Role Swap: When the Follower Leads and the Leader Follows

Jan Libich
La Trobe University and VŠB-TU Ostrava

Martin Macháček
VŠB-TU Ostrava

Dat Thanh Nguyen
The University of Danang

ONLINE APPENDIX

Appendix A. Proof of Proposition 1 for the 3×3 Game

The steps to derive the three types of equilibrium conditions for the 3×3 game, as well as the intuition, are analogous to those for the 2×2 game appearing in the main text. This is apparent when comparing Figure 4 below with Figure 2 in the main text. It is therefore unnecessary to provide detailed explanations here.

We will start again with the conditions for M's Dominance region, i.e. for $\left(M_1^S M_2^S, F_1^S F_2^S\right)$ to be the unique subgame perfect equilibrium based on elimination of strictly dominated strategies. F's Yielding condition is composed of four inequalities, which can be compactly written as follows.

(28)
$$\overbrace{\underbrace{(1-m)\,t}_{M_2^S} + \underbrace{my}_{\text{Switch to } M_2^B}}^{F_2^S} > \max\left\{ \overbrace{\underbrace{(1-m)\,v}_{M_2^S} + \underbrace{mr}_{\text{Switch to } M_2^B}}^{F_2^B}, \overbrace{\underbrace{(1-m)\,x}_{M_2^S} + \underbrace{mw}_{\text{Switch to } M_2^B}}^{F_2^C} \right\},$$
$$\overbrace{\underbrace{(1-m)\,t}_{M_2^S} + \underbrace{mz}_{\text{Switch to } M_2^C}}^{F_2^S} > \max\left\{ \overbrace{\underbrace{(1-m)\,v}_{M_2^S} + \underbrace{mu}_{\text{Switch to } M_2^C}}^{F_2^B}, \overbrace{\underbrace{(1-m)\,x}_{M_2^S} + \underbrace{ms}_{\text{Switch to } M_2^C}}^{F_2^C} \right\}.$$

Combining these conditions and rearranging implies

$$m < \bar{m}_{3\times 3} = \min\left\{ \frac{t-v}{t-v+r-y}, \frac{t-v}{t-v+u-z}, \frac{t-x}{t-x+s-z} \right\} \overset{(3)}{=} \frac{2}{7}. \tag{29}$$

Assuming (29) holds and moving backwards, M's Sticking condition is similarly composed of four constraints

(30)
$$\overbrace{\underbrace{(1-f)\,e}_{F_2^B} + \underbrace{fa}_{\text{switch to } F_2^S}}^{M_2^S} > \max\left\{ \overbrace{\underbrace{(1-f)\,c}_{F_2^B} + \underbrace{fi}_{\text{switch to } F_2^S}}^{M_2^B}, \overbrace{\underbrace{(1-f)\,h}_{F_2^B} + \underbrace{fg}_{\text{switch to } F_2^S}}^{M_2^C} \right\},$$
$$\overbrace{\underbrace{(1-f)\,d}_{F_2^C} + \underbrace{fa}_{\text{switch to } F_2^S}}^{M_2^S} > \max\left\{ \overbrace{\underbrace{(1-f)\,j}_{F_2^C} + \underbrace{fi}_{\text{switch to } F_2^S}}^{M_2^B}, \overbrace{\underbrace{(1-f)\,b}_{F_2^C} + \underbrace{fg}_{\text{switch to } F_2^S}}^{M_2^C} \right\}.$$

They can be rearranged into

$$f > \max\left\{\frac{c-e}{a-i+c-e}, \frac{b-d}{a-g+b-d}\right\} \overset{(3)}{=} \frac{3}{8}. \tag{31}$$

Finally, assuming (31) is satisfied, M's Contest condition can be expressed as

$$\begin{aligned} \overbrace{\underbrace{(1-f)\,e}_{F_1^B \text{ as can't revise}} + \underbrace{fa}_{\text{switch to } F_2^S}}^{M_1^S M_2^S} &> \overbrace{\underbrace{c}_{F_1^B}}^{M_1^B M_2^B}, \\ \overbrace{\underbrace{(1-f)\,d}_{F_1^C \text{ as can't revise}} + \underbrace{fa}_{\text{switch to } F_2^S}}^{M_1^S M_2^S} &> \overbrace{\underbrace{b}_{F_1^C}}^{M_1^C M_2^C}. \end{aligned} \tag{32}$$

After some manipulations we obtain

$$f > \hat{f}_{3\times3} = \max\left\{\frac{c-e}{a-e}, \frac{b-d}{a-d}\right\} \overset{(3)}{=} \frac{3}{5}. \tag{33}$$

It is apparent that like in the 2×2 game reported in the main text, in the 3×3 game the Contest condition in (33) is stronger than the Sticking condition in (31) for all general parameter values. This means that the necessary and sufficient conditions for M's Dominance region are jointly (29) and (33).

