Evolution of vulnerability to pain in interpersonal relations as a strategic trait aiding cooperation

Rtischev, Dimitry

20 September 2010

_____

Evolution of vulnerability to pain in interpersonal relations as a strategic trait aiding cooperation

Dimitry Rtischev

Faculty of Economics
Gakushuin University
Tokyo, Japan

September 23, 2010

Abstract Why are humans so vulnerable to pain in interpersonal relations and can so easily hurt others physically and emotionally? We theoretically examine whether being offensively strong but defensively weak can evolve as a strategic trait that fosters cooperation. We study a population comprised of "thick-skinned" and "thin-skinned" agents by using an indirect evolution model that combines rational choice in strategic interactions with evolutionary selection across generations. We find that (a) the relatively vulnerable and cooperative thin-skins cannot evolve under purely random matching, (b) with some assortment thin-skins evolve and can take over the entire population, (c) vulnerability to greater pain makes it easier for thin-skins to evolve, and (d) proximate pain which merely feels bad but does not lower fitness helps thin-skins evolve even more than pain which accurately reflects fitness consequences. We draw contrast with the Hawk-Dove model and identify several ways in which rationality hinders the evolution of the relatively vulnerable and peaceful type of agent.

Keywords cooperation, conflict avoidance, Hawk-Dove, pain, emotions, vulnerability, indirect evolution, assortative matching

JEL Classification C73, D03, D74

1. Introduction

Human vulnerability to suffer psychological and physical pain from interacting with others is so obvious that it is easy to overlook how remarkable a trait it is. But upon some reflection, it appears remarkable that almost anyone can easily hurt anyone else nearby, physically or emotionally. Whether they know each other or not, any two people near each other are only seconds away from inflicting or feeling pain through words, gestures, or physical violence. Despite the fact that fingernails, fists, teeth, sticks, and stones have been ubiquitous in the environment in which humans have evolved, human skin and flesh are not well adapted to protect against such weapons. Human anatomy and posture leave digestive and reproductive organs exposed to blows and projectiles. Common greeting rituals such as bowing, hugging, and kissing momentarily escalate the vulnerability to even greater heights. Psychologically, one can deeply hurt another merely by uttering a few words. Even a child can drive a parent or teacher to tears. Evidently, interpersonal relations are carried on in the shadow of much potential pain. We will generically refer to this widespread human trait – the vulnerability to physical and psychological pain in interpersonal relations – as "thin skin."

Thin skin is puzzling in the Darwinian framework, since it is obvious how this trait can hinder an individual in its struggle to acquire resources necessary for survival and reproduction, but it is not clear how the trait can help.[1] Thin skin is at odds with the tendency noted by evolutionary biologists that "evolution produces shields to its own weapons" (Hamilton, 1971, p. 218), *i.e.*, that evolution usually gives rise to arms' races that escalate but balance offensive and defensive capabilities.

Our goal is to explore an evolutionary logic that can account for thin skin in connection with another distinctive human trait, the propensity to cooperate. In principle, thin skin and cooperation could be causally related in both directions. It may be the case that thin skin is a byproduct of the high degree of cooperation achieved in most interpersonal relations most of the time, making the fitness cost of thick skin not worth the fitness benefit it could provide. Although we do not pursue this possibility, it appears doubtful as the main reason for the evolution of thin skin given that pain in many interpersonal relations, even (especially?) among genetically related individuals, is hardly uncommon. For instance, not being so sensitive to hurtful words arguably could be a useful and perhaps not prohibitively expensive shield; and yet many people must go through life without one, always at the mercy of others who might and sometimes do utter such words at them. This suggests that vulnerability to pain may be more of a cause rather than consequence of cooperation. We pursue this possibility and seek to identify a strategic role played by thin skin that yields its bearers a fitness advantage.

We conduct our analysis using a game-theoretic model in which there is both evolutionary selection and rational choice. The indirect evolutionary approach (Guth and Kliemt, 1998) that we follow allows us to explicitly treat in an integrated way both a long-run population-level selection process as well as short-

---

[1] In the physiological context, vulnerability to pain has a straightforward evolutionary explanation as a mechanism that induces efforts to alleviate hunger, avoid burns and falls, etc. In light of the many examples of evolution adapting pre-existing organs to new functions, it is conceivable that physiological pain mechanisms have evolved to also become engaged in the social context, which is our exclusive focus.

run individual-level rational choice behavior.[2]   We accomplish this by taking vulnerability to pain to be an immutable lifelong trait of an individual, a trait which is selected by an evolutionary process and which in turn affects payoffs facing the individual as he chooses payoff-maximizing actions in specific strategic interactions that arise during his lifetime.   The prospect of pain causes vulnerable individuals to aggress less, and under certain conditions this raises the average fitness of vulnerable individuals relative to the non-vulnerable ones who get into more fights.   Pain can play this strategic role only if it is credible, that is, only if individuals cannot rationally override their vulnerability to pain. This is our explanation for why evolution not only avoids producing shields to its own weapons, but even makes it difficult for individuals to develop such shields, such as by training oneself to not be sensitive to hurtful words or physical pain.

In our model, there is a population composed of thick-skinned and thin-skinned agents.   Each generation, the agents are paired, and each dyad plays a demand game to divide a resource.   Each agent in a dyad may either aggress by demanding all or seek to compromise by demanding half.   Bilateral aggression leads to a fight which wastes some of the resource and also inflicts pain on agents with thin skin.   Bilateral compromise leads to fair division without waste or pain. A unilateral effort to compromise backfires in the sense that the agent who aggressed gets the entire resource, but there is no waste or pain.   Payoffs earned in dyads are treated as fitness that determines population shares of thin- and thick-skinned agents in the next generation via a replicator dynamics.

We fully specify this model in the next section and then proceed to study conditions under which thin-skinned agents evolve.   We explore several variants of the model as follows.   In Section 3 we analyze the case of complete type information, *i.e.*, we assume that each agent can observe the skin of the opponent in its dyad.   In Section 4 we decouple proximate payoffs from ultimate fitness, assuming that the prospect of feeling pain influences an agent's decision-making but actually experiencing pain does not reduce fitness.   In Section 5 we drop the assumption that types are observable within dyads and instead assume that agents only know population shares of each type and the degree of assortment.   In Section 6 we assume the demand game in dyads is played sequentially rather than simultaneously.   In Section 7 we discuss our results in comparison with the Hawk-Dove model of conflict and the evolutionary analysis of the trust game.

We find that (a) thin-skins cannot evolve under purely random matching, (b) with some assortment thin-skins evolve and can even take over the entire population, and (c) the greater the pain, the lesser the degree of assortment needed to sustain any given positive population share of thin-skins in equilibrium. Comparing the case of proximate pain to fitness-reducing pain, we find that thin skin can evolve with a lower degree of assortment if the pain merely causes bad feelings but does not actually reduce fitness.   Based on this finding, we argue that evolution may favor pain which exaggerates fitness consequences.   Comparing the case in which agents can observe others' skin thickness to the case in which they only know population shares and degree of assortment, we find the "veil of uncertainty" of the latter case helps thin-skins evolve.   Based on these findings and the comparison with the Hawk-Dove model, we argue that more information and greater rationality hinder the evolution of the type of agent who relies on vulnerability to earn a peace dividend.

_____

[2] Our model lies in the middle range of the spectrum between rational-choice "teleology" and zero-intelligence "direct evolution," as discussed by Berninghaus, Guth and Kliemt (2003).

_____

2. The model

A large population consists of two types of agents: type A agents who have thick skin and type B agents who have thin skin. Alternatively, the agents may be visualized as anatomically identical but differently armed: type A carrying both a sword and a shield and type B only a sword.

Each generation, agents are paired according to a random process which we will specify below. After being paired, agents in each dyad first observe each other's type[3] and then play a demand game to divide one unit of a resource. In the demand game, each agent can demand either ALL or HALF of the resource. Agents announce their demands simultaneously.[4] Three outcomes are possible. First, "compromise" is the outcome that ensues if both agents demand HALF. In this outcome each agent gets half of the resource. Second, "exploitation" is the outcome that ensues if one agent demands ALL and the other agent demands HALF. In this outcome, the agent who demanded all gets the entire resource and the agent who tried compromising gets nothing. Third, a "fight" ensues if both agents demand ALL. The expected net payoff that an agent of type i gets if it fights an agent of type j will be denoted by F_{ij}. This amount includes portion of the resource obtained by agent i at the end of the fight minus the cost of the fight itself such as energy spent and injury or pain suffered. The payoff matrix of the demand game is shown in Figure 1. The payoffs of the demand game represent fitness that determines the composition of the next generation via a standard replicator dynamics, except in Section 4 where we explicitly decouple fitness from demand game payoffs.

