identical productive skills* ($f_A(L) = f_B(L)$ *for all* $L \in [0, T]$) *and identical fighting skills* ($p_A = p_B$)*, then* $\Delta_A > 0$ *and* $\Delta_B > 0$.

Proof. It is straightforward to verify that if the players are identical, then in the NE no fighting occurs. The corollary then follows from Lemma 3. ☐

An immediate implication of Corollary 1 is that in order for there to exist conflict in the players' private incentives to establish the basic property rights (i.e., in order for $\Delta_i < 0$ for some i), the players have to be different in some respects (such as in their productive and/or fighting skills). That is:

- *Heterogeneity in the players' fighting skills and/or productive skills is necessary for it to be the case that* $\Delta_i < 0$ *for some* i ($i = A$ or $i = B$).

It is trivial to verify that for any parameter values such that $\Delta_i < 0$, it must be the case that $p_i f_j(L_j^N) > p_j f_i(L_i^N)$ — that is, in the NE player i's WDA is to fight and player j's WDA is not to fight. This makes intuitive sense. The following corollary implies, in particular, that the set of parameter values under which the players' private incentives to establish property rights are in conflict is *non-empty.*

Corollary 2 (Conflicting Private Incentives). *Fix $i, j = A, B$ with $i \neq j$. If p_j is arbitrarily close to zero and $p_i > 0$ but bounded away from one, then $\Delta_i < 0$.*

Proof. If p_j is arbitrarily close to zero, then the difference $L_i^F - L_i^N$ is arbitrarily close to zero, which, in turn, implies that the left-hand side of (7) is arbitrarily close to zero. Hence $\Delta_i < 0$ provided that $p_i > 0$ but bounded away from one — where the latter condition ensures that $L_j^N > 0$. □

The intuition behind Corollary 2 is as follows. If p_j is arbitrarily close to zero, then it is "as if" player i has property rights over his output, while $p_i > 0$ means that player j does not. Furthermore, since p_i is bounded away from one, player j produces some output. Hence, player i strictly prefers the NE over the PRE.

3.3 An Example: Identical Preferences and Linear Productive Skills

In order to develop some further insights concerning the role of the parameters on the players' private incentives to establish the basic property rights, in this subsection I study the case in which the players have identical preferences and linear productive skills. In particular, assume that for each $i = A, B$, $v_i(l) = \sqrt{l}$. Furthermore, assume that each player's productive skills are "linear", in the sense that for each $i = A, B$, there exists $\lambda_i > 0$ such that $f_i(L_i) = \lambda_i L_i$. This implies that player i's marginal productivity is constant and equals λ_i. The condition

$$(1 - p_j)\lambda_i > \frac{1}{2\sqrt{T}}, \tag{8}$$

which we assume is satisfied by the parameters, ensures that Assumption 1 is satisfied. For example, this condition is satisfied for T sufficiently large. It follows from Proposition 1 that in the NE, $L_i^N = T - [1/4\lambda_i^2(1 - p_j)^2]$, and from Lemma 2 that in the PRE, $L_i^F = T - (1/4\lambda_i^2)$. Applying (7), I obtain (after some simplification) that

$$\Delta_i \gtreqless 0 \Longleftrightarrow \lambda_i p_j - \lambda_j p_i \gtreqless \frac{1}{4T}\left[\frac{p_j}{\lambda_i(1 - p_j)} - \frac{p_i}{\lambda_j(1 - p_i)^2}\right].$$

It follows immediately that if $\lambda_i p_j > \lambda_j p_i$ (respectively, $\lambda_i p_j < \lambda_j p_i$) and T is sufficiently large, then $\Delta_i > 0$ (resp., $\Delta_i < 0$). The following claim is an immediate consequence of this observation.

Claim 1 (Heterogeneity and Conflict). *Fix any p_A, p_B, λ_A and λ_B such that (8) holds for both $(i, j) = (A, B)$ and $(i, j) = (B, A)$.*
(i) If $\lambda_B p_A > \lambda_A p_B$ then there exists a $\overline{T}$ such that for any $T > \overline{T}$, $\Delta_A < 0$ and $\Delta_B > 0$.
(ii) If $\lambda_B p_A < \lambda_A p_B$ then there exists a $\overline{T}$ such that for any $T > \overline{T}$, $\Delta_A > 0$ and $\Delta_B < 0$.

The expression $\lambda_i p_j$ is the output per unit of labour input that player i expects to loose from a fight. It may be interpreted as the (constant) unit cost of fighting to player i. The result contained in the above claim states that if the players' unit costs of fighting differ (i.e., $\lambda_B p_A \neq \lambda_A p_B$), then (provided T is sufficiently large) at least one player prefers the NE over the PRE — that is, the players' private incentives to establish the property rights are in conflict.

Assume that T is sufficiently large so that Claim 1 is applicable. Suppose that player i is more productive than player j (i.e., $\lambda_i > \lambda_j$), and that player i is weaker than player j (i.e., $p_i < p_j$). It follows from Claim 1 that player i prefers the PRE over the NE, while player j prefers the NE over the PRE. Thus, when one player is stronger and the other more productive, the stronger player prefers the NE while the more productive prefers the PRE. Now suppose that player i is also the more stronger of the two players (i.e., $\lambda_i > \lambda_j$ and $p_i > p_j$). It follows from Claim 1 that if $p_i/p_j > \lambda_i/\lambda_j$ then player i prefers the NE over the PRE, but if $p_i/p_j < \lambda_i/\lambda_j$ then he prefers the PRE over the NE. Thus, when player i has *absolute* advantage over player j in both production and fighting, his preference between the PRE and the NE depends on his *comparative* advantage. If his comparative advantage lies in fighting, then he prefers the NE over the PRE. But if his comparative advantage lies in production, then he prefers the PRE over the NE. I summarise these insights in the following claim.

Claim 2 (Comparative Advantage and Conflict). *Assume that T is sufficiently large so that Claim 1 is applicable. Furthermore, fix $i, j = A, B$ with $i \neq j$.*
(i) If $\lambda_i > \lambda_j$ and $p_j > p_i$, then $\Delta_i > 0$ and $\Delta_j < 0$.
(ii) If $\lambda_i > \lambda_j$, $p_i > p_j$ and $p_i/p_j > \lambda_i/\lambda_j$, then $\Delta_i < 0$ and $\Delta_j > 0$.
(iii) If $\lambda_i > \lambda_j$, $p_i > p_j$ and $p_i/p_j < \lambda_i/\lambda_j$, then $\Delta_i > 0$ and $\Delta_j < 0$.

3.4 The Effects of Fighting Skills on Incentives

Since fighting skills have no effect on each player's PRE payoff, it follows immediately from Proposition 2 that an improvement in player i's fighting skill enhances *both* players' private incentives to establish the property rights *provided* that player i is sufficiently strong.

An interesting question — which I now address — is whether or not *both* players' private incentives to establish the property rights are enhanced if the fighting skills of *both* players were to improve? Fix, therefore, an arbitrary configuration of the players' fighting and productive skills, and suppose that *both* players' fighting skills improve by an *identical, small* amount. In that case, the (total) change in player i's private incentives is formally captured by the sum of the partial derivative of Δ_i *w.r.t.* p_i and the partial derivative of Δ_i *w.r.t.* p_j. Given the expressions for the derivatives of V_i^N *w.r.t.* to p_i and p_j stated above in subsection 2.3, it follows (since the derivatives of V_i^F *w.r.t.* to p_i and p_j are both zero) that for each $i, j = A, B$ with $i \neq j$,

$$\frac{\partial \Delta_i}{\partial p_i} + \frac{\partial \Delta_i}{\partial p_j} = f_i(L_i^N) - f_j(L_j^N) - p_i f_j'(L_j^N) \frac{\partial L_j^N}{\partial p_i}.$$

In general, this expression can be positive or negative. However, if the players are identical, then this expression is strictly positive.[19] This means that if the players are identical, then a small (and identical) improvement in both players' fighting skills enhances their respective private incentives to establish the property rights.

A key implication of this result is that in a state-of-nature with identical players, the players' private incentives to establish the property rights — and thus, not engage in war — are maximised when each player is militarily as strong as he can be (i.e., their common level of fighting skill is set equal to 1/2). As indicated above in subsection 2.3, this conclusion can be interpreted as a formal statement of the *Mutual Assured Destruction* doctrine.

Another interesting question that is especially relevant when (in section 5) I extend the base model by allowing the players to communicate with each other and negotiate over inter-player transfers of output is whether or not the players' *collective* incentive is enhanced when one or both players become more strong? The "collective" incentive is formally captured by the sum $\Delta_A + \Delta_B$; it defines the *aggregate net surplus* from having property rights. It is straightforward to show that for any p_A, p_B, f_A and f_B,

[19] This is because if the players are identical (in the obvious sense as, for example, specified above in Corollary 1), then (since $L_i^N = L_j^N$ and $f_i \equiv f_j$) the first two terms cancel out.

and for each $i = A, B$,

$$\frac{\partial(\Delta_A + \Delta_B)}{\partial p_i} = -p_i f_j'(L_j^N)\frac{\partial L_j^N}{\partial p_i} \quad (j \neq i),$$

which is strictly positive. This result makes intuitive sense, since the net effect of a marginal increase in p_i is that in the NE, player j produces less output (the strategic effect), which decreases both players' NE payoffs, and hence, enhances the collective incentive. It immediately follows that (whether or not the players are identical) the collective incentive to establish the property rights are increasing in each player's fighting skill. This result, which is related to the *Mutual Assured Destruction* doctrine, implies that the collective incentive is maximised *only if* $p_A + p_B = 1$ (i.e., war should have a victor).

