

A Online Supplement to "The Myopic Stable Set for Social Environments"

A.1 Examples

Example A.1. In this example, we describe a social environment which has no MSS when iterated external stability is used in the definition instead of asymptotic external stability.

Consider the social environment $\Gamma = (\{1\}, (X, d), E, \succeq_1)$, where the state space is given by $X = \{1/k | k \in \mathbb{N}\} \cup \{0\}$ and $d(x, y) = |x - y|$. Note that X is compact. Preferences \succeq_1 are defined by $x \succeq_1 y$ if and only if $x = y$ or $x < y$. The effectivity correspondence E is such that $\{1\} \in E(1/k, 1/(k + 1))$ for every $k \in \mathbb{N}$ and $E(x, y) = \emptyset$ otherwise. It follows that

$$f\left(\frac{1}{k}\right) = \left\{\frac{1}{k}, \frac{1}{k+1}\right\}.$$

Observe that $0 \in f^\infty(x)$ for every $x \in X$ and that $f(0) = \{0\}$. It now follows easily that $\{0\}$ is an MSS.

We show that there is no closed set satisfying iterated external stability together with deterrence of external deviations and minimality. Towards a contradiction, suppose that the closed set $M \subseteq X$ satisfies these properties. Since, for every $k \in \mathbb{N}$, $0 \notin f^\mathbb{N}(1/k)$, the set $\{0\}$ does not satisfy iterated external stability. Given that $M \neq \{0\}$ and M is non-empty, there is $k \in \mathbb{N}$ such that $1/k \in M$. Let k be the smallest such number. From deterrence of external deviations, we have that also $1/(k + 1) \in M$. Based on the corresponding properties of M , it is easy to verify that the closed, non-empty set $M' = M \setminus \{1/k\}$ satisfies deterrence of external deviations and iterated external stability. Since M' is a proper subset of M , M violates the minimality property.

Example A.2. In the next example, we consider an infinite social environment for which there is more than one MSS.

Consider the social environment $\Gamma = (\{1\}, (X, d), E, \succeq_1)$, where

$$X = \left\{0, \frac{1}{2}, 1\right\} \cup \left\{\frac{1}{k} \mid k \in \mathbb{N} \setminus \{1, 2\}\right\} \cup \left\{1 - \frac{1}{k} \mid k \in \mathbb{N} \setminus \{1, 2\}\right\}$$

and the metric is given by $d(x, y) = |x - y|$.

The effectivity correspondence is such that the individual can move from both states 0 and 1 to state $1/2$ and, for every $k \in \mathbb{N} \setminus \{1, 2\}$, from state $1 - 1/k$ to state $1/k$ and from state $1/k$ to state $1 - 1/(k + 1)$. The individual cannot make any other moves. The preferences of the individual are such that

$$\frac{2}{3} \prec_1 \frac{1}{3} \prec_1 \frac{3}{4} \prec_1 \frac{1}{4} \prec_1 \frac{4}{5} \prec_1 \frac{1}{5} \prec_1 \cdots \prec_1 1 \prec_1 0 \prec_1 \frac{1}{2}.$$

We claim that both $\{0, 1/2\}$ and $\{1/2, 1\}$ are myopic stable sets. Since the effectivity correspondence admits no move outside the respective sets, both $\{0, 1/2\}$ and $\{1/2, 1\}$ satisfy deterrence of external deviations. For asymptotic external stability, observe that for every $k \in \mathbb{N} \setminus \{1, 2\}$ it holds that $\{0, 1\} \subset f^\infty(1/k)$ and $\{0, 1\} \subset f^\infty(1 - 1/k)$. Moreover, we have $1/2 \in f(0) = f^\infty(0)$ and $1/2 \in f(1) = f^\infty(1)$. For minimality, the sets $\{0\}$ and $\{1\}$ violate deterrence of external deviations since $1/2 \in f(0)$ and $1/2 \in f(1)$. The set $\{1/2\}$ violates asymptotic external stability as $1/2 \notin f^\infty(x)$ for any $x \in X$ different from 0, $1/2$ and 1.

Example A.3. This example provides a social environment in which the effectivity correspondence is lower hemi-continuous and the preferences are continuous, but where the weak dominance MSS is not unique.

Consider the social environment $\Gamma = (\{1, 2\}, (X, d), E, (\succeq_1, \succeq_2))$, where

$$X = \{(0, 0), (1, 0), (2, 0)\} \cup \{(0, \frac{2}{k}), (1, \frac{1}{k}), (2, \frac{1}{k}) | k \in \mathbb{N}\}$$

and d is the Euclidean metric on X , so $d(x, y) = \|x - y\|_2$. It clearly holds that X is compact.

Individual 1 only cares about the first component of the state while individual 2 only cares about the second component. Both individuals prefer states where the component they care about is lower over states where it is higher. Note that these preferences are continuous.

The effectivity correspondence is as follows. For every $k \in \mathbb{N}$, the singleton $\{1\}$ can move from state $(2, 1/k)$ to state $(1, 1/k)$ and the singleton $\{2\}$ can move from state $(1, 1/k)$ to state $(2, 1/(k + 1))$. Moreover, for every $k \in \mathbb{N}$, the singleton

$\{2\}$ can move from state $(0, 2/k)$ to state $(1, 1/k)$. Coalition $\{1, 2\}$ can move from states $(1, 0)$ and $(2, 0)$ to state $(0, 0)$ and, for every $k \in \mathbb{N}$, from states $(1, 1/k)$ and $(2, 1/k)$ to state $(0, 2/k)$. No other moves are possible.

To see that the effectivity correspondence is lower hemi-continuous, let the sequence $(x^k)_{k \in \mathbb{N}}$ in X be such that $x^k \rightarrow x$. There are only three relevant sequences of states in X : the sequence $((0, 2/k))_{k \in \mathbb{N}}$, the sequence $((1, 1/k))_{k \in \mathbb{N}}$, and the sequence $((2, 1/k))_{k \in \mathbb{N}}$. The first converges to $(0, 0)$, the second to $(1, 0)$, and the third to $(2, 0)$.

Let some $x \in \{(0, 0), (1, 0), (2, 0)\}$ be given. Since $G_{\{1\}}(x) = \{x\}$ and $G_{\{2\}}(x) = \{x\}$, it is immediate that $G_{\{1\}}$ and $G_{\{2\}}$ are lower hemi-continuous.

For $G_{\{1,2\}}$, the only non-trivial cases are $x = (1, 0)$ and $x = (2, 0)$. We give the argument for state $x = (1, 0)$ explicitly. The argument for state $(2, 0)$ follows by symmetry. For every $y \in G_{\{1,2\}}(1, 0)$ we have to find a sequence $(y^k)_{k \in \mathbb{N}}$ such that $y^k \in G_{\{1,2\}}(1, 1/k)$ and $y^k \rightarrow y$. If $y = (0, 0)$, we take the sequence $((0, 2/k))_{k \in \mathbb{N}}$. If $y = (1, 0)$, we take the sequence $((1, 1/k))_{k \in \mathbb{N}}$.

