# A MODEL OF THE ORIGINS OF BASIC PROPERTY RIGHTS

ABHINAY MUTHOO

Abstract. This paper studies the *origins* of one of the most basic of property rights, namely, the right of an individual or an organization to the fruits of its labour. My objective is to address the questions of *why, when* and *how* this property right can emerge and be made secure. I develop a model of the strategic interaction between two players in the *state-of-nature*, which is an environment characterized by the absence of any laws and institutions (including property rights and the state). My analyses explores, in particular, the roles of the players' fighting and productive skills on the emergence and security (or otherwise) of this property right.

"... 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

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. Similarly, if the output of an organization (such as a nation-state or a local community) is not secure from pillage, then the organization's incentives to produce

*Date*: May 15, 2002.

*Key words and phrases.* property rights, state-of-nature, natural equilibrium, self-enforcement, inter-player transfers of output, incentive-compatible agreements.

I acknowledge, with gratitude, the financial support kindly provided by the UK Economic and Social Research Council (ESRC) through a research grant (No: R000222995). For helpful comments by Avinash Dixit, Ken Shepsle, the seminar participants at New York University, University College London, University of Essex, University of Aarhus, and the 2000 World Congress of the Econometric Society and the joint World Bank and Princeton University's Center for International Studies Workshop on *The Economics of Political Violence*, I am grateful.

such output in the first place would be adversely affected. Indeed, the fundamental importance of secure property rights for the 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*.

This paper is a theoretical contribution to the study of the *origins* of property rights. Specifically, I focus attention on the origins of one of the most basic of property rights, namely, the right of an individual or an organization to the fruits of its labour. My objective is to address the questions of *why, when* and *how* this property right can emerge and be made secure.

I develop a model of the strategic interaction between two players in the *state-of-nature*, which is an environment characterized by the absence of any laws and institutions (including property rights and the state). My (base) model constitutes the infinite repetition of a two-stage game, where at the first stage each player allocates his (per-period) time endowment between work and leisure, and then at the second stage decides whether or not to fight in an attempt to steal the other player's output. My answers to the questions posed above are in terms of the core parameters of my model, which comprise the players' fighting skills, productive skills and discount factors.

A key aspect of my model and analyses is to entertain *heterogeneity* in the players' fighting and productive skills, the importance of which for the emergence (or otherwise) of secure property rights is developed. For example, I show that there exist configurations of the players' fighting and productive skills — such as when one player is unproductive but strong, while the other player is productive but weak — under which the players' private incentives from establishing the said property rights are in conflict. This insight challenges the not uncommon viewpoint — see, for example, Taylor (1987) — that the Prisoners' Dilemma game adequately 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 secure property rights are established, a viewpoint that my analyses shows can only be sustained when players have *similar* fighting and productive skills.

This paper makes two main contributions. The first is the model *itself*. In particular, as I discuss in the concluding section, my model 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. The second main contribution lies in the results that I obtain, to which I now turn.

I obtain three main sets of results. First, I derive and analyze the set of parameter values under which the property right in question can emerge and be made secure. This parameter set has the property that the players' private incentives are not in conflict. A number of insights are obtained that inform the questions posed above. For example, it is shown that improvements in the fighting skill of a strong player enhances the likelihood of the emergence of secure property rights, while improvements in the fighting skill of a weak player enhances the likelihood that existing secure property fights become insecure. Second, I derive and analyze the set of parameter values under which the property right in question can *never* emerge. Third, I show that the said property right can emerge and be made secure for a relatively wider set of parameter values when players can resort to the mechanism of *inter-player transfers of output*. For example, for some parameter values such that the players' private incentives are in conflict, an appropriate incentive-compatible per-period transfer of output from one player to the other would enable the property right to emerge and be made secure; but not so in the absence of such a transfer of output.

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 is typically 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) 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 seminal contribution; a main focus of that literature is on the issue of the optimal distribution of property rights. 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 *assumes* — sometimes explicitly, but more often only implicitly — that any specified distribution of property rights can be almost *costlessly enforced* by a third party such as the courts/state. In particular, 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.

Although therefore mainstream economists have largely ignored the issue of the origins of secure property rights (by taking them as exogenously given), there is a large literature in political and moral philosophy that is concerned with the origins of the state and conceptions of a just society, which does indirectly (if not directly) concern the issue of the origins of property rights. Early notable contributions to this literature were made by Hobbes (1651), Locke (1690), Hume (1739) and Rousseau (1762). Many key contributions were made in the twentieth century by 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)). Athough this literature contains a wealth of ideas, it does not provide an analytical model of the state-of-nature.

Buchanan (1975) presents an informal framework and analysis of the emergence of property rights. But he does not provide an analysis of the conditions under which property rights can be made secure — which is a crucial issue in understanding the origins of property rights, and which lies at the heart of my model and analyses. In the absence of any third party in a state-of-nature with two players, property rights have to be *self-enforcing* in order to be secure, which requires adopting a dynamic perspective, and constructing a repeated interaction model of the state-of-nature. I should emphasize that most scholars 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. For a thought-provoking recent commentary on issues that impinge on the study of the origins of property rights, see Basu (2000).

