Journal of Economic Theory and Econometrics, Vol. 24, No. 3, Sep. 2013, 306–319

A Stag Hunt Game with Social Norms*

Seung Han Yoo†

Abstract This paper studies a stag hunt game in which each player's payoff depends on the level of social norms in a community. First, we establish conditions under which without social norms, there exists a unique equilibrium regarding the cooperation level, but with social norms, there can be multiple equilibria. Second, we provide the comparative statics analysis such that for a statics framework, the local stability of an interior equilibrium works as a sufficient condition under which the cooperation level increases as the degree of social norms increases; and for a dynamic framework, the cooperation level always increases with social norms.

Keywords Stag hunt game, Social norms

JEL Classification D71

---

*I am grateful to Kaushik Basu and Sung-Ha Hwang for their valuable comments. I would also like to thank two anonymous referees of the journal for helpful suggestions. Of course, all remaining errors are mine.

†Department of Economics, Korea University, Seoul, Republic of Korea 136-701 (e-mail: shyoo@korea.ac.kr).

Received June 25, 2013, Revised September 13, 2013, Accepted September 17, 2013

1. INTRODUCTION

Understanding social norms is important in economics because they may have direct relationships with various economic activities (see Elster (1989) pp. 100–101 for a list of economy-related social norms). Lindbeck, Nyberg and Weibull (1999) is one of the influential papers that adopt social norms in a model formally, and the paper shows that under a social norm, e.g., social stigma, a society can have multiple equilibria regarding unemployment rates.[1]

In particular, they find that there are two types of equilibria: a "good" equilibrium and a "bad" equilibrium. In the former, anticipating that others will choose to work, and that the social stigma associated with not working will thus be high, more workers will find it optimal to work, whereas in the latter, anticipating that others will choose *not* to work, and that the social stigma will thus be low, more workers will find it optimal to depend on unemployment benefits.

In this paper, we extend their framework to a stag hunt game with *incomplete information* and provide the characterization of equilibria regarding the "cooperation level" of a community, the fraction of hunters who choose to hunt a stag, and the comparative statics on the changes in the cooperation level with respect to the changes in the degree of social norms. If a hunter chooses to hunt a hare instead of a stag, for instance, he or she may experience disutility from social stigma associated with the behavior in the village.

We first establish conditions under which without social norms, there exists a unique equilibrium regarding the cooperation level, whereas with social norms, there can be multiple equilibria as in Lindbeck, Nyberg and Weibull (1999): one equilibrium in which more hunters choose to hunt a stag and the other in which fewer hunters choose to hunt a stag. For the first result, we find that it is necessary to have hunter types whose dominant strategy is to choose either a stag or a hare, and sufficient to assume a strictly decreasing net benefit function of a "threshold hunter type" with a *strongly* convex form of social norms.

The second main contribution of this paper is to introduce the comparative statics analysis, which has not been discussed thoroughly in the literature.[2] It is commonly accepted that the greater the degree of social norms, the higher the cooperation level, and indeed, this is probably the reason that a community "in-

[1] "Individuals may, however, also experience disutility from accepting the transfer due to embarrassment or social stigma associated with living on public transfers rather than on one's own work. Such embarrassment is likely to be weaker, the greater the number of individuals in society who live on the transfer" (p.7, Lindbeck, Nyberg and Weibull (1999)).

[2] The comparative statics hold without assuming the conditions related to either uniqueness or multiplicity.

tends" to implement social norms. However, this paper shows that social norms may not enhance the cooperation level. Hence, it is important to examine conditions under which a society can have a higher level of cooperation or a lower level. First, for a static analysis, we establish that the local stability of an interior equilibrium works as a sufficient condition under which the cooperation level increases as the degree of social norms increases.[3] In addition, if an equilibrium is locally *unstable*, then the cooperation level can strictly decrease with the degree of social norms. We also present the comparative statics from a dynamic framework in which it is shown that as the degree of social norms increases, the cooperation level always increases.

We define social norms as an equilibrium phenomenon as in Basu and Weibull (2004), who show, using a coordination game, that a norm is *not innate* in human nature but is an outcome of an equilibrium. The theory of repeated games also explains how social norms in a "large" population are born, evolve and may die. Kandori (1992) provides a model in which a cooperative equilibrium can be sustained as an equilibrium either with a behavioral rule on people without information processing or with a system transmitting information in a society. Ghosh and Ray (1996) suggests a model without information flows and proves that a cooperation level in a community can be explained in terms of people following a certain set of rules with regard to how to develop their relationship with others. Both papers suppose either an information processing system or a pattern of behavior to show a cooperative action, which is defined as a social norm. In this paper, a cooperative equilibrium arises from the amount of negative payoff that a hunter receives as a result of social norms, such as social stigma, when the hunter chooses to hunt a hare, instead of a stag.