Since $f^\infty(0, 0) = \{(0, 0)\}$, $f^\infty(1, 0) = \{(1, 0)\}$, and $f^\infty(2, 0) = \{(2, 0)\}$, it follows from asymptotic external stability that $\{(0, 0), (1, 0), (2, 0)\}$ is a subset of any MSS. Since this set satisfies deterrence of external deviations and asymptotic external stability, it follows from minimality that the unique MSS is equal to $\{(0, 0), (1, 0), (2, 0)\}$.

On the other hand, both sets $\{(0, 0), (1, 0)\}$ and $\{(0, 0), (2, 0)\}$ are a weak dominance MSS. Indeed, from both $(1, 0)$ and $(2, 0)$, the coalition $\{1, 2\}$ can deviate to $(0, 0)$ if only weak dominance is imposed. To satisfy asymptotic external stability, it is sufficient that on top of the state $(0, 0)$, either the state $(1, 0)$ or the state $(2, 0)$ should be present. By minimality, it follows that only one of these states is included.

Example A.4. This example demonstrates that the stochastic approach to infinite environments based on irreducibility of the Markov chain can deliver predictions that differ drastically from those of the MSS.

Consider the social environment $\Gamma = (N, (X, d), E, (\succeq_i)_{i \in N})$, where $N = \{1, 2\}$,

$X = [0, 1] \times [0, 1]$, and the metric is $d(x, y) = \|x - y\|_1 = |x_1 - y_1| + |x_2 - y_2|$. The effectivity correspondence is such that individual 1 can change the first component of the state and individual 2 the second component, so $\{1\} \in E(x, y)$ if and only if $x_2 = y_2$ and $\{2\} \in E(x, y)$ if and only if $x_1 = y_1$. The coalition $\{1, 2\}$ is never effective. The preferences of the individuals are such that

$$\begin{aligned} x \succeq_1 y & \text{ if and only if } 2x_1x_2 - x_1 - x_2 \geq 2y_1y_2 - y_1 - y_2, \\ x \succeq_2 y & \text{ if and only if } 2x_1x_2 - x_1 - x_2 \leq 2y_1y_2 - y_1 - y_2. \end{aligned}$$

It is not hard to see that this social environment corresponds to the normal-form game of matching pennies, where x_1 is the probability of the row player choosing “up” and x_2 is the probability of the column player choosing “left”. The unique Nash equilibrium of this game is equal to $x^* = (1/2, 1/2)$.

For every $x \in X$, we define

$$\begin{aligned} f_1(x) &= \begin{cases} \{y \in X \mid y_1 \leq x_1 \text{ and } y_2 = x_2\} & \text{if } x_2 < \frac{1}{2}, \\ \{x\} & \text{if } x_2 = \frac{1}{2}, \\ \{y \in X \mid y_1 \geq x_1 \text{ and } y_2 = x_2\} & \text{if } x_2 > \frac{1}{2}, \end{cases} \\ f_2(x) &= \begin{cases} \{y \in X \mid y_1 = x_1 \text{ and } y_2 \geq x_2\} & \text{if } x_1 < \frac{1}{2}, \\ \{x\} & \text{if } x_1 = \frac{1}{2}, \\ \{y \in X \mid y_1 = x_1 \text{ and } y_2 \leq x_2\} & \text{if } x_1 > \frac{1}{2}, \end{cases} \end{aligned}$$

so we can express the dominance correspondence as

$$f(x) = f_1(x) \cup f_2(x).$$

We consider the better-response dynamics where each element of $f(x)$ is selected with equal probability. To do so, we define $\rho_1 : X \rightarrow [0, 1]$ and $\rho_2 : X \rightarrow [0, 1]$ as the functions that project x on its first and second coordinate, respectively. We use λ to denote the Lebesgue measure. Let $\mathcal{B}(X)$ denote the Borel σ -algebra on X . The transition probability kernel resulting from better-response dynamics

is obtained by defining, for every $x \in X$, and for every $A \in \mathcal{B}(X)$,

$$Q(x, A) = \begin{cases} 0 & \text{if } x = (\frac{1}{2}, \frac{1}{2}) \notin A, \\ 1 & \text{if } x = (\frac{1}{2}, \frac{1}{2}) \in A, \\ 2\lambda(\rho_1(A \cap f_1(x))), & \text{if } x_1 = \frac{1}{2}, x_2 \neq \frac{1}{2}, \\ 2\lambda(\rho_2(A \cap f_2(x))), & \text{if } x_1 \neq \frac{1}{2}, x_2 = \frac{1}{2}, \\ \frac{\lambda(\rho_1(A \cap f_1(x))) + \lambda(\rho_2(A \cap f_2(x)))}{\lambda(\rho_1(f_1(x))) + \lambda(\rho_2(f_2(x)))}, & \text{if } x_1 \neq \frac{1}{2}, x_2 \neq \frac{1}{2}. \end{cases}$$

The first and second equality above show that the better-response dynamics never leaves the Nash equilibrium once reached. The third equality concerns the case where only player 1 likes to move. Observe that if $x_1 = 1/2$ and $x_2 \neq 1/2$, then $\lambda(\rho_1(f_1(x))) = 1/2$, which explains the multiplication by 2. A similar remark applies to the fourth equality above. For the last equality, notice that $x_1 \neq 1/2$ and $x_2 \neq 1/2$ implies that $\lambda(\rho_1(f_1(x))) > 0$ or $\lambda(\rho_2(f_2(x))) > 0$, so there is no division by zero.

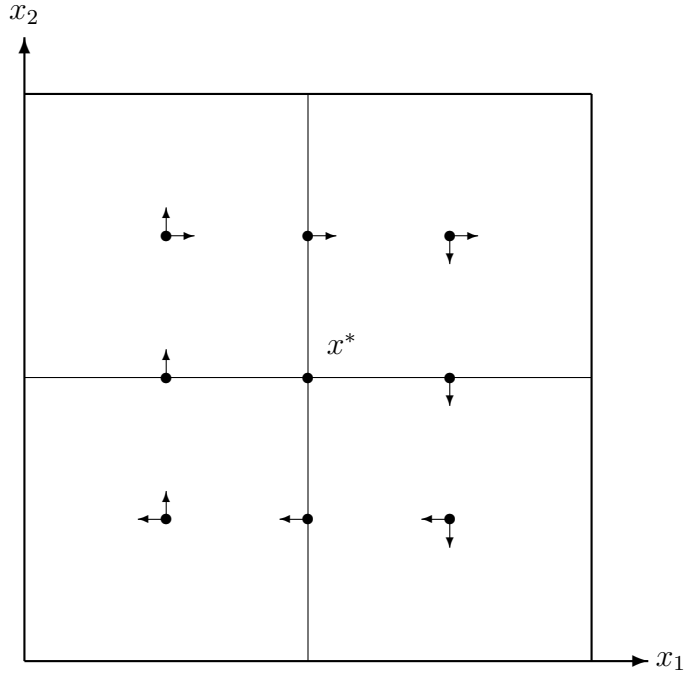


Figure 1: Better-response dynamics for the game of matching pennies.