I should briefly mention the literature on *static* models of conflict, which, though not directly concerned with the origins of property rights, shares some aspects with my model such as the notion that players may resolve conflict through warfare rather than through peaceful negotiations. Some notable examples from this literature include Skaperdas (1992), Grossman and Kim (1995), Hirshleifer (1995), Esteban and Ray (1999), and Marceau and Myers (2000).

The remainder of the paper is organized as follows. Section 2 lays down my base model and studies its unique stationary subgame perfect equilibrium, which I call the *natural equilibrium.*[1] The private net benefits to the players from establishing the property right in question, which define their respective private incentives to do so, are defined and studied in section 3. The study of the appropriate incentive-compatibility conditions that are required to hold for the property right to be *self-enforcing* (or secure) is the subject of section 4. The focus of the analysis here is to characterize the critical values of the discount factors, and then to analyze the impact on these critical values of the players' fighting and productive skills.[2] I extend the base model in section 5 by providing the players with the option to bargain over inter-player transfers of output. I conclude in section 6.

## 2. The 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 $f_i$ satisfies the following conditions: $f_i(0) = 0$, $f_i'(0) = +\infty$, $f_i' > 0$ and $f_i'' < 0$.

*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 in an attempt to steal the other player's output.[3] 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:

[1]The "natural equilibrium" terminology, which is taken from Buchanan (1975), aptly denotes the (equilibrium) outcome in the absence of secure property rights.

[2]The property right is "self-enforcing" if it is sustainable in a subgame perfect equilibrium of the base model, which occurs *if* the players' discount factors lie above their respective critical values.

[3]It may be noted that my analyses and results are unaffected by the alternative assumption that the players decide sequentially whether or not to fight.

• 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 $p_A + p_B < 1$. 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 $T - L_j$.[4]

• With probability $1 - p_A - p_B$ no one wins the fight and no one steals anything. In this case player $i$'s levels of consumption and leisure in this period are respectively $f_i(L_i)$ and $T - L_i$.[5]

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 assume that $U_i$ takes the following (quasi-linear) form: $U_i(c, l) = c + v_i(l)$, where $v_i$ satisfies the following conditions: $v_i(0) = 0$, $v_i'(0) = +\infty$, $v_i' > 0$ and $v_i'' < 0$.[6] 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 — the *state-of-nature* — 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 skill). Hence, the relevance (in each period) of stage 2 when each player entertains the possibility of stealing the other player's output.

A natural interpretation of the base model is implicit in its formal description: it represents the interaction between two individuals in the state-of-nature. The following alternative interpretation is also applicable: the model represents the interaction in the state-of-nature between two organizations or two groups of individuals (such as two nation-states or two local communities or two regions within a single nation-state). While the former interpretation is perhaps more useful

---

[4]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 *ex-ante* incentives to work.

[5]It may be noted that fighting does not lead to any loss (or destruction) of output. I adopt this assumption partly because it makes the emergence of secure property rights that much harder. I should also emphasize that a key cost of fighting is that it affects the players' *ex-ante* incentives to work. This cost of fighting — which is indirect but fundamental — is a key element of my model and analyses.

[6]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.

from a theoretical perspective, the latter has much relevance to the world in which we currently live.

I use the subgame perfect equilibrium concept (SPE, for short) to analyze the base model.[7] My analysis of the base model — which occupies the remainder of this section and sections 3 and 4 — explores, in particular, the roles of 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 $f_B$) on the emergence and security (or otherwise) of the property rights in question. In section 5 I study an extension of the base model in which the players negotiate over inter-player transfers of output.[8]

2.2. **The Natural Equilibrium.** I begin by characterizing the unique *stationary* SPE of the base model. Since a stationary SPE of the base model is the repeated play of a SPE of the two-stage game $\mathcal{G}$, I now derive the unique SPE of this two-stage game. The following lemma characterizes the unique equilibrium outcome at the second stage (for any possible actions chosen at the first stage):

**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 is not to fight if and only if $p_i f_j(L_j) \le p_j f_i(L_i)$.*

*Proof.* In the Appendix. □

The intuition for Lemma 1 follows by noting that $p_i f_j(L_j)$ 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 $p_i$; and that $p_j f_i(L_i)$ is his expected net loss from the fight — since in the fight he would lose all his output with probability $p_j$.

It follows immediately from Lemma 1 that for any pair $(L_A, L_B)$ chosen at stage 1 player $i$'s (equilibrium) expected payoff is $\Pi_i(L_i, L_j)$,

[7]In order to somewhat simplify the analysis, but without any significant loss of generality, I rule out SPE in which a player uses a weakly dominated strategy.

