On the Origins of Property Rights: Conflict and Production in the State of Nature

Catherine Hafer
Dept. of Politics, NYU

this version: October 2003

Abstract

I analyze the emergence of property rights in a model of conflict and production in the absence of institutions enforcing property rights. $n$ heterogeneous agents, who are drawn from a two-dimensional continuum which defines their corresponding productive abilities with and without the scarce resource, allocate their time between producing with/without the resource or fighting for its possession with other randomly matched agents. The game is an incomplete-information, infinite-period continuous-time war of attrition. I prove the existence of equilibrium and present its characterization. I show that in the steady-state of the game, attained in finite time, the population of agents converges to an equilibrium, in which agents have been sorted into two stable groups (resource "haves" and "have-nots") in accordance with the known function of their primitive types. Once the separation is sufficiently complete, no agent finds it in her interest to challenge another agent for possession of the resource, and so the security of possession aspect of property rights over the resource is established. However, I show that the pattern of ownership that emerges in the steady-state is stably inefficient, in that the introduction of trade would undermine not only the transferability, but also the security of possession.

1 Introduction

The critical importance of economic and political institutions for economic performance is one of the best known lessons of contemporary institutional analysis in political economy [e.g., North 1990, Barro 1997, Rodrik 1999]. One of the earliest entries on the long list of such performance-enhancing institutions is property rights, historically regarded as the institutional core of market economy [e.g., Davis and North 1971, North and Thomas 1973]. The development of a rigorous theory of property rights, and in particular, of the emergence of property rights, has, however,
lagged well behind the analysis of market interactions, and, in a number of important respects, continues to be nascent.

This paper contributes to the development of such a theory by analyzing a dynamic model of the emergence of property rights in a decentralized political and economic environment, i.e., an environment in which there are no external authorities or pre-existing socio-political institutions that define and enforce property rights and which bears considerable resemblance to the Hobbesian “state of nature.” Its distinguishing feature is the causal and analytical prominence of costly conflict in determining control rights over scarce inputs to production. Analytically, this determination supplies two critical elements for the analysis of the emergence of property rights: the nature of the conflict and production equilibrium, and the allocation of factor goods across agents as a function of their primitive skills. In addition, the model and results constitute a novel contribution to the literature on incomplete information second-price all-pay auctions (wars of attrition). Unlike extant models, an agent’s type consists of both her valuation of the prize and her cost of competing for it. These two dimensions of an agent’s type are independently drawn and both are private information. I completely characterize the solution to the two-player game when the distribution of types is the same for both players (i.e. when the players have common beliefs) and prove existence and some properties of the equilibrium when the players’ beliefs differ. I then present a large-population infinite-horizon dynamic game in which players are randomly matched in each period to play the war of attrition and, as a consequence of those interactions, the distribution of types among the winners and the losers changes with each round of conflict.

I show that in the steady-state equilibrium of this game, attained in finite time, the population of heterogeneous agents separates into two permanent groups (of resource-“haves” and “have-nots”) in accordance with a known function of their primitive types. Moreover, once the separation is sufficiently complete, no agent finds it in her interest to challenge another agent for possession of the resource, i.e., in the steady-state equilibrium, there is no conflict and the agents enjoy security of property. These results, which depart from the Hobbesian sensibilities of “perpetual war of every man against every man” shared by the recent economic literature on conflict and appropriation, suggest a causal model for the emergence of property rights proper. Furthermore, the allocation of factor goods in the steady-state equilibrium is inefficient, and although agents enjoy security, they cannot readily exploit gains from trade (security of property, thus, does not imply its alienability).

From the standpoint of constructing a theory of property rights, the model with the decentralized\(^1\) political and economic context has a key relative advantage over the centralized (third-part enforcer) model in giving prima facie theoretical priority to an integrated and systematic explanation of the existence of the enforcers of contracts and rights and their incentives in the model economy. The process whereby the political mechanism enforcing property rights itself comes about is a mystery.

---

\(^1\)Here and throughout, I am using “decentralized” as synonymous with “self-enforcing.”
too complex to assume away by fiat. The assumption of decentralized enforcement also underscores the empirical relevance of the present model for studying developing economies with weak or ineffective governments and instances where external enforcers lack information about or the authorization of the direct participants.\(^2\) Not only are such descriptions of the role of government fair depictions of the status quo in a large number of less developed countries, but many of these countries have managed to develop complex and relatively stable networks of customary property rights [e.g., Sub-Saharan and West Africa - see Ensminger 1992, Firmin-Sellers 1996, Kapilow and Shavell 2001] or conduct large proportions of economic transactions in the informal economic sector [e.g., post-communist and post-authoritarian transition economies - see De Soto 1989, Frye and Shleifer 1997, etc.]. Similar conditions characterized various economic domains of medieval and early modern European societies and the economies of the American frontier as late as the first part of 19th century [see, e.g., Berman 1983, De Roover 1965, Umbeck 1977, Libecap 1978]. What enables economic transactions and what conditions support the norms of property exploitation under such circumstances cannot be addressed within the third-party enforcer model.

Given the assumption that there is no third-party enforcer of rights, any emergent property right must be self-enforcing. Repeated game-theoretic models of contractual transactions [see e.g., Milgrom, North and Weingast 1990, Greif 1994a and b, Calvert 1995, Kranton 1996a and b; Rauch and Watson 1999, etc.] adopt the notion of an institution as an equilibrium, and so permit the analysis of the self-enforcing aspects of institutions in terms of agents’ economic attributes and other primitive features of the economic environment. However, they assume that the economy is characterized by the existence of at least control rights (or security of possession), with the subject of analysis inevitably shifting to the nature of exchange. From the standpoint of developing a theory of the emergence of property rights, this move implicitly puts the “rabbit” of the market (however imperfect) into the hat, while imposing a \textit{prima facie} constraint on individually rational behavior.

“State of nature” models of conflict and production [Skaperdas 1992; Grossman and Kim 1995; Skaperdas and Syropoulos 1996; Muthoo 2003] capture the basic notion of an “institutionless” state of nature, in that there is no enforcement of claims to property except individually provided coercive force, and thus redistribution can be achieved through aggression. This approach, thus, provides the appropriate context for the study of the emergence of property rights. Previous models pursuing this approach are two-agent complete information models, in which agents simultaneously and irrevocably divide their endowments between producing one or more consumption goods and increasing their ability to appropriate such goods, i.e., increasing their success in conflict, as determined by a commonly known conflict technology function [Dixit 1987; Hirshleifer 1991; Skaperdas 1991, 1992; Grossman and Kim

\(^2\) See, for example, the empirical studies of institutions in such contexts by de Soto [1989], Ellickson [1990], Ostrom [1990], and Umbeck [1977, 1981]. The theoretical approaches explicitly addressed to these contexts include Shleifer [1994] and Skaperdas and Syropoulos [1995].
or between producing a consumption good and leisure in expectation of conflict with exogenous probabilities of winning [Muthoo 2003]. In all cases, actions and the probabilities of winning are common knowledge before conflict ensues.\(^3\)

Muthoo [2003] combines a “state of nature” model with the repeated games approach to studying institutional stability. The explicit dynamic element of his model allows him to determine the conditions under which a de facto property right is stable, i.e. under which each player is able to keep, in its entirety, the output of her own production. His analysis of the subgame perfect equilibria of that game provides sufficient conditions for the emergence of security of property as well as conditions that guarantee that no equilibrium supporting such rights exists. However, the use of an infinitely-repeated stage game, in which the strategic environment is identical in each period, prevents the examination of the dynamically evolving cumulative effects of conflict on the distribution of resources across players and the consequences of that change for the emergence of property rights.

The present paper analyzes the effects of just such an evolving dynamics in relation to the control of physical capital inputs to production - a setting where the importance of such dynamics is, intuitively, most immediate. Whereas the models discussed above address conflict over the (re-)allocation of a consumption good, conflict over the control of factor goods is, at once, theoretically unexplored and empirically prominent [see, for example, de Soto 1989, Umbeck 1981]. The distribution of the factor good across agents at the beginning of each period reflects the outcomes of the previous period’s conflicts. Because agents dispute the possession of capital goods, rather than consumption goods, the outcomes of conflicts in one period affect the expected payoffs that agents face in the next period. This evolving dynamic element of the model proves critical in determining the emergence of property rights. Modeling it explicitly in the context of conflict over and production with factor goods permits the direct analysis of changes in the allocation of time to conflict and in the posterior distributions of goods. In so doing, it provides a natural way to explore the emergence of de facto property rights, intimated by Skaperdas and Syropoulos [1995] and Sugden [1986] and allows a measure of the extent of (unrealized) potential gains from trade, and so of the long-term effects of the inability to secure property and, separately, of the inability to exchange it.\(^4\)

In order to be able to study the changing distribution of property rights over

---

\(^3\)In a variation on this approach, Grossman and Kim [1995] employ a model of conflict technology that distinguishes between offensive and defensive measures, and the capital stock, rather than the consumption good, is the prize. The two agents have the same production function, however, and their interactions are not repeated, so the redistribution of capital stock resulting from conflict is moot.

\(^4\)This distinction comports with the “old” and “new” institutionalist [e.g. Coase 1960, Libecap 1989, Eggertsson 1990, Furubotn and Richter 1997] economic analyses of the multi-facetedness of property rights (security of possession, freedom of voluntary transfer or alienation, and use rights) as a socio-economic phenomenon.
the physical resource and to provide potential gains from trade, making the efficient distribution of factor goods a meaningful issue, this model incorporates several features that distinguish it from the models of conflict over consumption goods discussed above. First, unlike the extant models, I assume that the population of agents is large and heterogeneous in two different types of abilities. These abilities pertain to two different modes of production, one of which requires a complementary capital input ("land") and one of which requires only labor. Conflict occurs over the possession and use of land, and the variation in agents' skills, combined with the complementarity of land and labor, ensures that agents value the prize differently. Further, because skill in using one technology is uncorrelated with skill in using another, the opportunity cost of devoting time to conflict is uncorrelated with the agents' valuation of the prize.

Second, I model conflict over land as simultaneous incomplete information wars of attrition between randomly matched pairs of players. Both agents attempt to claim the good by committing their time to "conflict" over it; the conflict ends when one of the agents surrenders her claim to the prize by devoting her time to something else. The remaining player wins the prize and can then also devote her time to another activity (e.g. production). This model of conflict contrasts with the "conflict technology" approach [Hirshleifer 1991] in several respects. First, even if the winning agent is willing to commit more resources to the conflict, she only pays the loser's bid; superfluous resources devoted to conflict can be diverted to production. Second, each agent knows her own type and the distribution of types from which the population is drawn; agents do not know their opponents' types, and therefore do not know how long their opponents will fight. The winner cannot be determined ex ante. The importance of uncertainty about other players' types and of the costliness of discovering that information through conflict becomes evident in the dynamic setting.

The introduction of types that vary on two dimensions rather than one complicates the usual incomplete information war of attrition. I prove the existence of a stationary perfect Bayesian equilibrium and a partial characterization of its properties; for the one-shot game, I provide a complete solution. The closest extant model is that of Fudenberg and Tirole [1986], in which firms are heterogenous in both the value they place on winning and in their costs of competing, as in the present work, but in which each firm's value of winning and cost of competing is derived from the same underlying feature of the firm. Assuming such a relationship between the costs of conflict and the rewards of winning is unwarranted here, because in the state of nature, the means of acquiring land is entirely independent of the means of using it. Thus the model presented here is more general in the sense that the costs of engaging in conflict need not be correlated to the value of the prize; indeed, here they are taken to be independent. Krishna and Morgan [1997] find sufficient conditions

---

5 Sugden [1986] examines a one-shot war of attrition game as one of several possible metaphors for conflict in the state of nature, but does not discuss the choice to engage in conflict as opposed to production.
for equilibrium existence in an incomplete information war of attrition when player’s
types are affiliated and symmetrically distributed. The game presented here violates
both conditions, and in particular, because of the dynamic evolution of the distribu-
tion of the physical resource across types, the distribution of types is asymmetric in
all but the first period of play. Athey [2001] obtains sufficient conditions for equi-
librium existence in a variety of incomplete information games, including auctions,
but restricts her attention to auctions in which the value of winning is independent
of one’s opponents’ actions. This assumption is not satisfied in the war of attrition,
in which the winner must pay the loser’s bid. The proof of existence in the dynamic
game with two-dimensional types given here follows the limiting approach taken by
Athey [2001].