The Markov process is illustrated in Figure 1. The arrows indicate the direction in which a state changes. A typical state can change in two directions, either west

or east and either north or south, thereby generating two line segments on which the next state lies.

For every $A \in \mathcal{B}(X)$, $Q(\cdot, A)$ is a measurable function on X , but it is in general not continuous. For instance, if $A = \{x^*\}$, then $Q(x, A) = 1$ if $x = x^*$ and $Q(x, A) = 0$, otherwise. Indeed, the state x^* does not belong to $f(x)$ unless $x = x^*$ and in that case $f(x^*) = \{x^*\}$.

In this setting and other settings with an infinite state space, the Markov chain returns to a given state with probability zero, so the concept of a recurrent state is of less use and importance. Instead, for infinite settings, the property of irreducibility is often studied, which expresses that all parts of the state space can be reached by the Markov chain, no matter what the starting point is. Given a state $x \in X$ and a set A in the Borel σ -algebra $\mathcal{B}(X)$ on X , let $L(x, A)$ denote the probability that the Markov chain has a realization belonging to A at some point in the future when starting from x . Let φ be the measure on X that assigns to each set in $\mathcal{B}(X)$ its Lebesgue measure. A Markov process (X, Q) is called φ -irreducible if for every $A \in \mathcal{B}(X)$ such that $\varphi(A) > 0$ it holds that $L(x, A) > 0$ for every $x \in X$.

The Markov process (X, Q) in Example A.4 is such that X can be decomposed in two parts, namely $\{x^*\}$ and $X \setminus \{x^*\}$. There is no transition between these two sets of states and the restriction of the Markov process to each set is irreducible. This is obvious for $\{x^*\}$. The next result shows this for $X \setminus \{x^*\}$.

Theorem A.5. The restriction of the Markov process (X, Q) in Example A.4 to $X \setminus \{x^*\}$ is φ -irreducible.

Proof. According to Proposition 4.2.1 of Meyn and Tweedie (1993), we have to show that for every $x \in X \setminus \{x^*\}$, for every $A \in \mathcal{B}(X \setminus \{x^*\})$ such that $\varphi(A) > 0$, there exists $k \in \mathbb{N}$ such that $Q^k(x, A) > 0$, where $Q^k(x, A)$ denotes the probability of reaching A from x in k transitions.

It is convenient to partition the set $X \setminus \{x^*\}$ in four subsets,

$$\begin{aligned} X^1 &= \{x \in X \mid x_1 \leq \frac{1}{2}, x_2 > \frac{1}{2}\}, \\ X^2 &= \{x \in X \mid x_1 > \frac{1}{2}, x_2 \geq \frac{1}{2}\}, \\ X^3 &= \{x \in X \mid x_1 \geq \frac{1}{2}, x_2 < \frac{1}{2}\}, \\ X^4 &= \{x \in X \mid x_1 < \frac{1}{2}, x_2 \leq \frac{1}{2}\}. \end{aligned}$$

Let some $x \in X^4$ and some $A \in \mathcal{B}(X \setminus \{x^*\})$ such that $\varphi(A) > 0$ be given. We partition A in the four subsets $A^1 \subseteq X^1$, $A^2 \subseteq X^2$, $A^3 \subseteq X^3$, and $A^4 \subseteq X^4$. At least one of these four sets has positive Lebesgue measure. From x , the probability to reach a point in the set $Y^1 = \{y^1 \in X^1 \mid y_1^1 = x_1\}$ is at least $1/3$ and the probability distribution over Y^1 is uniform. From $y^1 \in Y^1$, the probability to reach a point in the set $Y^2(y^1) = \{y^2 \in X^2 \mid y_2^2 = y_2^1\}$ is at least $1/3$ and the probability distribution over $Y^2(y^1)$ is uniform. Thus, the probability to reach a point in X^2 after 2 transitions is at least $1/9$ and, conditional on reaching X^2 , the distribution of this point is uniform on X^2 . It now follows that $Q^2(x, A) \geq \varphi(A^2)/9$. Repeating this argument, we find that $Q^3(x, A) \geq \varphi(A^3)/27$, $Q^4(x, A) \geq \varphi(A^4)/81$, and $Q^5(x, A) \geq \varphi(A^1)/243$. Since at least one of A^1 , A^2 , A^3 , and A^4 has strictly positive Lebesgue measure, we have shown that the restriction of the Markov process to $X \setminus \{x^*\}$ is φ -irreducible. An analogous argument holds for $x \in X^i$, where $i \neq 4$. \square

Example A.4 shows that for the social environment corresponding to the normal-form game of matching pennies, none of the strategy profiles is singled out by the stochastic better-response dynamics. In contrast, we show in Section A.6 that the MSS is unique and consists of the Nash equilibrium x^* .

Example A.6. In this example, we show that the coalition structure core does not satisfy iterated external stability.

Let (N, v) be a coalition function form game such that $N = \{1, 2, 3\}$, $v(\{1, 2\}) = 1$, and $v(\{2, 3\}) = 1$. All other coalitions have a coalitional value of 0. Thus, player 2 can choose to form a coalition with either player 1 or player 3 to form a two-person coalition generating a surplus equal to one. The coalition structure core therefore consists of only two states, y and y' , with equal payoffs,

$u(y) = u(y') = (0, 1, 0)$, and coalitional structures $\pi(y) = \{\{1, 2\}, \{3\}\}$ and $\pi(y') = \{\{1\}, \{2, 3\}\}$.

Consider an initial state $x^0 \in X$ such that $\pi(x^0) = \{\{1\}, \{2\}, \{3\}\}$ and $u(x^0) = (0, 0, 0)$. Under our notion of a myopic improvement, where all players involved in a move have to gain strictly, a state $x^1 \neq x^0$ belongs to $f(x^0)$ if and only if either $\pi(x^1) = \{\{1, 2\}, \{3\}\}$ and $u(x^1) = (\varepsilon, 1 - \varepsilon, 0)$ for some $\varepsilon \in (0, 1)$ or $\pi(x^1) = \{\{1\}, \{2, 3\}\}$ and $u(x^1) = (0, 1 - \varepsilon, \varepsilon)$ for some $\varepsilon \in (0, 1)$. It follows that x^1 is a state where either player 1 or player 3 receives a payoff of zero and the other two players receive a strictly positive payoff summing up to 1.

Now consider any state x^k such that either player 1 or player 3 receives 0 and the other two players receive a strictly positive payoff summing up to 1. We claim that any state $x^{k+1} \in f(x^k)$ has the same properties. Without loss of generality, assume that $u_3(x^k) = 0$. Let x^{k+1} be an element of $f(x^k)$ different from x^k . Since $u_1(x^k) + u_2(x^k) = 1$, the moving coalition is $\{2, 3\}$ and it holds that $\pi(x^{k+1}) = \{\{1\}, \{2, 3\}\}$. Moreover, it must also hold that $u_2(x^{k+1}) > u_2(x^k) > 0$ and $u_3(x^{k+1}) > u_3(x^k) = 0$, which proves the claim. Thus, for every $k \in \mathbb{N}$, if $x^k \in f^k(x^0) \setminus \{x^0\}$, then x^k is such that there are two players with a strictly positive payoff. It follows that there is no $k \in \mathbb{N}$ such that x^k belongs to the coalition structure core.