[8]A player's exogenously given fighting and productive skills may be interpreted as being determined in part from his inherent abilities and in part from (unmodelled) costly investments undertaken at the beginning of time. I should like to note that many of the main qualitative insights obtained in this paper are robust to an extension of my base model in which players have the opportunity to make further investments in such skills. Some further comments on this issue are made in the concluding section.

where[9]

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

which may be interpreted as follows. In the case of a fight, player $i$'s total expected 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 $p_A f_B(L_B) = p_B f_A(L_A)$), he consumes all of his output and none of player $j$'s output. Given (1), it is straightforward to establish the following proposition:

**Proposition 1** (The Natural Equilibrium (NE)). *The base model has a unique stationary SPE. In this natural equilibrium, player $i$ ($i = A, B$) always (in any period) sets $L_i = L_i^N$, where $L_i^N$ is the unique solution to the first-order condition*

$$(1 - p_j)f_i'(L_i) = v_i'(T - L_i) \quad (j \neq i). \tag{2}$$

*Player $i$'s equilibrium expected payoff in each period is $V_i^N = \Pi_i(L_i^N, L_j^N)$, where $\Pi_i$ is defined above in (1).*

It may be noted that the left-hand side of (2) is player $i$'s marginal benefit from working, while the right-hand side is his marginal cost from doing so. Notice 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 $L_i^N$.[10]

[9] This is established as follows. First consider a pair $(L_A, L_B)$ such that $p_B f_A(L_A) = p_A f_B(L_B)$. In that case it follows from Lemma 1 that no fight occurs, and hence the expected payoff to player $i$ is $f_i(L_i) + v_i(T - L_i)$, which, however, equals $\Pi_i(L_i, L_j)$. Now consider a pair $(L_A, L_B)$ such that $p_B f_A(L_A) \neq p_A f_B(L_B)$. In that case it follows from Lemma 1 that a fight occurs, and hence the expected payoff to player $i$ is (after simplifying) $\Pi_i(L_i, L_j)$.

[10] For some parameter values (such as when the players are identical), no fighting occurs in the NE. Interestingly, however, even then the equilibrium levels of work are influenced by the players' fighting skills. This is because even then player $i$'s equilibrium marginal benefit from working is equal to the left-hand side of equation 2, since if he unilaterally deviates and chooses $L_i \neq L_i^N$, then a fight occurs. 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 actually takes place; it is the *fear* of war that ultimately determines each player's incentives to work.

2.3. **Comparative-Statics on Natural Equilibrium Payoffs.** First, notice that $\partial V_i^N/\partial p_j = -f_i(L_i^N)$. That is, a marginal increase in player $j$'s fighting skill decreases player $i$'s NE payoff, and *vice-versa.* Now consider the impact of a marginal change in $p_i$ on player $i$'s NE payoff. I obtain that

$$\frac{\partial V_i^N}{\partial p_i} = f_j(L_j^N) + p_i f_j'(L_j^N)\frac{\partial L_j^N}{\partial p_i}. \tag{3}$$

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 (3), 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 (3), which is strictly negative, may be called the *strategic 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. I now establish that if $p_i$ is sufficiently large, then a marginal increase in $p_i$ decreases $V_i^N$; otherwise, the opposite holds.[11]

**Proposition 2.** *If $f_j''' \leq 0$ and $v_j''' \geq 0$, then there exists $p_i^* \in (0,1)$ such that*

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

*Proof.* In the Appendix. □

In summary, if a 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 a player would have a relatively greater incentive to establish the property rights the smaller is his NE payoff, these comparative-static results suggest that secure property rights are most likely to emerge when both players are sufficiently strong. However, these results also suggest that a player would not wish to be too strong; for otherwise his opponent would have little incentive to produce any output.

Now 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

[11]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.

his total output higher for any level of labour input, but also his marginal product is higher.[12] It is straightforward to show (using (2)) 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. It is straightforward to verify that both players' NE payoffs increase as player $i$ becomes more productive.

## 3. The Costs and Benefits of Basic Property Rights

3.1. **The Property Rights Equilibrium.** In order to derive a player's *private* incentive to establish the property rights in question, I compare his NE payoff with his payoff in the *property rights equilibrium* (PRE, for short), 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 PRE, player $i$ chooses to work for $L_i^F$ units of time, where $L_i^F$ is the unique solution to the first-order condition

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

In the PRE player $i$'s payoff in each period is $V_i^F = f_i(L_i^F) + v_i(T - L_i^F)$.

It is straightforward to show that $L_i^N < L_i^F$ $(i = A, B)$. This "under-investment" result may be 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 affect *ex-ante* incentives to make such investments.