We provide a stag hunt game without social norms in Section 2, a stag hunt game with social norms in Section 3, comparative statics in Section 4, and, finally, an illustrative example in Section 5.

2. MODEL

Consider a variant of a stag hunt game with incomplete information. Every two hunters in a village are randomly matched and go on a hunt. They must cooperate in order to hunt a stag, and each hunter's payoff from capturing a stag is B. However, each can hunt a hare by himself or herself, and hunter i's payoff

[3] See Definition 1 for a formal definition of the local stability.

from a hare is C_i. The game for every two hunters is represented as below:

	Stag	Hare
Stag	B,B	$0,C_j$
Hare	$C_i,0$	C_i,C_j

(1)

We assume that the payoff from choosing to hunt a hare C_i is only known to each hunter i, and the other hunters know that C_i is independently and identically drawn from a differentiable distribution function F with its density $F' \equiv f > 0$ where its support is given as $\left[\underline{C}, \overline{C}\right]$, satisfying $\overline{C} > \underline{C},\ B > \underline{C}$ and $\overline{C} > 0$. The conditions, $\overline{C} > \underline{C},\ B > \underline{C}$ and $\overline{C} > 0$, ensure that there exists an interval $[C_1, C_2]$ with $C_1 < C_2$ such that for each $C_i \in [C_1, C_2]$, $B > C_i > 0$. As in a stag hunt game with complete information, the best response of a hunter with $C_i \in [C_1, C_2]$ is to choose Stag if the hunter expects his opponent to choose Stag, and to choose Hare if the hunter expects his opponent to choose Hare. However, we also allow hunters to have types whose *dominant strategy* is to choose either Stag or Hare.[4]

Since each hunter has two actions with a strictly monotonic C_i in the payoff matrix, in equilibrium, hunter i's strategy $h_i : [\underline{C}, \overline{C}] \to \{\text{Stag, Hare}\}$ must be a cut-off strategy with a threshold type k_i such that

$$h_i(C_i) = \begin{cases} \text{Stag} & \text{if} \quad C_i < k_i, \\ \text{Hare} & \text{if} \quad C_i > k_i. \end{cases}$$

Hunter i with type C_i greater than a threshold k_i chooses a hare, and hunter i with type C_i smaller than a threshold $k_i \in [\underline{C}, \overline{C}]$ chooses a stag. Then, hunter i's expected payoff from choosing a stag is given as $F(k_j)B$, where k_j is a threshold for hunter j's strategy, and the expected payoff from choosing a hare is C_i. A Bayesian-Nash equilibrium is defined as $(k_1^*, k_2^*) \in [\underline{C}, \overline{C}]^2$ such that for each $i \in \{1,2\}$,

$$\begin{array}{lll} F(k_j^*)B \leq k_i^* & \text{if} & k_i^* = \underline{C}, \\ F(k_j^*)B = k_i^* & \text{if} & k_i^* \in (\underline{C}, \overline{C}), \\ F(k_j^*)B \geq k_i^* & \text{if} & k_i^* = \overline{C}. \end{array} \tag{2}$$

We first show that an equilibrium must be symmetric with $k_i^* = k_j^* = k^*$.

[4]It is not exceptional that one can find this kind of variant of a well-known game by introducing incomplete information such that some types always choose a certain designated action. For examples, see the classical paper by Kreps, Milgrom, Roberts and Wilson (1982) in which "An alternative way to model this is to assume that ROW has available *all* the strategies above, but that with probability δ, ROW's payoffs are not as above but rather make playing Tit-for-Tat strongly dominant" (p. 247).

Lemma 1. *An equilibrium must be symmetric such that* $k_i^* = k_j^* = k^*$.

Proof. Suppose $k_j^* \neq k_i^* \in [\underline{C}, \overline{C}]$. WLOG, let $k_j^* > k_i^*$. From (2),

$$F\left(k_i^*\right)B \geq k_j^* \ (= \text{if } k_j^* < \overline{C}), \text{ and } F(k_j^*)B \leq k_i^* \ (= \text{if } k_i^* > \underline{C}).$$

Then, $F\left(k_i^*\right)B > F(k_j^*)B$ and $F\left(k_i^*\right) > F(k_j^*)$, which implies $k_i^* > k_j^*$, a contradiction. □

Then, given a threshold type k, $F\left(k\right)B$ is each hunter's incentive to choose a stag, and $F\left(k\right)B - k$ is *the threshold hunter type's net benefit to choose a cooperative action*. We assume a class of supports for the CDF $[\underline{C}, \overline{C}]$ that is *sufficient* for an interior equilibrium to exist.