A.2 Three-player Simple Games and the vNM Stable Set

Let (N, v) be a coalition function form game with $N = \{1, 2, 3\}$ corresponding to a proper simple game. Let Γ be the social environment induced by the γ -model. We compare the prediction of the MSS with the vNM stable set of Γ . The coalition function form game keeps track of the partition of the set of players and imposes that a coalition fully distributes its surplus between its members. The model of coalition function form games is therefore different from the one of transferable utility games. Hence, we cannot rely on the description of the vNM stable sets for three-player simple games as given in Lucas (1992), but have to derive them from scratch instead.

We restrict ourselves to the three most interesting cases: there is one winning two-player coalition, without loss of generality $\{1, 3\}$; there are two winning two-player coalitions, without loss of generality $\{1, 2\}$ and $\{2, 3\}$; all two-player coalitions are winning. The second case is known as the three-person veto-power game and the third case as the three-person simple majority game.

The first example shows that if $\{1, 3\}$ is the only winning two-player coalition, then the MSS and the vNM stable set are unique and equal to the coalition structure core. The prediction is therefore that either coalition $\{1, 3\}$ or coalition $\{1, 2, 3\}$ forms and payoffs are such that the entire surplus is shared between players 1 and 3.

Example A.7. Assume coalition $\{1, 3\}$ is the only winning two-player coalition and singletons are not winning. By direct computation or by Step 2 of the proof of Theorem 4.4, it holds that the core of Γ is equal to the coalition structure core of (N, v) , so

$$Y = \{y \in X \mid \{1, 3\} \in \pi(y) \text{ or } [\{1, 2, 3\} \in \pi(y) \text{ and } u_2(y) = 0]\}.$$

By Theorem 4.4, the MSS of Γ is unique and equal to Y . So either coalition $\{1, 3\}$ forms or the grand coalition forms and payoffs are such that the entire surplus is shared between players 1 and 3.

We argue that the vNM stable set is unique and equal to Y as well. Let V be a vNM stable set. For every $y \in Y$ it holds that $f(y) = \{y\}$, so by external stability $Y \subseteq V$. We show that Y satisfies external stability. Let $x \notin Y$ be given. If $\{1, 2, 3\} \notin \pi(x)$, then x does not contain a winning coalition, so $u(x) = (0, 0, 0)$, and $y \in Y$ defined by $\pi(y) = \{\{1, 3\}, \{2\}\}$ and $u(y) = (1/2, 0, 1/2)$ satisfies $y \in f(x)$. If $\{1, 2, 3\} \in \pi(x)$, then $x \notin Y$ implies $u_2(x) > 0$. Now $y \in Y$ defined by $\pi(y) = \{\{1, 3\}, \{2\}\}$ and $u(y) = (u_1(x) + u_2(x)/2, 0, u_3(x) + u_2(x)/2)$ satisfies $y \in f(x)$. We have shown that the core of Γ satisfies external stability. It must therefore be the unique vNM stable set.

We now turn to the three-person veto-power game, with player 2 being the veto player. The MSS is unique and equal to the coalition structure core, so one

of the winning coalitions forms and player 2 gets the entire surplus. The MSS therefore has three elements, depending on the winning coalition that forms. We argue that there are two vNM stable sets, both having a continuum of elements and containing the MSS as a proper subset. In an element of the vNM stable set it holds that a winning coalition forms and the entire surplus is either shared between players 1 and 2 or between players 2 and 3. It is not excluded that the veto player gets a payoff of 0.

Example A.8. Assume singletons are not winning and $\{1, 2\}$ and $\{2, 3\}$ are the winning two-player coalitions. By direct computation or by Step 2 of the proof of Theorem 4.4, it holds that the core of Γ is equal to the coalition structure core of (N, v) , so to the set

$$Y = \{y \in X \mid \pi(y) \cap \mathcal{W} \neq \emptyset \text{ and } u_2(y) = 1\}.$$

There are three states in Y . One of the winning coalitions $\{1, 2\}$, $\{2, 3\}$, and $\{1, 2, 3\}$ forms and players 1 and 3 receive a payoff of 0. By Theorem 4.4, the MSS of Γ is unique and equal to Y .

We argue that there are two vNM stable sets, both having a continuum of elements and containing the MSS as a proper subset. Let V be a vNM stable set. To satisfy external stability, it must hold that $Y \subseteq V$. Since states in Y do not dominate any other state, it follows by external stability that V contains Y as a proper subset. States $x \in X$ such that $\pi(x) \cap \mathcal{W} = \emptyset$ or $\pi(x) = \{\{1, 2, 3\}\}$ do not dominate any state where a two-player winning coalition forms. It therefore follows from external stability that V contains a state $x^1 \in X \setminus Y$ such that $\{1, 2\} \in \pi(x^1)$ or $\{2, 3\} \in \pi(x^1)$. Without loss of generality, assume that $\{1, 2\} \in \pi(x^1)$. Notice that $u_1(x^1) > 0$ since $x^1 \in X \setminus Y$. We distinguish between two cases: Case 1. There is $x^2 \in V \setminus Y$ such that $\{2, 3\} \in \pi(x^2)$. Case 2. For every $x \in V \setminus Y$ it holds that $\{2, 3\} \notin \pi(x)$.

Case 1. Since $x^2 \in V \setminus Y$, it holds that $u_3(x^2) > 0$. In order to satisfy internal stability, it must hold that $u_2(x^1) = u_2(x^2)$ and therefore $u_1(x^1) = u_3(x^2)$. Internal stability implies that there cannot be any other $x \in V \setminus Y$ such that $\pi(x)$ contains $\{1, 2\}$ or $\{2, 3\}$. Now $x^3 \in X \setminus V$ such that $\pi(x^3) = \{\{1, 2\}, \{3\}\}$ and $u_2(x^1) <$

$u_2(x^3)$ is not dominated by an element of V , so V does not satisfy external stability, and we have obtained a contradiction.

Case 2. None of the states $x \in X$ such that $\pi(x) = \{\{1, 2\}, \{3\}\}$ is dominated by a state in V , so every such state must belong to V to satisfy external stability. The same applies to a state $x \in X$ such that $\pi(x) = \{\{1, 2, 3\}\}$ and $u_3(x) = 0$. We have that the set

$$V' = \{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_3(x) = 0\}$$

is a subset of V . Notice that V' contains a single element with $\{2, 3\}$ as the winning coalition, a continuum of elements with $\{1, 2\}$ as the winning coalition, and a continuum of elements with $\{1, 2, 3\}$ as the winning coalition. It is easily verified that all states in $X \setminus V'$ are dominated by an element that belongs to V' . None of the elements in V' dominate each other. We have therefore shown that

$$V = \{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_3(x) = 0\}.$$

By symmetry it follows that

$$\{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_1(x) = 0\}$$

is a vNM stable set as well. This exhausts all possibilities.