The analogous circumstances for F's Dominance region, namely M's Yielding and F's Contest conditions, can again be derived by symmetry

$$\begin{aligned} & f < \bar{f}_{3\times3} = \min\left\{\tfrac{c-e}{c-e+a-i}, \tfrac{c-e}{c-e+d-j}, \tfrac{c-h}{c-h+b-j}\right\} \overset{(3)}{=} \tfrac{2}{7}, \text{ and} \\ & m > \hat{m}_{3\times3} = \max\left\{\tfrac{t-v}{r-v}, \tfrac{s-u}{r-u}\right\} \overset{(3)}{=} \tfrac{3}{5}. \end{aligned} \tag{34}$$

This completes the proof of Proposition 1 for the 3×3 game.

Appendix B. Proof of Proposition 2 for the 3×3 Game

Using the same logic as in the 2×2 Battle of the sexes, in the 3×3 version the condition for M to secure his preferred outcome despite being the Stochastic follower is

$$\hat{f}_{3\times3} = \max\left\{\frac{c-e}{a-e}, \frac{b-d}{a-d}\right\} < \bar{m}_{3\times3} = \min\left\{\frac{t-v}{t-v+r-y}, \frac{t-v}{t-v+u-z}, \frac{t-x}{t-x+s-z}\right\}.$$

Analogously, for F this occurs if

$$\hat{m}_{3\times3} = \max\left\{\frac{t-v}{r-v}, \frac{s-u}{r-u}\right\} < \bar{f}_{3\times3} = \min\left\{\frac{c-e}{c-e+a-i}, \frac{c-e}{c-e+d-j}, \frac{c-h}{c-h+b-j}\right\}.$$

To offer a numerical example, consider the specific payoffs in (3) and change the female's (B, B) payoff from $r = 5$ to $r = 11$. This implies $\hat{m}_{3\times3} = \frac{3}{11}$, which is smaller than $\bar{f}_{3\times3} = \frac{2}{7}$. This means that F's Dominance region crosses the 45-degree line, so F can ensure its Stackelberg payoffs even from the position of the Stochastic follower.