It is instructive to pursue this issue a bit further by determining the values of p_A and p_B that maximise $\Delta_A + \Delta_B$. Given the results established above, it follows that the (first-order) conditions that determine these values are:[20]

$$p_A f_B'(L_B^N)\frac{\partial L_B^N}{\partial p_A} = p_B f_A'(L_A^N)\frac{\partial L_A^N}{\partial p_B} \tag{9}$$

$$p_A + p_B = 1. \tag{10}$$

It is easy to derive closed-form solutions to these equations under the assumptions of the example studied in the previous subsection, namely, assuming "linear" productive skills ($f_i(L_i) = \lambda_i L_i$) and identical preferences ($v_i(l) = \sqrt{l}$). It is straightforward to show that the unique solution in that case is:

$$p_A^* = \frac{\gamma}{1+\gamma} \quad \text{and} \quad p_B^* = \frac{1}{1+\gamma}, \quad \text{where} \quad \gamma = \left[\frac{\lambda_B}{\lambda_A}\right]^{1/4}.$$

Thus, perhaps not surprisingly, if the players have identical productive skills, then the collective incentive is maximised when the players have identical and maximal fighting skills. However, if player i is relatively more productive than player j, then the collective incentive is maximised when player j has relatively better fighting skills. Notice that the ratio p_A^*/p_B^* is strictly increasing in the ratio λ_B/λ_A. As player B's productive skill relative to that of player A's traverses from zero to infinity, player A's "collective-incentive maximising" level of fighting skill relative to that of player B's traverses *monotonically* from zero to one. The intuition behind these results follows

[20]Intuitively, condition 9 equates the marginal change on $\Delta_A + \Delta_B$ through p_A to the marginal change on $\Delta_A + \Delta_B$ through p_B, while condition 10 has been established above.

by noting that the levels of the fighting skills which maximise the collective incentive are the same which minimize the *sum* of the players' NE payoffs.

3.5 The Effects of Productive Skills on Incentives

Now suppose that player i's productive skill improves (in the manner formalized above in subsection 2.4). It is straightforward to show that the increase in his PRE payoff is larger than the increase in his NE payoff. As such, as he becomes more productive, his private incentive to establish the property rights enhances; and this makes intuitive sense. However, since player i's productive skills have no effect on player j's PRE payoff it follows from the results established above in subsection 2.4 that player j's private incentive to establish the property rights diminishes, which also makes much intuitive sense.

Given the above conclusion, one can ask whether or not *both* players' private incentives to establish the property rights are enhanced if *both* players were to become more productive? I address this question in the context of the example studied above in subsection 3.3 (with linear productive skills and identical preferences), by computing the sum of the partial derivative of Δ_i *w.r.t.* λ_i and the partial derivative of Δ_i *w.r.t.* λ_j. Straightforward computations reveal that for each $i, j = A, B$ with $i \neq j$,

$$\frac{\partial \Delta_i}{\partial \lambda_i} + \frac{\partial \Delta_i}{\partial \lambda_j} = p_j \left[T + \frac{1}{4\lambda_i^2(1-p_j)} \right] - p_i \left[T + \frac{1}{4\lambda_j^2(1-p_i)^2} \right].$$

In general, this expression can be positive of negative. However, if the players are identical, then this expression is strictly negative. This means that if the players are identical, then a small (and identical) improvement in both players' productive skills diminishes their respective private incentives to establish the property rights.

If, as seems plausible, low productive skills are interpreted as signs of poverty while high productive skills as signs of economic prosperity, then this result suggests that in a state-of-nature with identical players, the private incentives of the players to establish the property rights are relatively greater when they are both poor than when they are both economically prosperous.

It is straightforward to show that whether or not the players are identical, their *collective* incentive to establish the property rights is diminished as one (or both) players become more productive. This conclusion is an immediate consequence of the result that for each $i, j = A, B$ with $i \neq j$, the partial derivative of the sum $\Delta_i + \Delta_j$ *w.r.t* λ_i is strictly negative.

In a nutshell, then, the results established here suggest that *economic prosperity adversely affects the players' collective incentive and their respective private incentives to establish the property rights.* All of these results concerning the effects of improvements in productive skills on private and collective incentives indicate that improvements in productive skills increase the cost of establishing the property rights *more* than they increase the benefit of having such rights — where the cost comes from the fact that a player can no longer steal his opponent's output, and the benefit from increased levels of outputs.

Remark 1 (A fundamental Insight concerning Incentives). It is instructive to bring together the main insight obtained in this subsection with that obtained in the previous subsection: namely, while economic prosperity adversly affects incentives to establish property rights, military strength affects them positively. As such, in order to promote and maintain such incentives, it would seem that improvements in the players' productive skills (or economic properity) should go hand-in-hand with improvements in their fighting skills (or military technologies). This fundamental insight appears, at least by casual observation, to be borne out by the historical path of human development. Furthermore, as I will explore further in later sections of this paper, it is an insight that bears upon the issue of the emergence of secure property rights in today's world.

4 On the Emergence of Secure Property Rights

I now turn attention to the main issue of concern in this paper, namely, the issue of the emergence of secure property rights. Formally, this issue is addressed by studying the conditions under which there exists a (necessarily, *non-stationary*) SPE whose equilibrium path is identical to the PRE path — namely, in each period, $L_i = L_i^F$ $(i = A, B)$ and no fight occurs. Since the PRE is *not* an SPE of the base model, the PRE path can potentially be sustained as a SPE path only by the threat of appropriate (and credible) punishment should any player unilaterally deviate from the PRE path.[21] It is worth emphasizing that property rights are therefore made secure (when they can be) in a *self-enforcing* manner, without any third-party enforcement. Indeed, since there are no third parties in the world under consideration, any enforcement mechanism can only involve the two players.

[21] This is, of course, a familiar idea from the *Theory of Repeated Games*.

Remark 2 (Issues of Interpretation). In subsections 4.1 and 4.3 below, I study the issue of the existence of two different non-stationary SPE whose equilibrium paths coincide with the PRE path. The analysis focuses on the study of the appropriate incentive-compatibility conditions that are required to be satisfied for these equilibria to exist. In particular, I will characterize the configurations of the players' fighting and productive skills under which these conditions are and are not satisfied. Since there exists many other SPE in the base model — in particular, the NE is always an SPE — how should one interpret the existence of the two non-stationary SPE described below? A most appropriate interpretation is that the existence of such equilibria implies that it is *possible* for secure property rights to emerge in the state-of-nature. Given the multiplicity of the SPE, I cannot (and do not) make the claim that secure property rights are bound to emerge; for after all, the players may be stuck in the NE. However, if the parameters (including the players' fighting and productive skills, and their respective discount factors) are such that the existence of one of these SPE is assured, then the players could implicitly (if not explicitly) establish secure property rights by moving play from the NE to this non-stationary SPE. In subsection 4.2 below, on the other hand, I characterize a set of parameter values under which there cannot exist *any* SPE whose equilibrium path is the PRE path. The natural interpretation of this result is that for such a set of parameter values, secure property rights can *never* emerge in the state-of-nature; here I can (and do) make the claim that secure property rights are bound *not* to emerge when the parameters belong to this set.

4.1 The Trigger-Strategy Equilibrium

I begin the analysis, in this section, by using the relatively simple, and fairly well-known "trigger-strategy" approach. The idea is to sustain the PRE path as a SPE path by moving play to the NE if any player ever (unilaterally) deviates from the PRE path. That is, a deviation from the PRE path triggers reversion to the NE. Since this punishment path is a SPE path, it is relatively easy to derive the appropriate incentive-compatibility conditions under which these trigger strategies constitute an SPE.[22]

[22]More precisely, the *trigger* strategy for player i ($i = A, B$) is as follows. The amount of work L_i^1 that he chooses in period 1 equals L_i^F. In period n ($n = 2, 3, \dots$), $L_i^n = L_i^F$ *if* in each preceding period (i.e., in periods $1, 2, \dots, n-1$) players A and B respectively worked for L_A^F and L_B^F units of time, and no fight occurred. Otherwise $L_i^n = L_i^N$. Furthermore, in period n ($n = 1, 2, \dots$), $\phi_i^n = nf$ *if* in each preceding period and in this period (i.e., in periods $1, 2, \dots, n$) players A and B respectively worked for L_A^F and L_B^F units of time, and in each preceding period (i.e., in periods $1, 2, \dots, n-1$) no fight occurred. Otherwise $\phi_i^n \equiv \phi_i^N$.