We finally turn to the three-player simple majority game. The MSS is unique and equal to the set of states such that a two-player winning coalition forms. We argue that there are four vNM stable sets, none of them being a subset of the MSS or containing the MSS as a subset. Every vNM stable set contains elements where the grand coalition forms. The union of the four vNM stable sets contains the MSS as a proper subset.

Example A.9. Assume all two-player coalitions are winning, whereas all singletons are not winning. It follows from Theorem 4.5 that the MSS is unique and equal to the set

$$F(X) = \{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{1, 3\}, \{2, 3\}\} \neq \emptyset\}.$$

In particular, it is excluded that the grand coalition forms.

Let V be a vNM stable set. Let $x \in X$ be a state such that $\pi(x) \cap \mathcal{W} = \emptyset$, or $\pi(x)$ contains a two-player winning coalition with one of the players in that coalition having a payoff of 1, or $\pi(x) = \{\{1, 2, 3\}\}$. Since x does not dominate any state in $F(X)$, it follows by external stability that V contains an element $x^1 \in F(X)$ with payoffs being strictly positive for both players in the winning coalition. Without loss of generality, assume that $\{1, 2\} \in \pi(x^1)$. We distinguish between two cases: Case 1. There is $x^2 \in V$ such that $\{1, 3\}$ or $\{2, 3\}$ belongs to $\pi(x^2)$ and $u_3(x^2) > 0$. Case 2. For every $x \in V$ it holds that if $\{1, 3\}$ or $\{2, 3\}$ belongs to $\pi(x)$, then $u_3(x) = 0$.

Case 1. Without loss of generality, assume $\{1, 3\} \in \pi(x^2)$. To satisfy internal stability, it must hold that $u_1(x^1) = u_1(x^2)$ and therefore $u_2(x^1) = u_3(x^2)$. There cannot be a state $x \in V \setminus \{x^1, x^2\}$ such that $\{1, 2\} \in \pi(x)$ or $\{1, 3\} \in \pi(x)$ and $u_1(x) < 1$ since otherwise internal stability would be violated.

Suppose, in order to derive a contradiction, that there is no state $x \in V$ with $\{2, 3\} \in \pi(x)$. If $u_1(x^1) \geq 1/2$, then the state $x \in X$ such that $\pi(x) = \{\{2, 3\}, \{1\}\}$ and $u(x) = (0, 1/2, 1/2)$ is not dominated by an element of V . If $u_1(x^1) < 1/2$, then the state $x \in X$ such that $\pi(x) = \{\{1, 2\}, \{3\}\}$ and $u(x) = (1/2, 1/2, 0)$ is not dominated by an element of V . Since V satisfies external stability, we have obtained a contradiction. Consequently, there is a state $x^3 \in V$ such that $\{2, 3\} \in \pi(x)$.

In order not to violate internal stability, it must hold that $u_2(x^3) = u_2(x^1)$ and $u_3(x^3) = u_3(x^2)$. Since $u_2(x^1) = u_3(x^2)$, this is only possible if $u(x^3) = (0, 1/2, 1/2)$. It follows that $u(x^1) = (1/2, 1/2, 0)$ and $u(x^2) = (1/2, 0, 1/2)$. We define $x^4, x^5, x^6 \in X$ by $\pi(x^4) = \pi(x^5) = \pi(x^6) = \{\{1, 2, 3\}\}$ and $u(x^4) = u(x^1)$, $u(x^5) = u(x^2)$, and $u(x^6) = u(x^3)$. It is easily verified that all states in $X \setminus \{x^1, \dots, x^6\}$ are dominated by x^1, x^2 , or x^3 . The states x^1, \dots, x^6 do not dominate each other. This yields $V = \{x^1, \dots, x^6\}$ as the unique vNM stable set satisfying the assumptions of Case 1.

Case 2. None of the states $x \in X$ such that $\pi(x) = \{\{1, 2\}, \{3\}\}$ is dominated by a state in V , so every such state must belong to V to satisfy external

stability. The same applies to states $x \in X$ such that $\pi(x) = \{\{1, 3\}, \{2\}\}$, $\pi(x) = \{\{2, 3\}, \{1\}\}$, or $\pi(x) = \{\{1, 2, 3\}\}$ and $u_3(x) = 0$. We have that the set

$$V' = \{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_3(x) = 0\}$$

is a subset of V . It is easily verified that all states in $X \setminus V'$ are dominated by an element that belongs to V' . It follows that

$$V = \{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_3(x) = 0\}.$$

We can easily check that V satisfies internal stability as well.

By symmetry it follows that

$$\{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_1(x) = 0\},$$

$$\{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_2(x) = 0\},$$

are vNM stable sets as well. This exhausts all possibilities.

A.3 Proper Simple Games and the δ -Model

Let (N, v) be a coalition function form game. The effectivity correspondence E is said to be induced by the δ -model if it satisfies coalitional sovereignty and for every $x, y \in X$, for every $S \in E(x, y)$, for every $T \in \pi(x)$ such that $T \setminus S \neq \emptyset$, it holds that $T \setminus S \in \pi(y)$. The latter condition simply expresses that residual players in some coalition stay together after coalition S leaves. Typically, it is assumed that the change in payoffs of the residual players in a given coalition has the same sign. When we restrict the analysis to proper simple games, we can obtain a characterization of the MSS without any such additional assumptions.

Let (N, v) be a coalition function form game such that v is a proper simple game with an empty core. We define the subset $F'(X)$ of X as the set of states such that its partition contains a winning coalition different from the grand coalition:

$$F'(X) = \{x \in X \mid \pi(x) \cap (\mathcal{W} \setminus \{N\}) \neq \emptyset\}.$$

It holds that $F(X) \subseteq F'(X)$, where $F(X)$ is defined in the main text. The only difference between these two sets is that $F'(X)$ does not require the non-winning coalitions in $\pi(x)$ to be singletons. In terms of payoff vectors that can be supported, there is no difference between $F(X)$ and $F'(X)$.

Theorem A.10. Let (N, v) be a coalition function form game such that v is a proper simple game with an empty core and let $\Gamma = (N, (X, d), E, (\sum_i)_{i \in N})$ be the social environment induced by the δ -model. It holds that the MSS of Γ is unique and equal to $F'(X)$.

Proof. The proof that $F'(X)$ satisfies deterrence of external deviations follows exactly the same steps as the corresponding part of the proof of Theorem 4.5. The same is true for the proof of asymptotic external stability of $F'(X)$. The only difference is that there is no need to verify that the non-winning coalitions are singletons.

We complete the proof by showing that for every $x \in F'(X)$ it holds that $f^\infty(x) = F'(X)$. Let some $x, y \in F'(X)$ be given and denote the winning coalition in $\pi(y)$ by W . We have to show that for every $\varepsilon > 0$ there exists $k' \in \mathbb{N}$ and $z \in f^{k'}(x)$ such that $z \in B_\varepsilon(y)$.