The remainder of the paper is organized as follows. Section 2, which presents the
analysis of the model, begins with the consideration of the model of individual conflict,
a variant of the one-shot incomplete-information war of attrition game, in isolation
(Section 2.1). This game is not a stage game, since the evolution of the distributions
of agent types among those who hold land and those who do not changes the payoffs
agents face from period to period. However, the sequence of play described in the one-
shot game is identical to the sequence of play repeated in each period of the dynamic
infinite sequential game, which is, then, analyzed in Section 2.2. In Section 3, I
discuss the implications of these results for understanding the emergence of property
rights and the roles of these institutions in promoting aggregate economic production.
Section 4, then, contains the formal proofs of the results in the paper.

2 The Model

2.1 The One-Period Game

Nature selects a fixed population of agents from a two-dimensional continuum of
types and endows some of them, chosen randomly, with identical, indivisible parcels
of land. Agent $i$’s possession of land is indicated by $l_i = 1$, where $l_i \in \{0, 1\}$. To
further simplify the initial analysis of the game, I assume here that exactly half the
agents possess land; I relax this assumption in the infinite sequential game. Agent
$i$’s type is characterized by the vector $(\alpha_i, \beta_i)$, where $\alpha_i$ is the marginal productivity
of the agent’s labor using technology $A$, and $\beta_i$ is the marginal productivity of the
agent’s labor in technology $B$.

Agents know the distribution of types from which the population is drawn, but
they do not know the distribution of types in the realized finite population. Agents
believe that the distribution of types in the population mirrors the distribution of
types from which the population was drawn: $\alpha$ and $\beta$ are independently distributed
over the intervals \([a, b]\) and \([c, d]\), respectively, with constant probability densities

\[
p(\alpha) = \frac{1}{b - a} \\
p(\beta) = \frac{1}{d - c}.
\]  

(1)

Each agent is endowed with \(T\) units of time per period, \(t\) of which can be allocated to appropriating or securing land, and \(T - t\) of which to producing the consumption good. Two production technologies are available. Only technology \(A\) requires land; both \(A\) and \(B\) require labor, although they require labor of different kinds, as signified by the distinction between \(\alpha\) and \(\beta\).

\[
A(\alpha, t, l) = \alpha l(T - t) \\
B(\beta, t) = \beta(T - t)
\]  

(2)

The lowest possible marginal productivity with technology \(A\) is no greater than the lowest possible marginal productivity with \(B\), \(a \leq c\), and the highest possible marginal productivity with \(A\) is greater than that with \(B\), \(b > d\).

At the beginning of each period, Nature randomly matches agents who do and do not possess land. The agent who does not possess land may attempt to acquire it by taking it from the other agent; in the absence of any institutions to enforce the claim of one or the other to the land, the agents must settle the dispute through open conflict, here modeled as a war of attrition, or second-price all-pay auction. Each agent chooses her strategy \(t\), the maximum amount of time she is willing to commit to an effort to obtain land, i.e. her bid for the land. The agents invest their time in the conflict starting at the beginning of the period. When one agent quits, the other agent is able to observe her opponent’s exit and stops devoting time to the conflict; thus the winner pays her opponent’s bid, which is less than her own. After the possession of land is settled, agents devote their remaining time to production. Because technology \(A(\alpha, t, l)\) requires land, only the winner can use it. Note that only agents for whom \(\alpha > \beta\) have incentives to engage in contests for land; agents with \(\alpha \leq \beta\) concede defeat immediately.\(^6\) It is assumed that parcels of land cannot be divided among agents and that there are no further challenges to the winner’s possession of the land until the next period, when Nature randomly matches landholders and landless agents again.\(^8\)

---

\(^6\)The substantive results are robust to relaxing this assumption, which is made purely for the sake of mathematical tractability.

\(^7\)If Nature endows an agent for whom \(\alpha < \beta\) with land and then matches that agent with another individual for whom \(\alpha < \beta\), the land lies fallow for the first period. Both agents use technology \(B()\). The final distribution of the land between them is irrelevant.

\(^8\)The assumption that these parcels are of a fixed size should be understood in conjunction with the form of the production functions. The implicit idea embedded in the form of these assumptions is that having more land reduces the amount of labor per unit of land, producing the same outcome. A single individual with a limited amount of labor farms her plot; increasing the size of the plot
Agents obtain utility only from consumption; agents desire land only as a means of producing the consumption good. No saving is possible. The payoff for agent $i$ when matched against agent $j$ is

$$u_i = \begin{cases} 
\beta_i(T - t_i) & \text{if } t_i < t_j \\
\alpha_i l_i(T - t_j) & \text{if } t_i > t_j
\end{cases}$$

(3)

2.1.1 Bayesian Equilibrium in the War of Attrition

Agent $i$ chooses the maximum amount of time that she is willing to commit to an attempt to obtain land, $t_i$. Letting $p(t)$ represent the expected probability that her randomly selected opponent, agent $j$, will quit at time $t$, her optimal strategy, given her own type $(\alpha_i, \beta_i)$ and the strategies played by other agents in the economy, satisfies

$$t_i \in \arg \max \left( \Pr(t_j > t_i) [\beta_i(T - t_i)] + \int_0^{t_i} p(t) [\alpha_i l_i(T - t)] dt \right).$$

Let $t(\alpha, \beta)$ denote the symmetric Bayesian equilibrium of this game, i.e. the solution to this optimization problem for every agent in the economy, and let $p(\alpha, \beta)$ denote the expected probability density of type $(\alpha, \beta)$ in the population of agents, which is the probability density of type $(\alpha, \beta)$ in the distribution from which the population was drawn. Then agent $i$’s optimization problem can be rewritten as

$$\max_{t_i} \Pr \left( (\alpha_j, \beta_j) : t(\alpha_j, \beta_j) > t_i \right) [\beta_i(T - t_i)]$$

$$+ \int_0^{t_i} p((\alpha_j, \beta_j) : t(\alpha_j, \beta_j) = t) [\alpha_i l_i(T - t)] dt).$$

The first order condition is

$$p((\alpha_j, \beta_j) : t(\alpha_j, \beta_j) = t_i) [\alpha_i l_i(T - t_i) - \beta_i(T - t_i)]$$

$$- \beta_i \Pr((\alpha_j, \beta_j) : t(\alpha_j, \beta_j) > t_i) = 0.$$  

The symmetric strategy $t(\alpha, \beta)$ cannot be solved from this expression because multiple—in fact, infinitely many—vectors $(\alpha, \beta)$ correspond to a given $t$. A key step in solving this problem is finding the expression that relates $\alpha$ to $\beta$ along level curves of $t$. Essentially, for two individuals pursuing the same strategy $t$ in equilibrium, we must understand how their differences in $\beta$ compensate for their differences in $\alpha$. Agents are willing to commit resources to obtaining land if the product of that land

---

reduce the amount of labor she can devote to each unit of land, reducing the per unit productivity of land. These assumptions are consistent with the economic activities described by Umbeck [1977], de Soto [1989] and others.
is great enough to compensate them for the resources expended, the value of which is measured in terms of their next-best use. Thus the relevant relationship between \( \alpha \) and \( \beta \) is the ratio of the additional product per unit of labor with land to the opportunity cost of obtaining land. I denote this ratio by \( k \).

\[
 k_i = k(\alpha_i, \beta_i) = \frac{\alpha_i - \beta_i}{\beta_i} \tag{4}
\]

As shown in the first Proposition below, agents with the same ratio of the marginal product of land to the opportunity cost of conflict allocate the same amount of time to attempts to obtain land, even though they are of different underlying marginal productivities. Agents with higher ratios of the marginal product of land to the opportunity cost of conflict will allocate more time to attempts to obtain land. These properties ultimately allow the derivation of a unique solution for the symmetric equilibrium strategy.

Agents’ strategies as functions of this ratio are represented by \( s(k) \); from \( s(k) \) we can later recover the agents’ strategies in terms of \((\alpha, \beta)\) through the substitution \( s(k(\alpha, \beta)) \). For notational convenience, henceforth \( s(k) \) refers to the equilibrium strategies with respect to \( k \). The expected probability density of \( k \) in the population of agents is denoted \( p(k) \) and is derived in Appendix 4.4 from the distributions of \( \alpha \) and \( \beta \) and (4).

The necessary condition for Bayesian equilibrium in terms of \( k \) is

\[
 s_i \in \arg \max (\Pr(s(k_j) > s_i)(T - s_i) \\
 + \int_{\{k_j: s(k_j) < s_i\}} p(k)[(k_i + 1)(T - s(k))]dk).
\]

Define \( \bar{k} \) as the minimum \( k \) such that \( s(\bar{k}) = T \) or the maximum value of \( k \), whichever is lower; likewise, \( s = s(\bar{k}) \). We can now state the following proposition:

**Proposition 1** \( s(k) \) is strictly increasing on the interval \((0, \bar{k})\), i.e.,

1. if \( k(\alpha', \beta') = k(\alpha'', \beta'') \), then \( s(k(\alpha', \beta')) = s(k(\alpha'', \beta'')) \) even though \( (\alpha', \beta') \neq (\alpha'', \beta'') \);
2. for \( k \in (0, \bar{k}) \), if \( k(\alpha', \beta') > k(\alpha'', \beta'') \), then \( s(k(\alpha', \beta')) > s(k(\alpha'', \beta'')) \).

**Proof of Proposition 1.** See Appendix 4.1.

The following lemma is instrumental in proving the uniqueness of the equilibrium (Proposition 2 below).

**Lemma 1** If an equilibrium exists, then the equilibrium strategy \( s(k) \) has the following properties:
1. \( s(0) = 0 \).

2. \( s(k) \) is continuous on \((0, \bar{k})\).

3. The inverse of \( s(k) \), \( K(s) \), exists and is differentiable over \((0, \bar{s})\).

**Proof of Lemma 1.** See Appendix 4.2.

This lemma allows us to rewrite agent \( i \)'s optimization problem, (5), in the form

\[
\begin{align*}
\max_{s_i} & \quad (1 - P(K(s_i)))(T - s_i) \\
& \quad + \int_0^1 p(K(s))K'(s) [(k_i + 1)(T - s)] ds.
\end{align*}
\]

The revised first-order condition

\[-(1 - P(K(s_i))) + p(K(s_i))K'(s_i)(k_i(T - s_i)) = 0,\]

yields the nonhomogeneous first-order linear differential equation

\[
s(k) + \frac{1 - P(k)}{kp(k)} s'(k) = T,\tag{6}
\]

the solution of which takes the following form on \( k \in (0, \bar{k}) \):

\[
s(k) = T - T \exp \left( - \int_0^k \frac{xp(x)}{1 - P(x)} dx \right).
\]

**Proposition 2** The equilibrium is unique, and the equilibrium strategy is characterized by

\[
s(\alpha, \beta) = \begin{cases} 
T \left( 1 - \left( 1 - \frac{c+d}{2b-c-d} \frac{\alpha-\beta}{\beta} \right)^{\frac{2\beta}{c+d}} \right)^{-1} e^{\frac{\alpha-\beta}{\beta}} & \text{if } \alpha \in \left[ \frac{a}{c}, \frac{b}{d} \beta \right] \\
T \left( 1 - \frac{\alpha}{\beta} \left( 1 - \frac{c}{b-c} \frac{\alpha-\beta}{\beta} \right)^{2\left(\frac{\beta}{c}-1\right)} \right)^{-1} e^{\frac{\alpha-\beta}{\beta}} & \text{if } \alpha > \frac{b}{d} \beta
\end{cases}
\]

**Proof of Proposition 2.** See Appendix 4.3.

Note that agents' equilibrium strategies are pure strategies. Recall that agents who differ markedly in productivity may play the same strategy and thus have the same probability of winning.

**Proposition 3** The equilibrium strategy \( s(\alpha, \beta) \) is increasing in \( \alpha \) and decreasing in \( \beta \).
Proof of Proposition 3. See Appendix 4.5.

Ceteris paribus, increasing an agent’s marginal productivity in the production process that requires land increases the amount of time she allocates to conflict (and thus decreases the amount of time she allocates to productive activities); and increasing an agent’s marginal productivity in the process that requires only labor decreases the amount of time she allocates to conflict (and thus increases the amount of time she allocates to production).