The following argument establishes the condition under which player i ($i = A, B$) cannot benefit from a (*one-shot*, unilateral) deviation from the PRE path.[23]

Fix an arbitrary period n (where $n = 1, 2, \dots$), and suppose that up until the end of period $n-1$ neither of the two players deviated from the PRE path. Player i considers (at stage 1 of this period) the net benefit from a (one-shot, unilateral) deviation in which he chooses $L_i^n \neq L_i^F$ (and conforms to the trigger strategy thereafter). His payoff from not conducting such a deviation (and thus conforming to the trigger strategy) is, of course, $V_i^F/(1-\delta_i)$. On the other hand, his payoff from the (one-shot) deviation of setting $L_i^n = L_i$, where $L_i \neq L_i^F$, equals $\Pi_i(L_i, L_j^F) + \delta_i V_i^N/(1-\delta_i)$. Since the maximum value of this payoff (across all possible values of $L_i \neq L_i^F$) equals $\Pi_i(L_i^N, L_j^F) + \delta_i V_i^N/(1-\delta_i)$, it follows that player i cannot benefit from a (one-shot) deviation to *any* $L_i^n \neq L_i^F$ if and only if[24]

$$\delta_i \Delta_i \geq (1-\delta_i)[\Pi_i(L_i^N, L_j^F) - V_i^F]. \tag{11}$$

Now suppose that player i conforms at stage 1 of period n (by setting $L_i^n = L_i^F$), but considers whether or not to conduct a (one-shot) deviation at stage 2. It is trivial to verify that he will conform to his trigger strategy (and not fight) if and only if $\delta_i \Delta_i \geq (1-\delta_i)[\Pi_i(L_i^F, L_j^F) - V_i^F]$.[25] Since $\Pi_i(L_i^N, L_i^F) > \Pi_i(L_i^F, L_i^F)$, it follows that this inequality is implied by (11).[26] I have thus established the following proposition.[27]

[23] From the *One-Shot Deviation Property* it thus follows that he cannot benefit from deviating to *any* strategy — that may involve more than just a *one-shot* deviation. The *One-Shot Deviation Property*, which is also known by other terms, is essentially the principle of optimality for discounted dynamic programming. A pair of strategies is a SPE if and only if each player's strategy is immune to profitable *one-shot* (unilateral) deviations. For a precise statement of this result, see, for example, Abreu (1988, Proposition 1), Fudenberg and Tirole (1991, Theorem 4.2), and Osborne and Rubinstein (1994, Exercise 123.1).

[24] The left-hand side of inequality 11 is player i's (long-run) average cost from the (optimal, one-shot) deviation; this is because from *next* period onwards, his *per-period* loss is $V_i^F - V_i^N$ (i.e., Δ_i). Note that, as is standard, a player's *average* payoff is $1-\delta_i$ times his present discounted value. Furthermore, the right-hand side of inequality 11 is player i's (short-run) average benefit from the (optimal, one-shot) deviation; this is because his (one-period) gain from this deviation is $\Pi_i(L_i^N, L_j^F) - V_i^F$.

[25] If the pair (L_A^F, L_B^F) satisfies equation 2, then $\Pi_i(L_i^F, L_j^F) = V_i^F$. But if this pair does not satisfy (2) then we have the following result. If $p_i f_j(L_j^F) > p_j f_i(L_i^F)$ (where $i, j = A, B$ with $i \neq j$), then $\Pi_i(L_i^F, L_j^F) > V_i^F$ and $\Pi_j(L_i^F, L_j^F) < V_j^F$.

[26] This conclusion makes sense, since (by being able to adjust L_i) a one-shot deviation at stage 1 is (in principle) relatively more beneficial than a one-shot deviation at stage 2.

[27] It may be noted that since $\Pi_i(L_i^N, L_j^F) = V_i^N + p_i[f_j(L_j^F) - f_j(L_j^N)]$, inequality 11 becomes (after substituting for $\Pi_i(L_i^N, L_j^F)$, using this expression, and then simplifying) $\Delta_i \geq p_i(1-\delta_i)[f_j(L_j^F) - f_j(L_j^N)]$.

Proposition 3 (The Trigger-Strategy Equilibrium (TSE)). *The pair of trigger strategies described above is a SPE if and only if the following two inequalities hold:*

$$\Delta_A \geq p_A(1-\delta_A)[f_B(L_B^F) - f_B(L_B^N)] \tag{12}$$

$$\Delta_B \geq p_B(1-\delta_B)[f_A(L_A^F) - f_A(L_A^N)]. \tag{13}$$

This SPE will be called the trigger-strategy equilibrium (TSE, for short). Since the TSE path is the PRE path, in equilibrium no fight ever occurs, and in each period player i ($i = A, B$) works for L_i^F units of time. Furthermore, of course, player i's TSE payoff in each period equals V_i^F.

I now explore some of the implications of this proposition for the emergence of secure (or self-enforcing) property rights. Since the right-hand side of both (12) and (13) are strictly positive, it immediately follows, not surprisingly, that if the players' private incentives to establish these rights are in conflict — that is, $\Delta_i < 0$ for some i ($i = A$ or $i = B$), then the PRE path cannot be sustained as a SPE using the trigger strategies. For future reference, I state this result in the following corollary:

Corollary 3 (Conflicting Private Incentives and Non-Existence of the TSE). *If the parameters are such that the players' private incentives to establish the basic property rights are in conflict (i.e., $\Delta_i < 0$ for some i), then the TSE does not exist.*

Indeed, the TSE exists *only if* the parameters are such that both players prefer the PRE over the NE — that is, the parameters are such that $\Delta_A \geq 0$ and $\Delta_B \geq 0$. Perhaps not surprisingly, the incentive-compatibility conditions ((12) and (13)) which ensure that the property rights are self-enforcing are much more severe than the conditions ($\Delta_A \geq 0$ and $\Delta_B \geq 0$) which ensure that the players would like to have these rights established.

It is useful to rewrite (12) and (13) respectively as follows:

$$\delta_A \geq \underline{\delta}_A \tag{14}$$

$$\delta_B \geq \underline{\delta}_B, \tag{15}$$

$$\text{where } \underline{\delta}_i = 1 - \frac{\Delta_i}{p_i[f_j(L_j^F) - f_j(L_j^N)]}, \quad (j \neq i). \tag{16}$$

Thus, the TSE exists if and only if player i's ($i = A, B$) discount factor lies above the critical value stated in (16), which depends upon the parameters. It is trivial to note that $\underline{\delta}_i < 1$ if and only if $\Delta_i > 0$. Furthermore, $\underline{\delta}_i > 0$ if and only if $p_i[f_j(L_j^F) - f_j(L_j^N)] > \Delta_i$, which, in turn, is if and only if $\Pi_i(L_i^N, L_j^F) - V_i^F > 0$. It follows immediately from the results stated in footnote 25 that if the pair (L_A^F, L_B^F) satisfies equation 2 then $\underline{\delta}_i > 0$ for both $i = A$ and $i = B$; but if $p_i f_j(L_j^F) > p_j f_i(L_i^F)$ (where $i, j = A, B$ with $i \neq j$) then $\underline{\delta}_i > 0$ while $\underline{\delta}_j$ could be positive or negative. In conclusion, then, I have established the following result.

Corollary 4 (Existence of the TSE). *For any parameter values such that $\Delta_A > 0$ and $\Delta_B > 0$ the TSE exists provided that, in addition, the parameters are such that player i's ($i = A, B$) discount factor $\delta_i \geq \underline{\delta}_i$, where $\underline{\delta}_i$ (defined above in 16) is strictly less than one. Furthermore, for at least one i ($i = A$ or $i = B$) — if not for both i ($i = A$ and $i = B$) — $\underline{\delta}_i > 0$.*

I now examine the impact of the players' fighting skills on the critical discount factors $\underline{\delta}_A$ and $\underline{\delta}_B$. It is straightforward to verify that for each $i, j = A, B$ with $j \neq i$,

$$\frac{\partial \underline{\delta}_i}{\partial p_j} < 0.$$

Thus, a marginal increase (for example) in player j's fighting skill makes player i more likely to "cooperate" (in the sense of respecting the basic property rights). This result makes intuitive sense: a marginal increase in p_j reduces V_i^N (cf. Proposition 2(a)), which, in turn, makes deviation from the PRE path less attractive to player i. As I now show, a marginal change in p_i, on the other hand, affects player i's willingness to cooperate in a not-so-simple manner. In the Appendix, I establish that for each $i = A, B$,

$$\frac{\partial \underline{\delta}_i}{\partial p_i} \gtreqless 0 \iff (1 - \underline{\delta}_i) f_j(L_j^F) + \underline{\delta}_i \frac{\partial V_i^N}{\partial p_i} \gtreqless 0 \quad \text{where } j \neq i. \tag{17}$$

Proposition 2(b) above states that player i's NE payoff V_i^N is not monotonic in p_i, and hence, it immediately follows (from (17)) that $\underline{\delta}_i$ is not monotonic in p_i either. However, using Proposition 2(b) it is straightforward to establish the following result:[28]

[28] For future reference, I also state in this corollary the result established above concerning the effect of p_j on $\underline{\delta}_i$.

Corollary 5 (Fighting Skills and TSE). *Fix $i, j = A, B$ with $i \neq j$.*
(a) Player i's critical discount factor $\underline{\delta}_i$ is strictly decreasing in p_j.
(b) If $f_i''' \leq 0$ and $v_i''' \leq 0$, then there exists $\widehat{p}_i \in (0,1)$ — where $\widehat{p}_i$ depends on p_j — such that

$$\frac{\partial \underline{\delta}_i}{\partial p_i} \gtreqqless 0 \quad \text{if} \quad p_i \lesseqqgtr \widehat{p}_i.$$

Proof. While Corollary 5(a) has been established above, Corollary 5(b) can be derived by using Proposition 2(b) and (17).[29] □

These (comparative-static) results may be interpreted as follows. As player i's fighting skill increases (for example), his opponent becomes more likely to cooperate, while he himself becomes less or more likely to cooperate depending on whether or not he is sufficiently weak. These results have a number of implications, which, in general terms, may be put as follows:

• If player i is sufficiently strong and secure property rights do not exist, then a marginal increase in player i's strength may create the conditions for the emergence of secure property rights. But, if player i is sufficiently strong and secure property rights do exist, then a marginal decrease in player i's strength may create conditions for the property rights to be no longer secure.

• If player i is sufficiently weak, then, whether or not secure property rights exist, a marginal change in player i's strength may create the conditions for the non-existence of secure property rights.