Assumption 1. *The CDF F has a support* $[\underline{C}, \overline{C}]$ *that satisfies*

$$[F\left(\underline{C}\right)B - \underline{C}][F\left(\overline{C}\right)B - \overline{C}] = -\underline{C}[B - \overline{C}] < 0.$$

There can be two cases: (i) $0 < \underline{C} < \overline{C} < B$ so that for each hunter, the payoff from a stag is greater than the payoff from a hare while the payoff from a hare is greater than 0, and (ii) $\underline{C} < 0 < B < \overline{C}$ so that the lowest type's payoff from a hare is so low that it is negative, and the highest type's payoff is so high that it is even greater than the payoff from a stag. Since $F\left(k\right)B - k$ is continuous on $[\underline{C}, \overline{C}]$ and $[F\left(\underline{C}\right)B - \underline{C}][F\left(\overline{C}\right)B - \overline{C}] < 0$, there exists an interior equilibrium $k^* \in (\underline{C}, \overline{C})$ satisfying $F\left(k^*\right)B = k^*$.

Note that if $0 < \underline{C} < \overline{C} < B$, there exist boundary equilibria with $k^* = \underline{C}$ and $k^* = \overline{C}$. For $k = \underline{C}$, $F(\underline{C})B - \underline{C} = -\underline{C} < 0$, and for $k = \overline{C}$, $F(\overline{C})B - \overline{C} = B - \overline{C} > 0$, which satisfies the boundary equilibrium conditions in (2). It follows that if $0 < \underline{C} < \overline{C} < B$, there exist at least three equilibria, $k^* = \underline{C}$, $k^* = \overline{C}$ and an interior equilibrium, which in turn implies that $\underline{C} < 0 < B < \overline{C}$ is a "necessary condition" for uniqueness. Furthermore, Proposition 1 provides a sufficient condition under which there exists a unique equilibrium, and it is an interior $k^* \in \left(\underline{C}, \overline{C}\right)$.

Proposition 1. *If* $F\left(k\right)B - k$ *is a strictly decreasing function of k on* $[\underline{C}, \overline{C}]$*, then there exists a unique equilibrium* $k^* \in \left(\underline{C}, \overline{C}\right)$*, and the probability of hunting a stag is* $F\left(k^*\right)$.

Proof. It is clear that a strictly monotone $F\left(k\right)B - k$ yields a unique interior threshold $k^* \in \left(\underline{C}, \overline{C}\right)$. Furthermore, since $F\left(k\right)B - k$ is strictly decreasing and $F\left(k^*\right)B - k^* = 0$, for $\underline{C} < k^*$, we have $F(\underline{C})B - \underline{C} > 0$, and for $\overline{C} > k^*$, $F(\overline{C})B - \overline{C} < 0$, which violates the boundary equilibrium conditions in (2). □

For the comparative statics in Section 4, we classify social norm equilibria into different groups by introducing a simple dynamic process, similar to the *Cournot adjustment process*. Denote by λ the fraction of hunters in the population who choose to hunt a stag, and in this dynamic process, each hunter's expectation about the fraction at period t, λ_t, is given as the previous period's fraction. Hence,

$$\lambda_t = F\left(k_{t-1}\right) \text{ for } t = 0,1,2,..., \tag{3}$$

which implies

$$k_t = \lambda_t B = F\left(k_{t-1}\right) B \text{ for } t = 0,1,2,.... \tag{4}$$

The definition of local stability in this context is introduced.

Definition 1. *An equilibrium threshold $k^* \in (\underline{C}, \overline{C})$ is locally stable if there exists an open interval (a,b) such that $k^* \in (a,b)$, and for each $k_0 \in (a,b)$ with $k_0 \neq k^*$, k_t converges to k^* as $t \to \infty$.*

We establish the condition for an equilibrium threshold k^* to be locally stable in the following proposition.