3.2. **Private Incentives.** 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 establishing the property rights, only benefits. However, if for some $i$ and $j$ $(i \neq j)$ $\Delta_i > 0$ and $\Delta_j < 0$, then there is a cost *and* a benefit from establishing these rights; the cost is to player $j$ and the benefit to player $i$. It should be noted that both $\Delta_A$ and $\Delta_B$ cannot be negative, since $\Delta_A + \Delta_B > 0$.

[12] An example is when $f_i(L_i) = \lambda_i L_i$ and $\widehat{f}_i(L_i) = \widehat{\lambda}_i L_i$, where $\widehat{\lambda}_i > \lambda_i > 0$.

A key insight obtained below is that the set of parameter values such that $\Delta_i < 0$ (for some $i$) is *non-empty*. This insight challenges the not uncommon viewpoint that the Prisoners' Dilemma game adequately 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 secure property rights are established. My insight is most transparent when one player is strong but unproductive, while the other player is weak but productive. In such a case the strong player would loose out when the property rights are established (since he would then not be able to steal any output from the other, more productive player).[13] 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 or otherwise of secure property rights.

It is straightforward to show that

(5)

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

If the 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 it follows from Lemma 1 that no fighting occurs in the NE. That is, the right-hand side of the second inequality in (5) is zero. Since, by definition, the left-hand side of the second inequality in (5) is strictly greater than zero, it thus follows that if the players are identical then $\Delta_A > 0$ and $\Delta_B > 0$. This result implies that in order for there to exist conflict in the players' private incentives to establish the basic property rights, 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 $i$. It is trivial to show that the set of parameter values under which conflicting private incentives exist is *non-empty*.[14]

[13] As I discuss in section 5, since the *sum* of the payoffs in the PRE exceeds the *sum* of the payoffs in the NE (i.e., since $\Delta_A + \Delta_B > 0$), secure property rights might be established in this case if, for example, the productive but weak player transfers some output (in each period) to the strong but unproductive player.

[14] For example, if $p_j$ is arbitrarily close to zero and $p_i > 0$ but bounded away from one, then $\Delta_i < 0$. This is because 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

3.3. **Comparative-Statics 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 in question provided that player $i$ is sufficiently strong. Now suppose that *both* players' fighting skills improve. In that case the total change in player $i$'s private incentives is captured by the sum of the partial derivative of $\Delta_i$ *w.r.t.* $p_i$ and the partial derivative of $\Delta_i$ *w.r.t.* $p_j$. I obtain that

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

In general, this expression can be positive or negative. However, if the players are identical, then this expression is strictly positive. This means that if the players are identical, then improvements in *both* players' fighting skills enhances their respective private incentives to establish the property rights.

The *collective* incentive, which is captured by the sum $\Delta_A + \Delta_B$, defines the *surplus* from establishing the property rights. I obtain that

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

which is strictly positive. Thus, the collective incentive to establish the property rights is increasing in each player's fighting skill. 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.

Now suppose that player $i$'s productive skill improves (in the manner formalized above, at the end of section 2.3). 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. However, since player $i$'s productive skills have no effect on player $j$'s PRE payoff, player $j$'s private incentive to establish the property rights diminishes. It is straightforward to show that the collective incentive to establish the property rights is diminished as one (or both) players become more productive.

The comparative-static results obtained above indicate that improvements in productive skills adversely affect incentives to establish property rights, while improvements in fighting skills affects them positively.

---

left-hand side of (5) 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 $L_j^N > 0$.

This insight suggests that in order to promote and maintain such incentives, investments in productive skills should go hand-in-hand with investments in fighting skills.

## 4. On the Emergence of Secure Property Rights

I now derive and analyze the *incentive-compatibility* conditions that are required to hold in order for there to exist a (necessarily, *non-stationary*) SPE whose equilibrium path is identical to the PRE path (in which, in each period, $L_A = L^F_A$, $L_B = L^F_B$ 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. It is worth emphasizing therefore that the property rights in question are made secure (when they can be) in a *self-enforcing* manner, *without* any third-party enforcement.[15]

Straightforward application of appropriate *Folk Theorems* from the *Theory of Repeated Games* would allow us to dispense with this issue without too much fuss. This is because such results can be applied to show that the PRE path is sustainable as a SPE path of the base model. *But*, these results apply *if and only if* players are sufficiently patient. In particular, if and only if players are infinitely patient (i.e., in the limit as both $\delta_A$ and $\delta_B$ tend to one) is it possible to show that for any configuration of the players' fighting and productive skills, the PRE path is sustainable as a SPE path. Although this (limiting) result provides a benchmark result, it is not particularly useful, since the assumption of negligible discounting is not particularly plausible. As such it is important to conduct an analyses when players are not sufficiently patient. The focus of my analyses below is to characterize *and* analyze the critical values of the discount factors (which have the property that the PRE path is sustainable as a SPE path *if* the players' discount factors lie above their respective critical values).

I adopt the well-known "trigger-strategy" approach, in which the PRE path is sustained as a SPE path by moving play to the NE if any player ever (unilaterally) deviates from the PRE path. 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, and thus establish:

[15] Indeed, since there are no third parties in the environment under consideration, any enforcement mechanism can only involve the two players.