Figure 1 illustrates, in more specific terms, the effect of a small *increase* in p_i on the emergence of secure property rights (via the TSE). The initial critical discount factors are $\underline{\delta}_i$ and $\underline{\delta}_j$, while the critical discount factors after a small increase in p_i are $\widehat{\underline{\delta}}_j$ for player j, $\widehat{\underline{\delta}}_i^s$ for player i if he is strong and $\widehat{\underline{\delta}}_i^w$ if he is weak.[30] I now provide some interpretation of the various regions in Figure 1.

In regions 1–6 the parameters are such that both initially and after a small increase in p_i, at least one of the player's critical discount factor lies above his discount factor. Thus, at least one of the player's discount factor is so small that a small increase in p_i has no effect on the issue of the existence of the TSE; secure property rights do not exist initially, and do not exist after a small increase in p_i.

[29] It may be noted that p_i^*, which is defined in Proposition 2(b), is less than $\widehat{p}_i$.
[30] He is strong if the initial $p_i < \widehat{p}_i$, and weak if $p_i > \widehat{p}_i$, where $\widehat{p}_i$ is defined in Corollary 5(b).

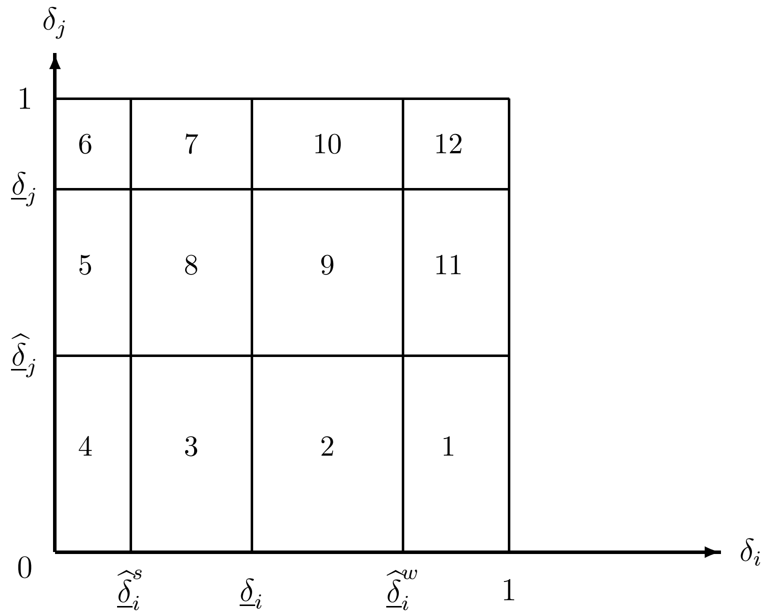

Figure 1: An illustration of the effect of a small increase in p_i on the existence of the TSE.

In regions 7–9 the parameters are such that initially δ_i and/or δ_j are too small (lie below their respective initial critical discount factors), and the TSE does not exist. However, after a small increase in p_i player j is willing to cooperate, and player i is willing to cooperate if and only if he is strong. An important insight provided by this result may be put as follows. An improvement in the military technology of a militarily strong player (such as the USA) can improve the likelihood of the emergence of secure property rights.

In region 10 secure property rights do exist initially, and remain in place after a small increase in p_i if and only if player i is strong. An important insight provided by this result may be put as follows. An improvement in the military technology of a militarily weak player may increase the likelihood that existing secure property rights become insecure.

In region 11 the TSE does not exist initially, but does after a small increase in p_i (by essentially making player j willing to cooperate). Finally, in region 12 the parameters are such that the players' discount rates are so high that they always lie above the relevant critical discount factors.

I now briefly examine the impact of the players' productive skills on the critical discount factors $\underline{\delta}_A$ and $\underline{\delta}_B$, where an improvement in a player's productive skills are formalized in the manner specified above in subsection 2.4. Since, as has been established in subsection 3.5, an improvement in player i's productive skill enhances his private incentives (i.e., increases Δ_i), it immediately follows from (16) that $\underline{\delta}_i$ decreases following an improvement in player i's productive skill. Thus, the more productive a player, the more willing is he to cooperate (and respect the property rights). What about his opponent? As player i's productive skills improve, it has been shown in subsection 3.5 that Δ_j decreases. Furthermore, it is easy to verify that the difference between player i's output levels in the PRE and NE is higher the more productive he is. It then immediately follows that $\underline{\delta}_j$ increases following an improvement in player i's productive skill. Thus, the more productive a player, the less willing is his opponent to cooperate.

4.2 A General Negative Result

It has been shown above that there exists a non-empty set of parameter values under which the TSE does not exist. In particular, if the parameters are such that the players' private incentives to establish the property rights are in conflict (i.e., $\Delta_i < 0$ for some i), then the TSE does not exist. Does this mean that for such parameter values, secure property rights cannot emerge in the state-of-nature? In order to answer this question, one needs to, in principle, study the set of all non-stationary SPE of the base model. The point is that there might exist a non-stationary SPE (different from the TSE) that does sustain the PRE path as a SPE path for a relatively wider set of parameter values. However, in this section, I characterize a set of parameter values under which there does not exist an SPE that sustains the PRE path.[31]

Let $\underline{V}_i$ denote the *worst* SPE (average) payoff to player i. Thus, by definition, the average payoff to player i in *any* SPE is greater than or equal to $\underline{V}_i$. The following result lies at the heart of the analysis to follow:[32]

Lemma 4. *If either one of the two inequalities stated below fails to hold, then there*

[31] In the next subsection, I construct a non-stationary SPE (different from the TSE) that does sustain the PRE path as SPE path even when $\Delta_i < 0$ for some i (*provided* that the players, effectively, do not discount future payoffs).

[32] This result is a straightforward application of some rather powerful results contained in Abreu (1988).

does not exist an SPE of the base model in which the PRE path is the equilibrium path:

$$V_A^F \geq (1-\delta_A)\Pi_A(L_A^N, L_B^F) + \delta_A \underline{V}_A$$
$$V_B^F \geq (1-\delta_B)\Pi_B(L_B^N, L_A^F) + \delta_B \underline{V}_B,$$

where $\underline{V}_i$ denotes player i's worst SPE payoff.

Proof. The argument can be made by contradiction. Thus, suppose (to the contrary) that there exists an SPE in which the PRE path constitutes the equilibrium path of play, *and* in which, for example,

$$V_A^F < (1-\delta_A)\Pi_A(L_A^N, L_B^F) + \delta_A \underline{V}_A. \tag{18}$$

Now suppose player A considers making a one-shot, unilateral deviation from the equilibrium path of play. His (average) payoff from not doing so is, of course, V_A^F. His (average) payoff from doing so is greater than or equal to the right-hand side of inequality 18. This is because in the period in which he unilaterally deviates, his optimal deviation is to set $L_A = L_A^N$, and thus, in that period his payoff is $\Pi_A(L_A^N, L_B^F)$. His continuation equilibrium average payoff (from the next period onwards) must, by definition, be greater than or equal to $\underline{V}_A$. Consequently, given inequality 18, it is optimal for player A to conduct the one-shot, unilateral deviation. But this is a contradiction. □

As is well-known, it is not easy to characterize the worst SPE payoffs to players in infinitely repeated games. And this is also the case here; I have not been able to characterize $\underline{V}_i$. However, one lower bound on it can be easily characterized; and that is player i's minimax payoff. It follows from an application of a well-known result that in any Nash equilibrium of the base model, player i's average payoff is greater than or equal to his minimax payoff. I now derive player i minimax payoff. The worst (from player i's perspective) possible strategy that player j could adopt is the one in which in each period, player j chooses not to work at all, and chooses to always fight (for any choices made in the past). The payoff per period to player i if player j adopts this (minimax) strategy is $(1-p_j)f_i(L_i) + v_i(T - L_i)$, which is maximised at $L_i = L_i^N$. Hence, player i's minimax payoff is

$$\underline{w}_i = \Pi_i(L_i^N, 0) = V_i^N - p_i f_j(L_j^N). \tag{19}$$

The following result is an immediate consequence of Lemma 4 and the observation that $\underline{V}_i \geq \underline{w}_i$.[33]

Proposition 4 (Non-Emergence of Secure Property Rights). *If the parameters are such that either one of the two inequalities stated below fails to hold, then there does not exist an SPE of the base model in which the equilibrium path is the PRE path:*

$$\Delta_A \geq p_A[(1-\delta_A)f_B(L_B^F) - f_B(L_B^N)] \tag{20}$$

$$\Delta_B \geq p_B[(1-\delta_B)f_A(L_A^F) - f_A(L_A^N)]. \tag{21}$$

It is useful to rewrite (20) and (21) respectively as follows:

$$\delta_A \geq \underline{\delta}_A^* \tag{22}$$

$$\delta_B \geq \underline{\delta}_B^*, \tag{23}$$

$$\text{where } \underline{\delta}_i^* = 1 - \left[\frac{\Delta_i + p_i f_j(L_j^N)}{p_i f_j(L_j^F)}\right] \quad (j \neq i). \tag{24}$$

It follows from Proposition 4 that the PRE path cannot be sustained in *any* SPE of the base model if some player's discount factor, say player i's, lies below the critical value stated in (24), which depends upon the parameters. It is straightforward to note that since $\Delta_i + p_i f_j(L_j^N) > 0$, it follows $\underline{\delta}_i^* < 1$. Furthermore, note that $\underline{\delta}_i^* > 0$ if and only $\underline{\delta}_i > 0$. Hence, from the results stated in the previous section (immediately after (16)), if the pair (L_A^F, L_B^F) satisfies equation 2 then $\underline{\delta}_i^* > 0$ for both $i = A$ and $i = B$; but if $p_i f_j(L_j^F) > p_j f_i(L_i^F)$ (where $i, j = A, B$ with $i \neq i$) then $\underline{\delta}_i^* > 0$ while $\underline{\delta}_j^*$ could be positive or negative. In conclusion, then, I have established the following result.