Let us first consider a dyad with two thick-skinned agents. Because they are well-protected, we will assume that a fight between two type-A agents is in effect like a Tullock-style rent-seeking contest that dissipates some but not all of the prize and leaves both agents with a positive expected net payoff.

Assumption 1. $\quad 0 < F_{AA} < \frac{1}{2}$

Under this assumption, the demand game of Figure 1 played between two thick-skins is a Prisoners' Dilemma with the unique equilibrium (ALL, ALL) and payoff $\pi_{AA} = F_{AA}$. If all agents in the population are thick-skinned, every dyad will always allocate the resource via fighting and thereby dissipate $1 - 2F_{AA}$. Against this baseline of constant and wasteful fighting between thick-skins, we want to explore conditions under which a relatively more vulnerable and peaceful type of agent can evolve. Our thin-skinned type B agent is unprotected and therefore suffers pain and injury if he gets into a fight, regardless of the type of his opponent. We model this by assuming that the pain outweighs the potential gain from the fight:

Assumption 2. $\quad F_{Bi} < 0 \quad \forall i \in \{A, B\}$

Under this assumption, the demand game of Figure 1 played between two thin-skins is a game of Chicken. To simplify notation without substantively changing the results, we will assume that the pain a thin-skinned agent suffers in a fight is the

_____

[3] We will relax the observability assumption in Section 5.
[4] We will study the case of sequentially announced demands in Section 6.

same regardless of the type of its opponent and denote this pain by $F \equiv F_{BA} = F_{BB}$. In the symmetric mixed-strategy Nash equilibrium, a thin-skinned agent paired with another such agent plays ALL with probability[5]

$$\theta = \frac{1}{1-2F} \tag{1}$$

and earns expected payoff

$$\pi_{BB} = \frac{F}{2F-1} \in (0, \tfrac{1}{2}) \tag{2}$$

Since $\partial \pi_{BB} / \partial F < 0$ and more negative values of F correspond to greater pain, the expected payoff of a thin-skinned agent in a homogeneous dyad is monotonically increasing in the degree of pain to which it is vulnerable.

Finally, we must consider the demand game played in a heterogeneous dyad. If a fight breaks out, the thick-skinned agent will waste some effort fighting but he is likely to get most of the resource and will not suffer pain. We model this via

Assumption 3. $\qquad \tfrac{1}{2} < F_{AB} < 1$

Under this assumption, exploitation is the unique Nash equilibrium in the demand game of Figure 1 played by a heterogeneous dyad. In the equilibrium, the thick-skinned agent demands ALL and the thin-skinned agent demands HALF, there is no fighting or pain, and the payoffs are $\pi_{AB} = 1$ and $\pi_{BA} = 0$.

Since thin-skinned agents have a comparative disadvantage in heterogenous dyads ($\pi_{BA} < \pi_{AB}$), unless they have a comparative advantage when paired homogenously their evolution would be impossible. Thus, a necessary condition for the evolution of thin-skins is for them to earn more when paired together than what thick-skins earn in their homo-dyads; i.e., $\pi_{BB} > \pi_{AA}$. We will assume this condition holds. In particular, using $\pi_{AA} = F_{AA}$ and (2), we express the condition in the form of the following minimum pain threshold:

Assumption 4. $\qquad F < \dfrac{F_{AA}}{2F_{AA}-1}$

Notably, if thin-skins are not sufficiently vulnerable to pain, they cannot evolve. And since the threshold in Assumption 4 is decreasing in F_{AA}, the less wasteful the fighting is among thick-skins, the more pain is necessary for thin-skins to have some comparative advantage.

Having specified the details of the interaction in each of the three possible dyads, we can summarize the payoffs to each type of agent in each type of dyad using the reduced-form payoff matrix of Figure 2. This payoff matrix represents

_____

[5] Alternatively, we could interpret θ as the fraction of thin-skins that always demand ALL in homo-dyads and $1 - \theta$ as the fraction of thin-skins that always demand HALF in homo-dyads.

the stage game of the evolutionary game that we will analyze below. For all values of F_{AA} and F possible under Assumptions 1 through 3, B is a strictly dominated strategy of the stage game and AA is the unique Nash equilibrium. If the thin-skinned agents are not sufficiently vulnerable to satisfy Assumption 4, then AA is the Pareto-efficient outcome of the stage game, thin-skin offers no fitness gain, and therefore such thin–skinned agents cannot evolve. If the thin-skinned agents are sufficiently vulnerable to satisfy Assumption 4, the stage game is a Prisoners' Dilemma with equilibrium outcome AA being Pareto-dominated by outcome BB.

3. Evolutionary analysis

Let $p \in [0, 1]$ be the fraction of thin-skinned type B agents in the population in a given generation. We first assume that in each generation agents are paired purely at random and show that only thick-skins survive in equilibrium.

Proposition 1.
Under purely random pairing of agents, the only evolutionarily stable population is thick-skin monomorphic, p = 0.

Proof.
A thin-skinned mutant cannot invade a population of all thick-skins since $0 = \pi_{BA} < \pi_{AA} \in (0, \frac{1}{2})$. Thus p = 0 is evolutionarily stable. A thick-skinned mutant can invade a population of all thin-skins since $1 = \pi_{AB} > \pi_{BB} \in (0, \frac{1}{2})$. Hence, p = 1 is not evolutionarily stable. The average fitness of the two types in a polymorphic population $p \in (0, 1)$ is given by $V(A) = (1 - p)\pi_{AA} + p\pi_{AB} = (1 - p)\pi_{AA} + p$ and $V(B) = p\pi_{BB} + (1 - p)\pi_{BA} = p\pi_{BB}$. Since $V(A) - V(B) = (1 - p)\pi_{AA} + p(1 - \pi_{BB}) > 0$ for all $p \in (0, 1)$, thick-skins have greater average fitness given any population share and therefore a polymorphic equilibrium does not exist. □

Since thin-skinned agents cannot exist in evolutionary equilibrium under purely random matching, let us consider their evolution under positive assortment. We adopt the assortment model of Hamilton (1971) as elaborated by Fagen (1980), with $a \in [0, 1]$ as the exogenous parameter that represents the degree of assortment. [6] The assortment model is as follows. Let *N(d)* denote the proportion of dyads of type $d \in \{AA, BB, AB\}$ in the population of all dyads formed in a given generation. Given population shares of thin-skins $p \in [0, 1]$ and thick-skins $q \equiv 1 - p$, our assumed assortment process produces the following proportions of the three possible dyad types:

$$N(BB) = p^2 + apq \tag{3}$$

$$N(AA) = q^2 + apq \tag{4}$$

_____

[6] Related formulations of assortment include Bergstrom's (2003) "index of assortativity" and Taylor and Nowak's (2006) "interaction rates."

$$N(AB) = 2pq(1-a) \tag{5}$$

This specification of the assortment process is technically attractive because it continuously and monotonically connects the boundary cases of purely random matching ($a = 0$) and perfect assortment ($a = 1$).[7]

Given this assortment process, the probability that a given thick-skinned agent is paired with a thin-skinned agent is

$$p^A = \frac{N(AB)}{2N(AA) + N(AB)} = (1-a)p \tag{6}$$

and the probability that a given thin-skinned agent is paired with another thin-skinned agent is

$$p^B = \frac{2N(BB)}{2N(BB) + N(AB)} = p + a(1-p) \tag{7}$$

After being matched, each dyad plays the demand game as described in Section 2. The payoff earned by each type of agent is the fitness that determines the composition of the next generation of agents via a standard replicator dynamics. The average fitness of thick- and thin-skinned agents are respectively

$$V(A) = (1 - p^A)\pi_{AA} + p^A = F_{AA} + p(1-a)(1 - F_{AA}) \tag{8}$$

$$V(B) = p^B \pi_{BB} = (p + a(1-p))\frac{F}{2F-1} \tag{9}$$

The conditions under which thin-skinned agents evolve are as follows.