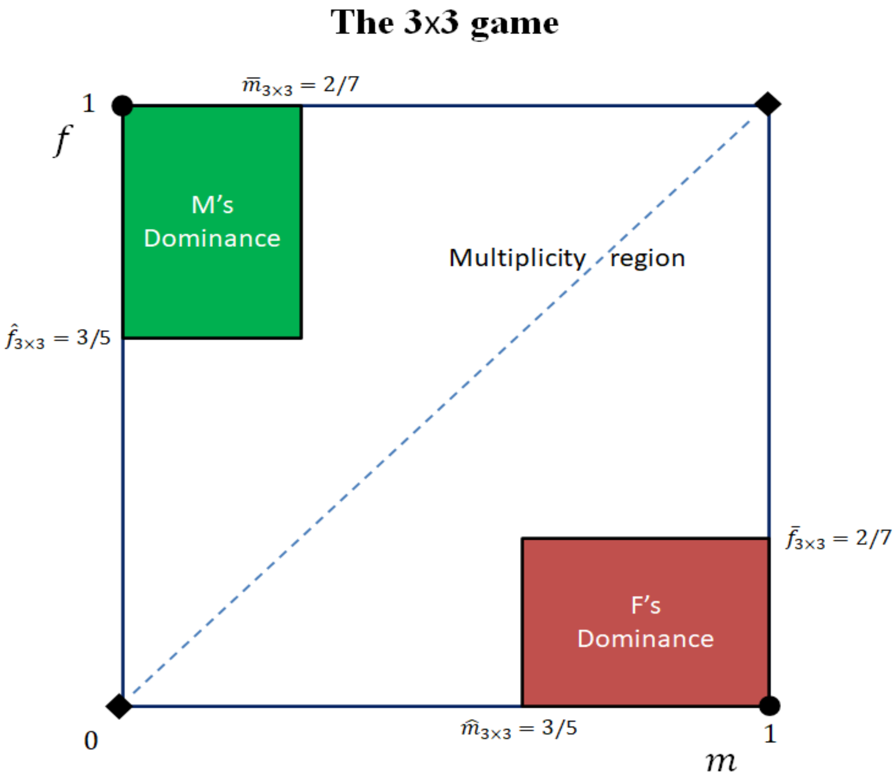

FIGURE 1. Equilibrium regions in the 3×3 Battle of the sexes under Stochastic leadership, featuring revision probabilities m and f. The reported numerical thresholds apply for the specific payoffs in (3).

APPENDIX C. PROOF OF PROPOSITION 3 FOR THE 3×3 GAME

Let us now examine the 3×3 Stag hunt, focusing again on M's Dominance region. F's Yielding condition is composed of the following

(35)

$$\overbrace{\underbrace{(1-m)\,\mathcal{R}}_{M_2^S \text{ as can't revise}} + \underbrace{m\mathcal{Z}}_{\text{switch to } M_2^H}}^{F_2^S} > \max\left\{\overbrace{\underbrace{(1-m)\,\mathcal{T}}_{M_2^S} + \underbrace{m\mathcal{W}}_{\text{switch to } M_2^H}}^{F_2^H}, \overbrace{\underbrace{(1-m)\,\mathcal{S}}_{M_2^S} + \underbrace{m\mathcal{Y}}_{\text{switch to } M_2^H}}^{F_2^B}\right\},$$

$$\overbrace{\underbrace{(1-m)\,\mathcal{R}}_{M_2^S} + \underbrace{m\mathcal{X}}_{M_2^B}}^{F_2^S} > \max\left\{\overbrace{\underbrace{(1-m)\,\mathcal{T}}_{M_2^S} + \underbrace{m\mathcal{V}}_{M_2^B}}^{F_2^H}, \overbrace{\underbrace{(1-m)\,\mathcal{S}}_{M_2^S} + \underbrace{m\mathcal{U}}_{M_2^B}}^{F_2^B}\right\}.$$

Rearranging and combining these inequalities implies

(36)

$$m < \min\left\{\frac{\mathcal{R}-\mathcal{T}}{\mathcal{R}-\mathcal{T}+\mathcal{W}-\mathcal{Z}}, \frac{\mathcal{R}-\mathcal{S}}{\mathcal{R}-\mathcal{S}+\mathcal{Y}-\mathcal{Z}}, \frac{\mathcal{R}-\mathcal{T}}{\mathcal{R}-\mathcal{T}+\mathcal{V}-\mathcal{X}}, \frac{\mathcal{R}-\mathcal{S}}{\mathcal{R}-\mathcal{S}+\mathcal{U}-\mathcal{X}}\right\}.$$

Using (17) suggests that the first, second and fourth fractions in (36) can be the smallest, which means we can streamline it as

$$m < \bar{m}_{3\times3}^{SH} = \min\left\{\frac{\mathcal{R}-\mathcal{T}}{\mathcal{R}-\mathcal{T}+\mathcal{W}-\mathcal{Z}}, \frac{\mathcal{R}-\mathcal{S}}{\mathcal{R}-\mathcal{S}+\mathcal{Y}-\mathcal{Z}}, \frac{\mathcal{R}-\mathcal{S}}{\mathcal{R}-\mathcal{S}+\mathcal{U}-\mathcal{X}}\right\} \overset{(15)}{=} \frac{1}{2}. \tag{37}$$

Moving backwards and assuming (37) is satisfied M's Sticking condition can be written as

$$\begin{aligned} &\overbrace{\underbrace{(1-f)\,\mathcal{J}}_{F_2^H}+\underbrace{f\mathcal{A}}_{F_2^S}}^{M_2^S} > \max\left\{\overbrace{\underbrace{(1-f)\,\mathcal{G}}_{F_2^H}+\underbrace{f\mathcal{C}}_{F_2^S}}^{M_2^H}, \overbrace{\underbrace{(1-f)\,\mathcal{I}}_{F_2^H}+\underbrace{f\mathcal{B}}_{F_2^S}}^{M_2^B}\right\},\\ &\overbrace{\underbrace{(1-f)\,\mathcal{H}}_{F_2^B}+\underbrace{f\mathcal{A}}_{F_2^S}}^{M_2^S} > \max\left\{\overbrace{\underbrace{(1-f)\,\mathcal{E}}_{F_2^B}+\underbrace{f\mathcal{C}}_{F_2^S}}^{M_2^H}, \overbrace{\underbrace{(1-f)\,\mathcal{D}}_{F_2^B}+\underbrace{f\mathcal{B}}_{F_2^S}}^{M_2^B}\right\}. \end{aligned} \tag{38}$$