**Theorem 1** (Trigger-Strategy Equilibrium (TSE))**.** *The PRE path can be sustained as a SPE path (using the trigger strategies) if and only if for each* $i = A, B$*,* $\delta_i \geq \underline{\delta}_i$*, where*

$$\underline{\delta}_i = 1 - \frac{\Delta_i}{p_i[f_j(L_j^F) - f_j(L_j^N)]} \quad (j \neq i). \tag{6}$$

*Proof.* In the Appendix.[16] □

As would be expected, the critical values of the discount factors $\underline{\delta}_A$ and $\underline{\delta}_B$ depend, in particular, on the players' fighting and productive skills. Notice that if the players' private incentives to establish the property rights are in conflict — that is, $\Delta_i < 0$ for some $i$ — then $\underline{\delta}_i > 1$ and hence, it follows from Theorem 1 that the PRE path cannot be sustained as a SPE (using the trigger strategies). Indeed, the TSE exists *only if* the parameters are such that both players *strictly* prefer the PRE over the NE — that is, $\Delta_A > 0$ and $\Delta_B > 0$.[17]

It is straightforward to verify that $\underline{\delta}_i$ is decreasing in $p_j$. 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 property rights); this is because a marginal increase in $p_j$ reduces $V_i^N$, 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 $i = A, B$,

$$\frac{\partial \underline{\delta}_i}{\partial p_i} \gtreqqless 0 \iff (1 - \underline{\delta}_i) f_j(L_j^F) + \underline{\delta}_i \frac{\partial V_i^N}{\partial p_i} \gtreqqless 0 \ \ (j \neq i). \tag{7}$$

Proposition 2 above states that player $i$'s NE payoff $V_i^N$ is not monotonic in $p_i$, and hence, it immediately follows (from (7)) that $\underline{\delta}_i$ is not monotonic in $p_i$ either. However, using Proposition 2, it is straightforward to establish that if $f_i''' \leq 0$ and $v_i''' \geq 0$, then there exists

[16] Player $i$'s incentive constraint $\delta_i \geq \underline{\delta}_i$ is equivalent to $\delta_i \Delta_i \geq (1 - \delta_i)[\Pi_i(L_i^N, L_j^F) - V_i^F]$, which is interpreted as follows. The left-hand side of this inequality is his (long-run) average cost by deviating from the PRE path; this is because from *next* period onwards, his *per-period* loss is $V_i^F - V_i^N$ ($\equiv \Delta_i$). The right-hand side is his (short-run) average benefit from the (optimal) deviation; this is because his (one-period) gain from this deviation is $\Pi_i(L_i^N, L_j^F) - V_i^F$.

[17] Perhaps not surprisingly, the incentive-compatibility conditions which ensure that the property rights are *self-enforcing* are much more severe than the conditions which ensure that the players do have private incentives to establish such property rights.

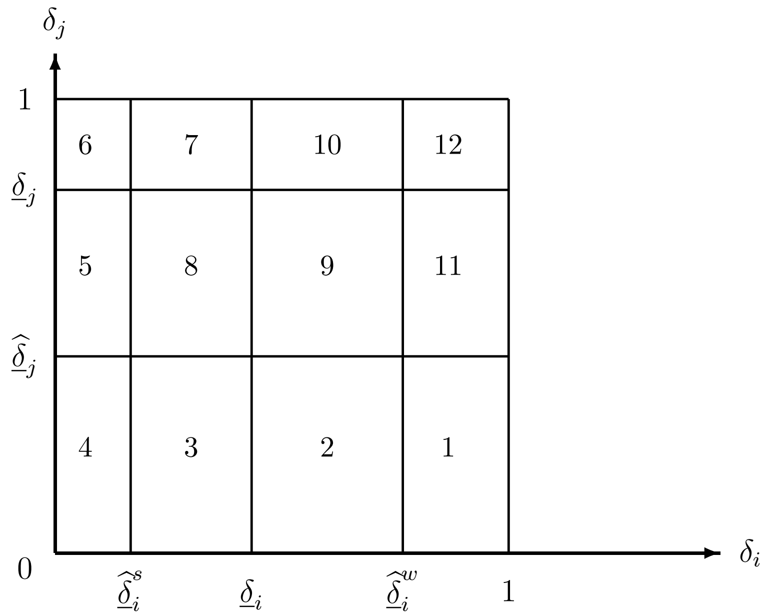


FIGURE 1. An illustration of the effect of a small increase in $p_i$ on the existence of the TSE.

$\widehat{p}_i \in (0,1)$ such that

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

In summary, as player $i$'s fighting skill increases, 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 comparative-static 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. 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.[18] 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 $p_i$.

In regions 7–9 the parameters are such that initially $\delta_i$ and/or $\delta_j$ 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 fighting skill of a 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 fighting skill of a 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.