Proposition 2.
The unique evolutionarily stable population corresponding to assortment of degree *a* is

$$\hat{p} = 0 \quad \text{(thick-skin monomorphic) if } 0 \le a \le a_{\min}$$
$$\hat{p}(a) \in (0, 1) \quad \text{(polymorphic) if } a_{\min} < a < a_{\max}$$
$$\hat{p} = 1 \quad \text{(thin-skin monomorphic) if } a_{\max} \le a \le 1$$

where

$$a_{\min} = \frac{\pi_{AA}}{\pi_{BB}} = \frac{F_{AA}(2F-1)}{F} > 0 \tag{10}$$

$$a_{\min} < a_{\max} = \frac{1 - \pi_{BB}}{1 - \pi_{AA}} = \frac{1 - F}{(1 - F_{AA})(1 - 2F)} < 1 \tag{11}$$

---

[7] Conceptually, the assortment parameter *a* can be interpreted as a summary statistic of a "courtship" process by which pairs of agents imperfectly ascertain each other's vulnerability to pain prior to entering into collaboration that produces gains to be divided. Given the demand game, thin-skins are preferred as partner by both thin- and thick-skins; intermediate values of *a* correspond to the situation in which some thin-skins mistakenly enter into collaboration with a thick-skinned partner. After pairing but before division of gains, our agents find out the true type of partner, or, as analyzed in Section 5, estimate the probability of the partner's type from *p* and *a*.

$$\hat{p}(a) = \frac{a\pi_{BB} - \pi_{AA}}{(1-a)(1-\pi_{AA}-\pi_{BB})} = \frac{F_{AA} + aF - 2F_{AA}F}{(1-a)(F + F_{AA}(1-2F))-1} \tag{12}$$

Proof.

If a polymorphic population $\hat{p} \in (0,1)$ is evolutionarily stable, then *V(A) = V(B)* when $p = \hat{p}$. Setting (8) equal to (9) and solving for *p* gives (12). Differentiating (12) shows that $\hat{p}(a)$ is monotonically increasing on (0, 1). Setting (12) equal to 0 and solving for *a* yields (10). Setting (12) equal to 1 and solving for *a* yields (11). It is straightforward to confirm that $0 < a_{min} < a_{max} < 1$. For any $a \in [0, a_{min})$ and $p \in [0,1]$, *V(A) > V(B)*, which implies $p = 0$ is the only stable population for any $a \in [0, a_{min})$. For any $a \in (a_{max}, 1]$ and $p \in [0,1]$, *V(A) < V(B)*, which implies $p = 1$ is the only stable population for any $a \in (a_{max}, 1]$. For any $a \in (a_{min}, a_{max})$, *V(A) > V(B)* if $p > \hat{p}(a)$ and *V(B) > V(A)* if $p < \hat{p}(a)$, which implies that $\hat{p}(a)$ is a stable and unique equilibrium. □

Figure 3 illustrates how the fraction of thin-skins in the evolutionarily stable population depends on the degree of assortment. Since $a_{min} > 0$, some degree of assortment is needed for thin-skins to exist in equilibrium. Since $a_{max} < 1$, even with imperfect assortment thin-skins can completely displace thick-skins in equilibrium. At intermediate degrees of assortment, a polymorphic population is stable, with stronger assortment corresponding to a higher population share of thin-skins.

Comparative statics shows that the greater the pain the less the reliance on assortment for the evolution of thin-skin. Since $\partial a_{min} / \partial F > 0$, more pain (*i.e.*, more negative *F*) corresponds to less assortment needed to admit thin-skins into evolutionarily stable populations. Since $\partial a_{max} / \partial F > 0$, more pain corresponds to less assortment needed to exclude thick-skins from evolutionarily stable populations. Finally, since $\partial \hat{p}(a) / \partial F < 0$ for all $a \in (a_{min}, a_{max})$, an incremental increase in pain corresponds to an incremental rise in the share of thin-skins in any polymorphic equilibrium.

In the extreme case of thin-skins vulnerable to infinitely intense pain, it follows from (1) that such thin-skins never demand ALL and therefore never fight. For such extremely vulnerable and peaceful thin-skins, the minimum degree of assortment needed to survive in a polymorphic equilibrium is

$$\lim_{F \to -\infty} a_{min} = 2F_{AA} \tag{13}$$

and the minimum degree of assortment needed to exclude all thick-skins is

$$\lim_{F \to -\infty} a_{max} = \frac{1}{2(1-F_{AA})} \tag{14}$$

In the further extreme case of thick-skins that completely dissipate the resource through fighting ($F_{AA} \to 0$), it follows from the above limits that an arbitrarily small degree of assortment would give rise to a polymorphic equilibrium with some extremely vulnerable/peaceful thin-skins present and if assortment exceeds $a \geq \frac{1}{2}$,

the extremely vulnerable/peaceful thin-skins would completely drive out the extremely wasteful/violent thick-skins.

4. Proximate pain without ultimate injury

Vulnerability to pain raises the average fitness of thin-skins because in trying to avoid the pain they fight less often when paired together. This reduction in the frequency of fighting outweighs the pain that thin skins suffer when they do fight. But the pain need not necessarily be coupled to fitness as we have been assuming. If the pain is merely perceived in the mind and has no fitness-relevant ramifications, then the instrumental role of pain can be more powerful. Let us consider the case when pain is solely a proximate mechanism that operates at the level of agents' strategic interaction in the demand game but not at the level of fitness and replication.

To model this, we assume that both thin- and thick-skinned agents play the demand game as if thin-skins are vulnerable to pain F_{Bi}, but when the fitness of thin-skin agents is calculated, F_{Ai} is used instead of F_{Bi}, for all $i \in \{A, B\}$. The only type of dyad in which it matters whether pain is fitness-relevant or proximate is the dyad in which pain actually occurs in equilibrium, namely the thin-skin homo-dyad. In particular, when two thin-skins meet, each agent demands ALL with the same equilibrium mixing probability given by (1), but instead of obtaining the pain fitness F when a fight breaks out, each gets the pain-free fitness F_{AA}. Thus, instead of (2) above, the fitness-relevant expected payoff to a thin-skinned agent in a homogenous dyad mixing with probability θ is:

$$\pi_{BB}^{prox} = \theta^2 F_{AA} + \tfrac{1}{2}(1-\theta)^2 + \theta(1-\theta) = \frac{F_{AA} + 2F(F-1)}{(1-2F)^2}$$

(2')

It is straightforward to verify that $\pi_{BB}^{prox} > \pi_{AA}$ even without Assumption 4, which means that thin-skins vulnerable to any degree of proximate pain, however minor, obtain higher fitness when paired together than the fitness obtained by thick-skins when they are paired together. It can also be readily verified that $\pi_{BB}^{prox} > \pi_{BB}$; thus thin-skins enjoy more fitness if pain is proximate than if it is fitness-relevant. Yet as the next proposition establishes, even with this additional fitness thanks to pain being proximate thin-skins still cannot evolve under purely random matching.

Proposition 3.
If pain is proximate, under purely random pairing of agents the only evolutionarily stable population is thick-skin monomorphic.

Proof.
From (2') it follows that $\pi_{BB}^{prox} \in (0, \tfrac{1}{2})$. The proof is analogous to the proof of Proposition 1 with π_{BB}^{prox} used in place of π_{BB}. □

The additional fitness enjoyed by thin-skins due to pain being merely proximate does make it easier for thin-skinned agents to evolve under positive assortment. The next two propositions make this precise. The dotted curve in Figure 3 illustrates the effect of proximate pain.

Proposition 4.
Under Assumptions 1-3 and for any degree of proximate pain $F < 0$, if positive assortment is sufficiently strong, then there is a unique evolutionarily stable equilibrium in which thin-skinned agents are present. Specifically, the unique evolutionarily stable population corresponding to assortment of degree a is

$$\hat{p}^{prox} = 0 \quad \text{if} \ \ 0 \le a \le a_{min}^{prox}$$
$$\hat{p}^{prox}(a) \in (0, 1) \quad \text{if} \ \ a_{min}^{prox} < a < a_{max}^{prox}$$
$$\hat{p}^{prox} = 1 \quad \text{if} \ \ a_{max}^{prox} \le a \le 1$$

where

$$a_{min}^{prox} = \frac{F_{AA}(2F-1)^2}{F_{AA} + 2F(F-1)} > 0 \tag{10'}$$

$$a_{min}^{prox} < a_{max}^{prox} = \frac{1 - F_{AA} + 2F(F-1)}{(1-F_{AA})(1-2F)^2} < 1 \tag{11'}$$

$$\hat{p}^{prox}(a) = \frac{F_{AA}(a - (2F-1)^2) + 2aF(F-1)}{(1-a)(1-2F_{AA})(1+2F(F-1))} \tag{12'}$$

Proof.
The proof is analogous to the proof of Proposition 2 with π_{BB}^{prox} used in place of π_{BB}. The minimum pain Assumption 4 is not necessary since (2') implies $\pi_{BB}^{prox} > \pi_{AA}$ for any $F < 0$. $\quad\square$

Proposition 5.
For a given degree of assortment, the share of thin-skinned agents in an evolutionarily stable population is weakly higher if the pain is proximate than if the pain is fitness-relevant.