Proposition 2. *An equilibrium threshold $k^* \in (\underline{C}, \overline{C})$ is locally stable if and only if there exists an open interval (a,b) such that $k^* \in (a,b)$ and for each $k_0 \in (a,k^*)$, $F(k_0)B - k_0 > 0$ and for each $k_0 \in (k^*,b)$, $F(k_0)B - k_0 < 0$.*

Proof. Part 1. Suppose that there exists an open interval (a,b) such that $k^* \in (a,b)$ and for each $k_0 \in (a,k^*)$, $F(k_0)B - k_0 > 0$ and for each $k_0 \in (k^*,b)$, $F(k_0)B - k_0 < 0$. It follows from (4) and $f > 0$ that if $k_t > k_{t-1}$ ($k_t < k_{t-1}$) for all $t = 0,1,2,...,$

$$\begin{aligned} k_{t+1} - k_t &= F\left(k_t\right) B - k_t = F\left(k_t\right) B - F\left(k_{t-1}\right) B \\ &= [F\left(k_t\right) - F\left(k_{t-1}\right)] B > 0 \ (<0). \end{aligned}$$

Case 1. For each $k_0 \in (a,k^*)$, $F(k_0)B - k_0 > 0$, so $k_1 > k_0$. Hence, $k_{t+1} > k_t$ for all $t = 0,1,2,...,$ and $\lim_{t\to\infty} k_t = k^*$.

Case 2. For each $k_0 \in (k^*,b)$, $F(k_0)B - k_0 < 0$, so $k_1 < k_0$. Hence, $k_{t+1} < k_t$ for all $t = 0,1,2,...,$ and $\lim_{t\to\infty} k_t = k^*$.

Part 2. Now, suppose that an equilibrium threshold $k^* \in (\underline{C}, \overline{C})$ is locally stable, and WLOG, for each $a < k^*$, there exists $k_a \in (a,k^*)$ such that $F(k_a)B - k_a \leq 0$. Since $k^* \in (\underline{C}, \overline{C})$ is locally stable, for some $k_0 < k^*$, k_t converges to k^* as $t \to \infty$. We can show a contradiction for each case below.

Case 1. For k_0, $F(k_0)B - k_0 = 0$. Then, k_t converges to k_0 as $t \to \infty$.

Case 2. For k_0, $F(k_0)B - k_0 > 0$. Then, by the intermediate value theorem, there exists $k' \in (k_0, k_a]$ such that $F(k')B - k' = 0$. Using the same argument as above, k_t converges to k' as $t \to \infty$.

Case 3. For k_0, $F(k_0)B - k_0 < 0$. Using the same argument as above, k_t converges to $k'' < k_0$ as $t \to \infty$ if it converges. □

If $f(k^*)B < 1$, the threshold type's net benefit function $F(k)B - k$ is a strictly decreasing function of k at k^*, which implies the condition in Proposition 2. However, the reverse relationship may *not* be true. Suppose that $F(k)B - k$ is given as $-k^3$. Then, $k^* = 0$ and k^* is locally stable given Proposition 2, but

$$\left.\frac{d(-k^3)}{dk}\right|_{k=k^*} = 0. \tag{5}$$

A simple sufficient condition for the local stability can be derived in Remark 1.

Remark 1. *If $f(k^*)B < 1$, there exists an open interval (a,b) such that $k^* \in (a,b)$ and for each $k_0 \in (a,k^*)$, $F(k_0)B - k_0 > 0$ and for each $k_0 \in (k^*,b)$, $F(k_0)B - k_0 < 0$.*

Proof. Note that $f(k^*)B < 1$ can be rewritten as $d[F(k^*)B - k^*]/dk < 0$, and $F(k^*)B - k^* = 0$. □

In the following section, we introduce social norms and examine how the community can have multiple equilibria.

3. STAG HUNT WITH SOCIAL NORMS

Let a differentiable function $s : [0,1] \times \mathbb{R}_+ \to \mathbb{R}_+$ be the amount of negative payoff that a hunter receives from social norms when hunting a hare with $s(\lambda, \alpha)$, where α *measures the degree of social norms*. We assume that s is a strictly increasing function of λ for $\alpha > 0$, and a strictly increasing function of α for $\lambda > 0$ such that for each λ, as α increases, the degree of social norms increases, in other words, $s(\lambda, \alpha') > s(\lambda, \alpha)$ for all $\alpha' > \alpha$. If $\alpha = 0$, there is no negative payoff from social norms, that is, $s(\lambda, \alpha) = 0$ for all λ. Denote by $s_\lambda(\lambda, \alpha)$ the partial derivative with respect to λ and $s_\alpha(\lambda, \alpha)$ the partial derivative with respect to α.