Now consider an improvement in a player's productive skills (as formalized above, at the end of section 2.3). Since an improvement in player $i$'s productive skill enhances his private incentives (i.e., increases $\Delta_i$), it immediately follows from (6) 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 that $\Delta_j$ 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

[18]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 above.

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.

Since, in fact, the base model has a multiplicity of SPE, one cannot rule out the possibility that the PRE path is sustainable for a relatively wider set of parameter values by some SPE other than the TSE. However, rather than pursue that issue here, I now conclude this section by characterizing a set of parameter values under which there does not exist an SPE that sustains the PRE path; that is, for such parameter values secure property rights can *never* emerge.[19]

**Theorem 2** (Non-Emergence of Secure Property Rights)**.** *If the parameters are such that either $\delta_A < \underline{\delta}_A^*$ or $\delta_B < \underline{\delta}_B^*$, then there does not exist an SPE of the base model in which the equilibrium path is the PRE path, where*

$$\underline{\delta}_i^* = 1 - \left[ \frac{\Delta_i + p_i f_j(L_j^N)}{p_i f_j(L_j^F)} \right] \quad (j \neq i). \tag{8}$$

*Proof.* In the Appendix. □

It may be noted that, not surprisingly, for any parameter values, $0 < \underline{\delta}_i^* < 1$ and $\underline{\delta}_i^* < \underline{\delta}_i$, where $\underline{\delta}_i$ is defined in Theorem 1. Using (3) and Proposition 2, it can be shown that there exists $p_i'$ such that

$$\frac{\partial \underline{\delta}_i^*}{\partial p_i} \gtreqless 0 \quad \text{if} \quad p_i \lesseqgtr p_i', \qquad \textit{and} \qquad \frac{\partial \underline{\delta}_j^*}{\partial p_i} \lesseqgtr 0 \quad \text{if} \quad p_i \lesseqgtr p_i'.$$

Thus, a marginal change in player $i$'s fighting skill has, in general, an ambiguous effect on the range of discount factors under which secure property rights can never emerge. 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 $\underline{\delta}_A^* > \underline{\delta}_B^*$. It follows from Theorem 2 that in this case, secure property rights can never emerge if the players' common discount factor $\delta < \underline{\delta}_A^*$. Hence, it follows from the above comparative-static results that an improvement in player $A$'s fighting skill increases the likelihood that secure property rights

[19] In the next section I return to the issue of whether or not the PRE path can be sustained for a relatively wider set of parameter values than for which the TSE exists, but by allowing for inter-player transfers of output.

can never emerge *if* player $A$ is weak; while the opposite is the case *if* player $A$ is strong. This insight reinforces the following key message of this paper: *an improvement in the fighting skill of a weak player is inimical to the emergence of secure property rights, while the opposite is the case with an improvement in the fighting skill of a strong player.*

## 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 negotiate over whether or not to establish the property rights under consideration. The key aspect of such negotiations is that a player can offer the other player (or, as the case may be, to demand from him) some output in return for agreeing to establish these 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 the negotiations. A main objective of the analysis to follow is to explore the extent to which such inter-player transfers of output enhance 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 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 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 per-period output to the other player. Let the *game form* that formally encapsulates this bargaining process be denoted by $\phi$, which has two types of outcomes: (i) the players reach an agreement $t \in \Re$, with the interpretation that in each period, player $A$ gives $t$ units of output to player $B$ and both players will not fight,[20] 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, 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.[21]

I assume that if the players fail to reach an agreement in $\phi$, 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, if the players do not reach an agreement in $\phi$, then player $i$'s payoff in each period is $V_i^N$.

5.2. **Existence of Bargaining Equilibrium with Agreement.** An agreement in $\phi$, 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

[20]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 $A$.

[21]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 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 negotiate only at the beginning of time.

unilaterally deviates from this extended PRE path (i.e., violates the agreement), then immediately play proceeds according to the NE.[22] Through a straightforward extension of the arguments used in the proof of Theorem 1, I obtain that an agreement to establish the property rights with player $A$ giving $t \in \Re$ units of output to player $B$ in each period is *incentive-compatible* (or self-enforcing) if the following two inequalities hold:[23]

$$t \leq \Delta_A - p_A(1-\delta_A)[f_B(L_B^F) - f_B(L_B^N)] \tag{9}$$

$$t \geq p_B(1-\delta_B)[f_A(L_A^F) - f_A(L_A^N)] - \Delta_B. \tag{10}$$

Since it is straightforward to verify that an agreement $t$ is incentive-compatible *only if* it is feasible in the sense that $f_A(L_A^F) \geq t \geq -f_B(L_B^F)$, and Pareto-dominates disagreement in the sense that $V_A^F - t > V_A^N$ and $V_B^F + t > V_B^N$, it follows that if there *exists* an incentive-compatible agreement, then agreement to establish secure property rights is struck.[24] It thus follows immediately that:

**Theorem 3** (Existence of a Bargaining Equilibrium with Agreement)**.** *In the extended base model with bargaining, the players strike an agreement to establish the property rights in question with some inter-player transfer of output $t^*$ if and only if the parameters are such that*

$$\Delta_A + \Delta_B \geq p_A(1-\delta_A)[f_B(L_B^F) - f_B(L_B^N)] + p_B(1-\delta_B)[f_A(L_A^F) - f_A(L_A^N)]. \tag{11}$$

Since the right-hand side of (11) is strictly positive, this means that in order for agreement to be struck, the *surplus* (or collective incentive), which is the left-hand side of (11), must be sufficiently large; it is not enough that there just exists some surplus, i.e., $\Delta_A + \Delta_B > 0$ — which,

[22] 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.

[23] These two inequalities are respectively player $A$'s and player $B$'s incentive constraints. For example, the first of these inequalities is player $A$'s incentive constraint, which can rewritten as $\delta_A(\Delta_A - t) \geq (1-\delta_A)[\Pi_A(L_A^N, L_B^F) - (V_A^F - t)]$, which is interpreted as follows. The left-hand side of this inequality is his (long-run) average cost by deviating from the extended PRE path; this is because from *next* period onwards, his *per-period* loss is $(V_A^F - t) - V_A^N$ $(\equiv \Delta_A - t)$. The right-hand side is his (short-run) average benefit from the (optimal) deviation; this is because his (one-period) gain from this deviation is $\Pi_A(L_A^N, L_B^F) - (V_A^F - t)$.

[24] This is based on the (plausible) assumption that the equilibrium outcome in $\phi$ is Pareto-efficient.

recall, holds for any parameter values.[25] Not surprisingly, (11) holds in the limit as both $\delta_A$ and $\delta_B$ tend to one, but not when $\delta_A = \delta_B = 0$. The key issue, however, is whether or not it holds for a relatively wider range of parameter values than for which the TSE exists. It is straightforward to show that there exists parameter values under which the TSE does not exist, but the bargaining equilibrium with agreement does exist.[26] 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 as when one player is strong but unproductive, while the other player is weak but productive), inter-player transfers of output can be a mechanism through which secure property rights get established.

Stated more generally, Theorem 3 implies 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 (to the emergence of secure property rights) 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: Characterization of the Agreement.** The insights derived above 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 $\phi$ is Pareto-efficient. Theorem 3 states the condition on the parameters under which the bargaining equilibrium with agreement exists. But it is silent on the characteristics of the equilibrium negotiated transfer. This is not surprising since that depends on the details of the bargaining process $\phi$. I now assume that the bargaining process is such that the equilibrium

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

[26] It is of course trivial to verify that if a TSE exists then the bargaining equilibrium with agreement also exists.

negotiated transfer $t^*$ can be characterized by the Nash bargaining solution (NBS, for short) with the *disagreement* point $(V_A^N, V_B^N)$.[27]

Ignoring for a moment that the equilibrium negotiated transfer must be incentive-compatible (i.e., $t^*$ must satisfy (9) and (10)), the NBS is in general captured by the *Split-the-Difference* rule, whose application here implies that player $i$ $(i = A, B)$ receives a utility payoff $u_i$ that equals his disgreement payoff $V_i^N$ plus one-half of the surplus $\Delta_A + \Delta_B$. Letting $\widehat{t}$ denote the associated transfer, this means that

$$u_A \equiv V_A^F - \widehat{t} = V_A^N + \frac{\Delta_A + \Delta_B}{2} \quad \text{and} \quad u_B \equiv V_B^F + \widehat{t} = V_B^N + \frac{\Delta_A + \Delta_B}{2}.$$

This, in turn, implies that

$$\widehat{t} = \frac{\Delta_A - \Delta_B}{2}.$$

It follows immediately from the definition of the NBS that the equilibrium negotiated transfer $t^* = \widehat{t}$ *provided* that $\widehat{t}$ is incentive-compatible, which (after substituting for $\widehat{t}$ in (9) and (10), and simplifying) means that $t^* = \widehat{t}$ if the parameters are such that

$$\Delta_A + \Delta_B > 2 \max\{\xi_A, \xi_B\},$$

where $\xi_i = p_i(1 - \delta_i)[f_j(L_j^F) - f_j(L_j^N)]$. Thus, when the surplus is sufficiently large, then the NBS implies that the equilibrium negotiated transfer is made by the player whose private incentive (or net benefit) from establishing the property rights in question is relatively higher. The exact level of the transfer depends on the amount by which his net benefit exceeds that of the other player.