In aggregate, a population of agents that is more adept with respect to technology A, relative to technology B, devotes more time to conflict overall. Similarly, an improvement in technology A, which is equivalent to a global increase in skill $\alpha$, results in a greater amount of time being devoted to conflict, both because conflicts between agents last longer and because more conflicts occur. Fewer agents have $\alpha \leq \beta$, so fewer agents choose to concede possession of the land to their opponents without a fight. Thus an improvement in the technology that requires the capital input decreases the amount of labor allocated to both forms of production.

### 2.2 The Infinite Dynamic Game

Repeating the sequence of play of the one-shot game discussed above produces a dynamic infinite sequential game, where agents’ payoffs, and thus their optimal strategies, change from period to period in response to the evolution of the expected distribution of agent types among those who possess land and those who do not. The evolution of the distribution of agent types is, however, itself a consequence of the strategies agents choose. Modeling this dynamic element explicitly permits the direct analysis of the long-term effects of the absence or failure of rights-enforcing institutions on the level of conflict and on the distributions of both capital goods and income.

The initial conditions of the infinite sequential game are essentially the conditions described in the one-period game. Before play begins, Nature selects a fixed and finite population of agents, $N + L$, from a two-dimensional continuum of types distributed with constant density. Nature then chooses $L < N$ of these $N + L$ agents and endows them with land. Henceforth, agents who possess land will be said to be members of set $L$, and the remaining agents will be called members of set $N$. Once chosen, the population remains fixed throughout. Agents’ skill types do not change, but agents can move from set $L$ to set $N$ and vice versa. The skill dimensions (1) and production technologies (2), and hence the per-period payoffs (3), are the same ones described earlier; however, agents now consider the sums of their discounted expected future payoffs when allocating their time between production and conflict, not just their single-period payoffs.

At the beginning of each period, Nature randomly pairs agents in set $L$ with agents in set $N$. Note that $N - L$ agents will not be matched with opponents; these agents have no opportunity to acquire land that period and allocate the entire period...
to producing the consumption good with technology B. Conflicts over the possession of land are modeled as wars of attrition, as described in the one-shot game. Once again, agents do not know their opponents’ types, but they do know the distribution of types from which the population was drawn, and thus they can form expectations about the distributions of types in N and L in each period. Initially, the expected distributions of agent types are identical in N and L, but after the first round of contests, both the expected and the actual distributions of types are different among agents who possess land (the winners), L, than they are among those who do not, N. Consequently, agents of the same type k in different groups may have different optimal strategies after the first round of play, since they assign different probabilities to their opponents’ being of any given type. An agent without land who challenges a landholder knows that her opponent won in the previous period; similarly, the landholder knows that, if her opponent fought in the previous period, she must have lost. To avoid the proliferation of unnecessary notation, I first assume that agents remember neither the identities nor the actions of past opponents, and later argue that relaxing this assumption to allow perfect recall does not affect the main substantive results. After the possession of a parcel of land has been determined, the agents who contested it use the remainder of the period to produce using whatever technology is available to them.

9Recall that any agent may use technology B; possessing land does not bar an agent from choosing technology B over technology A. Only agents with $\beta \geq \alpha$ would make such a choice, however, and they are not likely to possess land after the first period, since it is not in their interest to engage in conflict for it.

10This assumption captures the notion that the capital good is already in some agent’s possession; a challenger must actually wrest control of the resource from its current holder in order to make use of it herself. The common assumption of a probabilistic tie-breaker (flipping a coin) is less appropriate in this context because it treats the resource as if it were entirely in the public domain. The substantive results in this paper are robust to allowing conflicts to continue into subsequent periods. The continuation of conflict is made possible by matching contestants who fight until the end of the period with each other again in the next period, rather than with new opponents. Because this extension complicates the recursive definitions of the probability distributions of agent types considerably, while yielding the same substantive results as the basic model, the simpler assumption is used here.

**Theorem 1** There exists a stationary Perfect Bayesian Equilibrium in non-decreasing strategies.

Proof of Theorem 1. See Appendix 4.6.

Before proceeding with the analysis of the equilibrium strategies, it is useful to confirm that some of the basic properties of the optimal strategies in the one-period war of attrition are also true of the optimal strategies in the infinite sequential game, and to define formally agents’ expectations about the evolving distributions of agent types among landholding and landless agents. The strategy of an agent of type $k$ in...
group \( L \) in period \( t \) is denoted \( s^L_t(k) \). Once again, \( k \) is defined as the minimum \( k \) such that \( s(k) = T \) or the maximum value of \( k \), whichever is lower; conversely, \( \bar{s} = s(\bar{k}) \).

**Proposition 4** The equilibrium strategies \( s^L_t(k) \) and \( s^N_t(k) \) have the following properties:

1. \( s^L_t(0) = s^N_t(0) = 0 \).
2. Within each group, for all \( t \), strategy \( s \) is strictly monotonically increasing in \( k \) on \([0, \bar{k}]\).
3. Within each group, for all \( t \), strategy \( s \) is continuous on the interval \([0, \bar{k}]\).
4. The inverses of \( s^L_t(k) \) and \( s^N_t(k) \), \( K^L_t(s) \) and \( K^N_t(s) \), exist and are differentiable over \((0, \bar{s})\).

**Proof of Proposition 4.** See Appendix 4.7.

In period \( t+1 \), the probability that a landless agent \( j \) assigns to the event that a landholding agent is of type \( k \) can be defined recursively. Recall that the expected probability density of \( k \) in the first period is given (4.4). The probability that a landholding agent is of type \( k \), \( p^L_{t+1}(k) \), is the sum of the probability that a landholding agent was of type \( k \) in \( t \) and retained possession of the land and the probability that a landless agent was of type \( k \) in \( t \) and acquired land. Consider a single landholder of type \( k \) who possesses land. She retains possession of that land if her opponent concedes before she does; the expected probability that her landless opponent is of a type that concedes before she does is \( P^N_t(K^N_t(s^L_t(k))) \). An agent of type \( k \) who does not have land wins it if she is matched with a landholding agent of a type that concedes before she does; this occurs with probability \( P^L_t(K^L_t(s^N_t(k))) \). The expression for \( p^N_{t+1}(k) \) is derived similarly but complicated slightly by the fact that \( L < N \).

\[
\begin{align*}
p^L_{t+1}(k) &= p^L_t(k) P^N_t(K^N_t(s^L_t(k))) + p^N_t(k) P^L_t(K^L_t(s^N_t(k))) \\
p^N_{t+1}(k) &= \frac{L}{N} p^L_t(k) \left[ 1 - P^N_t(K^N_t(s^L_t(k))) \right] \\
&\quad + \frac{L}{N} p^N_t(k) \left[ 1 - P^L_t(K^L_t(s^N_t(k))) \right] + \frac{N - L}{N} p^N_t(k) 
\end{align*}
\]

Let \( E[U^L_t(k)] \) represent the expected sum of discounted future payoffs for agent \( k \).
evaluated in period $t$, conditional on her beginning period $t+1$ in group $L$:

$$
E[U_t^L(k)] = \delta((1 - P_{t+1}^N(s_{t+1}^L(k)))(T - s_{t+1}^L(k) + E[U_{t+1}^N(k)])
+ \int_0^1 p_{t+1}^N(s)((k + 1)(T - s) + E[U_{t+1}^L(k)])ds)
= \delta(1 - P_{t+1}^N(s_{t+1}^L(k)))(T - s_{t+1}^L(k))
+ \delta \int_0^1 p_{t+1}^N(s)(k + 1)(T - s)ds
+ \delta(1 - P_{t+1}^N(s_{t+1}^L(k)))E[U_{t+1}^N(k)] + \delta P_{t+1}^N(s_{t+1}^L(k))E[U_{t+1}^L(k)].
$$

Similarly, $E[U_t^N(k)]$ represents expected payoffs conditional on her beginning period $t+1$ in $N$:

$$
E[U_t^N(k)] = \frac{L}{N}((1 - P_{t+1}^L(s_{t+1}^N(k)))(T - s_{t+1}^N(k) + E[U_{t+1}^N(k)])
+ \int_0^1 p_{t+1}^L(s)((k + 1)(T - s) + E[U_{t+1}^L(k)])ds)
= \delta\frac{L}{N}[(1 - P_{t+1}^L(s_{t+1}^N(k)))(T - s_{t+1}^N(k))]
+ \delta \frac{L}{N} \int_0^1 p_{t+1}^L(s)(k + 1)(T - s)ds
+ \delta(1 - \frac{L}{N})E[U_{t+1}^N(k)] + \frac{L}{N} P_{t+1}^N(s_{t+1}^N(k))E[U_{t+1}^L(k)]
+ \delta \frac{L}{N} P_{t+1}^N(s_{t+1}^N(k))E[U_{t+1}^L(k)].
$$

The agents’ optimization problems are

$$
\begin{align*}
 s_t^N & \in \arg \max_{s_t^N}((1 - P_{t}^L(s_t^N))(T - s_t^N) + E[U_t^N(k_t)]) \\
 & + \int_0^1 p_{t}^L(s)((k_t + 1)(T - s) + E[U_t^L(k_t)])ds
\end{align*}
$$

(8)
and

\[ s^L_t \in \arg \max((1 - P^N_t(s^L_t))((T - s^L_t) + E[U_t^N(k_i)]) + \int_0^\infty p^N_t(s)((k_i + 1)(T - s) + E[U_t^L(k_i)]) \, ds). \] (9)

Having established that the inverse functions \( K^L_t(s) \) and \( K^N_t(s) \) exist and that the agents’ beliefs about the distributions of types they face evolve according to (7), the optimization problem can be re-written to produce first-order conditions that result in the following system of nonhomogeneous differential equations:

\[ s^N_t(k) = T + \frac{\Delta^N_t(k)}{k} - \frac{1 - P^L_t(K^L_t(s^N_t(k)))}{kp^L_t(K^L_t(s^N_t(k))) K^L_t(s^N_t(k))} \]
\[ s^L_t(k) = T + \frac{\Delta^L_t(k)}{k} - \frac{1 - P^N_t(K^N_t(s^L_t(k)))}{kp^N_t(K^N_t(s^L_t(k))) K^N_t(s^L_t(k))}, \] (10)

where \( \Delta_t(k) = E[U^L_t(k)] - E[U^N_t(k)] \).

**Proposition 5** For \( k \in (0, \bar{k}_t) \) and \( t > 1 \), \( s^L_t(k) > s^N_t(k) \).

**Proof of Proposition 5.** See Appendix 4.8.

After the first round of play, a wedge is driven between the strategies of agents of the same type in different groups. Let \( \bar{k}_t \) represent the smallest \( k \) such that \( s^N_t(k) = T \) or the largest possible value of \( k, \frac{b+c}{c} \), whichever is smaller. A player of a given type \( (0, \bar{k}_t) \) is willing to commit more time to retaining possession of land than she is to obtaining it.

Thus the expected (and actual) distributions of types that result from the first round of play further advantage those agents who won in the first period. Over time, only agents of progressively higher types can expect to acquire land if they do not already possess it, but agents of progressively lower types can expect to retain land once they acquire it. After a sufficient number of periods, no agents in \( N \) choose to challenge landholders.

**Theorem 2** In a finite number of periods,

1. \( L \) and \( N \) converge to some \( L^* \) and \( N^* \) such that, for every \( i \) in \( L^* \) and every \( j \) in \( N^* \), \( s^L_t(k_i) > s^N_t(k_j) \).
2. For every type \( k \) in \( N^* \), \( s^N_t(k) = 0 \).

**Proof of Theorem 2.** See Appendix 4.9.

Thus the allocation of land across agents becomes fixed, and in the steady-state equilibria there is virtually no conflict in the state of nature. Active conflict dies away.
after the allocation of land becomes fixed because agents’ strategies are a function of their beliefs about the populations of types rather than of the actual populations of types. Landholders have secure possession of their land, and agents who do not possess land appear to “respect” the security of the landholders’ property.

It is important to note that the sorting of agents across the two groups (N and L) by k need not be complete: it is quite possible that in the steady state there are two agents, i and j, such that i ∈ L, j ∈ N and k_i < k_j; the trick, as Proposition 2 implies, is that such situations are sufficiently uncommon to not make it worthwhile for j ∈ N to challenge a randomly matched and unknown to him member of group L. The exact composition of L* and N* depends not only on the primitives of the economy, which determine the strategies the agents choose, but also on the initial allocation of land across agents and the realizations of the random matches. All the equilibria, however, share the qualitative characteristics described in Propositions 2 and 3.