Proof.
From equations (10) through (12) and (10') through (12') it follows that
$a_{min}^{prox} < a_{min}$, $\quad a_{max}^{prox} < a_{max}$, and $\hat{p}^{prox}(a) > \hat{p}(a) \ \forall a \in (a_{min}, a_{max}^{prox})$. $\quad\square$

　　　More generally, if a thin-skinned agent fights another agent, its proximate payoff can be expressed as $F = F_{AA} - y$ and its fitness-relevant payoff can be expressed as $U = F_{AA} - z$, where we can think of $y > F_{AA}$ as "*pain*" and $z \ge 0$ as "*injury*." In Section 3, we analyzed the case $y = z$. In this section, we analyzed the case $z = 0$. The two cases are extremes: the former assumes that pain precisely tracks injury whereas the latter assumes there is no injury but only pain. In the intermediate and more realistic case, thin-skins suffer injuries ($z > 0$) but the pain they feel (or anticipate feeling) may either overstate ($y > z$) or understate ($y < z$) it. The greater the pain, the less frequent the fighting between thin-skins, but the fitness gained due to less fighting is reduced by injury that fighting causes. Thus, for a given level of pain, the more the pain overstates injury, the easier it is for thin-skins to evolve, in the sense of lesser reliance on positive assortment. Conversely, if pain understates injury, the more difficult it is for thin-

skins to evolve. This line of reasoning leads us to conjecture that evolution is likely to favor pain that overstates rather than understates injury.

5. Imperfect type information: skin thickness not observable

A critical but unrealistic assumption underlying our analysis so far is that agents can distinguish those who are vulnerable to pain from those who are not. In this section we relax the assumption that skin thickness is observable. Instead, we assume that besides knowing one's own vulnerability, an agent only has population-level type information, namely the population share of thin-skins and the degree of assortment.

The key to analyzing this case is the fact that for a thick-skinned agent, HALF is a dominated strategy regardless of the type of opponent in its dyad. Thus, whether it knows the type of its dyad opponent or not, a thick-skinned agent maximizes its payoff by demanding ALL. A thin-skinned agent, however, prefers to demand HALF against a thick-skin and mix against another thin-skin. Since it cannot see the type of opponent, the thin-skinned agent estimates of type of its dyad opponent according to (7) and chooses the payoff maximizing strategy based on this estimate.

Proposition 6.
If matching is random, thickness of skin unobservable, and agents know the population share of thin-skins, then the only evolutionarily stable population is thick-skin monomorphic.

Proof.
Same as proof of Proposition 7(i) below with a = 0. □

Thus, whether type information is available at the dyad level (Proposition 1) or at the population level (Proposition 6), thin-skins cannot evolve under purely random matching. We next introduce assortment and find that ignorance at the dyad level actually helps thin-skins, in the sense that less assortment is needed to allow thin-skins to evolve under population-level information than under dyad-level information. The following proposition states the details and Figure 4 illustrates. We use the superscript "pop" to distinguish the population-level type information case.

Proposition 7.
If thickness of skin is not observable but population share of thin-skins p and the degree of assortment a are common knowledge, then the evolutionarily stable population is as follows:
(i)   For any $a \in [0, a_{\min}^{pop}]$, the unique evolutionarily stable population is thick-skin monomorphic, $\hat{p} = 0$.
(ii)   For any $a \in (a_{\min}^{pop}, \overline{a}]$, the unique evolutionarily stable population is polymorphic, with the share of thin-skins given by

$$\hat{p}_I^{pop}(a) = \frac{1}{1-a} - \frac{1}{1-2F_{AA}} \tag{15}$$

Thin-skinned agents always demand HALF.

(iii)   For any $a \in (\bar{a}, a_{max}^{pop})$, the unique evolutionarily stable population is polymorphic, with the share of thin-skins given by

$$\hat{p}_{II}^{pop}(a) = \frac{1}{2} - \frac{F_{AB} - F - T}{2(1-a)(1-2F)(F_{AB} - F_{AA})} \tag{16}$$

where

$$T = [((1-a)F_{AA} + aF_{AB} + F(2(1-a)(F_{AB} - F_{AA}) - 1))^2$$
$$- 4a(1-2F)(F_{AA}(1-2F) + F)(F_{AB} - F_{AA})]^{\frac{1}{2}} \tag{17}$$

Thin-skin agents play the mixed strategy demanding ALL with probability

$$\theta^* = 1 + \frac{2F}{(1-2F)(\hat{p} + a(1-\hat{p}))} \tag{18}$$

(iv)   For any $a \in [a_{max}^{pop}, 1]$, the unique evolutionarily stable population is thin-skin monomorphic, $\hat{p} = 1$. Thin-skin agents play the mixed strategy demanding ALL with probability $\theta^* = 1/(1-2F)$.

The boundaries of the above intervals are given by

$$a_{min}^{pop} = 2F_{AA} \tag{19}$$

$$\bar{a} = \frac{F_{AA} - F}{(1-2F)(1-F_{AA})} \tag{20}$$

$$a_{max}^{pop} = \frac{F_{AB} - F}{(F_{AB} - F_{AA}) - 2F(1-F_{AA})} \tag{21}$$

Proof.
See Appendix.

Comparing the results of this section to the full-information case of Section 3 reveals that thin-skins evolve with less assortment and behave less aggressively when type information is available at the population level than when it is known at the dyad-level.

Proposition 8.
For a given degree of assortment, the share of thin-skinned agents in an evolutionarily stable population is weakly higher under population-level type information than under perfect observability of skin thickness.

Proof.
See Appendix.

Intuitively, a veil of uncertainty helps thin-skins because the higher the likelihood that a thin-skinned agent estimates that the opponent will demand ALL, the less is the payoff it can expect from demanding ALL itself. Therefore, uncertainty about the type of opponent induces thin-skins to demand ALL less often or not at all, which helps them avoid fights with each other. Thick-skins, however, always demand ALL whether they know the type of opponent or not, and therefore their behavior is the same under either information regime.

Figure 4 illustrates how a veil of uncertainty helps thin-skins to evolve. A numerical example also helps illustrate. Suppose fighting between thick-skins wastes 80% of the resource ($F_{AA} = 0.1$), domination of a thin-skin by a thick-skin wastes 20% of the resource ($F_{AB} = 0.8$), and pain is commensurate with the entire resource ($F = -1$). Under population-level type information, thin-skins cannot evolve if assortment is weaker than $a_{\min}^{pop} = 0.2$. When assortment is between $a_{\min}^{pop} = 0.2$ and $\overline{a} = 0.4$, population includes only peaceful thin-skins that always demand HALF and whose share can be as high as $\hat{p}^{pop}(\overline{a}) = 44\%$ of the population. If the degree of assortment exceeds $\overline{a} = 0.4$, thin-skins mix between demanding HALF and ALL, and their population share rises to 100% as the degree of assortment rises to $a_{\max}^{pop} = 0.72$. If, however, agents can observe others' skin thickness, peaceful thin skins never evolve, and mixing thin skins evolve only if the degree of assortment exceeds $a_{\min} = 0.3$ and dominate the population only if assortment exceeds $a_{\max} = 0.74$.

6. Sequential play in demand game

In this section, we re-specify the demand game in extensive form, letting agents in each dyad announce their demands sequentially with the first-mover chosen via an unbiased random draw. Although a much more nuanced picture of how vulnerable agents evolve emerges, overall we confirm that the main qualitative aspects of our findings do not critically depend on the assumption of perfectly simultaneous play. Doing away with both simultaneity and observability assumptions, Section 6.2 details a surprisingly intricate and non-monotonic relationship between assortment and evolution of vulnerable agents.

6.1 Perfect type information

Under perfect type information and sequential demands, there are four types of dyads that must be distinguished: A_1A_2, B_1B_2, A_1B_2, and B_1A_2 (subscript indicates move order). Backward induction on the game trees shown in Figure 5 reveals that the payoffs to agents in AA and AB dyads are independent of the order in which agents move and equal the payoffs in the simultaneous-move demand game; that is, each A in a homo-dyad earns $\pi_{AA} = F_{AA}$, an A in a hetero-dyad earns $\pi_{AB} = 1$, and a B in a hetero-dyad earns $\pi_{BA} = 0$. The order of play only makes a difference in the B_1B_2 dyad: the payoffs are 1 for the first-moving agent and 0 for the second-moving agent. Assuming play order is determined by an unbiased random draw, a thin-skinned agent when paired with another earns $\overline{\pi}_{BB} = \frac{1}{2}$ on average, which is strictly larger than the expected payoff in simultaneous-move BB

dyads given by (2). Figure 6 summarizes all the payoffs in all the dyads in a payoff matrix. This matrix represents the stage game of the evolutionary game and is a Prisoners' Dilemma, just like before except that the minimum pain threshold Assumption 4 is not needed. The following propositions specify conditions under which thin-skinned agents evolve.