The stag hunt game with the social norms is given as:

	Stag	Hare
Stag	B, B	$0, C_2 - s(\lambda,\alpha)$
Hare	$C_1 - s(\lambda,\alpha), 0$	$C_1 - s(\lambda,\alpha), C_2 - s(\lambda,\alpha)$

As in the case without social norms, hunter i's expected payoff from choosing to hunt a stag is $F(k)B$, but here, the expected payoff from choosing a hare with the social norms is $C_i - s(\lambda,\alpha)$.

A social norm equilibrium requires two conditions; each group has a Bayesian-Nash equilibrium, and each hunter's belief about λ must be correct in an equilibrium (*rational expectations*) such that $\lambda = F(k^*)$ where k^* is an equilibrium threshold. Then, a symmetric equilibrium is defined as $k^* \in [\underline{C},\overline{C}]$ such that

$$\begin{array}{lll} F(k^*)B+s(F(k^*),\alpha) \leq k^* & \text{if} & k^* = \underline{C}, \\ F(k^*)B+s(F(k^*),\alpha) = k^* & \text{if} & k^* \in (\underline{C},\overline{C}), \\ F(k^*)B+s(F(k^*),\alpha) \geq k^* & \text{if} & k^* = \overline{C}. \end{array} \tag{6}$$

We show in the following Proposition that even if $F(k)B-k$ is strictly decreasing as in Proposition 1, a strictly convex function $s(\lambda,\alpha)$ with respect to λ can result in multiple equilibria. With multiple equilibria, two hunters in one group may have expectations about the equilibrium strategy that are different from those of the two hunters in the other group. Hence, we assume that members in each group have "common beliefs" about k.[5]

Proposition 3. *If $F(k)B-k$ is a strictly decreasing function of k on $[\underline{C},\overline{C}]$ and $s(\lambda,\alpha)$ is a strictly convex function of λ satisfying conditions $\lim_{\lambda\to 0} s_\lambda(\lambda,\alpha) = 0$ and $\lim_{\lambda\to 1} s_\lambda(\lambda,\alpha) = +\infty$, then for $\underline{C}$ sufficiently close to 0, there are multiple equilibria.*

Proof. Since $F(k)B-k$ is strictly decreasing and $F(k^*)B-k^* = 0$, for $\underline{C} < k^*$, we have

$$F(\underline{C})B-\underline{C} = F(\underline{C})B+s(F(\underline{C}),\alpha)-\underline{C} > 0. \tag{7}$$

Since $F(k)B-k$ is strictly decreasing and $\lim_{\lambda\to\underline{C}} s_\lambda(\lambda,\alpha) = 0$, we have

$$\left.\frac{d\,[F(k)B+s(F(k),\alpha)-k]}{dk}\right|_{k=\underline{C}} < 0,$$

which implies that there exists a (small) interval containing $\underline{C}$ such that $F(k)B+s(F(k),\alpha)-k$ is a strictly decreasing function of k on the interval. Now, if $\underline{C}$ is sufficiently close to 0, there exists $k' > \underline{C}$ such that

$$F(k')B+s(F(k'),\alpha)-k' < 0. \tag{8}$$

[5]In Proposition 3, $\lim_{\lambda\to\underline{C}} s_\lambda(\lambda,\alpha) = 0$ and $\lim_{\lambda\to\overline{C}} s_\lambda(\lambda,\alpha) = +\infty$ can be called "reversed" *Inada conditions*.

From (7) and (8), there exists an equilibrium threshold $k_1^* \in (\underline{C}, k')$. On the other hand, $\lim_{\lambda \to \overline{C}} s_\lambda(\lambda, \alpha) = +\infty$ entails that for k'' sufficiently close to $\overline{C}$,

$$F(k'')B + s(F(k''), \alpha) - k'' > 0. \tag{9}$$

Hence, it follows from (8) and (9) that there exists an equilibrium threshold $k_2^* \in (k', k'')$. □

The local stability in this section is the one with social norms, but its definition is easily derived from the local stability without social norms in Definition 1 and Proposition 2, by replacing $F(k_0)B - k_0$ with $F(k_0)B + s(F(k_0), \alpha) - k_0$.