The following proposition establishes the relationship between the extent of the steady-state sorting and the speed of convergence to the steady state. Before formulating this proposition, however, we need to introduce a measure of agent sorting in this economy. Define k*(a, b, c, d, N, L) such that the cumulative probability function \( P(k^*) = \frac{N}{N+L} \). Because k* is function of primitives and not agent type, I suppress its arguments. When agents sort completely in accordance with k, then all agents with types between k* and the highest value of k, \( \frac{b-c}{c} \), are in group L; likewise, all agents with types between k* and the lowest value of k, \( \frac{a-d}{d} \), are in group N. A measure of the incompleteness of sorting, then, is the number of agents in N whose k-types are between \( \frac{b-c}{c} \) and k*, \( \lvert N^* \cap (k^*, \frac{b-c}{c}) \rvert \).

**Theorem 3** The more rapidly \( s^L(k) \) and \( s^N(k) \) diverge,

1. the sooner actual conflict ceases entirely, i.e., the sooner \( s^N(k) = 0 \) for every type k in N;
2. the less complete the sorting of agents by k into L and N, i.e., the larger is \( \lvert N^* \cap (k^*, \frac{b-c}{c}) \rvert \).

**Proof of Theorem 3.** See Appendix 4.10.

An economy that converges more slowly devotes more time to conflict, and hence less time to productive activities. It also sorts more completely, so that in the steady-state the allocation of land across agents is more efficient, in the sense that the agents with greater differences in productivity with and without land are more likely to possess land. Thus there is an apparent trade-off between the improving efficiency on the path of convergence and in the steady-state. The implications of this result are discussed in more detail in the next section.

The results on the relationship between the extent of conflict on the transition path and the steady state growth have also somewhat unexpected implications for the
efficiency/distributive advantage divide that characterizes the institutionalist theories (including theories of property rights). Whereas such authors such as Demsetz [1967] and Davis and North [1971] argued that property rights emerge in response to the needs of increasing certainty (and decreasing transaction costs) of economic exchanges, Knight [1992] argued forcefully that institutions are best seen as responses to the distributive competition.\footnote{Although they are not explicitly committed to this position, repeated games models of rights have also focused more on the efficiency-increasing role of institutions.} As Proposition 3 indicates, however, the causal divide between efficiency and distributive advantage may be overdrawn. Efficiency may result because of the conflict (as distinct from a productive competition) over the finite resources.

2.3 Robustness

Recall now that the foregoing has assumed that agents do not have perfect recall - in particular, that they do not retain memory of the identities and actions of their own past opponents. The assumption of perfect recall complicates the analysis because an agent learns something about her opponent’s type from the outcome of the war of attrition, altering not only her beliefs about the type of her opponent when she is matched with someone she has encountered before, but also her expected future utility. To see that the substantive results embodied in Theorems 2 and 3 are robust to assuming perfect recall consider the following. The probability that $j$ assigns to the event that some agent with whom she has been matched before is of type $k$ depends on the information she obtained from observing that opponent’s strategy (if $j$ won) or from observing that that opponent’s strategy was greater than her own (if $j$ lost). Thus her updated beliefs about the type, and hence the current strategy, of an opponent she has fought before are specific to that opponent. Because every agent is an independent draw from the underlying distribution of types, and because agents do not observe conflicts (or their outcomes) other than those in which they are participants, information about one agent’s type does not alter beliefs about the type of any other agent. Thus we can consider the effects of the possibility of encountering known agents separately.

Denote the cumulative distribution function that represents $j$’s beliefs about agent $i$’s type in period $t$ as $P_t^i(k, h_j)$, where $h_j$ contains $j$’s information about past play. Let $\Delta_t(k, h) = E[U_t^L(k, h)] - E[U_t^N(k, h)]$, where $E[U_t^L(k, h)]$ is the expected utility of an agent of type $k$ with history $h$ who starts the next period in group $L$, and, similarly, $E[U_t^N(k, h)]$ is the expected utility of an agent of type $k$ and history $h$ who starts the next period in group $N$. Agent $j$’s strategy when facing a past opponent is the solution to a differential equation of the same form as (10), but using the opponent-specific $P_t^i(k, h_j)$ and $p_t^i(k, h_j)$ and a difference in expected utilities $\Delta_t(k, h_j)$ that accounts for the possibility of being matched with past opponents, with whom the strategies played in the war of attrition may be different.
Consider first the effects of \( j \)'s beliefs about \( i \)'s type on her strategy, bearing in mind that \( i \)'s beliefs about \( j \)'s type are also a function of the outcome of their previous encounter. If \( j \) lost to \( i \) before, then she assigns greater probability to \( i \) being of higher types than she would assign to someone whom she had not met before, and consequently she chooses a lower strategy. The parallel conclusion that \( i \), having beaten \( j \) before and hence having lower expectations of \( j \)'s type, chooses a higher strategy than she would have chosen had she been matched with someone whom she had not met before, reinforces \( j \)'s choice to concede sooner. Thus the direct effects of the contestants’ updated beliefs about each other’s type is consistent with the sorting by \( k \) type that occurs when agents do not remember the identities of past opponents.

Information about other agents’ garnered through past play also affects the difference between the agent’s expected utility if she wins in the current period and her expected future utility if she loses, \( \Delta_t(k) \). The substantive content of Proposition 5 and Theorems 2 and 3 is robust to these effects because the altered difference in expected utility is still increasing in \( k \). Consider separately the agent’s beliefs about her expected future interactions with unknown opponents and with known opponents. (Recall that each agent is an independent draw from the underlying distribution of types, and hence that information about one opponent’s type does not alter the agent’s beliefs about any other potential opponent’s type.) If the difference in her expected future utility is increasing in \( k \) for both unfamiliar potential opponents and for familiar (known) potential opponents, then the difference in her expected future utility when she has acquired knowledge of some but not all potential opponents’ types is also increasing in \( k \). It has already been demonstrated that the difference in expected future utility is increasing in \( k \) when the agent’s beliefs about her opponents are based solely on her knowledge of the underlying distribution of types and the equilibrium of the game, i.e. when all her potential opponents are unfamiliar; it remains to see that it is increasing in \( k \) when her potential opponents are familiar. When two contestants in a war of attrition know each other’s types, the lower type conceding immediately (and the higher type fighting as long as is necessary to win) is an equilibrium. It follows that, when the pool of known potential opponents is known, the difference in expected utility is \((P_{t+1}^N(k) - \frac{1}{N}P_{t+1}^L(k))(kT + \Delta_{t+1}(k))\), which is increasing in \( k \) and ranges from \((1 - \frac{1}{N})(kT + \Delta_2(k))\), achieved in \( t = 1 \), to \( kT + \frac{3kT}{1-3} \).

3 Discussion

3.1 Conflict and Security in the “State of Nature”

Although conflict is commonplace in the initial stage of this dynamic model, I show that the reallocation of land through conflict results in a situation where the expected success of challengers is so low that potential challengers are deterred from initiating
conflicts, and thus, in the steady-state, landholders enjoy security of property.

Prior to convergence, or whenever a shock results in the belief that enough land has been re-distributed away from agents of high type, i.e. agents with high ratios of net marginal productivity with land \((\alpha - \beta)\) to opportunity cost of conflict \((\beta)\), toward agents of low type, open conflict occurs. During this stage, increasing an agent’s marginal productivity in the production process that requires land increases the amount of time she allocates to conflict (and thus decreases the amount of time she allocates to productive activities); and increasing an agent’s marginal productivity in the process that requires only labor decreases the amount of time she allocates to conflict (and thus increases the amount of time she allocates to production). In aggregate, a population of agents that is more adept with respect to technology \(A\), relative to technology \(B\), devotes more time to conflict overall. Similarly, an improvement in technology \(A\), which is equivalent to a global increase in skill \(\alpha\), results in a greater amount of time being devoted to conflict. Because those agents who value the capital good value it more highly, they are willing to expend more resources to acquire it; hence the typical dispute lasts longer. Because more agents value the capital good, i.e. fewer agents have \(\alpha \leq \beta\), more agents are willing to expend some resources in an attempt to acquire it; thus more disputes occur. In the absence of adequate property rights institutions, improvements in the technology that requires the capital input increases conflict and decreases the amount of labor allocated to both forms of production.

Successive rounds of conflict result in very different distributions of agent types among those who hold land and those who do not. Although an agent cannot directly determine the type of her opponent, she knows whether or not her opponent possesses land. Since higher types, i.e. agents with high ratios of net marginal productivity with land \((\alpha - \beta)\) to opportunity cost of conflict \((\beta)\), become landholders disproportionately, the fact that an agent possesses land provides information about the likelihood that she will concede first. In the long run, agents who do not possess land cease altogether to challenge those who do.

3.2 The Allocation of Factor Goods

Note first that agents with the same ratios of net marginal productivity with land to opportunity cost of conflict, \(k = \frac{\alpha - \beta}{\beta}\), earn different payoffs from production because of the differences in their underlying productivities \((\alpha, \beta)\). Because agents’ success in conflict, and hence the possession and use of land, depends on \(\frac{\alpha - \beta}{\beta}\) rather than \(\alpha - \beta\), there is a discrepancy between the steady-state equilibrium allocation of land and labor and the efficient allocation of land and labor. Note that, although the form of \(k\) is a result of this particular model, e.g. of modeling conflict as a war of attrition and of the particular forms of the production and utility functions, the substantive conclusion is robust. For any means of determining who will use a resource that does not depend solely on the agents’ valuations of that resource, i.e. any means
that does not systematically direct the resource to its highest value use, allocational inefficiencies will result.

It is intuitive that agents’ optimal strategies in any contest in which their participation is costly will depend on both their valuation of the prize and the costs of their participation. In the model presented above, their valuations of the prize depend on their ability to derive enjoyment from consumption (their utility functions), the technology that transforms resources into consumption (the production functions), and their particular productivities with those technologies (their types $\alpha$ and $\beta$). The costs of their participation in the contest are determined by the manner in which participation consumes resources and the costs of those resources for the agents. The war of attrition requires the participants time, and the cost of that time for an agent is the utility she would derive from the next best use of that time, in this case, the utility she would derive from the consumption good she could produce using the $B$-technology with her productivity $\beta$. It is the costliness of the conflict over the resource that causes the discrepancy between the ultimate distribution of the resource that results from conflict and the efficient distribution of the resource. Note that the addition of a third dimension to the agents’ primitive types to capture any differences in their ability to use the conflict technology, e.g. differences in their primitive abilities to fight, would not affect the conclusion that inefficiencies are the natural result of allocating goods through conflict. It would simply introduce another factor, the differences in the rates at which agents’ participation in conflict consumes their resources, in addition to the differences in their opportunity costs of using that resource, into the determination of the costliness of conflict. The existence of agents’ optimal strategies for engaging in conflict is all that is necessary to imply that conflict will result in the sorting of agents into groups of “haves” and “have-nots” according to some composite type, a function of the primitive types, that determines their choice of strategy in conflict. It follows that allocational inefficiencies of the sort that result from the model presented here are robust to a wide range of conflict and production technologies and to the heterogeneity of the population in aspects relevant to them.

These allocational inefficiencies are exacerbated by the fact that, ultimately, agents do not sort completely even with respect to the composite type, $k$ (Theorem 2). Incomplete sorting exacerbates inefficiency in expectation because, as discussed above, agents’ strategies in conflict do depend (positively) on their valuation of the prize. Thus the allocation of the resource through conflict does correlate, however imperfectly, with the efficient allocation, whereas the assumed initial allocation was completely uncorrelated with it. The extent to which agents are sorted according to the composite type ($k$) in the steady-state depends on how quickly the gap between the conflict strategies of identical types in $L$ and $N$, $s^L(k)$ and $s^N(k)$, widens, and hence on how quickly the system converges (Theorem 3). The more quickly the gap widens, the more quickly the system converges and the more incomplete the sorting; thus the less severe are the losses due to the allocation of time to conflict and the more severe are the losses associated with the misallocation of land across agents.
The above analysis identifies two sources of potential gains in an individual agent’s production, and thus in her utility: reducing the time allocated to conflict (before convergence to the steady-state equilibrium) and engaging in mutually beneficial voluntary exchanges. The magnitude and commonality of these incentives depends directly on the primitive features of the economy (the technologies, the amount of land, the number of agents, and the distributions and ranges of $\alpha$ and $\beta$).