Proposition 9.
If each agent can observe the type of its opponent and the agents make demands sequentially, the only evolutionarily stable population under purely random matching is thick-skin monomorphic.

Proof.
The proof is analogous to the proof of Proposition 1, using $\overline{\overline{\pi}}_{BB} = \frac{1}{2}$ instead of π_{BB} given by (2). □

Proposition 10.
If each agent can observe the type of its opponent and the agents make demands sequentially, the unique evolutionarily stable population corresponding to assortment of degree a is

$$\hat{p} = 0 \quad \text{if} \quad 0 \leq a \leq a_{\min}$$
$$\hat{p}(a) \in (0, 1) \quad \text{if} \quad a_{\min} < a < a_{\max}$$
$$\hat{p} = 1 \quad \text{if} \quad a_{\max} \leq a \leq 1$$

where

$$a_{\min} = 2F_{AA} > 0 \tag{10''}$$

$$a_{\min} < a_{\max} = \frac{1}{2(1 - F_{AA})} < 1 \tag{11''}$$

$$\hat{p}(a) = \frac{1}{1 - a} - \frac{1}{1 - 2F_{AA}} \tag{12''}$$

Proof.
The proof is analogous to the proof of Proposition 2, using $\overline{\overline{\pi}}_{BB} = \frac{1}{2}$ instead of π_{BB} given in (2). □

Thus, if the demand game is played sequentially under perfect type information, the only parameter that matters for the evolution of thin-skins is thick-skins' level of efficiency when allocating the resource by fighting each other, namely F_{AA}. There is no fighting in equilibrium and therefore the extent of thin-skins' pain F is irrelevant. Any non-zero amount of pain, or even proximate pain in the sense of Section 4, lead to the same fitness and equilibria. The shape of the curve representing the fraction of thin-skins in the evolutionarily stable population as a function of the degree of assortment is qualitatively the same as in the case of simultaneous demands under full information (Figure 3). Moreover, since (12'') is identical to (15) and (10'') is identical to (19), it is straightforward to adapt the

proof of Proposition 8 to show that less assortment is needed for a given population share of thin-skins to evolve with sequential demands than with simultaneous demands. Intuitively, announcing demands sequentially facilitates the evolution of thin-skins because it lets them coordinate so as to avoid all fighting and raise their homo-dyad payoffs to $\overline{\pi}_{BB} = \frac{1}{2} > \pi_{BB}$.

6.2 Population-level type information

As we did in Section 5, we now suppose agents cannot observe others' skin thickness but population shares and degree of assortment are common knowledge. Examining the game trees in Figure 5 reveals that the best-response of the second-mover of any type in any dyad is the same whether the agent knows or doesn't know the type of the first-mover. Moreover, the first-moving thick-skinned agent does best by demanding ALL whether it knows the type of its opponent or not. The only player whose demand decision is impaired by unobservability of the type of its opponent is the first-moving thin-skinned agent. Figure 7 shows the game in which a thin-skinned agent moves first. If the probability that the opponent has thick-skin is high enough, the first-moving thin-skin prefers to play it safe by demanding HALF, but if the probability is low enough it prefers to demand ALL. As the next proposition and Figure 8 show, this leads to a nuanced picture of how pain and assortment enable the evolution of thin-skinned agents. Unlike our earlier findings, on some ranges greater assortment corresponds to lower population share of thin-skins.

Proposition 11.
If agents make demands sequentially, pain is strong enough to satisfy $F < 2F_{AA}/(2F_{AA}-1)$, thickness of skin is not observable but population share of thin-skins p and the degree of assortment a are common knowledge, then the evolutionarily stable populations are as follows:

(i)   For any $a \in [0, a_{\min}]$, the unique evolutionarily stable population is thick-skin monomorphic, $\hat{p} = 0$. Thin-skinned agents demand HALF when moving first.

(ii)   For any $a \in (a_{\min}, \overline{a})$, the unique evolutionarily stable population is polymorphic, with the share of thin-skins given by (12"). Thin-skinned agents demand HALF when moving first.

(iii)   For $a \in (\overline{a}, a_0)$, the unique evolutionarily stable population is polymorphic, with the share of thin-skins given by

$$\hat{p} = 1 - \frac{1}{(1-a)(1-F)}$$

(22)

The fraction of thin-skinned agents who demand ALL when moving first is

$$\hat{x} = \frac{2a(1-F)(1-F_{AA}) - 2F_{AA} + F}{a(1-F)(1-F_{AB}) - F_{AB}F}$$

(23)

the remaining $(1 - \hat{x})$ fraction of thin-skins demand HALF when

moving first.[8]

(iv) For any $a \in (a_0, a_2)$, an evolutionarily stable population is thick-skin monomorphic, $\hat{p} = 0$. Thin-skinned agents demand ALL when moving first.

(v) For any $a \in (a_1, 1]$, an evolutionarily stable population is thin-skin monomorphic, $\hat{p} = 1$. Thin-skinned agents demand ALL when moving first.

(vi) If $F < 2F_{AA} - F_{AB}$, then $a_1 < a_2$ and both monomorphic populations are evolutionarily stable for any $a \in [a_1, a_2]$

(vii) If $F > 2F_{AA} - F_{AB}$, then $a_1 > a_2$ and for any $a \in (a_2, a_1)$ the unique evolutionarily stable population is polymorphic, with the share of thin-skins given by

$$\hat{p} = \frac{a(F-1) + 2F_{AA} - F}{(1-a)(2F_{AA} - F_{AB} - F)}$$

(24)

Thin-skinned agents demand ALL when moving first.

The boundaries of the above intervals are given by

$$a_{\min} = 2F_{AA}$$

(25)

$$\overline{a} = \frac{2F_{AA} - F}{2(1-F)(1-F_{AA})}$$

(26)

$$a_0 = \frac{F}{F-1}$$

(27)

$$a_1 = \frac{F_{AB}}{1 - 2F_{AA} + F_{AB}}$$

(28)

$$a_2 = \frac{2F_{AA} - F}{1-F}$$

(29)

Proof.
See Appendix.

The case of sequential demands under population-level type information differs from all other cases that we have considered in that increasing assortment may hinder the evolution of thin-skins. This happens on the interval $(\overline{a}, a_0)$, where greater degrees of assortment correspond to a smaller share of thin-skins in stable polymorphic populations. This is a region of transition from "meek" thin-skins (those who demand HALF when moving first) to "assertive" thin-skins (those who demand ALL when moving first). Except for this relatively small region, stronger assortment helps thin-skins evolve, consistent with our earlier results.

---

[8] Equivalently, under a mixed-strategy interpretation, $\hat{x}$ is the probability with which every thin-skinned agent demands ALL when moving first.

Comparing to the perfect-information case of Section 6.1, it is notable that the unobservability of types does not affect the evolution of thin-skins at low degrees of assortment $a < \bar{a}$, impedes their evolution at intermediate degrees of assortment, and can help thin-skins achieve the monomorphic population in the sense that $a_1 < a_{max}$ (see Figure 8). Intuitively, the reason why not knowing the type of opponent helps thin-skins in the monomorphic equilibrium is because a thick-skinned mutant attempting to invade such a population earns F_{AB} against a first-moving assertive thin-skin, which is less than 1 that the mutant would have earned if he could display his thick skin and thereby intimidate the thin-skin to demand HALF. Thus, ignorance hurts the unlucky thin-skin who gets paired with the mutant but helps defend the population of thin-skins against invasion by the mutants.

7. Discussion: Relation to Hawk-Dove and the role of rational choice and information

Like the classic Hawk-Dove model (Maynard Smith and Parker, 1976), we have studied the evolutionary logic of conflict in a population composed of aggressive and peaceful agents. Both models share the same basic structure: two agent types, pairwise matching, action set {aggress, yield}, and replicator dynamics. Yet the equilibria in the two models are different: Doves evolve under random matching but thin-skins do not. Since the most important lesson of the Hawk-Dove model is that peaceful agents evolve alongside aggressive ones without any rational choice by the agents or bias in the pairing process, it is instructive to understand why this finding was not replicated in our model.