Corollary 6 (Non-existence of SPE with the PRE path). *For any parameter values such that for some i ($i = A$ or $i = B$) player i's discount factor $\delta_i < \underline{\delta}_i^*$ there does not exist an SPE of the base model whose equilibrium path is the PRE path, where for at least one i ($i = A$ or $i = B$) — if not for both i ($i = A$ and $i = B$) — $\underline{\delta}_i^* > 0$. Furthermore, for each i ($i = A, B$) $\underline{\delta}_i^* < 1$.*

[33] Since $\underline{V}_i \geq \underline{w}_i$, the inequalities stated in the proposition are effectively the inequalities stated in Lemma 4 but replacing $\underline{V}_i$ with $\underline{w}_i$, and then substituting for $\Pi_i(L_i^N, L_j^F)$ and $\underline{w}_i$, and then finally simplifying and rearranging terms.

I now examine the impact of the players' fighting skills on these critical discount factors $\underline{\delta}_A^*$ and $\underline{\delta}_B^*$. After substituting for Δ_i in the expression for $\underline{\delta}_i^*$ and then simplifying, I obtain that

$$\underline{\delta}_i^* = 1 - \left[\frac{\gamma_i + p_j f_i(L_i^N)}{p_i f_j(L_j^N)}\right], \quad \text{where} \quad \gamma_i = [f_i(L_i^F) + v_i(T - L_i^F)] - [f_i(L_i^N) + v_i(T - L_i^N)].$$

Using (5) it is straightforward to show that for each $i, j = A, B$ with $\neq j$:

$$\frac{\partial \underline{\delta}_i^*}{\partial p_i} \gtreqless 0 \iff \frac{\partial V_i^N}{\partial p_i} \gtreqless 0.$$

and that

$$\frac{\partial \underline{\delta}_i^*}{\partial p_j} \gtreqless 0 \iff \frac{\partial V_j^N}{\partial p_j} \lesseqgtr 0.$$

Hence, the following corollary is an immediate consequence of Proposition 2(b):

Corollary 7 (Fighting Skills and Non-Emergence of Secure Property Rights). *Fix $i, j = A, B$ with $i \neq j$.*
(a) If $f_j''' \leq 0$ and $v_j''' \leq 0$, then there exists $p_i^* \in (0,1)$ — where p_i^* is independent of p_j — such that

$$\frac{\partial \underline{\delta}_i^*}{\partial p_i} \gtreqless 0 \quad \textit{if} \quad p_i \lesseqgtr p_i^*.$$

(b) If $f_i''' \leq 0$ and $v_i''' \leq 0$, then there exists $p_j^* \in (0,1)$ — where p_j^* is independent of p_i — such that

$$\frac{\partial \underline{\delta}_i^*}{\partial p_j} \gtreqless 0 \quad \textit{if} \quad p_j \gtreqless p_j^*.$$

Corollary 7 implies that a marginal change in player i's fighting skill will, in general, have an ambiguous effect on the range of discount factors under which secure property rights cannot emerge in the state-of-nature. For example, if player i is weak (i.e., p_i is small), then a marginal increase in his fighting skill increases $\underline{\delta}_i^*$ but decreases $\underline{\delta}_j^*$. However, unambiguous results can be obtained in some special cases. Consider, for example, the case in which the players' discount factors are identical (i.e., $\delta_A = \delta_B = \delta$), and, in which their fighting and productive skills are such that $\delta_A^* > \delta_B^*$. It follows from Proposition 4 that in this case, secure property rights can never emerge in the state-of-nature if $\delta < \delta_A^*$. Hence, it immediately follows from Corollary 7 that an improvement in player A's fighting skill increases the likelihood that secure property rights can never emerge *if* player A is weak; while the opposite is the case *if* player A is strong.

4.3 The Mutual Minimax Equilibrium

While the above subsection establishes a "negative" result, in this subsection I establish a "positive" result; I shall show (by construction) that there exists a non-stationary SPE whose equilibrium path is the PRE path for *any* parameter values provided that the players are sufficiently patient.

The TSE is built upon the idea that the punishment path is the NE path. Two key properties of the TSE are as follows. Firstly, deviation by any player from the TSE path (which is the PRE path) entails the loss of the property rights *forever after*. And, secondly, in the punishment path of the TSE, player i's payoff in each period is V_i^N. In contrast, the two corresponding properties of the SPE constructed in this section are as follows. Deviation by any player from the PRE path entails the loss of property rights for only a *finite* number of periods, after which the property rights are re-established. Furthermore, in the punishment path player i's payoff is strictly less than V_i^N. It is this latter property which allows the PRE path to be sustainable (as a SPE path) for parameter values such that $\Delta_i < 0$ for some i.

It is convenient to describe the proposed non-stationary SPE — which, for reasons that will shortly become clear, I call the *mutual minimax equilibrium* (MME, for short) — by describing two "phases" (or paths of play), and transition rules between them. The initial phase (in which play begins) is called the *cooperative* phase; it constitutes the PRE path. The other phase (to which play moves to if any player unilaterally from the cooperative phase) is called the *punishment* phase.

Plays begins in the cooperative phase:

• **The Cooperative Phase**. Each player i ($i = A, B$) sets $L_i = L_i^F$, and does not fight.

• **Transition Rule 1**. If any player unilaterally deviates in the cooperative phase, then immediately play moves to the punishment phase.

• **The Punishment Phase**. For k periods $L_A = L_B = 0$, and both players choose to fight. Thereafter, play returns to the cooperative phase.

• **Transition Rule 2**. If any player unilaterally deviates from the punishment phase, then immediately the punishment phase starts all over again.[34]

I now derive the incentive-compatibility conditions under which the pair of strategies (implicitly) described above are in a SPE. Note that player i's equilibrium average payoff in the cooperative phase is V_i^F. Letting z_i denote player i's equilibrium average payoff at the start of the punishment phase, it is trivial to note that

$$z_i = (1 - \delta_i^k)v_i(T) + \delta_i^k V_i^F. \tag{25}$$

It thus follows that a unilateral, one-shot deviation from the cooperative phase by player i is not profitable for him if and only if

$$V_i^F \geq (1 - \delta_i)\Pi_i(L_i^N, L_j^F) + \delta_i z_i. \tag{26}$$

Furthermore, a unilateral one-shot deviation from the punishment phase by player i is not profitable for him if and only if $z_i \geq (1 - \delta_i)\underline{w}_i + \delta_i z_i$ — that is:

$$z_i \geq \underline{w}_i, \tag{27}$$

where $\underline{w}_i$ is defined above in 19. Thus, player i has no incentive to deviate from his proposed equilibrium strategy if and only if k is such that (26) and (27) hold. This makes intuitive sense. On the one hand, k should be sufficiently large (thereby making the punishment sufficiently severe) so as to provide player i with an incentive not to deviate from the PRE path. But, on the other hand, k should be sufficiently small (thereby making the punishment sufficiently lax) so as to provide player i with an incentive to take part in the punishment of either player and thereafter return to the cooperative phase. Thus, k has to determine both the thickness of the *stick* with which punishment is meted out, and the attractiveness of the *carrot* which is obtained only after the punishment is taken. I have thus established the following proposition.

Proposition 5 (The Mutual-Minimax Equilibrium (MME)). *The pair of non-stationary strategies (implicitly) described above (in terms of the cooperative and punishment phases and the two transition rules) is an SPE if and only if k satisfies (26)*

[34]It may be noted that in the punishment phase both players use their respective minimax strategies, which are, of course, identical (namely, "do no work, and fight"). Thus, whether it is player A or player B who unilaterally deviates from the cooperative phase, in the punishment phase each player minimaxes the other player for a certain number of periods. This idea of "mutual minimaxing" to sustain cooperation in repeated games was first used by Fudenberg and Maskin (1986), but, it may be noted, in the context of repeated *normal-form* (or static) games with the players having *identical* discount factors.

and (27) (for $i, j = A, B$ with $i \neq j$). This SPE will be called the mutual-minimax equilibrium (MME, for short). Since the MME path is the PRE path, in equilibrium no fight ever occurs, and in each period player i ($i = A, B$) works for L_i^F units of time. Furthermore, of course, player i's MME payoff in each period equals V_i^F.

The following corollary addresses the issue of the existence of the MME. It shows, in particular, that even if the players' private incentives to establish the property rights are in conflict, the MME exists provided that the players are sufficiently patient.

Corollary 8 (Existence of MME). *For any parameter values there exists $\widehat{\delta}_A \in (0,1)$ and $\widehat{\delta}_B \in (0,1)$ such that if $\delta_A \in (\widehat{\delta}_A, 1)$ and $\delta_B \in (\widehat{\delta}_B, 1)$ then the MME exists.*

Proof. In the Appendix. □

Thus, in the limit as both δ_A and δ_B tend to one, the MME exists for *any* parameter values. This result may be stated as follows: if each player does not discount future payoffs, then secure property rights can emerge in the state-of-nature. The intuition behind this result is that while the (short-run) benefit to a player by a unilateral deviation from an arrangement with secure property rights is finite, the (long-run) cost from doing so becomes unboundedly large as his discount factor becomes arbitrarily close to one. Although this is an important *benchmark* result, it is not particularly useful, since the assumption of negligible discounting is not particularly plausible. Nevertheless, this result emphasizes that making the future important to the players can help provide them with appropriate incentives to respect the property rights.