### 3.3 Security vs. Transferability

Although this *de facto* security of property has some of the desirable characteristics of a socially legitimated and enforced right to the security of property, such as providing incentives to engage in productive activities and even to invest, it has a critical shortcoming: it cannot support voluntary exchange. Thus the pattern of ownership that emerges spontaneously in the state of nature differs from a property right not only in its formal, legal status but in its immediate economic consequences. Inefficiency persists in the presence of a stable ownership pattern because, as already noted, factor goods are not allocated efficiently. Agents cannot remedy this misallocation through trade because they cannot credibly commit to refrain from expropriating the goods they have just “sold.” When an agent loses her land through conflict, she has no incentive to return to challenge the victor again; she has just discovered that that particular agent is willing to commit more time than she is to securing the land. However, when an agent agrees to cede possession of land to another agent without conflict, she may find it in her interest to renege by re-expropriating the land.

The attainment of security of property does not necessarily lead to the ability to engage in trade or to the increase in allocative efficiency trade implies; nor does the ability of agents acting in their own self-interest to create institutions that constrain their behavior necessarily promote the development of a system of property rights that enables trade or signal an improvement in macroeconomic performance. As is evident in Hafer [2002], if security depends on the inability to distinguish those who are able to defend their resources against incursion from those who are, exchange is severely hampered.

Consider, next, the steady-state equilibrium described in Theorem 2. Landholders with low $\alpha - \beta$ could benefit from voluntarily alienating their land to landless agents with high $\alpha - \beta$ in exchange for payments in the consumption good, if they were able to create an institution that could guarantee the contract. Landless agents with low $k$ would initially be willing to make such deals, because, given the commonly-shared beliefs about the distributions of types in $L$ and in $N$ in the steady state, they would anticipate that their possession of the land would not be challenged. After many such exchanges, the (perceived) distributions of types could be sufficiently altered to make challenges from members of the landless group profitable again, initiating a new stage of pervasive conflict. Note that all members of the landholding group would ultimately experience losses in utility as a result of these voluntary exchanges,
not only the “new” members or the members of lower types, because all landholders would be forced to commit time to conflict in order to defeat challengers. Thus agents in group $L$ have some incentive to prevent voluntary exchange altogether. The basis of the de facto security of property in the steady-state equilibrium not only fails to enable voluntary exchange, it actually produces incentives that may inhibit the transfer of property.

In order for agents to gauge the credibility of each other’s commitments, either as parties to one-to-one agreements or as participants in institutional arrangements, they must be able to determine the incentives of other agents to fulfill their obligations. For instance, a group of landholding agents might wish to monitor transfers of land from agents who won land in conflict to agents who did not. It is in their interest to screen new landholders publicly and convincingly on the basis of their willingness to fight if challenged; in this way, they can discourage future attacks on landholders by maintaining the group reputation for fighting. Some landless agents, however, have incentives to claim that they would be willing to fight in situations when they actually would not. They may value land highly enough to want to acquire it through trade when they anticipate that their possession of it will not be challenged, but if, at some point, the economy were shocked in such a way that future conflict seemed likely, they would not be willing to commit enough time to conflict to defeat challengers. Knowing that such types exist, landholding agents actively attempt to identify such agents and prevent them from obtaining land; to the extent that they are successful, some agents who are willing to purchase land are unable to do so. Thus the ability of agents to engage in exchange, as distinguished from their willingness, is a consequence of both their type and the means of securing property. The opportunity - or even the socially-determined right – to exchange property may be available to an endogenously determined subset of agents without being available to all, as demonstrated in the context of rental contracts in Hafer [2002]. Modeling the emergence of security explicitly from a “state of nature” ultimately permits the identification of the types of agents who have the ability to acquire property through voluntary exchange – a distinction between agents that cannot be made in the extant theories of institutions, which assume implicitly that agents who are willing to exchange property must be uniformly able to do so.

4 Appendix

4.1 Proof of Proposition 1

For notational convenience, let $s' = s(k(\alpha', \beta'))$ and $s'' = s(k(\alpha'', \beta''))$. By the definition of $s(k)$ as an equilibrium strategy, an agent of type $k(\alpha', \beta')$ must prefer $s'$ to $s''$; similarly an agent of type $k(\alpha'', \beta'')$ must prefer $s''$ to $s'$. From (5), the following
Because an agent knows with certainty that her opponent will not concede defeat in the interval \((s', s'')\) and conflict is costly, \(i\) does better conceding immediately after \(s'\) than at any \(s \in (s', s'')\). Furthermore, \(i\) does not concede immediately after \(s''\), because either \(p(s'')\) is negligible, in which case \(i\) is better off conceding at \(s'\), or \(p(s'')\) is non-negligible, in which case \(i\) is better off waiting until \(s > s''\), and hence, following this logic, no agent concedes until \(T\).

3. The existence, continuity and monotonicity of \(K(s)\), defined on \(s \in (0, s)\) to take values in the interval \((0, k)\), follow from the monotonicity and continuity of \(s(k)\). For \(K(s)\) continuous and monotonic and \(p(K(s)) > 0\), the differentiability of \(K(s)\) follows by Lemma 1, part 4 of Fudenberg and Tirole [1986].

\[\Pr(s(k_j) \geq s')(T - s') + \int_{\{k_j : s(k_j) < s'\}} p(k_j) \left[(k' + 1)(T - s(k_j))\right] dk_j\]

\[\geq \Pr(s(k_j) \geq s'')(T - s'') + \int_{\{k_j : s(k_j) < s''\}} p(k_j) \left[(k'' + 1)(T - s(k_j))\right] dk_j\]

Subtracting the smaller side of the second expression from the larger side of the first, and subtracting the larger side of the second expression from the smaller side of the first, I obtain

\[\int_{\{k_j : s'' \leq s(k_j) \leq s'\}} p(k_j) \left[(k' - k'')(T - s(k_j))\right] dk_j \geq 0.\]

Thus if \(k(\alpha', \beta') = k(\alpha'', \beta''),\) then \(s' = s''\); and if \(k(\alpha', \beta') > k(\alpha'', \beta'')\), then \(s' > s''\).

\[\blacksquare\]

### 4.2 Proof of Lemma 1

1. For \(k = 0, \alpha = \beta\). If the agent commits \(s > 0\), the payoff is \(\alpha l(T - s)\), which is less than \(\beta T\), the payoff for \(s = 0\).

   2. An agent’s expected utility, as is clear from (5), is continuous in \(k\).

   Suppose \(s(k)\) is not continuous in \(k\). Then, for some \(k\),

   \[\lim_{k \to k^+} s(k) = s' < s'' = \lim_{k \to k^-} s(k).\]

   Because an agent \(i\) knows with certainty that her opponent \(j\) will not concede defeat in the interval \((s', s'')\) and conflict is costly, \(i\) does better conceding immediately after \(s'\) than at any \(s \in (s', s'')\). Furthermore, \(i\) does not concede immediately after \(s''\), because either \(p(s'')\) is negligible, in which case \(i\) is better off conceding at \(s'\), or \(p(s'')\) is non-negligible, in which case \(i\) is better off waiting until \(s > s''\), and hence, following this logic, no agent concedes until \(T\).

   3. The existence, continuity and monotonicity of \(K(s)\), defined on \(s \in (0, s)\) to take values in the interval \((0, k)\), follow from the monotonicity and continuity of \(s(k)\). For \(K(s)\) continuous and monotonic and \(p(K(s)) > 0\), the differentiability of \(K(s)\) follows by Lemma 1, part 4 of Fudenberg and Tirole [1986].

\[\blacksquare\]
4.3 Proof of Proposition 2

The proof of uniqueness is by contradiction. Suppose that there exist at least two distinct equilibria, \(s(k)\) and \(\tilde{s}(k)\). For some \(k > 0\), \(s(k) \neq \tilde{s}(k)\). If \(s(k) > \tilde{s}(k)\), (6) implies that \(s'(k) < \tilde{s}'(k)\); likewise, if \(s(k) < \tilde{s}(k)\), then \(s'(k) > \tilde{s}'(k)\). But, from Lemma 1, \(s(0) = \tilde{s}(0) = 0\), and from Proposition 1, \(s(k)\) and \(\tilde{s}(k)\) are strictly increasing. Thus, \(s(k) > \tilde{s}(k)\) implies \(s'(k) > \tilde{s}'(k)\), a contradiction.

Recovering the equilibrium strategy in terms of \(\alpha\) and \(\beta\), \(s(k(\alpha, \beta))\), is complicated by the fact that infinitely many pairs of \((\alpha, \beta)\) correspond to the same value \(k\). To derive the distribution of \(k\) as a function of the underlying parameters \(\alpha\) and \(\beta\), note that the probability of a given \(k\) is the probability of all possible combinations of \((\alpha, \beta)\) that produce that value \(k\), i.e. \(p(k) = p((\alpha, \beta) : \frac{\alpha - \beta}{\beta} = k)\). The explicit formulae for the cumulative distribution of \(k\), \(P(k)\), and the probability density of \(k\), \(p(k)\), are derived in Appendix 4.4. They are

\[
P(k) = \begin{cases} 
\frac{(a-d(1+k))^2}{2(b-a)(d-c)(1+k)} & \text{if } k < \frac{a}{c} - 1 \\
\frac{(d+c)(k+1)-2a}{2(b-a)} & \text{if } k \in \left[\frac{a}{c} - 1, \frac{b}{d} - 1\right] \\
\frac{2bd(1+k)-b^2-(k+1)(c^2(1+k)+2a(d-c))}{2(b-a)(d-c)(1+k)} & \text{if } k > \frac{b}{d} - 1 
\end{cases}
\]

and

\[
p(k) = \frac{dP(k)}{dk} = \begin{cases} 
\frac{d^2(1+k)^2-a^2}{2(b-a)(d-c)(1+k)^2} & \text{if } k < \frac{a}{c} - 1 \\
\frac{c+d}{2(b-a)} & \text{if } k \in \left[\frac{a}{c} - 1, \frac{b}{d} - 1\right] \\
\frac{b^2-c^2(1+k)^2}{2(b-a)(d-c)(1+k)^2} & \text{if } k > \frac{b}{d} - 1 
\end{cases}
\]

Substituting \(P(k)\), \(p(k)\) and (4) into (6) yields the expression for \(s(\alpha, \beta)\) in the statement of the proposition.

4.4 Derivation of \(P(k)\) and \(p(k)\)

Let \(P(k)\) denote the cumulative probability function, i.e. \(P(k) = Pr(\kappa \leq k)\), and \(p(k)\) denote the probability density function, \(p(k) = \frac{dP(k)}{dk}\), where

\[ k = \frac{\alpha - \beta}{\beta} \]

From the definitions of \(P(k)\) and \(k\),

\[ P(k) = Pr((\alpha, \beta) : \frac{\alpha - \beta}{\beta} \leq k), \]
which can be expressed in terms of the known distributions of $\alpha$ and $\beta$, as follows:

\[
P(k) = \int_c^d \Pr\left(\frac{\alpha - \beta}{\beta} \leq k|\beta\right) p(\beta) \, d\beta = \int_c^d \Pr(\alpha \leq \beta(k+1)|\beta) p(\beta) \, d\beta
\]

\[
= \int_c^d \left( \int_a^{\beta(k+1)} p(\alpha) \, d\alpha \right) p(\beta) \, d\beta.
\]

(11)

Recall that $\alpha$ and $\beta$ are distributed with constant density over the intervals $[a, b]$ and $[c, d]$, respectively:

\[
p(\alpha) = \begin{cases} 
\frac{1}{b-a} & \text{if } \alpha \in [a, b] \\
0 & \text{if } \alpha \notin [a, b]
\end{cases}
\]

\[
p(\beta) = \begin{cases} 
\frac{1}{d-c} & \text{if } \beta \in [c, d] \\
0 & \text{if } \beta \notin [c, d]
\end{cases}
\]

Substituting these values in (11),

\[
P(k) = \begin{cases} 
\frac{1}{d-c} \int_a^{\beta(k+1)} \left( \int_a^{\frac{1}{b-a}} \, d\alpha \right) \, d\beta & \text{if } k < \frac{a}{c} - 1 \\
\frac{1}{d-c} \int_a^{\beta(k+1)} \left( \int_a^{\frac{1}{b-a}} \, d\alpha \right) \, d\beta & \text{if } k \in \left[\frac{a}{c} - 1, \frac{b}{d} - 1\right] \\
\frac{1}{d-c} \int_a^{\beta(k+1)} \left( \int_a^{\frac{1}{b-a}} \, d\alpha \right) \, d\beta & \text{if } k > \frac{b}{d} - 1
\end{cases}
\]