Comparing the payoffs in homogeneous and heterogeneous dyads reveals the same pattern in both models: two peaceful agents paired together earn higher fitness than two aggressive agents, but when the two different types meet, the aggressive one earns more than the peaceful. What is not the same is the relative size of the payoff earned by the aggressive type in a homo-dyad compared to what the peaceful type earns in a hetero-dyad. Specifically, whereas a Hawk who fights another Hawk earns less than a Dove who yields to a Hawk, a thick-skin who fights another thick-skin earns more than a thin-skin who yields to a thick-skin. As a consequence, the stage game in Hawk-Dove is Chicken and in our model it is Prisoners' Dilemma (Figure 2). The different nature of the stage game leads to different evolutionary properties.

If Hawks were not hard-wired to always escalate but could choose to go away without a fight, then in the Hawk-Hawk dyad Hawks would play a mixed strategy analogous to (1). Hawks would then earn a positive payoff in homo-dyads as in (2), the stage game would be a Prisoners' Dilemma, and Doves would not evolve under random matching. Thus, the celebrated finding that Doves evolve under random matching critically rests on the Hawks being greatly handicapped on two accounts: first, fighting between Hawks is so wasteful that its cost exceeds the value of the resource being divided, and second, Hawks are so stupid and inflexible that they cannot do anything but fight. It is not too surprising that Doves can evolve among such handicapped Hawks even without assortment.

Giving Hawks rational choice capability would thus undermine Dove's ability to evolve. In our model, both types are given rational choice capability. If we were to take away rational choice and instead assume that thick-skins are hard-wired to always aggress and thin-skins to always yield, the stage game would

_____

remain a Prisoners' Dilemma but the payoffs would change from the matrix in Figure 2 to the matrix in Figure 6. The reason is that since thick-skins always choose to aggress in all dyads, hard-wiring them has no effect. However, hard-wiring thin-skins to always yield actually helps them avoid the occasional fighting among themselves and thereby raises their payoffs in homo-dyads. Although this is not sufficient to enable thin-skins to evolve under random matching, hard-wiring helps them evolve in the sense of lowering the degree of assortment needed to achieve any given positive and stable population share. In terms of equilibrium play in dyads, assuming agents are hard-wired turns out to be equivalent to assuming the demand game is played sequentially and under perfect information as in Section 6.1, and the evolutionarily stable populations are as in Proposition 10.

Thus, both in Hawk-Dove and in our model, "upgrading" the agents from zero-intelligence to rational choice has the effect of impeding the evolution of the vulnerable-peaceful type. There are two other respects in which additional rationality and information hinder the evolution of vulnerable-peaceful agents. We found in Section 4 that if pain operates only at the proximate level, then thin-skinned agents can evolve with a lesser degree of assortment than otherwise, and conjectured that pain is more likely to evolve to overstate rather than understate ultimate consequences. Critical has been the assumption that agents' cognitive sphere is not so wide as to include calculation of ultimate fitness. If agents' self-knowledge and self-control were to expand to realize and act upon the fact that what really counts is not proximate feelings but just the "bottom-line" of ultimate fitness, the beneficial role of proximate pain would disappear.

In a similar vein, in Sections 5 and 6 we found that thin-skins can evolve with less assortment under the veil of uncertainty of population-level type information than under perfect observability of skin thickness within dyads. That better type information can hurt the evolution of cooperative agents is the opposite of what Guth and Kliemt (1998) found in the context of the trust game. The reason is that in the trust game better type information promotes more trusting and thereby raises the fitness of the trustworthy type. In our game, however, better type information promotes more aggression by thin-skins against other thin-skins, which reduces the fitness of the thin-skins. Conversely, a thin-skin operating behind the veil of uncertainty has a lower expected benefit of aggressing, which induces thin-skins to aggress less and earn more fitness.

8. Conclusion

We have analyzed whether agents who are vulnerable to pain when trying to divide a resource can evolve in a population that includes agents who do not feel pain. In a model that combines rational choice with evolutionary selection, we found that thin-skinned agents cannot evolve under random matching, but with sufficient assortment they can take over any positive share of the population. It is the prospect of pain that makes thin-skins relatively less aggressive. However, risking pain is sometimes worth it, and in equilibrium thin-skins sometimes aggress and sometimes suffer the pain. Ironically, this occasional aggression and suffering by thin-skins occurs because they rationally try to exploit the vulnerability of other thin-skins. Pain thus appears at the center of the interplay between evolutionary selection and rational decision-making: evolution selects individuals that are vulnerable to pain, since doing so raises relative fitness, and vulnerable individuals rationally take on some risk of actually suffering the pain, since doing so increases

payoffs.

One conclusion that comes out from our analysis is that assortment promotes the evolution of the vulnerable-peaceful type of agent. This agrees with Fagen's (1980) finding that assortment in the Hawk-Dove model leads to a "conspiracy of doves" equilibrium in which there are no Hawks and Bergstrom's (2003) finding that assortment enables cooperation in evolutionary Prisoners' Dilemma. It is also consistent with the general arguments of Hamilton (1971) and Skyrms (1994) that introducing correlation into the matching process can profoundly change the nature of an evolutionary game, yielding equilibria that differ from those under random matching and possibly with more cooperation and efficiency.

A second conclusion is that rational choice and information hinder the evolution of the vulnerable-peaceful type of agent. Thin-skins need less assortment to evolve if they are hard-wired to always compromise, or are kept in the dark about the proximate nature of the pain they feel, or are unsure about what kind of opponent they are facing.

Appendix

Proof of Proposition 7

For a thick-skinned agent, HALF is a dominated strategy regardless of the type of opponent in its dyad. Thus, agent of type A demands ALL. A thin-skinned agent prefers to demand HALF against A and mix against B, but he can only estimate which type he is facing in his dyad according to (7). Suppose each B agent plays ALL with probability $\theta \in [0,1]$. Then, the expected payoffs to a B agent from playing ALL and HALF are

$$\pi_B(ALL) = (1 - p^B)F + p^B(\theta F + (1 - \theta)) \tag{30}$$

$$\pi_B(HALF) = (1 - p^B)0 + p^B(\theta 0 + \tfrac{1}{2}(1 - \theta)) = \tfrac{1}{2}p^B(1 - \theta) \tag{31}$$

Setting $\pi_B(HALF) = \pi_B(ALL)$ and solving for p assuming $\theta = 0$ gives $\overline{p}(a) \equiv 1 - \dfrac{1}{(1 - a)(1 - 2F)}$. This curve partitions the space of all populations and assortment degrees into two regions, as follows.

Case I. If $0 \leq p < \overline{p}(a)$, then $\pi_B(HALF) > \pi_B(ALL)$ for any $\theta \in [0,1]$ and therefore B always demands HALF. This is the case in which there are few other thin-skins and/or assortment is weak, so that it's best for a thin-skin to play it safe by demanding HALF. The average fitness of the two types are:

$$V(A) = (1 - p^A)F_{AA} + p^A = F_{AA} + p(1 - a)(1 - F_{AA}) \tag{32}$$

$$V(B) = \tfrac{1}{2}p^B = \tfrac{1}{2}(a + p(1 - a)) \tag{33}$$

If a stable polymorphic equilibrium exists at $\hat{p}$, then V(A) = V(B), which implies (15). Setting $\hat{p} = 0$ and solving for a gives (19). Setting $\hat{p} = \overline{p}$ and solving for a gives (20).