Remark 2. *An equilibrium threshold $k^* \in (\underline{C}, \overline{C})$ is locally stable if and only if there exists an open interval (a, b) such that $k^* \in (a, b)$ and for each $k_0 \in (a, k^*)$, $F(k_0)B + s(F(k_0), \alpha) - k_0 > 0$ and for each $k_0 \in (k^*, b)$, $F(k_0)B + s(F(k_0), \alpha) - k_0 < 0$.*

The sufficient condition in Remark 1 can be extended to the one with social norms as below:

$$f(k^*)B + s_\lambda(F(k^*), \alpha) < 1. \tag{10}$$

The following Corollary shows that given a milder condition than Proposition 3, there exists a locally stable equilibrium. The example in Section 5 illustrates multiple equilibria with a stable equilibrium k_1^* and an unstable equilibrium k_2^*.

Corollary 1. *If $F(k)B - k$ is a strictly decreasing function of k on $[\underline{C}, \overline{C}]$ and $\lim_{\lambda \to 0} s_\lambda(\lambda, \alpha) = 0$, then for $\underline{C}$ sufficiently close to 0, there exists a locally stable equilibrium.*

Proof. Since $F(k)B + s(F(k), \alpha) - k$ is a strictly decreasing function of k at $k_1^* \in (\underline{C}, k')$ from the proof of Proposition 3, it satisfies the local stability condition. □

In addition, if $F(k)B + s(F(k), \alpha) - k$ is a strictly decreasing function of k on $[\underline{C}, \overline{C}]$, then we can show that there exists a *unique* stable equilibrium $k^* \in (\underline{C}, \overline{C})$, and the probability of hunting a stag is $F(k^*)$, similar to the uniqueness result in Proposition 1.[6]

[6] A strictly monotone $F(k)B + s(F(k), \alpha) - k$ yields a unique threshold $k^* \in (\underline{C}, \overline{C})$. Furthermore, since $F(k)B + s(F(k), \alpha) - k$ is strictly decreasing and $F(k^*)B + s(F(k^*), \alpha) - k^* = 0$, for $\underline{C} < k^*$, we have $F(\underline{C})B + s(F(\underline{C}), \alpha) - \underline{C} > 0$, and for $\overline{C} > k^*$, $F(\overline{C})B + s(F(\overline{C}), \alpha) - \overline{C} < 0$, which violates the boundary equilibrium conditions in (6).

4. COMPARATIVE STATICS

We study comparative statics with respect to changes in social norms by examining how each equilibrium threshold point changes as the degree of social norm α changes. Choose an *interior* equilibrium threshold $k^* \in (\underline{C}, \overline{C})$, which exists if α is sufficiently close to 0 as in Section 2 with assumption 1, and denote by $k(\alpha)$ an equilibrium threshold point given α.

The first main result in this section shows that if $k(\alpha)$ is a locally stable point with social norms, then as the degree of social norms increases from α to α', an equilibrium threshold point increases from $k(\alpha)$ to $k(\alpha')$. Note that the result is given for *any two points* α', α such that α' is sufficiently close to α so that we can apply a milder local stability condition in Remark 2, not with the *sufficient condition* in (10). Moreover, the following results hold without assuming the previous conditions related to either uniqueness or multiplicity.

Proposition 4. *If $k(\alpha) \in (\underline{C}, \overline{C})$ is a locally stable point, then for any pair $\alpha' > \alpha$ such that α' is sufficiently close to α, $k(\alpha') > k(\alpha)$.*

Proof. Since $k(\alpha) \in (\underline{C}, \overline{C})$ is locally stable with social norms, there exists an open interval (a,b) such that for each $k_0 \in (k,b)$, $F(k_0)B + s(F(k_0), \alpha) - k_0 < 0$. For $\alpha' > \alpha$, we have

$$F(k(\alpha))B + s(F(k(\alpha)), \alpha') > k(\alpha).$$

If α' is sufficiently close to α, there exists $k_0 \in (k(\alpha), b)$ such that

$$F(k_0)B + s(F(k_0), \alpha') < k_0.$$

Hence, the intermediate value theorem entails that there exists $k(\alpha') \in (k(\alpha), k_0)$ such that

$$F(k(\alpha'))B + s(F(k(\alpha')), \alpha') = k(\alpha').$$

Hence, $k(\alpha') > k(\alpha)$ for $\alpha' > \alpha$ such that α' is sufficiently close to α. □

We strengthen the above result by imposing the stronger condition, the sufficiency, in (10). Without this, if the function is given as $-k^3$ like the counter example in (5), the *implicit function theorem* cannot be applied. Hence, although we use the same notation $k(\alpha)$ for both Propositions 4 and 5, with a slight abuse of the notation, the former is for any given two points, whereas the latter is for an interval around α, where the implicit function theorem is applied.