(12)

Recalling that the probability density $p(k) = \frac{\partial P(k)}{\partial k}$, it follows from (12) that

\[
p(k) = \begin{cases} 
\frac{d^2(1+k)^2-a^2}{2(b-a)(d-c)(1+k)^2} & \text{if } k < \frac{a}{c} - 1 \\
\frac{c+d}{2(b-a)} & \text{if } k \in \left[\frac{a}{c} - 1, \frac{b}{d} - 1\right] \\
\frac{b^2-c^2(1+k)^2}{2(b-a)(d-c)(1+k)^2} & \text{if } k > \frac{b}{d} - 1
\end{cases}
\]

(13)

The conditional probability density of $\beta$, $p(\beta|k)$, is the inverse of the length of the projection of the $k$-isoquant onto the $\beta$-coordinate axis. The conditional probability of $\alpha$, $p(\alpha|k)$ can be found applying the same logic.
Suppose \( b - a > d - c \); then

\[
p(\beta|k) = \begin{cases} 
  \frac{k+1}{dk+d-a} & \text{if } k < \frac{a}{c} - 1 \\
  \frac{1}{d-c} & \text{if } k \in \left[\frac{a}{c} - 1, \frac{b}{d} - 1\right] \\
  \frac{k+1}{b-ck-c} & \text{if } k > \frac{b}{d} - 1
\end{cases}
\]

\[
p(\alpha|k) = \begin{cases} 
  \frac{1}{dk+d-a} & \text{if } k < \frac{a}{c} - 1 \\
  \frac{1}{(d-c)(k+1)} & \text{if } k \in \left[\frac{a}{c} - 1, \frac{b}{d} - 1\right] \\
  \frac{1}{b-ck-c} & \text{if } k > \frac{b}{d} - 1.
\end{cases}
\]

### 4.5 Proof of Proposition 3

We proceed by signing the derivatives of equilibrium \( s(\alpha, \beta) \) derived in From Proposition 2:

\[
\frac{\partial s(\alpha, \beta)}{\partial \alpha} = \begin{cases} 
  \frac{\partial}{\partial \alpha} \left[ T \left( 1 - \left( 1 - \frac{c+d}{2b-c-d} \frac{\alpha-\beta}{\beta} \right) e^{\frac{a}{c}} e^{-\frac{a}{c}} \right) \right] & \text{if } \alpha \in \left[\frac{a}{c}, \frac{b}{d}\right] \\
  \frac{\partial}{\partial \alpha} \left[ T \left( 1 - \frac{\alpha}{\beta} \left( 1 - \frac{c}{b-c} \frac{\alpha-\beta}{\beta} \right)^2 e^{\frac{a}{c}} e^{-\frac{a}{c}} \right) \right] & \text{if } \alpha > \frac{b}{d}\beta
\end{cases}
\]

\[
= \begin{cases} 
  -\frac{(c+d)(-2b+c+d)e^{\frac{a}{c}} T(1+\frac{(c+d)(\alpha-\beta)}{(b-c)d+\alpha}) \frac{2b}{\beta}(\alpha-\beta)}{(c\alpha+d\alpha-2b\beta)^2} & \text{if } \alpha \in \left[\frac{a}{c}, \frac{b}{d}\right] \\
  \frac{(b-c)e^{\frac{a}{c}} T(\alpha) \left( 1+\frac{(c+d)(\alpha-\beta)}{(b-c)d+\alpha} \right) \frac{2b}{\beta}(\alpha-\beta)}{(c\alpha+d\alpha-2b\beta)^2} & \text{if } \alpha > \frac{b}{d}\beta
\end{cases}
\]

The top expression is positive if and only if \(-2b + c + d < 0\) and therefore if and only if \(b > \frac{c+d}{2}\), which is always true since \(b > d > c\). The bottom expression is positive if and only if \(-c\alpha + b\beta > 0\), i.e., iff \(\frac{b}{c} > \frac{a}{\beta}\), which is always true since \(\alpha \leq b\) and \(\beta \geq c\).

Similarly,

\[
\frac{\partial s(\alpha, \beta)}{\partial \beta} = \begin{cases} 
  -\frac{(c+d)(-2b+c+d)e^{\frac{a}{c}} T(1+\frac{(c+d)(\alpha-\beta)}{(b-c)d+\alpha}) \frac{2b}{\beta}(\alpha-\beta)}{(c\alpha+d\alpha-2b\beta)^2} & \text{if } \alpha \in \left[\frac{a}{c}, \frac{b}{d}\right] \\
  -\frac{(b-c)e^{\frac{a}{c}} T(\alpha) \left( 1+\frac{(c+d)(\alpha-\beta)}{(b-c)d+\alpha} \right) \frac{2b}{\beta}(\alpha-\beta)}{(c\alpha+d\alpha-2b\beta)^2} & \text{if } \alpha > \frac{b}{d}\beta
\end{cases}
\]

As both expressions for \(\frac{\partial s(\alpha, \beta)}{\partial \beta}\) are negative if and only if the expressions in the corresponding intervals for \(\frac{\partial s(\alpha, \beta)}{\partial \alpha}\) are positive, the statement of the proposition follows.

\[ \blacksquare \]

### 4.6 Proof of Theorem 1

Because agents remember neither identities nor actions of their past opponents, all agents have common beliefs about the distribution of types in each group in each
period. The only impact of an agent’s play in the current period \( t \) on her expected payoffs in subsequent periods is through the determination of her group membership in \( t + 1 \). Expected future payoffs are a function of \( k_i, t \), and membership in \( N \) or \( L \) but not \( s_i^*(k_i) \) directly. Thus, \( E[U^L_i(k_i, t)] \) and \( E[U^N_i(k_i, t)] \) may be treated as constraints in period \( t \), given future sequentially rational behavior. Therefore, equilibrium strategy must satisfy (8) and (9).

I first prove the existence of a fixed point in the best response correspondence for period \( t \), given \( P^L_t(\cdot), P^N_t(\cdot), E[U^L_t(\cdot), E[U^N_t(\cdot)] \). The stationary Perfect Bayesian Equilibrium consists of that behavioral strategy profile in each period, with beliefs updated after each period according to Bayes’ Rule. As in Athey [2001], I prove the existence of the equilibrium behavioral strategy profile using a limiting argument, by first establishing equilibrium existence in a finite action game, then establishing that the limit of the equilibria of a sequence of finite-action games that converge to the continuous game is the equilibrium of the continuous action game. Equilibrium existence in the finite-action game is readily established using results from Athey [2001], but establishing that the limit of the sequence of equilibria is itself an equilibrium requires extending the arguments made in Athey [2001]. The chief complication here is the dependence of the value of winning on the opponent’s action. Notation introduced in the proof follows that of Athey [2001] to facilitate comparison.

Consider a sequence of games \( \{\Gamma^n\} \) with successively finer action spaces, and which are otherwise identical to the interaction in period \( t \). For each game in the sequence the distribution of types \( k \) is bounded and has no mass points on \([\frac{a-d}{d}, \frac{b-c}{c}]\). The expected utility of \( i \) is well-defined and finite \( \forall i, \forall \) subsets of \( i \)’s opponent’s possible types, and \( \forall \) possible strategies of \( i \)’s opponent. The game satisfies the Single Crossing Condition for games of incomplete information, and therefore each game in the sequence of finite-action games has a Pure Strategy Nash Equilibrium, in which each player’s equilibrium strategy \( s_{i,n}(k_i) \) is non-decreasing [Athey 2001, Theorem 1]. From the Nash condition, \( \forall n, \forall k_i \leq 0, s_{i,n}(k_i) = 0 \).

Let \( \tau_{i,n}^*(s') \) be the event that agent \( i \) ties and wins the tie playing \( s' \), and let \( \tau_{i,n}^*(s') \) be the event that agent \( i \) ties and loses playing \( s' \).

**Lemma 2**  
Construct \( s^*(k) \) as the limit of the sequence of nondecreasing equilibrium strategies in finite games, \( s_{i,n}(k) \). Then \( s^*(k) \) has no mass points on \( k > 0 \), i.e., \( \forall i \in L, \forall s' \in [0, T], \Pr(s_i^*(k_i) = s') \Pr(\tau_{i,n}^*(s')) = 0 \) and \( \forall j \in N, \forall s' \in [0, T], \Pr(s_j^*(k_j) = s') \Pr(\tau_{j,n}^*(s')) = 0 \).

Proof of Lemma 2. We prove by contradiction. Suppose \( \exists \) player \( j \) and action \( s' \in [0, T] \), such that \( \Pr(s_j^*(k_j) = s') \Pr(\tau_{j,n}^*(s')) > 0 \). By assumption, \( \Pr(\tau_{j,n}^*(\cdot)) = 0 \) \( \forall i \in L \). Therefore, \( j \in N \) and \( \Pr(\tau_{j,n}^*(b)) = \Pr(s_j^*(k_i) = s'), \Pr(s_j^*(k_j) = s') \Pr(\tau_{j,n}^*(s')) = 0 \).

\( s_{i,n}(k_i) \) and \( s_{j,n}(k_j) \) are measurable and converge almost everywhere to \( s_i^*(k_i) \) and \( s_j^*(k_j) \). Therefore, the sequences converge uniformly to \( s_i^*(k_i) \) and \( s_j^*(k_j) \) except on a set of arbitrarily small measure. Thus, \( s_j^*(k_j) = s' \) on an open interval \( S_j = \)
\( \{ k_j : s_j^*(k_i) = s' \} \). It follows that \( \forall \eta > 0, \exists N_d \) such that \( \forall n > N_d, \forall k_j \in S_j \setminus E_j \), 
\( |s_{j,n}(k_j) - s'| < d \). The same argument follows for \( s_i^*(k_i) = s' \) on an open interval \( S_i \).

Choose \( d \). Choose \( N_d \) such that \( \forall n > N_d, \forall k \in E, \ |s_{i,n}(k_i) - s_i^*(k_i)| < d \) and 
\( |s_{j,n}(k_j) - s_j^*(k_j)| < d \). Therefore, \( \forall n > N_d, \forall k_i \in S_i \setminus E_i, \ s_{i,n}(k_i) \in (s' - d, s' + d) \), and 
\( \forall k_j \in S_j \setminus E_j, \ s_{j,n}(k_j) \in (s' - d, s' + d) \). Suppose now that each player’s action space in the 
the \( n \)th finite game is \( \{ 0 + \frac{m}{10^n} T | m = 0, \ldots, 10^n \} \), i.e., the increment between actions 
is \( \frac{T}{10^n} \). Given \( s_n(k) \) is an equilibrium, the FOC for \( j \) implies

\[
\Pr\left( \{ k_i : s_{j,n}(k_j) - \frac{T}{10^n} \leq s_{i,n}(k_i) < s_{j,n}(k_j) \} \right) 
\frac{T}{10^n} (k_j(T - s_{j,n}(k_j)) + \Delta(k_j, t)) 
\geq 
\Pr\left( \{ k_i : s_{j,n}(k_j) \leq s_{i,n}(k_i) < s_{j,n}(k_j) + \frac{T}{10^n} \} \right) 
\frac{T}{10^n} (k_j(T - s_{j,n}(k_j) - \frac{T}{10^n}) + \Delta(k_j, t))
\]

\( s' > 0 \) and \( s_{j,n}(k_j) \) a best response implies \( (k_j(T - s_{j,n}(k_j)) + \Delta(k_j, t)) > 0 \). If it
were negative, \( k_j \) would be better off choosing lower action, since she must pay her
bid even when she has 0 probability of winning. As \( d \to 0 \) and \( n \to \infty \),

\[
\Pr\left( \{ k_i : s_{j,n}(k_j) \leq s_{i,n}(k_i) < s_{j,n}(k_j) + \frac{T}{10^n} \} \right) \to \Pr(k_i \in S_i \setminus E_i),
\]

which is positive and \( \frac{T}{10^n} \to 0 \). This establishes a contradiction with the Nash conditions. Therefore, \( \forall \) players \( j \), \( \forall s' \in (0, T), \ Pr(s_j^*(k_j) = s') \Pr(\tau_j^{L*}(s')) = 0 \).  