Let some $\varepsilon \in (0, 1/n)$ and $S \in \pi(y) \setminus \{W\}$ be given. By following exactly the same steps as in the proof of Theorem 4.5 it can be shown that there is $k \in \mathbb{N}$ and $x^k \in f^k(x)$ such that for every $i \in N \setminus S$, $u_i(x^k) < \varepsilon/n$ and $\pi(x^k)$ contains a winning coalition. Since $\sum_{i \in N \setminus S} u_i(x^k) < 1$, it follows that $\sum_{i \in S} u_i(x^k) > 0$ and that S has a non-empty intersection with the winning coalition in $\pi(x^k)$.

Write $\pi(y)$ as $\{S^1, \dots, S^{\ell'}\}$ with $S^1 = S$ and $S^{\ell'} = W$. For $\ell = 1, \dots, \ell' - 2$, let $x^{k+\ell} \in X$ be such that $\pi(x^{k+\ell}) = \{S^1, \dots, S^\ell, S^{\ell+1} \cup \dots \cup S^{\ell'}\}$, $u_{S^{\ell+1} \cup \dots \cup S^{\ell'}}(x^{k+\ell}) \gg u_{S^{\ell+1} \cup \dots \cup S^{\ell'}}(x^{k+\ell-1})$ and for every $i \in W$, $u_i(x^{k+\ell}) < \varepsilon/n$. In step ℓ , coalition $S^{\ell+1} \cup \dots \cup S^{\ell'}$ forms and increases the payoffs of its members, whereas the payoffs of the players in W are kept strictly below ε/n . Coalition S^ℓ becomes part of $\pi(x^{k+\ell})$ as a residual set of players. Since $\sum_{i \in S^\ell} u_i(x^{k+\ell-1}) > 0$, such a state $x^{k+\ell}$ exists.

We define the possibly empty set $W^0 = \{i \in W \mid u_i(y) \leq \varepsilon/n\}$. Let $w \in W$ be a player such that $u_w(y) \geq 1/n$. Let $z \in X$ be such that $\pi(z) = \pi(y) = \{S^1, \dots, S^{\ell'}\}$

and

$$u_j(z) = \begin{cases} \frac{\varepsilon}{n}, & j \in W^0, \\ u_j(y), & j \in W \setminus (W^0 \cup \{w\}), \\ u_w(y) - \sum_{i \in W^0} (\frac{\varepsilon}{n} - u_i(y)), & j = w. \end{cases}$$

For every $j \in W^0$ it holds that $u_j(z) = \varepsilon/n > u_j(x^{k+\ell'-2})$, for every $j \in W \setminus (W^0 \cup \{w\})$ it holds that $u_j(z) = u_j(y) > \varepsilon/n > u_j(x^{k+\ell'-2})$, and

$$u_w(z) = u_w(y) - \sum_{i \in W^0} (\frac{\varepsilon}{n} - u_i(y)) \geq \frac{1}{n} - \frac{n-2}{n} \frac{\varepsilon}{n} > \frac{\varepsilon}{n} > u_w(x^{k+\ell'-2}),$$

so $u_w(z) \gg u_w(x^{k+\ell'-2})$. It follows that $z \in f(x^{k+\ell'-2})$ and therefore $z \in f^{k+\ell'-1}(x)$. We have that $\pi(y) = \pi(z)$, for every $j \in W^0$ it holds that $|u_j(y) - u_j(z)| \leq \varepsilon/n$, for every $j \in W \setminus (W^0 \cup \{w\})$ it holds that $|u_j(y) - u_j(z)| = 0$, and $|u_w(y) - u_w(z)| \leq (n-2)\varepsilon/n$, therefore $z \in B_\varepsilon(y)$, so z has all the desired properties.

It follows by Theorem 3.9 that $F'(X)$ is a subset of the MSS and since $F'(X)$ satisfies deterrence of external deviations and asymptotic external stability, it must be equal to the MSS. \square

We now turn to the case where (N, v) is a proper simple game with a non-empty core and show that the analogue of Theorem 4.4 for the δ -model holds.

Theorem A.11. Let (N, v) be a coalition function form game such that v is a proper simple game with a non-empty core and let $\Gamma = (N, (X, d), E, (\sum_i)_{i \in N})$ be the social environment induced by the δ -model. It holds that the MSS of Γ is unique and equal to the coalition structure core Y of (N, v) .

Proof. The coalition structure core Y is the set of states such that a winning coalition forms, the other players are partitioned in arbitrary coalitions, and the veto players are the only ones with a positive payoff. Step 2 in the proof of Theorem 4.4 can be used to show that also for the δ -model the core CO of Γ is equal to Y . Since CO is closed, the remark below Theorem 3.13 implies that we only have to show that Γ satisfies the weak improvement property.

We need to show that for every $x \in X$, $f^\infty(x) \cap Y \neq \emptyset$. This is trivial if $x \in Y$. Assume $x \in X \setminus Y$. We have to show that for every $\varepsilon > 0$ there exists $k \in \mathbb{N}$,

$z \in f^k(x)$, and $y \in Y$ such that $z \in B_\varepsilon(y)$. Let some $\varepsilon \in (0, 1/n)$ be given. It holds that either $\pi(x) \cap \mathcal{W} = \emptyset$ or there is $i' \in N \setminus S^*$ such that $u_{i'}(x) > 0$.

If $\pi(x) \cap \mathcal{W} = \emptyset$, then choose a winning coalition $W \in \mathcal{W}$ and a veto player $w \in W \cap S^*$. Let $z \in X$ be such that $W \in E(x, z)$ and

$$\begin{aligned} z_j &= 0, & j &\in N \setminus W, \\ z_j &= \frac{\varepsilon}{n}, & j &\in W \setminus \{w\}, \\ z_j &= 1 - \sum_{i \in W \setminus \{w\}} \frac{\varepsilon}{n}, & j &= w. \end{aligned}$$

It holds that $z \in f(x)$. Let $y \in Y$ be such that $\pi(y) = \pi(z)$ and $u_w(y) = 1$. It holds that $z \in B_\varepsilon(y)$. This shows that z has the desired properties.

If there is $i' \in N \setminus S^*$ such that $u_{i'}(x) > 0$, then let W be the unique element in $\pi(x) \cap \mathcal{W}$. We show first that there exists $k \in \mathbb{N}$ and $z \in f^k(x)$ such that, for every $i \in N \setminus S^*$, $u_i(z) < \varepsilon$. If for every $j \in W \setminus S^*$ it holds that $u_j(x) < \varepsilon$, then take $z = x$. Otherwise, there is $j \in W \setminus S^*$ such that $u_j(x) \geq \varepsilon$. Since j is not a veto player, it holds that $N \setminus \{j\} \in \mathcal{W}$. Let $x^1 \in X$ be such that $\pi(x^1) = \{N \setminus \{j\}, \{j\}\}$, $u_{N \setminus \{j\}}(x^1) \gg u_{N \setminus \{j\}}(x)$, and, for every $i \in N \setminus \{j\}$ such that $u_i(x) < \varepsilon$, it holds that $u_i(x^1) < \varepsilon$. Since $u_j(x) \geq \varepsilon$, such an element x^1 exists. It holds that $x^1 \in f(x)$ and $u_j(x^1) = 0$. If there is $j^1 \in W \setminus S^*$ such that $u_{j^1}(x^1) \geq \varepsilon$, then we repeat this argument using j^1 . Since the set $W \setminus S^*$ is finite, we reach a state z with the desired properties in a finite number of steps. Clearly, there is $y \in Y$ with $\pi(y) = \pi(z)$ and $z \in B_\varepsilon(y)$. \square

A.4 The vNM Stable Set for the Tamura Example of the Knuth Model

Let us reconsider the graph on page 316 of Tamura (1993). There is a total of 24 matchings, denoted by M_1, \dots, M_{24} . The core of the social environment induced by the Knuth (1976) model is equal to $\text{CO} = \{M_1, M_8, M_{10}, M_{19}, M_{24}\}$.