Proposition 5. *If $k(\alpha) \in (\underline{C}, \overline{C})$ is a locally stable point satisfying (10), then there exists a neighborhood of α such that an equilibrium threshold k^* is a strictly increasing function of α in the neighborhood of α.*

Proof. Suppose that $k(\alpha) \in (\underline{C}, \overline{C})$ is a locally stable point satisfying

$$f(k(\alpha))B + s_\lambda(F(k(\alpha)), \alpha) < 1, \tag{11}$$

which is sufficient for an equilibrium to be locally stable with social norms. Then, the implicit function theorem entails that there exists a neighborhood of α such that k^* is a function of α, and

$$k'(\alpha) = -\frac{s_\alpha(F(k(\alpha)), \alpha)}{f(k(\alpha))B + s_\lambda(F(k(\alpha)), \alpha) - 1} > 0. \tag{12}$$

□

In sum, in a static analysis, the local stability of an equilibrium point is sufficient for more hunters to be cooperative with greater social norms. On the other hand, if the condition is not satisfied, social norms can yield a *worse outcome* for a society such that fewer hunters become cooperative with greater social norms.

Corollary 2. *If $k(\alpha)$ is a locally unstable point satisfying*

$$f(k(\alpha))B + s_\lambda(F(k(\alpha)), \alpha) > 1, \tag{13}$$

then there exists a neighborhood of α such that an equilibrium threshold k^* is a strictly decreasing function of α in the neighborhood of α.

Proof. The implicit function theorem entails that

$$k'(\alpha) = -\frac{s_\alpha(F(k(\alpha)), \alpha)}{f(k(\alpha))B + s_\lambda(F(k(\alpha)), \alpha) - 1} < 0.$$

□

Finally, we examine the dynamic implications of social norms by using the dynamic process in (4) as the degree of social norms increases from α to α', i.e., *for any pair* $\alpha' > \alpha$. In contrast to the result from the static analysis in Proposition 4, *dynamically*, it is always the case that more hunters will be cooperative with greater social norms.

Proposition 6. *If $k(\alpha) \in (\underline{C}, \overline{C})$ is an equilibrium point, then for any pair $\alpha' > \alpha$, the adjustment process dynamically converges to $k(\alpha') > k(\alpha)$.*

Proof. Given $t = 0$, $k_0(\alpha') = k(\alpha)$. For $\alpha' > \alpha$, we have

$$F(k_0(\alpha'))B + s(F(k_0(\alpha')), \alpha') > k_0(\alpha'),$$

so $k_1(\alpha') > k_0(\alpha')$. By (3), if $k_t(\alpha') > k_{t-1}(\alpha')$ for all $t = 0, 1, 2, \ldots$,

$$\begin{aligned}
&k_{t+1}(\alpha') - k_t(\alpha') \\
&= F(k_t(\alpha'))B + s(F(k_t(\alpha')), \alpha') - k_t(\alpha') \\
&= F(k_t(\alpha'))B + s(F(k_t(\alpha')), \alpha') - [F(k_{t-1}(\alpha'))B + s(F(k_{t-1}(\alpha')), \alpha')] \\
&= [F(k_t(\alpha')) - F(k_{t-1}(\alpha'))]B + [s(F(k_t(\alpha')), \alpha') - s(F(k_{t-1}(\alpha')), \alpha')] > 0,
\end{aligned}$$

where the strict inequality follows from $f > 0$, and $s_\lambda > 0$. Hence, $k_{t+1}(\alpha') > k_t(\alpha')$ for all $t = 0, 1, 2, \ldots$. If there exists $k(\alpha') \leq \overline{C}$ such that $F(k(\alpha'))B + s(F(k(\alpha')), \alpha') = k(\alpha')$, then $\lim_{t\to\infty} k_t(\alpha') = k(\alpha') > k(\alpha)$. If there is no such $k(\alpha') \leq \overline{C}$, then $k(\alpha') = \overline{C}$, since it satisfies the third condition in (6), and $\lim_{t\to\infty} k_t(\alpha') = \overline{C}$. □

5. AN EXAMPLE

We provide an example in this section. Let C_i be drawn from a uniform distribution $U[\underline{C}, \overline{C}]$. Then, $F(k)B - k$ is given as:

$$F(k)B - k = \left(\frac{k - \underline{C}}{\overline{C} - \underline{C}}\right)B - k = \left(\frac{B - \overline{C} + \underline{C}}{\overline{C} - \underline{C}}\right)k - \frac{B\underline{C}}{\overline{C} - \underline{C}}.$$