\textbf{Lemma 3} \( \forall i \) \textit{and almost all} \( k_i \), \textit{such that} \( s_{i,n}(k_i) \) \textit{converges to} \( s_i^*(k_i) \),

(1) \( E[u_i(s_i, s_i^*(\cdot), k_i)] \) \textit{is continuous at} \( s_i = s_i^*(k_i) \),

(2) \( E[u_i(s_{i,n}(k_i), s_{j,n}(\cdot), k_i)] \) \textit{converges to} \( E[u_i(s_i^*(k_i), s_i^*(\cdot), k_i)] \).

\textbf{Proof of Lemma 3.} (1) The payoffs from winning, \( ((k_i + 1)(T - s_j) + E[U^L(k_i, t)]) \),
are continuous in \( s_i \). The payoffs from losing, \( (T - s_i + E[U^N(k_i, t)]) \) are continuous in 
\( s_i \). Since \( s_j^*(k) \) has no mass points for \( k_j > 0 \), it follows that \( \forall j \in N, \Pr(\tau_j^{L*}(s_j^*(k_j))) = 0 \). Since, by assumption, \( \Pr(\tau_j^{L*}(\cdot)) = 0 \), it must be that the probability of winning
is continuous at \( s_j^*(k_j) \) \( \forall k_j > 0, \forall j \in N \).

\( \forall j \in N, \Pr(s_j^*(k_j) = s') \Pr(\tau_j^{L*}(s') = 0 = 0 \forall s', \forall i \in L \).
By assumption, \( \Pr(\tau_i^{L*}(s')) = 0 \forall s', \forall k_i \in L \). Hence, the probability of winning is 
continuous \( \forall k_i > 0, \forall i \in L \). \( s_{i,n}(k_i) \) converges to \( s_i^*(k_i) \) \( \forall i \) and approximately \( k_i \).
Hence part (1) of the lemma follows.

(2) Consider \( k_i \) such that \( s_{i,n}(k_i) \to s_i^*(k_i) \) and the probability of winning with 
action \( s_i \) is continuous at \( s_i = s_i^*(k_i) \) (from part (1)).

\[
E[u_i(s_i^*(k_i), s_j^*(k_j), k_i)] - E[u_i(s_{i,n}(k_i), s_j,n(k_j), k_i)] 
= (E[u_i(s_i^*(k_i), s_j^*(k_j), k_i)] - E[u_i(s_{i,n}(k_i), s_j^*(k_j), k_i)]) 
+ (E[u_i(s_{i,n}(k_i), s_j^*(k_j), k_i)] - E[u_i(s_{i,n}(k_i), s_j,n(k_j), k_i)]).
\]
The first difference on the right-hand side converges to 0 as \( n \to \infty \) (by part (1)). It remains to prove that \( E[u_i(s'_i, s_{j,n}(k_j), k_i)] \) converges uniformly to \( E[u_i(s'_i, s^*_j(k_j), k_i)] \) in the neighborhood of \( s'_i = s^*_i(k_i) \). Pick \( \eta > 0, d > 0 \). There exists \( E \) with measure \(< \eta \) and \( N_d \) such that \( \forall k \notin E, \forall n > N_d, \) and \( \forall \) players \( i, |s^*_i(k_i) - s_{i,n}(k_i)| < d \). Then, \( \forall n > N_d, \)

\[
\max_{s'_i \in (s^*_i(k_i) - d, s^*_i(k_i) + d)} \left| E[u_i(s'_i, s_{j,n}(k_j), k_i)] - E[u_i(s'_i, s^*_j(k_j), k_i)] \right|
\]

\[
< \max_{s'_i \in (s^*_i(k_i) - d, s^*_i(k_i) + d)} \left| \Pr(s'_i \in (s_{j,n}(k_j), s^*_j(k_j))|k_j \notin E) \Pr(k_j \notin E)(k_i(T - (s^*_j(k_j) - s)) + \Delta(k_i, t)) + \Pr(s'_i \in (s_{j,n}(k_j), s^*_j(k_j))|k_j \in E) \Pr(k_j \in E)(k_i(T - (s^*_j(k_j) - s)) + \Delta(k_i, t)) \right|
\]

\[
+ \int_{k_j \notin E} |s_{j,n}(k_j) - s^*_j(k_j)| p(k_j) dk_j + \int_{k_j \in E} |s_{j,n}(k_j) - s^*_j(k_j)| p(k_j) dk_j
\]

As \( d \to 0, \)

\[
\Pr(s'_i \in (s_{j,n}(k_j), s^*_j(k_j))|k_j \notin E) \to \Pr(s^*_i(k_i) = s^*_j(k_j)) = 0,
\]

and \( |s_{j,n}(k_j) - s^*_j(k_j)| \) → 0.

Since \( d \) can be chosen so that first two terms are arbitrarily close to 0, and \( \eta \) can be chosen so that the last two terms are arbitrarily close to 0, it follows that \( |E[u_i(s'_i, s_{j,n}(k_j), k_i)] - E[u_i(s'_i, s^*_j(k_j), k_i)]| \) is bounded above by a value arbitrarily close to 0 for \( s'_i \) near \( s^*_i(k_i) \), and part (2) of the lemma follows. \( \blacksquare \)

By Lemma 2, \( s^*_i(k_i) \) is a best response to \( s^*_j(k_j) \) for almost all \( k_i \). \( \forall n \) and almost all \( k_i, \)

\[
E[u_i(s_{i,n}(k_i), s_{j,n}(.), k_i)] \geq E[u_i(s'_i, s_{j,n}(.), k_i)] \forall s' \in \{0 + \frac{m}{10^n}, T|m = 0, ..., 10^n\}
\]

Let \( D^i \) be set of all actions \( s' \) such that for large enough \( N, s' \in \{0 + \frac{m}{10^n}, T|m = 0, ..., 10^n\} \) \( \forall n > N \). From Lemma 3, if \( s' \in D^i, E[u_i(s^*_i(k_i), s^*_j(., k_i)] \geq E[u_i(s'_i, s^*_j(., k_i)] \). Suppose \( s' \notin D^i \). If \( E[u_i(s'_i, s^*_j(., k_i)] \) is continuous at \( s' \), then \( \exists \{s^k\}, s^k \in D^i, \) that converges to \( s' \). Hence \( E[u_i(s_k, s^*_j(., k_i)] \to E[u_i(s'_i, s^*_j(., k_i)] \). If \( Pr(s^*_j(k_j) = s'_j) > 0 \) and \( k_i > 0 \), then \( \exists \delta > 0 \) such that \( s'_i + \delta \) is used on a set of opponent’s types of measure 0 in the limit and such that \( s'_i + \delta \in D^i \). Since \( s'_i + \delta \) wins against \( j \) at \( s'_i \), it follows that for sufficiently small \( \delta, E[u_i(s'_i + \delta, s^*_j(., k_i)] \to E[u_i(s'_i, s^*_j(., k_i)] \). Since \( s'_i + \delta \in D^i \), it must be that \( E[u_i(s^*_i(k_i), s^*_j(., k_i)] \geq E[u_i(s_i + \delta, s^*_j(., k_i)] \), which completes the proof. \( \blacksquare \)

4.7 Proof of Proposition 4

1. This follows from the fact that an agent of type \( k \leq 0 \) is at least as productive using the \( B \) technology as the \( A \) technology and thus only reduces her productive output by allocating a positive amount of time to conflict.
2. The proof proceeds in the same manner as the proof of Proposition 1. The
conditions for the optimal strategies are:

\[ s^N_t(k_i) \in \arg\max((1 - P^N_t(s^N_t))((T - s^N_t) + E[U^N_t(k_i)])) \]
\[ + \int_{\{k_j: s^L_t(k_j) < s^N_t\}} p^L_t(k_j) [(k_i + 1)(T - s^L_t(k_j)) + E[U^L_t(k_i)]] dk_j \]

and

\[ s^L_t(k_i) \in \arg\max((1 - P^N_t(s^L_t))((T - s^L_t) + E[U^N_t(k_i)])) \]
\[ + \int_{\{k_j: s^N_t(k_j) < s^L_t\}} p^N_t(k_j) [(k_i + 1)(T - s^N_t(k_j)) + E[U^L_t(k_i)]] dk_j, \]

where \( E[U^N_t(k_i)] \) is the expected sum of discounted future payoffs for agent \( k_i \) evaluated in period \( t \), conditional on her beginning period \( t + 1 \) in group \( N \), and \( E[U^L_t(k_i)] \) is her expected sum of discounted future payoffs if she begins period \( t + 1 \) in group \( L \). Let \( \Delta_t(k) \) represent the difference \( E[U^L_t(k)] - E[U^N_t(k)] \). As in the proof of Proposition 1, the conditions for the optimal strategies can be manipulated to produce expressions of the form

\[ \int_{\{k_j: s^N_t(k_j) < s'\}} p(k_j) (k' - k'')(T - s(k_j)) dk_j \]
\[ + [\Pr(s(k_j) < s' - \Pr(s(k_j) < s'')] (\Delta(k') - \Delta(k'')) \geq 0 \]

for each group, \( N \) and \( L \). Using the additional fact that \( \Delta(k) \) is increasing in \( k \), if \( k' > k'' \), it must be that \( s' > s'' \) and vice versa.

3. From (9) and (8), the expected utility of agent \( i \) is continuous in \( k_i \). Given that payoffs are continuous with respect to agent type, the equilibrium strategy is continuous with respect to agent type by the same argument as in the proof of Lemma 1, part 2.

4. The existence, continuity and monotonicity of \( K^L_t(s) \), defined on \( s \in (0, \bar{s}^L_t) \) to take values in the interval \((0, \bar{k}^L_t)\), follow from the monotonicity and continuity of \( s^L_t(k) \). For \( K^L_t(s) \) continuous and monotonic and \( p(K^L_t(s)) > 0 \), the differentiability of \( K^L_t(s) \) follows from Lemma 1, part 4 of Fudenberg and Tirole [1986]. The same argument holds for \( K^N_t(s) \).

\[ \boxed{4.8 \text{ Proof of Proposition 5}} \]

It is useful to establish the following lemma before proceeding with the proof of the proposition. Let \( k^* = k \) such that \( P_1(k^*) = \frac{1}{2} \).
Lemma 4  In the second period,
\begin{enumerate}
\item $P_2^N(k) = \frac{N+L}{N} P_1(k) - \frac{L}{N} (P_1(k))^2$ and $P_2^L(k) = (P_1(k))^2$
\item $p_2^N(k) = \frac{N+L}{N} p_1(k) - 2 \frac{L}{N} p_1(k) P_1(k)$ and $p_2^L(k) = 2p_1(k) P_1(k)$
\item $\frac{\partial p_2^N(k)}{\partial k} < 0$
\item $\frac{\partial}{\partial k} \left( \frac{1 - p_2^N(k)}{p_2^L(k)} \right) > 0.$
\end{enumerate}

Proof of Lemma 4. Because the actual population of types is fixed, $\forall t, \forall k$,
\[(N + L) p_1(k) = N p_t^N(k) + L p_t^L(k)\]
and
\[(N + L) P_1(k) = N P_t^N(k) + L P_t^L(k)\]
Solving for $P_2^N(k)$, $P_2^L(k)$, $p_2^N(k)$, and $p_2^L(k)$ yields the expressions given in the statement of the lemma.

Substituting the known expressions for $P_1(k)$ and $p_1(k)$ derived in Appendix 4.4 and taking the appropriate derivatives, it is evident that $\frac{\partial p_2^N(k)}{\partial k} < 0$ and $\frac{\partial}{\partial k} \left( \frac{1 - p_2^N(k)}{p_2^L(k)} \right) > 0$.

Proof of Proposition 5. From Proposition 4, $s_2^N(0) = s_2^L(0) = 0$ and $s_2^N(k)$ and $s_2^L(k)$ are each monotonically increasing in $k$; therefore it must be that either $\exists k \in (0, \bar{k})$ such that $s_2^N(k) = s_2^L(k)$ or either $s_2^N(k) > s_2^L(k)$ or $s_2^N(k) < s_2^L(k)$ for all $k \in (0, \bar{k})$.