Now suppose that $\widehat{t}$ is *not* incentive-compatible. This can happen because this level of transfer is either too high or too low, in the follow senses. It is too high (which, more precisely, means it fails to satisfy player $A$'s incentive constraint (9)) in the sense that agreement to it would give player $A$ a utility level that lies below what he can obtain from reneging on this agreement. Symmetrically, it is too low (which, more precisely, means it fails to satisfy player $B$'s incentive constraint (10)) in the sense that agreement to it would give player $B$ a utility level that lies below what he can get by reneging on this agreement. In these cases, the NBS is a corner solution. In the former case, $t^* = \Delta_A - \xi_A$, which is the equilibrium agreed transfer when the parameters are such that $\xi_A + \xi_B < \Delta_A + \Delta_B < 2\xi_A$; while in the latter case, $t^* = \xi_B - \Delta_B$, which is the equilibrium agreed transfer when the parameters are such that $\xi_A + \xi_B < \Delta_A + \Delta_B < 2\xi_B$.

[27]As is well-known — see, for example, Muthoo (1999) — this may be justified by assuming that $\phi$ is the Rubinsteinian, alternating-offers bargaining process.

## 6. Summary and Concluding Remarks

A main 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.

The second main contribution lies in the results I have obtained. Some of the most fundamental insights provided by these results 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 strong but unproductive, while the other player is weak but productive — under which the players' private incentives are in conflict.

• In order to promote and maintain incentives, 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 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, resorting to the mechanism of *inter-player transfers of output* may be required for the property rights to emerge and be made secure.

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

• Establishing secure property rights is "costly", in the sense that costly but unproductive investments in fighting skills are required for their emergence and security.

I now turn to a discussion of the main limitations of my model as a model of the state-of-nature with two players.[28] First, a simplifying but somewhat restrictive assumption that underlies my model is that the players cannot make any (further) investments in their fighting and productive skills. Although many of my main qualitative insights would be robust to allowing for such investments, it would be useful to formally address this issue partly because new insights may be obtained. 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

[28] I should like to note that addressing some of these limitations — by extending and/or modifying my model — should 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. In Muthoo (2002), 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.

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 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 Lemma 1.** Consider 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 (after substituting for the assumed quasi-linear form of player $i$'s utility function) that

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

Without loss of generality, but 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: if $E_i^f \leq E_i^{nf}$ then player $i$'s weakly dominant action is not to fight, while if $E_i^f > E_i^{nf}$ then player $i$'s weakly dominant action is to fight.

**Proof of Proposition 2.** It follows from (2) that

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

Hence, after substituting for the derivative of $L_j^N$ with respect to $p_i$, and then simplifying, (3) becomes:

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

Notice that[29]

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

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

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

Under the hypotheses of Proposition 2 the derivative of $\Upsilon$ 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 $p_i$.

**Proof of Theorem 1.** The following argument establishes the *incentive-compatible* condition under which player $i$ cannot benefit from a (*one-shot*, unilateral) deviation from the PRE path. Fix an arbitrary period $n$ (where $n = 1, 2, \ldots$), 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

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

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 $\delta_i \Delta_i \geq (1-\delta_i)[\Pi_i(L_i^F, L_j^F) - V_i^F]$. Since $\Pi_i(L_i^N, L_j^F) > \Pi_i(L_i^F, L_j^F)$, it follows that this inequality is implied by (A.1). Finally, note 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 A.1 becomes (after substituting for $\Pi_i(L_i^N, L_j^F)$, using this expression, and then simplifying) $\Delta_i \geq p_i(1-\delta_i)[f_j(L_j^F) - f_j(L_j^N)]$. Theorem 1 then follows immediately.

[29]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 (7).** 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 (7)) follows immediately (since $p_i \tau_j > 0$).

**Proof of Theorem 2.** Letting $\underline{V}_i$ denote the *worst* SPE (average) payoff to player $i$, it is straightforward to verify that if 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 PRE path is the equilibrium path:

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

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,

$$V_A^F < (1 - \delta_A) \Pi_A(L_A^N, L_B^F) + \delta_A \underline{V}_A. \tag{A.2}$$

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 A.2. 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 A.2, it is optimal for player $A$ to conduct the one-shot, unilateral deviation. But this is a contradiction.

I now derive player $i$'s 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

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

Theorem 2 follows immediately since $\underline{V}_i \geq \underline{w}_i$.

## References

**Basu, Kaushik** (2000), *Prelude to Political Economy: A Study of the Social and Political Foundations of Economics*, Oxford: Oxford University Press.

**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.

**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.

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

**Esteban, Joan** and **Debraj Ray** (1999), "Conflict and Distribution", *Journal of Economic Theory*, **82**, 379-415.

**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).

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

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

**Marceau, Nicolas** and **Gordon M. Myers** (2000), "From Foraging to Agriculture", unpublished mimeo, Simon Fraser University.

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

**Muthoo, Abhinay** (2002), "Coalition Formation in the State-of-Nature", In preparation, University of Essex.

**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.

**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.

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

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

Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ

*E-mail address*: muthoo@essex.ac.uk