5 The Role of Inter-Player Transfers of Output

It has been shown in the previous section that there exists a range of parameter values under which the PRE path cannot be sustained in *any* SPE of the base model. In particular, this is the case when the players' private incentives to establish the property rights are in conflict and they are not sufficiently patient. Does this mean that under such conditions the players are doomed to live in the state-of-nature without secure property rights? Or, perhaps more optimistically, does this mean that they will resort to some mechanism which (in this two-player environment) might enable the emergence of secure property rights? In this section I explore the potential role of one such mechanism, namely, *inter-player transfers of output*.

More precisely, I now extend the base model by allowing the players the option to communicate and negotiate with each other regarding whether or not to establish the property rights under consideration. In particular, the novel aspect is that a player will have the option to offer the other player some output in return for getting him to agree to the establishment of such property rights. Of course, since negotiated agreements are not automatically enforceable, each player will have the option to (*ex-post*) renege on his part of the agreement, whether or not the other player does so. This means that only those agreements which satisfy certain incentive-compatibility conditions are relevant to these negotiations. The objective of the analysis to follow is to explore the extent to which such inter-player transfers of output enhances the range of parameter values under which secure property rights can emerge.

5.1 An Extended Base Model with Bargaining

There are several alternative, plausible manners in which the opportunity to communicate and negotiate can be interlaced within the structure of the base model. As such there are several alternative, plausible extensions of the base model in which the mechanism of inter-player transfers of output exists. It turns out, however, that (because of the underlying stationary structure of the environment) the main analyses of all such extended models are identical. I now turn to a description of one such extended model. A key characteristic of this extension of the base model is that the players can communicate and negotiate only once, namely, at the beginning of time.

Before the beginning of period 1, the players meet to discuss whether or not to establish the property rights under consideration, and, whether or not one of the players should give some of his output to the other player in each period. Let the *game form* that formally encapsulates this bargaining and communication process be denoted by ϕ, which has two types of outcomes: (i) the players reach an agreement $t \in \Re$, with the interpretation that in each period, player A will give t units of output to player B and both players will not fight,[35] and (ii) the players fail to reach an agreement. If no agreement is struck, then play proceeds according to the base model. However, if an agreement is struck, then (since it is not automatically enforceable) play proceeds according to an extended version of the base model in which at the end of each period,

[35]Notice that t can be positive or negative. A negative t means that it is player B who will give $-t$ units of output to player A.

before stage 2 takes place, the player who has to make the payment decides whether or not to do so. Thus, notice that in each period, each player can choose whether or not to renege on his part of the agreement.

There are several assumptions implicitly built into the above described extension of the base model. First, if the players fail to reach an agreement at the beginning of time, then they cannot attempt to do so in any later period. Second, an agreement struck at the beginning of time cannot be renegotiated at any later date. Third, an agreed transfer of output is constant across all periods. It is straightforward to consider various alternative, plausible extensions of the base model in which one or more of such assumptions is relaxed. For example, one could consider an extended model in which the players communicate and negotiate at the beginning of each period over whether or not to establish the property rights (with some transfer of output) for just that period. As I mentioned above, because of the stationary structure that underlies the environment, it turns out that the main analyses of such alternative, plausible extended models are identical to the analysis of the extended model described above in which the players communicate and negotiate only at the beginning of time.

In what follows, I assume, as seems plausible, that if the players fail to reach an agreement in ϕ, then play (in the base model) proceeds according to the NE. The motivation for this assumption is that if they do not agree to establish the property rights when they are in face-to-face communication, then it is unlikely that secure property rights will emerge *implicitly* via some non-stationary SPE of the base model. Hence, it may be noted that if the players do not reach an agreement in ϕ, then player i's payoff in each period is V_i^N.

5.2 Incentive-Compatible Agreements

An agreement in ϕ, which is characterized by a real number t, is *self-enforcing* if in the extended base game that ensues there is a SPE whose equilibrium path is the following extended PRE path: in each period, player i ($i = A, B$) sets $L_i = L_i^F$, the agreed transfer t is implemented and no fight takes place. I assume that if any player unilaterally deviates from this extended PRE path (i.e., violates the agreement), then immediately play proceeds according to the NE. A player could deviate at any one of the three stages within each period: either at stage 1 by choosing $L_i \neq L_i^F$, or, if this is relevant to him, after stage 1 but before stage 2 by not transfering the agreed output t to the other player (I call this stage, stage 1.5), or, at stage 2 by choosing to fight. Through a straightforward extension of the arguments used in section 4.1, it follows that for player A, a (one-shot, unilateral) deviation at stages 1, 1.5 and 2, respectively, are not profitable if and only if the following three inequalities hold:

$$V_A^F - t \geq (1-\delta_A)\Pi_A(L_A^N, L_B^F) + \delta_A V_A^N$$
$$V_A^F - t \geq (1-\delta_A)\Pi_A(L_A^F, L_B^F) + \delta_A V_A^N$$
$$V_A^F - t \geq (1-\delta_A)[\Pi_A(L_A^F, L_B^F) - t] + \delta_A V_A^N.$$

Hence, I obtain the following result:[36]

Lemma 5 (Incentive-Compatible Agreements). *An agreement to establish the property rights with player A giving $t \in \Re$ units of output to player B in each period is self-enforcing (or incentive-compatible) if the following two conditions hold:*

$$V_A^F - t \geq (1-\delta_A)\Pi_A(L_A^N, L_B^F) + \delta_A V_A^N$$
$$V_B^F + t \geq (1-\delta_B)\Pi_B(L_A^F, L_B^N) + \delta_B V_B^N.$$

An immediate consequence of Lemma 5 is that an agreement t is incentive-compatible *only if* it is feasible in the sense that $f_A(L_A^F) \geq t \geq -f_B(L_B^F)$.

Another consequence of Lemma 5 is that (since $\Pi_i(L_i^N, L_j^F) > V_i^N$) an agreement is incentive-compatible *only if* $V_A^F - t > V_A^N$ and $V_B^F + t > V_B^N$. This means that each player strictly prefers reaching agreement on *any* incentive-compatible agreement to not reaching any agreement. Although, therefore, the players have a common interest to reach agreement on some incentive-compatible agreement (if one exists), they have conflicting (or divergent) interests over the set of such agreements (since each player's payoff from agreement is increasing in the amount of transfer of output that he receives from the other player). A key point that, therefore, emerges here is as follows:

- *If there exists an incentive-compatible agreement, then the players will reach agreement to establish secure property rights, provided (as is plausible to assume) that the equilibrium outcome in ϕ is Pareto-efficient.*

By using Lemma 5, I now derive the condition under which there exists an incentive-compatible agreement. After substituting for $\Pi_i(L_i^N, L_j^F)$ $(i, j = A, B$ with $i \neq j)$ in the inequalities stated in Lemma 5, and then simplifying, it follows that an agreement

[36]Notice that if the first of these three inequalities holds, then the latter two inequalities also hold. Furthermore, through a symmetric argument one can derive similar incentive-compatibility conditions which would ensure that player B also would not have an incentive to violate the agreement.

$t \in \Re$ is incentive-compatible if t satisfies the following inequalities:

$$t \leq \Delta_A - p_A(1-\delta_A)[f_B(L_B^F) - f_B(L_B^N)] \quad (28)$$

$$t \geq p_B(1-\delta_B)[f_A(L_A^F) - f_A(L_A^N)] - \Delta_B. \quad (29)$$

This implies that there exists an incentive-compatible agreement if and only if

$$\Delta_A + \Delta_B \geq p_A(1-\delta_A)[f_B(L_B^F) - f_B(L_B^N)] + p_B(1-\delta_B)[f_A(L_A^F) - f_A(L_A^N)]. \quad (30)$$

Since the right-hand side of (30) is strictly positive, this means that in order for there to exist an incentive-compatible agreement, the *aggregate net surplus* (or collective incentive) from having the property rights, which is the left-hand side of (30), must be sufficiently large; it is not enough that there just exists some aggregate net surplus, i.e., $\Delta_A + \Delta_B \geq 0$ — which, recall, holds for any parameter values.[37] This means that (30) holds in the limit as both δ_A and δ_B tend to one. At the other extreme, when $\delta_A = \delta_B = 0$, inequality 30 does not hold, which can be seen after some manipulation of this inequality. After substituting for Δ_A and Δ_B in (30), and then simplifying and re-arranging terms, it follows that (30) can be re-written as follows:

$$\Gamma_A + \Gamma_B \geq 0,$$

where for each $i, j = A, B$ with $i \neq j$,

$$\Gamma_i = \left[[1 - p_j(1-\delta_j)] f_i(L_i^F) + v_i(T - L_i^F) \right] - \left[[1 - p_j(1-\delta_j)] f_i(L_i^N) + v_i(T - L_i^N) \right].$$

By definition, if $\delta_j = 0$ then $\Gamma_i < 0$, and, if δ_j is sufficiently close to one then $\Gamma_i > 0$. Not surprisingly, an incentive-compatible agreement exists provided that the players do not discount future payoffs too much. The key issue, however, is whether or not such an agreement exists for a wider range of parameter values than for which the TSE exists. It is trivial to note that any parameter values which satisfy (12) and (13) must also satisfy (30). Hence, if the TSE exists, then an incentive-compatible agreement also exists. However, notice that there exists parameter values which satisfy (30), but which will not satisfy (12) and (13). Thus, there exists parameter values under which the TSE does not exist, but an incentive-compatible agreement does exist. This is the case, for example, for some parameter values such that $\Delta_i < 0$ for some i. Hence, when the players' private incentives to establish the property rights are in conflict (such

[37]As would be expected, the requirement that an agreement be incentive-compatible is somewhat constraining.

as when one player is quite strong but quite unproductive, while the reverse is the case regarding the other player), inter-player transfers of output can be a mechanism through which secure property rights get established.