Suppose that $s_2^N(k) = s_2^L(k)$ for some $k \in (0, \bar{k})$ and, therefore, that $K_2^L(s) = K_2^N(s)$ for the corresponding value of $s$. From (10), it follows that
\[
\frac{1 - P_t^L(k)}{p_t^L(k)} = \frac{1 - P_t^N(k)}{p_t^N(k)}. \tag{14}
\]

From parts (1) and (2) of Lemma 4,
\[
\frac{1 - P_2^L(k)}{p_2^L(k)} = \frac{1 - (P_1(k))^2}{2p_1(k)P_1(k)}
\]
and
\[
\frac{1 - P_2^N(k)}{p_2^N(k)} = \frac{N - (N + L)P_1(k) - L(P_1(k))^2}{(N + L)p_1(k) - 2Lp_1(k)P_1(k)}.
\]
Substituting into 14 and reducing yields $P_1(k) = 1$, which is false $\forall k \in (0, \bar{k})$. Therefore $s_2^N(k) \neq s_2^L(k)$ and either $s_2^N(k) > s_2^L(k)$ or $s_2^N(k) < s_2^L(k)$ $\forall k \in (0, \bar{k})$.

Given that $s_2^N(0) = s_2^L(0)$, both $s_2^N(k)$ and $s_2^L(k)$ are monotonically increasing in $k$, and there does not exist a $k \in (0, \bar{k}_2)$ such that $s_2^N(k) = s_2^L(k)$, it is enough to show that $s_2^L(k) > s_2^N(k)$ for some $k \in (0, \bar{k}_2)$ to demonstrate that $s_2^L(k) > s_2^N(k)$ for any $k \in (0, \bar{k}_2)$. I prove, by contradiction, that $s_{t+1}^L(k) > s_{t+1}^N(k)$. I then prove that $s_{t+1}^L(k) > s_t^L(k) > s_t^N(k)$. 

31
Suppose $s^L_2(k) < s_1(k)$. Consider $\tilde{k} > 0$ such that $s^L_1(\tilde{k}) = \tilde{s}$, where $\tilde{s}$ is sufficiently close to 0 that $P^N_2(\tilde{s}) > P_1(\tilde{s})$, and such that $\tilde{k}$ is sufficiently close to 0 that an agent of that type is more likely to lose than to win in the first period; thus $p^N_2(\tilde{k}) > p_1(\tilde{k})$. (From Lemma 4 and the continuity of $P^N_2(k)$ and $s^L_2(k)$, $\exists k > 0$ such that $P^N_2(\tilde{s}) > P_1(\tilde{s})$.) From the first-order conditions (10), it must be that

$$s^L_1(\tilde{k}) = \tilde{s} = -\frac{1 - P^N_1(\tilde{s})}{kp^N_1(\tilde{s})} + T + \frac{\Delta(\tilde{k})}{k}.$$ 

$P^N_2(\tilde{s}) > P_1(\tilde{s})$ and $s^L_2(k) < s_1(k)$ imply $p^N_2(\tilde{s}) < p_1(\tilde{s})$. Combining inequalities, $p^N_2(\tilde{s}) < p_1(\tilde{s}) = p_1(k) < p^N_2(\tilde{k})$. Given part (3) of Lemma 4, $p^N_2(\tilde{s}) < p^N_2(k)$ implies $s^N_2(k) < \tilde{s} = s^N_1(\tilde{k})$.

Given the assumption that $s^L_2(k) < s_1(k)$, from the first-order conditions (10), it must be that

$$\frac{1 - P^N_1(K^N_t(s^L_1(k)))}{p^N_1(K^N_t(s^L_1(k)))K^N_t(s^L_1(k))} > \frac{1 - P_1(k)}{p_1(k)}.$$ 

From Lemma 4, $P^N_2(k) > P_1(k) \forall k < \tilde{k}$ and $P^N_2(k) > P_1(k) \forall k < k^*$, where $k^*$ is such that $P_1(k^*) = \frac{1}{2}$. Hence

$$\frac{1 - P^N_2(k)}{p^N_2(k)} < \frac{1 - P_1(k)}{p_1(k)}.$$

Combined with the fact that $\frac{\partial}{\partial k} \left( \frac{1 - P^N_1(k)}{p^N_2(k)} \right) > 0$ from Lemma 4, this implies that $K^N_t(s^L_1(k)) > k$, i.e. that $s^N_2(k) > s_1(k)$, a contradiction.

Therefore, $s_1(k) < s^L_2(k)$. It is readily shown, using the same reasoning as in the preceding paragraph, that $s^N_2(k) < s_1(k)$ follows from $s_1(k) < s^L_2(k)$.

Because $s^L_2(k) < s_1(k) < s^L_2(k)$, $P^N_3(0) > P^N_2(0)$, $P^L_3(0) < P^L_2(0)$, and $p^L_3(k) < p^L_2(k)$ for $k < k^*$; hence, it can be shown that $s^L_2(k) > s^N_1(k)$ for all $k \in (0, \bar{k})$ for $t \geq 3$ using the same reasoning.

\textbf{4.9 Proof of Theorem 2}

1. Because the population of agents is fixed and finite, there is a finite number of combinations of $L$ agents, that is there is a finite number of divisions of agents between sets $N$ and $L$. Because, from Proposition 4, $\forall t$, $\frac{\partial p^L_t}{\partial k} > 0$ and $\frac{\partial s^N_t}{\partial k} > 0$, and from Proposition 5, $\forall t > 1$, $s^L_t(k) > s^N_t(k)$, agents in $L$ can be replaced only by agents in $N$ of higher type $k$. Therefore, in a finite number of periods, $L$ and $N$ converge to $L^*$ and $N^*$ such that, $\forall i \in L^*$ and $\forall j \in N^*$, $s^L_t(k_i) > s^N_t(k_j)$.$^{12}$

2. As $t \to \infty$, $\forall k$ such that $\lim_{t \to \infty} p^N_t(k) > 0$, $P^L_t(K^L_t(s^N_t(k))) \to 0$. Therefore, $\forall k$ such that $\lim_{t \to \infty} p^N_t(k) > 0$, $\frac{1 - P^L_t(K^L_t(s^N_t(k)))}{p^L_t(K^L_t(s^N_t(k)))} \to \infty$.

$^{12}$I thank David Austen-Smith for drawing my attention to the fact that the time required for convergence is finite.
Let $\hat{k}$ be the highest type in $N^*$. Because $\partial s^N_k \geq 0$ (from Proposition 4), if $s^N(\hat{k}) = 0$, then $s^N(k) \to 0 \ \forall k \in N^*$. From (10),

$$s_t^N(k) = T + \frac{\Delta(k)}{k} - \frac{1 - P^L_t(K^L_t(s^N_t(k)))}{kp^L_t(K^L_t(s^N_t(k))) K^L_t(s^N_t(k))}.$$ 

Therefore, $s^N(\hat{k}) = 0$ when

$$\hat{k}T + \Delta(\hat{k}) \leq \frac{1 - P^L_t(K^L_t(\hat{s}_t^N(\hat{k})))}{p^L_t(K^L_t(\hat{s}_t^N(\hat{k}))) K^L_t(\hat{s}_t^N(\hat{k}))}.$$

Given that $\hat{k}T + \Delta(\hat{k}) \leq \hat{k}T + \frac{\delta}{1-\delta} \hat{k}T = \frac{\hat{k}T}{1-\delta}$, it is sufficient that

$$\frac{\hat{k}T}{1-\delta} \leq \frac{1 - P^L_t(K^L_t(s^N(\hat{k})))}{p^L_t(K^L_t(s^N(\hat{k}))) K^L_t(s^N(\hat{k}))}.$$

Given that $\lim_{t \to \infty} \frac{1 - P^L_t(K^L_t(s^N_t(k)))}{p^L_t(K^L_t(s^N_t(k))) K^L_t(s^N_t(k))} = \infty$ and $\frac{\hat{k}T}{1-\delta}$ is finite and constant with respect to time, there must exist some $t^*$ such that

$$\frac{1 - P^L_t(K^L_{ts}(s^N_{ts}(\hat{k})))}{p^L_t(K^L_{ts}(s^N_{ts}(\hat{k}))) K^L_{ts}(s^N_{ts}(\hat{k}))} \geq \frac{\hat{k}T}{1-\delta} > \frac{1 - P^L_{t^*-1}(K^L_{t^*-1}(s^N_{t^*-1}(\hat{k})))}{p^L_{t^*-1}(K^L_{t^*-1}(s^N_{t^*-1}(\hat{k}))) K^L_{t^*-1}(s^N_{t^*-1}(\hat{k}))}.$$


4.10 Proof of Theorem 3

Consider two distinct economies such that

$$\hat{s}_t^L(k) - \hat{s}_t^N(k) > \hat{s}_t^L(k) - \hat{s}_t^N(k) \quad \forall t > 1, \forall k > 0,$$

where (suppressing subscripts and superscripts) $\hat{s}(k)$ and $\hat{s}(k)$ denote the respective equilibrium strategies played in these economies. By Proposition 5, this implies $\hat{s}_t^L(k) > \hat{s}_t^L(k)$ and $\hat{s}_t^N(k) < \hat{s}_t^N(k)$. Therefore

$$\hat{K}^L_t(\hat{s}_t^N(k)) < \hat{K}^L_t(\hat{s}_t^N(k)) < \hat{K}^N_t(\hat{s}_t^N(k)) < \hat{K}^N_t(\hat{s}_t^N(k)).$$

Hence

$$\hat{P}^L_t(\hat{K}^L_t(\hat{s}_t^N(k))) < \hat{P}^L_t(\hat{K}^L_t(\hat{s}_t^N(k))) \quad \forall t > 1,$$

i.e., the proportion of agents in $\hat{L}_t$ that $j \in \hat{N}_t$ of type $k$ would defeat is less than the proportion of agents in $\hat{L}_t$ that $j \in \hat{N}_t$ of type $k$ would defeat.
From (15), it follows that, \( \forall k \), such that \( \lim p_t^N(k) > 0 \), as \( t \to \infty \), \( \tilde{P}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right) \) goes to 0 faster than \( \tilde{P}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right) \). It follows that, as \( t \to \infty \),

\[
\frac{1 - \tilde{P}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right)}{\tilde{p}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right) \tilde{K}_t^L(\tilde{s}_t^N(k))} \to \infty \frac{1 - \tilde{P}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right)}{\tilde{p}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right) \tilde{K}_t^L(\tilde{s}_t^N(k))}. \]

Following the proof of part 2 of Theorem 2, \( \hat{t}^* \) such that

\[
\frac{1 - \tilde{P}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right)}{\tilde{p}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right) \tilde{K}_t^L(\tilde{s}_t^N(k))} \geq \frac{kT}{1 - \delta} \frac{1 - \tilde{P}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right)}{\tilde{p}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right) \tilde{K}_t^L(\tilde{s}_t^N(k))}.
\]

is less than \( \hat{t}^* \) such that

\[
\frac{1 - \tilde{P}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right)}{\tilde{p}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right) \tilde{K}_t^L(\tilde{s}_t^N(k))} \geq \frac{kT}{1 - \delta} \frac{1 - \tilde{P}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right)}{\tilde{p}_t L \left( \tilde{K}_t^L(\tilde{s}_t^N(k)) \right) \tilde{K}_t^L(\tilde{s}_t^N(k))}.
\]

From (15), the probability that \( j \) would move from \( \tilde{N}_t \) into \( \tilde{L}_{t+1} \) is less than the probability that \( j \in \tilde{N}_t \) of type \( k \) would move from \( \tilde{N}_t \) into \( \tilde{L}_t \). Likewise,

\[
1 - \tilde{P}_t^N \left( \tilde{K}_t^N(\tilde{s}_t^N(k)) \right) < 1 - \tilde{P}_t^N \left( \tilde{K}_t^N(\tilde{s}_t^N(k)) \right),
\]

i.e., the proportion of agents in \( \tilde{N}_t \) who defeat \( i \in \tilde{L} \) of type \( k \) is less than the proportion of agents in \( \tilde{N}_t \) who could defeat \( i \in \tilde{L} \) of type \( k \). Hence the probability of an agent of type \( k \) moving from \( \tilde{L}_t \) into \( \tilde{N}_{t+1} \) is less than the probability of an agent of type \( k \) moving from \( \tilde{L}_t \) into \( \tilde{N}_{t+1} \). Therefore,

\[
\left| \tilde{N}^* \cap (k^*, \frac{b - c}{c}) \right| > \left| \tilde{N}^* \cap (k^*, \frac{b - c}{c}) \right|.
\]

Department of Politics, New York University

References