Combining them implies

$$f > \max\left\{\frac{\mathcal{G}-\mathcal{J}}{\mathcal{G}-\mathcal{J}+\mathcal{A}-\mathcal{C}}, \frac{\mathcal{I}-\mathcal{J}}{\mathcal{I}-\mathcal{J}+\mathcal{A}-\mathcal{B}}, \frac{\mathcal{E}-\mathcal{H}}{\mathcal{E}-\mathcal{H}+\mathcal{A}-\mathcal{C}}, \frac{\mathcal{D}-\mathcal{H}}{\mathcal{D}-\mathcal{H}+\mathcal{A}-\mathcal{B}}\right\}. \tag{39}$$

It is apparent from (17) that the first, second and fourth fractions in (39) can be binding, which allows us to reduce it to

$$f > \hat{f}_{3\times3}^{SH} = \max\left\{\frac{\mathcal{G}-\mathcal{J}}{\mathcal{G}-\mathcal{J}+\mathcal{A}-\mathcal{C}}, \frac{\mathcal{I}-\mathcal{J}}{\mathcal{I}-\mathcal{J}+\mathcal{A}-\mathcal{B}}, \frac{\mathcal{D}-\mathcal{H}}{\mathcal{D}-\mathcal{H}+\mathcal{A}-\mathcal{B}}\right\} \overset{(15)}{=} \frac{1}{2}. \tag{40}$$

Assuming that both (37) and (40) hold M's Contest condition is

$$\overbrace{\underbrace{(1-f)\,\mathcal{J}}_{F_1^H}+\underbrace{f\mathcal{A}}_{F_2^S}}^{M_1^S M_2^S} > \overbrace{\mathcal{G}}^{M_1^H M_2^H} \quad\text{and}\quad \overbrace{\underbrace{(1-f)\,\mathcal{H}}_{F_1^B}+\underbrace{f\mathcal{A}}_{F_2^S}}^{M_1^S M_2^S} > \overbrace{\mathcal{D}}^{M_1^B M_2^B}. \tag{41}$$

Upon rearranging, we obtain

$$f > \max\left\{\frac{\mathcal{G}-\mathcal{J}}{\mathcal{A}-\mathcal{J}}, \frac{\mathcal{D}-\mathcal{H}}{\mathcal{A}-\mathcal{H}}\right\} \overset{(15)}{=} \frac{1}{3}. \tag{42}$$

The Contest condition is always weaker than the Sticking condition, as was the case in the 2×2 Stag hunt. As such, the necessary and sufficient conditions in the 3×3 game for M's dominance region are (37) and (40).

Similarly, the necessary and sufficient conditions for F's dominance region are

$$f < \bar{f}_{3\times3}^{SH} = \min\left\{\frac{\mathcal{A}-\mathcal{C}}{\mathcal{A}-\mathcal{C}+G-\mathcal{J}}, \frac{\mathcal{A}-\mathcal{B}}{\mathcal{A}-\mathcal{B}+\mathcal{I}-\mathcal{J}}, \frac{\mathcal{A}-\mathcal{B}}{\mathcal{A}-\mathcal{B}+\mathcal{D}-\mathcal{H}}\right\} \overset{(15)}{=} \frac{1}{2},$$

and

$$m > \hat{m}_{3\times3}^{SH} = \max\left\{\frac{\mathcal{W}-\mathcal{Z}}{\mathcal{W}-\mathcal{Z}+\mathcal{R}-\mathcal{T}}, \frac{\mathcal{Y}-\mathcal{Z}}{\mathcal{Y}-\mathcal{Z}+\mathcal{R}-\mathcal{S}}, \frac{\mathcal{U}-\mathcal{X}}{\mathcal{U}-\mathcal{X}+\mathcal{R}-\mathcal{S}}\right\} \overset{(15)}{=} \frac{1}{2}.$$