Stated more generally, the results established here imply that the mechanism of inter-player transfers of output does allow for the emergence of secure property rights in circumstances in which they would not otherwise emerge. The intuition for this conclusion follows by noting that when the players *negotiate* over whether or not to establish secure property rights with an appropriate, per-period transfer of some output between them, their respective *private* incentives are no longer relevant to the issue of the emergence of secure property rights; it is their *collective* incentive that takes on centre stage. More precisely, since their collective incentive (namely, $\Delta_A + \Delta_B$) is the *sum* of their respective private incentives, the players' private incentives do continue matter, but only to the extent that they determine the collective incentive. For example, conflicting private incentives no longer pose the kind of threat that they did in the analysis of section 4, since what matters to the analysis in this section is the collective incentive.

5.3 Nash Bargains

The insights derived in the previous subsection (concerning the role of the mechanism of inter-player transfers of output on the emergence of secure property rights) are based on teasing out the implications of the requirement that negotiated agreements must be incentive-compatible, and, on the plausible assumption that the outcome in ϕ will be Pareto-efficient. As such these insights are fairly general, since, in particular, they are independent of the exact nature of the agreed transfer (which would clearly depend on the details of the bargaining process ϕ).

However, in order to develop some further (albeit specific) insights, I now assume that the bargaining process ϕ is such that the equilibrium agreed transfer, denoted by t^*, can be characterized using the Nash bargaining solution (NBS, for short) with the *disagreement* point being the payoff pair (V_A^N, V_B^N). As is well-known — see, for example, Muthoo (1999) — this can easily be justified by assuming that ϕ is the Rubinsteinian, alternating-offers bargaining process.

Ignoring, for a moment, the requirement of incentive compatibility, it follows that:[38]

$$t^* = \frac{\Delta_A - \Delta_B}{2} \tag{31}$$

Thus, interestingly, in the NBS, the player who has to transfer some of his output in each period to the other player is the one whose private net benefit from having the property rights is higher. If these private net benefits are identical, then no transfer of output takes place.

Now, t^* will be the equilibrium agreed transfer provided that it is incentive-compatible. After substituting for $t = t^*$ (using (31)) into (28) and (29), and then simplifying and re-arranging terms, it thus follows that t^* is incentive-compatible if

$$\Delta_A + \Delta_B \geq 2 \max\{p_A(1-\delta_A)[f_B(L_B^F) - f_B(L_B^N)], p_B(1-\delta_B)[f_A(L_A^F) - f_A(L_A^N)]\}.$$

That is, if the parameters are such that

$$p_j(1-\delta_j)[f_i(L_i^F) - f_i(L_i^N)] \leq p_i(1-\delta_i)[f_j(L_j^F) - f_j(L_j^N)], \tag{32}$$

where $i, j = A, B$ with $i \neq j$, then t^* is incentive-compatible if

$$\delta_i \geq \underline{\delta}'_i,$$

where

$$\underline{\delta}'_i = 1 - \frac{\Delta_A + \Delta_B}{2p_i[f_j(L_j^F) - f_j(L_j^N)]}. \tag{33}$$

This immediately implies, for example, that the agreement with the Nash bargained transfer t^* *is* incentive-compatible when the players' private incentives to establish the property rights are in conflict and they do not have high discount factors. The analysis in section 4 showed that under such conditions secure property rights could not emerge. In contrast, the result obtained here shows that the mechanism of inter-player transfers of output can provide the basis for their emergence via self-enforcing, negotiated agreements.

---

[38] This is obtained by first noting that in the NBS the players split equally the aggregate net surplus $\Delta_A + \Delta_A$. Hence, player A's utility payoff in the NBS is $u_A^* = V_A^N + (\Delta_A + \Delta_B)/2$. Consequently, since (by definition) the transfer in the NBS is $t^* = V_A^F - u_A^*$, it follows that t^* is as stated in (31).

6 Summary and Concluding Remarks

A main, key contribution of this paper is my *model* of the state-of-nature. It captures the essential, basic elements of the strategic interaction between two players in the state-of-nature. Although, as I shall discuss below, the model contains several, restrictive assumptions, an important aspect of my model is that it provides a basic framework that can be extended and/or modified to capture various, omitted features of the strategic interaction between the two players in the state-of-nature.

Some of the most fundamental insights obtained in this paper concerning the origins of secure (or self-enforcing) property rights are as follows:

• *Heterogeneity* in the players' fighting and productive skills plays a crucial role in determining their incentives to establish secure property rights. In particular, there exist configurations of the players' fighting and productive skills — such as when one player is quite unproductive but very strong, while the other player is quite productive but very weak — under which the players' private incentives are in conflict.

• In order to promote and maintain incentives (private and collective), improvements in the players' productive skills (or economic prosperity) should go hand-in-hand with improvements in their fighting skills (or military technologies).

• If the players' fighting and productive skills are such that their private incentives to establish the property rights are in conflict, then in order for secure property rights to emerge, it is necessary that the players communicate with each other and *negotiate* over whether or not to establish the property rights with an appropriate, per-period transfer of output between them. That is, in such circumstances, the use of the mechanism of *inter-player transfers of output* is vital for the emergence of secure property rights.

• When the players' private incentives are not in conflict, then secure property rights might emerge without the mechanism of inter-player transfers of output. The likelihood of this happening is higher the more concerned are the players for their future payoffs, or the greater are their fighting skills, or the lower are their productive skills.

• Improvements in the fighting skills of a militarily strong player (such as the USA) enhances the likelihood of the emergence of secure property rights. On the other hand, improvements in the fighting skills of a militarily weak player enhances the likelihood that existing secure property rights become insecure.

I now turn to a discussion of the main limitations of my model as a model of the state-of-nature with *two* players. Before I do so, let me note upfront that addressing some of these limitations — by extending and/or modifying my model — will not necessarily require any conceptual innovations (although the analyses of some of these extensions may be technically demanding). In contrast, when extending my model to an environment with *three* (or more) players, new conceptual issues will necessarily arise such as the issue of who forms a coalition with whom.[39]

First, a simplifying, but restrictive assumption that underlies my model is that the players cannot make any investments in their fighting and productive skills. Clearly, this assumption should be relaxed in future research. One potential way of extending the base model to allow for such investments is as follows. At the beginning of each period, before stage 1, each player has the option to spend some time to improve his fighting skill and/or his productive skill. By incorporating the possibility of such costly investments, the equilibrium growth rates of the players' fighting and productive skills can be determined, which, in turn, will affect the dynamics of the likelihood of the emergence of secure property rights.

Second, it is important to relax the implicit assumption that a player cannot steal the other player's endowments (of productive and military technologies, and/or, of time) — in my model, a player can only steal the other player's final output. This, of course, raises some important issues. For example, stealing *all* of a player's endowments is like making that player one's slave — which is, however, not the case if one only steals that player's technologies. The issues of how a player might then make use of the stolen endowments, and, of how he would keep them over time — in the face of the other player's incentive to fight back — need to be addressed. A key question is whether or not there can exist configurations of the players' fighting and productive skills such that in equilibrium one player enslaves the other, and, maintains the slave's incentives not to break free by some appropriate, minimal lump-sum transfer of output.

Third, it would be interesting to explore the implications of relaxing some of the informational assumptions. For example, I have assumed that a player's output is observable by the other player. It would be interesting to explore the implications of an extension of my model in which a player receives an *imperfect* signal of his opponent's output level. A focal question is whether or not such imperfect observability, which seems like a reasonable assumption, adversely affects the likelihood of the emergence of secure property rights.

Fourth, there is no room for specialization and trade in my model. It would be

[39]In my companion paper, Muthoo (In Preparation), I construct and study a model of the state-of-nature with three players that is based upon the model studied in the current paper.

interesting and useful to extend my model by having, for example, two consumption goods. In such an environment, secure property rights might have a relatively better chance of emerging when each player specializes in the production of one good, and, then obtains the other good through trade with his opponent.

There are, of course, several other potentially fruitful extensions of my model; extensions, like those discussed above, which would enhance our understanding of the topic under consideration.

Appendix

PROOF OF PROPOSITION 2(b)

It follows from (4) that

$$\frac{\partial L_j^N}{\partial p_i} = \frac{f_j'(L_j^N)}{\Upsilon}, \quad \text{where } \Upsilon = (1-p_i)f_j''(L_j^N) - v_j''(T - L_j^N). \tag{34}$$

After substituting for the derivative of L_j^N with respect to p_i, using (34), into (5), and then simplifying, (5) becomes:

$$\frac{\partial V_i^N}{\partial p_i} = f_j(L_j^N) + \frac{p_i[f_j'(L_j^N)]^2}{\Upsilon}. \tag{35}$$

Notice that[40]

$$\lim_{p_i \to 0} \frac{\partial V_i^N}{\partial p_i} = f_j(L_j^F), \quad \text{and} \quad \lim_{p_i \to 1} \frac{\partial V_i^N}{\partial p_i} = -\infty.$$

Furthemore, notice that the derivative of V_i^N with respect to p_i is continuous in p_i, and that it is independent of p_j. Now, it is straightforward to verify that

$$\frac{\partial^2 V_i^N}{\partial p_i^2} < 0 \quad \text{provided that} \quad \frac{\partial \Upsilon}{\partial p_i} \geq 0,$$

where Υ is defined above in (34). Since (by the hypothesis stated in Proposition 2(b)) $f_j''' \leq 0$ and $v_j''' \leq 0$), the derivative of Υ with respect to p_i is positive. Hence, the second derivative of V_i^N with respect to p_i is strictly negative. The desired conclusion follows immediately from the above results concerning the nature of the derivative of V_i^N with respect to p_i.