Case II. If $\bar{p}(a) \leq p \leq 1$, then there exists a unique $\theta^* \in [0,1]$ such that $\pi_B(ALL) = \pi_B(HALF)$. Solving this equation for θ gives (18). The average fitness of the two types when thin skins mix with probability θ^* is:

$$V(A) = (1 - p^A)F_{AA} + p^A(\theta^* F_{AB} + (1 - \theta^*)) \tag{34}$$

$$V(B) = \tfrac{1}{2}p^B(1 - \theta^*) = \frac{F}{2F - 1} \tag{35}$$

If a stable polymorphic population exists at $\hat{p}$, then *V(A) = V(B)*, which yields (16)-(17). Setting $\hat{p} = \bar{p}$ and solving for *a* gives (20) and establishes the continuity of polymorphic equilibria on the boundary between Cases I and II. Setting $\hat{p} = 1$ and solving for *a* gives (21). It is straightforward to confirm that if *p = 1*, then *V(B) > V(A)* for all $a \in (a_{max}^{pop}, 1)$, which implies stability. □

Proof of Proposition 8

Subtracting (19) from (10) establishes that $a_{min}^{pop} < a_{min}$, which implies that on $[0, a_{min}^{pop}]$ there are no thin-skins under either information regime and on $(a_{min}^{pop}, a_{min}]$ there are thin-skins under the population-level information but not under perfect information. Subtracting (21) from (11) establishes that $a_{max}^{pop} < a_{max}$, which implies that on $[a_{max}, 1]$ thin-skins have 100% population share under either information regime and on $[a_{max}^{pop}, a_{max}]$ thin-skins have 100% population share under population-level information but not under perfect information. On $a \in (a_{min}, \bar{a})$, comparing (8) with (32) and (9) with (33) shows that thick-skins have the same average fitness under both regimes but thin-skins have strictly higher fitness under population-level information. By continuity, this implies that thin-skins hold higher population share for any $a \in (a_{min}, \bar{a})$. On $a \in (\bar{a}, a_{max}^{pop})$, comparing (8) with (34) shows that thick-skins have lower average fitness under population-level information than under perfect information; comparing and (9) with (35) shows that thin-skins have higher average fitness under population-level information than under perfect information. By continuity, this implies that thin-skins hold higher population share for any $a \in (\bar{a}, a_{max}^{pop})$. □

Proof of Proposition 11

A thick-skinned agent maximizes its payoff by demanding ALL under all circumstances. A thin-skinned agent moving second maximizes its payoff by demanding HALF in response to ALL and vice versa. A first-moving thin-skinned agent who demands HALF earns $\pi_B^{first}(HALF) = 0$; if it instead demands ALL, the expected payoff is $\pi_B^{first}(ALL) = (1 - p^B)F + p^B$. Setting $\pi_B^{first}(HALF) = \pi_B^{first}(ALL)$ and solving for proportion of thin-skins gives $\bar{p}(a) \equiv 1 - \frac{1}{(1-a)(1-F)}$. This curve partitions the space of all populations and assortment degrees into two regions, as follows. For all

$(p,a) \in \{(p,a) \mid p < \overline{p}(a)\}$ first-moving thin-skins demand HALF. For all $(p,a) \in \{(p,a) \mid p > \overline{p}(a)\}$ first-moving thin-skins demand ALL. On the boundary, $(p,a) \in \{(p,a) \mid p = \overline{p}(a)\}$ first-moving thin-skins are indifferent between demanding ALL and HALF. The proof proceeds by considering the two regions and the boundary as separate cases:

Case I. $(p,a) \in \{(p,a) \mid p < \overline{p}(a)\}$. This is the case in which there are few other thin-skins and/or the assortment is weak, so that it's best for the thin-skinned agent to play it safe by demanding HALF. The average fitness of the two types are the same as in the full-information case of Section 6.1:

$$V(A) = (1 - p^A)F_{AA} + p^A \qquad (36)$$

$$V(B) = \tfrac{1}{2} p^B \qquad (37)$$

If a stable polymorphic equilibrium exists at $\hat{p}$, then V(A) = V(B), which implies (12"). Differentiating (12") shows that $\hat{p}$ is monotonically increasing in *a*. Setting $\hat{p} = 0$ and solving for *a* gives (25). Setting $\hat{p} = \overline{p}$ and solving for *a* gives (26). The population $\hat{p} = 0$ is stable on $[0, a_{min})$ since V(A) > V(B) for all $a \in [0, a_{min})$ such that $p < \overline{p}(a)$. The population given by (12") is stable on $[a_{min}, \overline{a})$ since V(A) = V(B) for $p = \hat{p}(a)$, V(A) < V(B) for $p < \hat{p}(a)$, and V(A) > V(B) for $p > \hat{p}(a)$.

Case II. $(p,a) \in \{(p,a) \mid p > \overline{p}(a)\}$ This is the case in which there are many other thin-skins and/or the assortment is strong, so that the payoff-maximizing strategy for the thin-skinned agent is to assume the opponent in his dyad is another a thin-skin and try to exploit its vulnerability to pain by demanding ALL. The average fitness of the two types are:

$$V(A) = (1 - p^A)F_{AA} + \tfrac{1}{2} p^A (F_{AB} + 1) \qquad (38)$$

$$V(B) = \tfrac{1}{2}(1 - p^B)F + \tfrac{1}{2} p^B \qquad (39)$$

Monomorphic thick-skin population is stable if V(A) > V(B) when p = 0. This holds if $a < a_2$, where a_2 is given by (29). Monomorphic thin-skin population is stable if V(A) < V(B) when p = 1. This holds if $a > a_1$, where a_1 is given by (28). If $F > 2F_{AA} - F_{AB}$, then $a_1 > a_2$; solving V(A) = V(B) for *p* gives (24). The population given by (24) is stable since V(A) < V(B) for $p < \hat{p}(a)$, and V(A) > V(B) for $p > \hat{p}(a)$.

Case III. $(p,a) \in \{(p,a) \mid p = \overline{p}(a)\}$ This is the case in which the first-moving thin-skin is indifferent between demanding ALL or HALF. Let *x* be the fraction of thin-skins that demand ALL when moving first in a dyad. The average fitness of thick- and thin-skins are, respectively:

$$V(A) = (1 - p^A)F_{AA} + \tfrac{1}{2} p^A ((xF_{AB} + (1 - x)) + 1) \qquad (40)$$

$$V(B) = \tfrac{1}{2}(1 - p^B)xF + \tfrac{1}{2} p^B \qquad (41)$$

Solving V(A) = V(B) for *x* gives (23), which is the unique proportion of "assertive"

thin-skins necessary for evolutionary stability of a polymorphic population along the boundary between Cases I and II. Differentiating (23) shows that $\hat{x}$ is monotonically increasing in a, and is thus uniquely determined for each a or p. It is straightforward to confirm that $\hat{x} \in [0,1]$ for all $a \in [\overline{a}, a_0]$ and $\hat{x} < 0$ for all $a \in [0, \overline{a})$. Solving $\overline{p}(a) = 0$ gives (27) and establishes the upper bound on assortment in this Case III. This population is stable because for $p < \overline{p}(a)$ the average fitness of thin-skins (37) is higher than average fitness of thick-skins (36), and for $p > \overline{p}(a)$ the average fitness of thick-skins (38) is higher than average fitness of thin-skins (39). □

References

Bergstrom, T. C. 2003. The algebra of assortative encounters and the evolution of cooperation. International Game Theory Review, 5:3, 211-228

Berninghaus, S., W. Guth and H. Kliemt. 2003. From teleology to evolution: Bridging the gap between rationality and adaptation in social explanation. Journal of Evolutionary Economics, 13, 385-410

Fagen, R. M. 1980. When doves conspire: Evolution of nondamaging fighting tactics in a nonrandom-encounter animal conflict model. American Naturalist, 115:6, 858-869

Guth, W. and H. Kliemt. 1998. The indirect evolutionary approach: Bridging the gap between rationality and adaptation. Rationality and Society, 10:3, 377-399

Hamilton, W. D. 1971. Selection of selfish and altruistic behaviour in some extreme models. Reprinted in *Narrow Roads of Geneland*, Vol. 1, Freeman (New York), 1996, 198-227

Maynard Smith, J. and G. A. Parker. 1976. The logic of asymmetric contests. Animal Behaviour, 24:1, 159-175

Skyrms, Brian. 1994. Darwin meets *The Logic of Decision*: Correlation in evolutionary game theory. Philosophy of Science, 61, 503-528

Taylor, C. and M. Nowak. 2006. Evolutionary game dynamics with non-uniform interaction rates. Theoretical Population Biology, 69, 243-252
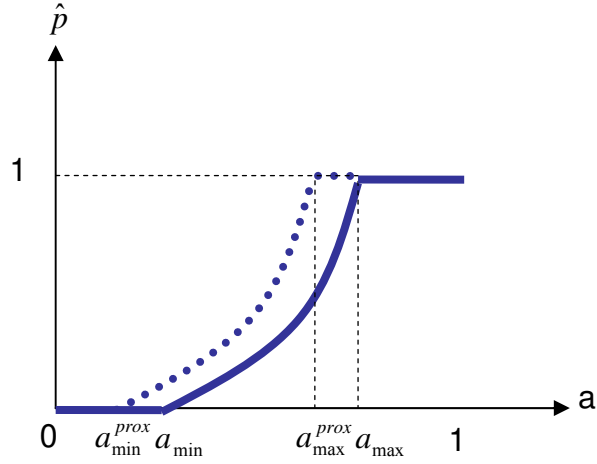
		Agent type j			
		ALL		HALF	
Agent type i	ALL	F_{ij}	F_{ji}	*1*	*0*
	HALF	*0*	*1*	$\frac{1}{2}$	$\frac{1}{2}$