The condition for M's Dominance region to cross the 45-degree line is

$$\hat{f}^{SH}_{3\times 3} = \max\left\{\tfrac{\mathcal{G}-\mathcal{J}}{\mathcal{G}-\mathcal{J}+\mathcal{A}-\mathcal{C}}, \tfrac{\mathcal{I}-\mathcal{J}}{\mathcal{I}-\mathcal{J}+\mathcal{A}-\mathcal{B}}, \tfrac{\mathcal{D}-\mathcal{H}}{\mathcal{D}-\mathcal{H}+\mathcal{A}-\mathcal{B}}\right\} <$$
$$\bar{m}^{SH}_{3\times 3} = \min\left\{\tfrac{\mathcal{R}-\mathcal{T}}{\mathcal{R}-\mathcal{T}+\mathcal{W}-\mathcal{Z}}, \tfrac{\mathcal{R}-\mathcal{S}}{\mathcal{R}-\mathcal{S}+\mathcal{Y}-\mathcal{Z}}, \tfrac{\mathcal{R}-\mathcal{S}}{\mathcal{R}-\mathcal{S}+\mathcal{U}-\mathcal{X}}\right\},$$

whereas the analogous one for F's Dominance region is

$$\hat{m}^{SH}_{3\times 3} = \max\left\{\tfrac{\mathcal{W}-\mathcal{Z}}{\mathcal{W}-\mathcal{Z}+\mathcal{R}-\mathcal{T}}, \tfrac{\mathcal{Y}-\mathcal{Z}}{\mathcal{Y}-\mathcal{Z}+\mathcal{R}-\mathcal{S}}, \tfrac{\mathcal{U}-\mathcal{X}}{\mathcal{U}-\mathcal{X}+\mathcal{R}-\mathcal{S}}\right\} <$$
$$\bar{f}^{SH}_{3\times 3} = \min\left\{\tfrac{\mathcal{A}-\mathcal{C}}{\mathcal{A}-\mathcal{C}+\mathcal{G}-\mathcal{J}}, \tfrac{\mathcal{A}-\mathcal{B}}{\mathcal{A}-\mathcal{B}+\mathcal{I}-\mathcal{J}}, \tfrac{\mathcal{A}-\mathcal{B}}{\mathcal{A}-\mathcal{B}+\mathcal{D}-\mathcal{H}}\right\}.$$

This completes the proof of Proposition 3 for the 3×3 game.

Appendix D. Proof of Proposition 4 for the 3×3 Game

Turning to the Hawk and dove game, let us again start with the conditions for M's Dominance region in the 3×3 game. The steps and intuition are again analogous to the 2×2 game (as well as to the Battle of the sexes) above. F's Yielding condition requires

$$\begin{aligned}
&\overbrace{\underbrace{(1-m)\,W}_{M_2^H} + \underbrace{mU}_{M_2^D}}^{F_2^D} > \max\left\{\overbrace{\underbrace{(1-m)\,Z}_{M_2^H} + \underbrace{mR}_{M_2^D}}^{F_2^H}, \overbrace{\underbrace{(1-m)\,Y}_{M_2^H} + \underbrace{mS}_{M_2^D}}^{F_2^O}\right\},\\
&\overbrace{\underbrace{(1-m)\,W}_{M_2^H} + \underbrace{mV}_{M_2^O}}^{F_2^D} > \max\left\{\overbrace{\underbrace{(1-m)\,Z}_{M_2^H} + \underbrace{mX}_{M_2^O}}^{F_2^H}, \overbrace{\underbrace{(1-m)\,Y}_{M_2^H} + \underbrace{mT}_{M_2^O}}^{F_2^O}\right\},
\end{aligned} \tag{43}$$

which can be rearranged into

$$m < \min\left\{\frac{W-Z}{W-Z+R-U}, \frac{W-Y}{W-Y+S-U}, \frac{W-Z}{W-Z+X-V}, \frac{W-Y}{W-Y+T-V}\right\}. \tag{44}$$

Once we utilize (22) we can note that the first and fourth fractions in (44) can be the smallest of the four. The required condition can thus be simplified into

$$m < \bar{m}^{HD}_{3\times 3} = \min\left\{\frac{W-Z}{W-Z+R-U}, \frac{W-Y}{W-Y+T-V}\right\} \overset{(23)}{=} \frac{5}{8}. \tag{45}$$