[40] These results are based on the results that $L_j^N \to L_j^F$ as $p_i \to 0$ and $L_j^N \to 0$ as $p_i \to 1$.

DERIVATION OF CONDITION (17)

After differentiating $\underline{\delta}_i$ with respect to p_i, it follows that

$$\frac{\partial \underline{\delta}_i}{\partial p_i} \gtreqless 0 \iff -p_i \tau_j \frac{\partial \Delta_i}{\partial p_i} + \Delta_i \left[\tau_j - p_i f_j'(L_j^N) \frac{\partial L_j^N}{\partial p_i} \right] \gtreqless 0,$$

where $\tau_j = f_j(L_j^F) - f_j(L_j^N)$. Thus,

$$\frac{\partial \underline{\delta}_i}{\partial p_i} \gtreqless 0 \iff p_i \tau_j \frac{\partial V_i^N}{\partial p_i} + \Delta_i \left[\tau_j - p_i f_j'(L_j^N) \frac{\partial L_j^N}{\partial p_i} \right] \gtreqless 0.$$

Thus, since the derivative of V_i^N with respect to p_i is identical to the derivative of $p_i f_j(L_j^N)$ with respect to p_i, it follows that

$$\frac{\partial \underline{\delta}_i}{\partial p_i} \gtreqless 0 \iff (p_i \tau_j - \Delta_i) \frac{\partial V_i^N}{\partial p_i} + \Delta_i f_j(L_j^F) \gtreqless 0.$$

Since (by the definition of $\underline{\delta}_i$) $\Delta_i = (1 - \underline{\delta}_i) p_i \tau_j$, it follows that

$$\frac{\partial \underline{\delta}_i}{\partial p_i} \gtreqless 0 \iff \underline{\delta}_i p_i \tau_j \frac{\partial V_i^N}{\partial p_i} + (1 - \underline{\delta}_i) p_i \tau_j f_j(L_j^F) \gtreqless 0.$$

The desired conclusion (namely, condition (17)) follows immediately (since $p_i \tau_j > 0$).

PROOF OF COROLLARY 8

The proof to follow is an adaption of the argument that underlies the proof of Theorem 1 in Fudenberg and Maskin (1986). Choose $\widehat{\delta}_A$ and $\widehat{\delta}_B$ close to one so that for each $i = A, B$,

$$V_i^F > (1 - \widehat{\delta}_i) \Pi_i(L_i^N, L_j^F) \tag{36}$$

and (27) holds when $k = 1$.[41] If, with $k = 1$, (26) is violated for some i, then raise k until (26) is satisfied for both i ($i = A$ and $i = B$) — which is possible since (i) z_i is continuous and decreasing in k, and, as k tends to infinity, z_i tends to $v_i(T)$ (which is strictly less than V_i^F), and (ii) (36) holds. Now, since z_i is increasing in δ_i, by taking $\widehat{\delta}_i$ close enough to one we can ensure that (27) will be satisfied for the first k for which (26) holds. The Corollary now follows immediately, since for any $\delta_A > \widehat{\delta}_A$ and $\delta_B > \widehat{\delta}_B$ there exists a k' such that for each $i = A, B$, (26) and (27) hold for such disount factors and $k = k'$.

[41] This is possible since $V^F > \underline{w}_i > 0$.

References

Abreu, Dilip (1988), "On the Theory of Infinitely Repeated Games with Discounting", *Econometrica*, **56**, 383-396.

Bates, Robert, Avner Greif and **Smita Singh** (2000), "Organizing Violence: Wealth, Power and Limited Government", mimeo, Department of Economics, Stanford University.

Besley, Timothy (1995), "Property Rights and Investment Incentives: Theory and Evidence from Ghana", *Journal of Political Economy*, **103**, 903-937.

Binmore, Ken (1994), *Game Theory and The Social Contract, Vol I: Playing Fair*, Cambridge, Massachusetts: MIT Press.

Binmore, Ken (1998), *Game Theory and The Social Contract, Vol II: Just Fair*, Cambridge, Massachusetts: MIT Press.

Buchanan, James (1975), *The Limits to Liberty*, University of Chicago Press.

Bush, Winston and **Lawrence Mayer** (1974), "Some Implications of Anarchy for the Distribution of Property", *Journal of Economic Theory*, **8**, 401-412.

Cheung, Steven (1963), "Private Property Rights and Sharecropping", *Journal of Political Economy*, **76**, 1107-1122.

Coase, Ronald (1960), "The Problem of Social Cost", *Journal of Law and Economics*, **3**, 1-44.

Collier Paul and **Anke Hoeffler** (1998), "On Economic Causes of Civil War", *Oxford Economic Papers*, **50**, 563-573.

Demirguc-Kunt, Asli and **Vojislav Maksimovic** (1998), "Law, Finance and Firm Growth", *Journal of Finance*, **53**, 2107-2137.

Demsetz, Harold (1967), "Towards a Theory of Property Rights", *Journal of Political Economy*, **57**, 347-359.

Fearon, James and **David Laitin** (1999), "Weak States, Rough Terrain, and, Large-Scale Ethnic Violence Since 1945", mimeo, Department of Political Science, Stanford University.

Fudenberg, Drew and **Eric Maskin** (1986), "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information", *Econometrica*, **54**, 533-554.

Fudenberg, Drew and **Jean Tirole** (1991), *Game Theory*, MIT Press, Cambridge, Massachusetts.

Garfinkel, Michelle and **Stergios Skaperdas** (2000), "Contract or War? On the Consequences of a Broader View of Self-Interest in Economics", mimeo, Department of Economics, University of California, Irvine.

Gauthier, David (1986), *Morals by Agreement*, Oxford: Oxford University Press.

Grossman, Sanford and **Oliver Hart** (1986), "The Costs and Benefits of Ownership: A Theory of Lateral and Vertical Integration", *Journal of Political Economy*, **94**, 691-719.

Grossman, Herschel and **Minseong Kim** (1995), "Swords or Plowshares? A Theory of the Security of Claims to Property", *Journal of Political Economy*, **103**, 1275-1288.

Hart, Oliver (1995), *Firms, Contracts, and Financial Structure*, Oxford: Oxford University Press.

Hirshleifer, Jack (1995), "Anarchy and Its Breakdown", *Journal of Political Economy*, **103**, 26-52.

Hobbes, Thomas (1651), *Leviathan.* Pengiun classics, London 1986 (edited by C.B. Macpherson).

Houba, Harold and **Hans-Peter Weikard** (1995), "Interaction in Anarchy and the Social Contract: A Game-Theoretic Perspective", Working Paper No: 95-186, Tinbergen Institute, The Netherlands.

Hume, David (1739), *A Treatise on Human Nature.* Oxford: Clarendon Press, 1978 (edited by L.A. Selby-Bigge).

Johnson, Simon, John McMillan and **Christopher Woodruff** (1999), "Property Rights, Finance and Entrepreneurship", mimeo, GSB, Stanford University (http://faculty-gsb.stanford.edu/mcmillan/personal/homepage.htm).

Locke, John (1690), *Two Treatises of Government*, Cambridge: Cambridge University Press, 1963.

Mauro, Paolo (1995), "Corruption and Growth", *Quarterly Journal of Economics*, **110**, 681-712.

Muthoo, Abhinay (1999), *Bargaining Theory with Applications*, Cambridge: Cambridge University Press.

Muthoo, Abhinay (In Preparation), "On the Foundations of Basic Property Rights, Part II: Coalition Formation in the State-of-Nature".

Neary, Hugh (1996), "To Fight or Not to Fight: Equilibrium Structure in a Conflict Model when Conflict is Costly", mimeo, Department of Economics, University of British Columbia (http://web.arts.ubc.ca/econ/neary/homepage.htm).

North, Douglass C. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge: Cambridge University Press.

Nozick, Robert (1974), *Anarchy, State and Utopia*, New York: Basic Books.

Osborne, Martin and **Ariel Rubinstein** (1994), *A Course in Game Theory*, MIT Press, Cambridge, Massachusetts.

Powell, Robert (1999) *In the Shadow of Power: States and Strategies in International Politics*, Princeton: Princeton University Press.

Rawls, John (1972), *A Theory of Justice*, Oxford: Oxford University Press.

Rousseau, Jean-Jacques (1762), *The Social Contract*, New York: Dutton, 1938.

Skaperdas, Stergios (1992), "Cooperation, Conflict, and Power in the Absence of Property Rights", *American Economic Review*, **82**, 720-739.

Sugden, Robert (1986), *The Economics of Rights, Cooperation and Welfare*, Oxford: Basil Blackwell.

Svensson, Jakob (1998), "Investment, Property Rights and Political Instability: Theory and Evidence", *European Economic Review*, **42**, 1317-1341.

Taylor, Michael (1987), *The Possibility of Cooperation*, Cambridge: Cambridge University Press.

Umbeck, John R. (1981), *A Theory of Property Rights*, Ames: The Iowa State University Press.

World Development Report (1997), *The State in Changing World*, Oxford University Press.