Hence, given a uniform distribution, the condition $\underline{C} < 0 < B < \overline{C}$ is sufficient for $F(k)B - k$ to be a strictly decreasing function, which satisfies the condition in Proposition 1. Now, assume that $s(\lambda, \alpha) = \alpha\lambda^2$ for social norms, and one can check that $\lim_{\lambda\to 0} s(\lambda, \alpha) = 0$ and as $\lambda \to 1$, $s(\lambda, \alpha)$ becomes sufficiently large if α is chosen appropriately, and these are milder conditions than those in Proposition 3. It follows that in equilibrium,

$$F(k)B + s(F(k), \alpha) - k = \left(\frac{k - \underline{C}}{\overline{C} - \underline{C}}\right)B + \alpha\left(\frac{k - \underline{C}}{\overline{C} - \underline{C}}\right)^2 - k = 0. \tag{14}$$

A calculation yields the following two equilibrium thresholds:

$$\begin{aligned}
k_1^* &= (\overline{C} - \underline{C})\left[\frac{-(B - \overline{C} + \underline{C}) - \sqrt{(B - \overline{C} + \underline{C})^2 + 4\alpha\underline{C}}}{2\alpha}\right] + \underline{C}, \\
k_2^* &= (\overline{C} - \underline{C})\left[\frac{-(B - \overline{C} + \underline{C}) + \sqrt{(B - \overline{C} + \underline{C})^2 + 4\alpha\underline{C}}}{2\alpha}\right] + \underline{C}.
\end{aligned}$$

To have a more concrete idea, let $\underline{C} = -1$, $B = 1$ and $\overline{C} = 4$. Then,

$$k_1^* = 5\left[\frac{2-\sqrt{4-\alpha}}{\alpha}\right] - 1;\, k_2^* = 5\left[\frac{2+\sqrt{4-\alpha}}{\alpha}\right] - 1. \tag{15}$$

For $\alpha = 3$, we have $k_1^* = \frac{2}{3}$ and $k_2^* = 4$. It can be shown that k_1^* is a strictly increasing function of α, and k_2^* is a strictly decreasing function of α, provided that there exist both equilibrium points from the quadratic equation in (14), which are related to the local stability condition in Proposition 5 and the non-stability condition in Corollary 2, respectively. To check the local stability using (11) and (13), we derive $f(k^*)B + s_\lambda(F(k^*), \alpha)$ such that

$$f(k^*)B + s_\lambda(F(k^*), \alpha) = \frac{B}{\overline{C} - \underline{C}} + 2\alpha\frac{(k - \underline{C})}{(\overline{C} - \underline{C})^2}.$$

Let $\underline{C} = -1$, $B = 1$, $\overline{C} = 4$ and $\alpha = 3$. Hence,

$$\frac{B}{\overline{C} - \underline{C}} + 2\alpha\frac{(k - \underline{C})}{(\overline{C} - \underline{C})^2} = \frac{1}{5} + \frac{6(k+1)}{25}.$$

For $\alpha = 3$, we have $k_1^* = \frac{2}{3}$ and $k_2^* = 4$, and one can check

$$\begin{aligned} \frac{1}{5} + \frac{6(k+1)}{25} &= \frac{3}{5} < 1 \text{ for } k_1^* = \frac{2}{3}; \\ \frac{1}{5} + \frac{6(k+1)}{25} &= \frac{7}{5} > 1 \text{ for } k_1^* = 4. \end{aligned}$$

Hence, k_1^* is a stable point, whereas k_2^* is an unstable point.

REFERENCES

Basu, K., and J. Weibull (2004). Punctuality: A cultural trait as equilibrium, Economics for an Imperfect World: Essays in Honor of Joseph Stiglitz, MIT Press, Cambridge.

Elster, J. (1989). Social norms and economic theory, Journal of Economic Perspectives 3, pp. 99–117.

Kandori, M. (1992), Social norms and community enforcement, Review of Economic Studies 59, pp. 63–80.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982). Rational cooperation in the finitely repeated prisoners' dilemma, Journal of Economic Theory 27, pp. 245–252.

Lindbeck, A., S. Nyberg, and J. Weibull (1999). Social norm and economic incentives in the welfare state, Quarterly Journal of Economics 114, pp. 1–35.

Ghosh, P., and D. Ray (1996). Cooperation in community interaction without information flows, Review of Economic Studies 63, pp. 491–519.