Assuming (5) is satisfied, M's Sticking condition requires

$$\begin{aligned}
&\overbrace{\underbrace{(1-f)\,J}_{F_2^H} + \underbrace{fA}_{F_2^D}}^{M_2^H} > \max\left\{\overbrace{\underbrace{(1-f)\,G}_{F_2^H} + \underbrace{fD}_{F_2^D}}^{M_2^D}, \overbrace{\underbrace{(1-f)\,I}_{F_2^H} + \underbrace{fB}_{F_2^D}}^{M_2^O}\right\},\\
&\overbrace{\underbrace{(1-f)\,H}_{F_2^O} + \underbrace{fA}_{F_2^D}}^{M_2^H} > \max\left\{\overbrace{\underbrace{(1-f)\,E}_{F_2^O} + \underbrace{fD}_{F_2^D}}^{M_2^D}, \overbrace{\underbrace{(1-f)\,C}_{F_2^O} + \underbrace{fB}_{F_2^D}}^{M_2^O}\right\},
\end{aligned} \tag{46}$$

which can be expressed as

$$f > \max\left\{\frac{G-J}{G-J+A-D}, \frac{I-J}{I-J+A-B}, \frac{E-H}{E-H+A-D}, \frac{C-H}{C-H+A-B}\right\}.$$

Using (22), we can check that the first, second and forth fractions can be highest. So we can rewrite this condition as

$$f > \hat{f}^{HD}_{3\times 3} = \max\left\{\frac{G-J}{G-J+A-D}, \frac{I-J}{I-J+A-B}, \frac{C-H}{C-H+A-B}\right\} \overset{(23)}{=} \frac{5}{6}. \tag{47}$$

Moving backwards and assuming the Yielding and the Sticking conditions both hold, M's Contest condition is as follows

$$\overbrace{\underbrace{(1-f)\,J}_{F_1^H} + \underbrace{fA}_{F_2^D}}^{M_1^H M_2^H} > \overbrace{G}^{M_1^D M_2^D} \quad \text{and} \quad \overbrace{\underbrace{(1-f)\,H}_{F_1^O} + \underbrace{fA}_{F_2^D}}^{M_1^H M_2^H} > \overbrace{C}^{M_1^O M_2^O}. \tag{48}$$

Rearranging this yields

$$f > \max\left\{\frac{G-J}{A-J}, \frac{C-H}{A-H}\right\}. \tag{49}$$

It is clear that in the Hawk and dove game, like in the Battle of the sexes, the Contest condition is stronger than the Sticking condition in both the 2×2 and 3×3 games. Similarly, the conditions for F's dominance region are:

$$f < \bar{f}^{HD}_{3\times 3} = \min\left\{\frac{G-J}{G-J+A-D}, \frac{G-I}{G-I+C-E}\right\} \overset{(23)}{=} \frac{5}{8},$$

and

$$m > \hat{m}^{HD}_{3\times 3} = \max\left\{\frac{W-Z}{W-Z+R-U}, \frac{Y-Z}{Y-Z+R-S}, \frac{T-X}{T-X+R-S}\right\} \overset{(23)}{=} \frac{5}{6}.$$

The condition for the role swap, in which M behaves as the Stackelberg leader despite being the Stochastic follower is

$$\hat{f}^{HD}_{3\times 3} = \max\left\{\tfrac{G-J}{G-J+A-D}, \tfrac{I-J}{I-J+A-B}, \tfrac{C-H}{C-H+A-B}\right\} <$$
$$\bar{m}^{HD}_{3\times 3} = \min\left\{\tfrac{W-Z}{W-Z+R-U}, \tfrac{W-Y}{W-Y+T-V}\right\}$$

and for F the analogous conditions is

$$\hat{m}^{HD}_{3\times 3} = \max\left\{\tfrac{W-Z}{W-Z+R-U}, \tfrac{Y-Z}{Y-Z+R-S}, \tfrac{T-X}{T-X+R-S}\right\} <$$
$$\bar{f}^{HD}_{3\times 3} = \min\left\{\tfrac{G-J}{G-J+A-D}, \tfrac{G-I}{G-I+C-E}\right\}$$

This completes the proof of Proposition 4 for the 3×3 game.