The MSS contains 13 matchings. In addition to the matchings in the core, we obtain 8 matchings in a closed cycle and find that the MSS is equal to

$$\text{CO} \cup \{M_2, M_{16}, M_{22}, M_{12}, M_7, M_9, M_3, M_4\}.$$

There are two different vNM stable sets in this example. The first vNM stable set V_1 is given by

$$V_1 = \text{CO} \cup \{M_4, M_5, M_9, M_{12}, M_{13}, M_{16}, M_{17}, M_{20}, M_{21}\}.$$

Another vNM stable set is equal to

$$V_2 = \text{CO} \cup \{M_2, M_3, M_7, M_{11}, M_{14}, M_{17}, M_{18}, M_{22}, M_{23}\}.$$

The prediction of the vNM stable sets seems rather unappealing. First, the dominated state M_{17} is part of each vNM stable set. Second, for the largest connected subgraph of the divorce digraph, half of the states is in V_1 , while the other half is in V_2 .

A.5 Shapley-Scarf Housing Markets

Another prominent matching model is the housing matching model of Shapley and Scarf (1974). This model can be represented by a tuple $(N, H, (P_i)_{i \in N})$, where N is a finite set of individuals, H is a finite set of houses with the same cardinality as the set of individuals, and each individual $i \in N$ has a strict preference relation P_i over H . The original paper by Shapley and Scarf (1974) does not require a strict preference relation. However, as shown in Roth and Postlewaite (1977), when preferences are strict, then the strong core, i.e., the core based on weak dominance, contains a unique element. The version with strict preferences therefore became popular in the literature. Without loss of generality, we assume that $N = H$ and that the initial endowment of individual i is house i . An allocation is represented by a permutation matrix A with rows indexed by elements of N and columns indexed by elements of H . All entries of A are 0 or 1 and both rows and columns of A sum up to 1. If for some $h \in H$, for some $i \in N$, entry $A_{ih} = 1$, then house h has been assigned to individual i . Row $i \in N$ of the matrix A is denoted by A_i .

In this setting, it is convenient to define the state space X as the set of all permutation matrices A . Since X is finite, we can endow it with the discrete metric $d(A, A') = \mathbb{1}_{\{A \neq A'\}}$.

agents	1	2	3
first choice	2	3	1
second choice	3	1	2
third choice	1	2	3

Table 1: A Shapley-Scarf housing matching market with cycling.

The preferences of the individuals $(\succeq_i)_{i \in N}$ over the set X are induced by their preferences over houses in the following way. Let some individual $i \in N$ be given as well as $A, A' \in X$. Let $h, h' \in H$ be such that $A_{ih} = A'_{ih'} = 1$. Notice that h and h' are uniquely determined. It holds that $A \succ_i A'$ if and only if $h P_i h'$.

A coalition $S \in \mathcal{N}$ can arbitrarily redistribute the initial endowments of houses of its members within the coalition. More formally, the effectivity correspondence satisfies the following two conditions:

1. For every $S \in \mathcal{N}$, for every $A, A' \in X$, if $S \in E(A, A')$ then for all $i \in S$, there is $h \in S$ such that $A'_{ih} = 1$.
2. For every $S \in \mathcal{N}$, for every $A \in X$, and for every bijection $\phi : S \rightarrow S$, there exists $A' \in X$ such that for all $i \in S$, $A'_{i\phi(i)} = 1$ and $S \in E(A, A')$.

The first condition requires that if S is effective in moving from state A to state A' , then at A' the initial endowments of members of S are reallocated within S . The second condition states that every reallocation of initial endowments of houses within a coalition is feasible. Observe that the conditions impose no restrictions on how the houses of members outside the deviating coalition are reallocated, so we allow for various reallocation processes here. This completes the description of the social environment.

We show first that the MSS may contain closed cycles that do not correspond to a core element. We consider the example illustrated in Table 1. Since the initial endowments correspond to every individual's worst choice, every allocation in X

is individually rational. We argue that the set $\{A^1, A^2, A^3\}$ is a closed cycle, where

$$A^1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, A^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, A^3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

At A^1 , individual 1 obtains his first best house. Since A^1 is individually rational, coalition $\{2, 3\}$ is the only coalition that can achieve a strict improvement. The only state that dominates A^1 is therefore A^2 . By the same argument, the only state that dominates A^2 is A^3 , and the only state that dominates A^3 is A^1 . We have obtained a closed cycle that does not contain a core element. By Theorem 3.9, the core of the housing market model is a proper subset of the MSS.

We now turn to the weak dominance MSS and show that it is equal to the strong core by using the top trading cycle algorithm of Shapley and Scarf (1974).

Theorem A.12. Let $(N, H, (P_i)_{i \in N})$ be a housing matching problem and let Γ be the induced social environment. The weak dominance MSS of Γ is equal to the strong core.

Proof. Since the strong core of the housing matching problem is unique and satisfies deterrence of external deviations, we only have to show that it satisfies iterated external stability. Let $S^1, \dots, S^{k'}$ be the coalitions that are successively formed by an application of the top trading cycle algorithm of Shapley and Scarf (1974).

Consider any allocation $A \in X$ that is not equal to the strong core allocation A^* of the housing matching model. Let \tilde{f} denote the weak dominance correspondence. We generate a sequence of allocations $A^1, A^2, \dots, A^{k'}$ such that for all $k \leq k'$, $A^k \in \tilde{f}^k(A)$ and $A^{k'} = A^*$. Let $A^0 = A$ and $k = 1$. We construct the sequence in the following way.

1. If $k = k' + 1$ stop.
2. If for every $i \in S^k$ it holds that $A_i^{k-1} = A_i^*$, then we set $A^k = A^{k-1}$. Increase k by one and go back to step 1.

If there is $i \in S^k$ such that $A_i^{k-1} \neq A_i^*$, then define $T^k = \cup_{j \leq k} S^j$, and let A^k be an allocation such that $T^k \in E(A^{k-1}, A^k)$ and, for every $i \in T^k$, $A_i^k = A_i^*$. Increase k by one and go back to step 1.