Figure 1 Payoff matrix of the demand game between agents of type $i, j \in \{A, B\}$

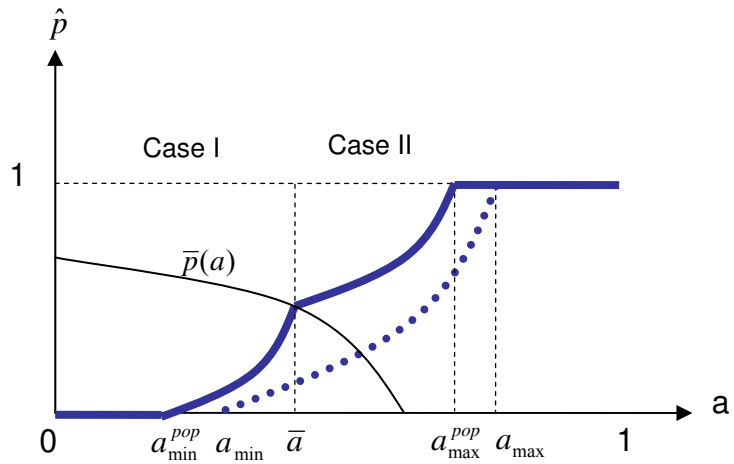|	A (thick-skin)		B (thin-skin)	
A	F_{AA}	F_{AA}	*1*	*0*
B	*0*	*1*	$\dfrac{F}{2F-1}$	$\dfrac{F}{2F-1}$

Figure 2 Payoff matrix of the stage game, derived from equilibria of the demand game played in the three types of dyads.

Figure 3 Fraction of thin-skinned agents in evolutionarily stable population as a function of the degree of assortment. The dotted curve shows the proximate pain case discussed in Section 4.
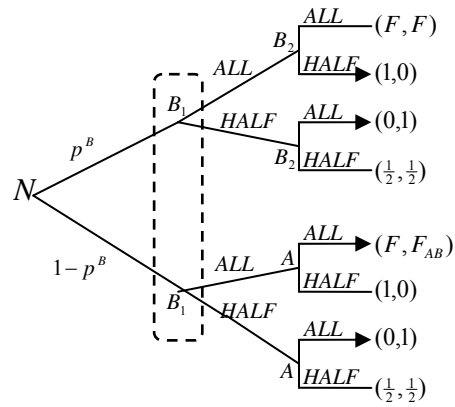

Figure 4 Fraction of thin-skinned agents in evolutionarily stable population as a function of the degree of assortment when agents cannot observe type of opponent in dyad but know population share of thin-skins and degree of assortment. In the Case I region thin-skins always demand HALF. In the Case II region, thin-skins mix between ALL and HALF. The dotted curve shows the full-information case from Section 3. In the case shown $a_{\min} < \overline{a}$; in the other possible case $\overline{a} < a_{\min} < a_{\max}^{pop}$.
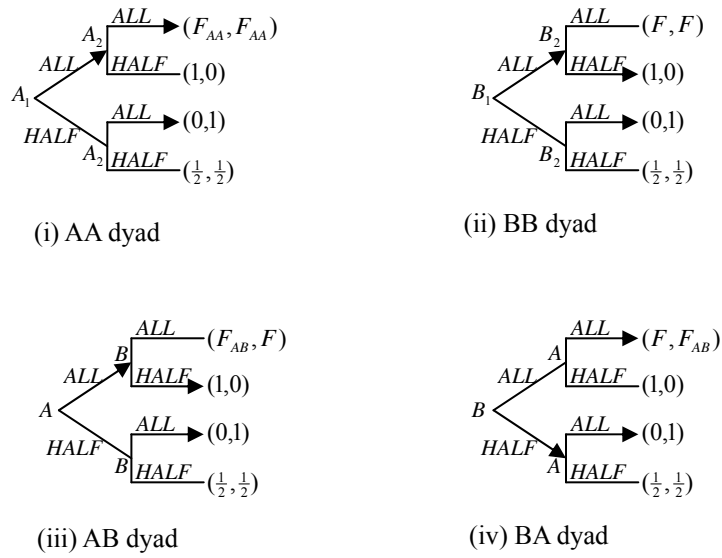
Figure 5 Extensive-form demand game played in dyads under perfect information. Arrows indicate subgame-perfect strategy at each node.

	A (thick-skin)		B (thin-skin)	
A	F_{AA}	F_{AA}	*1*	*0*
B	*0*	*1*	$\frac{1}{2}$	$\frac{1}{2}$

Figure 6 Payoff matrix of the stage game assuming the demand game is played sequentially, with order of play determined by a random draw.

Figure 7 Extensive-form demand game played under imperfect information in the dyad in which B moves first. Initial node N is "nature's" hidden move that determines whether the dyad is BA or BB. Arrows indicate dominant strategy for the second mover.
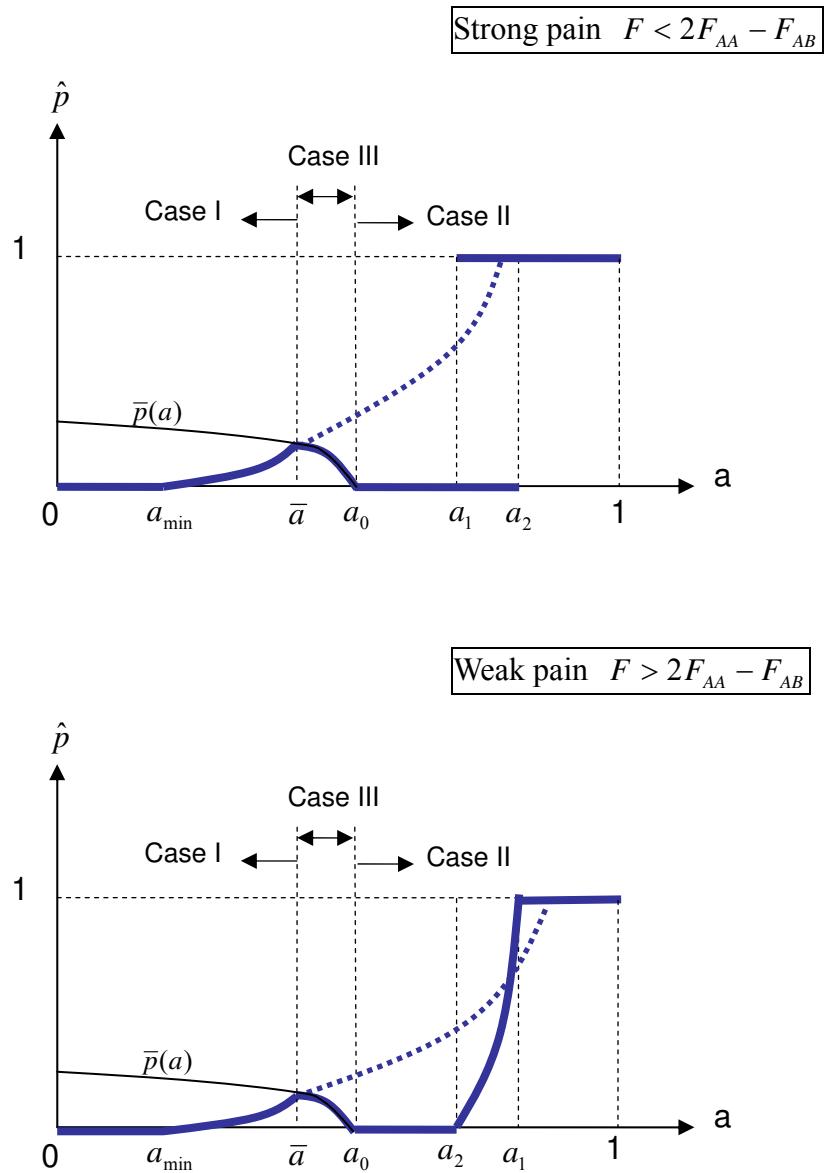
Figure 8   Fraction of thin-skinned agents in evolutionarily stable population as a function of the degree of assortment when agents make demands sequentially, cannot observe type of opponent in dyad, but know population shares and degree of assortment.   In the Case I region first-moving thin-skins demand HALF.   In the Case II region, first-moving thin-skins demand ALL.   In the Case III region, a fraction $\hat{x}$ of first-moving thin-skins demands ALL and the rest demand HALF.   The dotted curve shows the full-information case of Section 6.1.