We argue that for every $k = 1, \dots, k'$, $A^k \in \tilde{f}(A^{k-1})$. This is trivial if for every $i \in S^k$ it holds that $A_i^{k-1} = A_i^*$, since then $A^k = A^{k-1}$.

Let $k \in \{1, \dots, k'\}$ and $i \in S^k$ be such that $A_i^{k-1} \neq A_i^*$. By the rules of the top trading cycle algorithm, the house corresponding to A_i^* is the best house for i in the set of houses $N \setminus T^{k-1}$, so in particular it holds that $A_i^* \succ_i A_i^{k-1}$. It now follows that $A^k \in \tilde{f}(A^{k-1})$.

The proof is completed by observing that $A^{k'} = A^*$. □

In the proof of Theorem A.12, we have the union T^k of the coalitions S^1, \dots, S^k generated in the first k steps of the top trading cycle algorithm deviating in iteration k of our construction. The reason is that the assumptions on the effectivity correspondence are so weak that a deviation by S^k might upset the assignment of individuals in T^{k-1} . Under stronger assumptions on the effectivity correspondence, like those corresponding to the γ or the δ -model, it would be sufficient to have deviations by S^k in iteration k .

A.6 Mixed Environments

Let $G = (N, ((\Sigma_i, d_i), u_i)_{i \in N})$ be a finite normal-form game, so for each player $i \in N$ it holds that Σ_i is finite and $d_i(s_i, s'_i) = \mathbb{1}_{\{s_i \neq s'_i\}}$.

Let us now introduce the mixed extension $\tilde{G} = (N, ((\Delta_i, \delta_i), v_i)_{i \in N})$ of G , where Δ_i is the set of probability distributions on Σ_i . For $\sigma_i \in \Delta_i$, σ_{i, s_i} denotes the probability that player i uses pure strategy s_i . The metric δ_i on Δ_i is defined by

$$\delta_i(\sigma_i, \sigma'_i) = \max_{s_i \in \Sigma_i} |\sigma_{i, s_i} - \sigma'_{i, s_i}|.$$

We denote $\Delta = \prod_{i \in N} \Delta_i$ and endow Δ with the product metric $\delta(\sigma, \sigma') = \sum_{i \in N} \delta_i(\sigma_i, \sigma'_i)$. For a given strategy profile $\sigma \in \Delta$, we denote the probability

that pure strategy profile $s \in \Sigma$ is played by $\sigma_s = \prod_{i \in N} \sigma_{i,s_i}$. Let $v_i : \Delta \rightarrow \mathbb{R}$ be the expected utility associated to strategy profiles $\sigma \in \Delta$,

$$v_i(\sigma) = \sum_{s \in \Sigma} \sigma_s u_i(s).$$

Preferences $(\succeq_i)_{i \in N}$ are such that $\sigma \succeq_i \sigma'$ if and only if $v_i(\sigma) \geq v_i(\sigma')$. The social environment $\tilde{\Gamma} = (N, (\Delta, \delta), E, (\succeq_i)_{i \in N})$ corresponds to the game \tilde{G} where E only allows singletons to deviate and $\{i\} \in E(\sigma, \sigma')$ if and only if $\sigma_{-i} = \sigma'_{-i}$.

A strategy profile $\sigma \in \Delta$ is said to be a mixed strategy Nash equilibrium of G if it is a pure strategy Nash equilibrium of \tilde{G} . The core of $\tilde{\Gamma}$ coincides with the set of mixed strategy Nash equilibria of G . Additionally, note that the expected utility functions $(v_i)_{i \in N}$ are continuous on Δ and that E is lower hemi-continuous. As such, Theorems 3.7 and 3.13 give the following result.

Corollary A.13. Let \tilde{G} be the mixed extension of the finite normal-form game G and let $\tilde{\Gamma}$ be the social environment corresponding to \tilde{G} . The MSS of $\tilde{\Gamma}$ coincides with the set of mixed strategy Nash equilibria of G if and only if $\tilde{\Gamma}$ satisfies the weak improvement property.

Clearly, the pure strategy Nash equilibria of G are also mixed strategy Nash equilibria of G , so belong to the MSS of $\tilde{\Gamma}$. On the other hand, it is easy to find examples such that some profiles in the MSS of Γ are not in the MSS of $\tilde{\Gamma}$.

A finite two-player game $G = (N, ((\Sigma_i, d_i), u_i)_{i \in \{1,2\}})$ is zero-sum if for all strategy profiles $s \in \Sigma$, $u_1(s) + u_2(s) = 0$. The following result shows that for such games the MSS of $\tilde{\Gamma}$ coincides with the set of mixed strategy Nash equilibria of G .

Theorem A.14. Let \tilde{G} be the mixed extension of a finite two-player zero-sum game G and let $\tilde{\Gamma}$ be the social environment corresponding to \tilde{G} . Then the MSS of $\tilde{\Gamma}$ coincides with the set of mixed strategy Nash equilibria of G .

Proof. Using Corollary A.13, it remains to show that $\tilde{\Gamma}$ satisfies the weak improvement property, i.e., for every strategy profile $\sigma \in \Delta$, $f^\infty(\sigma)$ contains a mixed strategy Nash equilibrium of G . Let v denote the value of the game.

Let some $\sigma \in \Delta$ be given which is not a mixed strategy Nash equilibrium of G , i.e., there is a player i such that σ_i is not a minmax strategy. We distinguish between two cases.

Case 1: σ_1 and σ_2 are not minmax strategies.

1.1 If $v_1(\sigma) \neq v$, then there exists a player i who is below his minmax payoff.

Without loss of generality, let this be player 1, so $v_1(\sigma) < v$. Let (σ_1^*, σ_2^*) be a profile of minmax strategies. Note that $v_1(\sigma_1^*, \sigma_2) \geq v$. Since σ_2 is not a minmax strategy, there exists a pure strategy $s_1 \in \Delta_1$ such that $v_1(s_1, \sigma_2) > v$. Thus, for every $\varepsilon \in (0, 2]$, it holds that

$$v_1\left(\frac{\varepsilon}{2}s_1 + \left(1 - \frac{\varepsilon}{2}\right)\sigma_1^*, \sigma_2\right) > v.$$

It holds that

$$v_2\left(\frac{\varepsilon}{2}s_1 + \left(1 - \frac{\varepsilon}{2}\right)\sigma_1^*, \sigma_2^*\right) \geq -v,$$

so for every $\varepsilon > 0$, $f^2(\sigma)$ contains a state which is in an ε -neighborhood of a mixed strategy Nash equilibrium of G , and therefore $f^\infty(\sigma)$ contains a mixed strategy Nash equilibrium of G .

1.2 Suppose $v_1(\sigma) = v$. Then, there exists a pure strategy $s_1 \in \Delta_1$ such that

$$v_1(s_1, \sigma_2) > v,$$

since otherwise σ_2 would be a minmax strategy. If s_1 is a minmax strategy, then player 2 can deviate to a minmax strategy σ_2^* to obtain $v_2(s_1, \sigma_2^*) = -v$, i.e., $f^2(\sigma)$ contains a mixed strategy Nash equilibrium of G . If s_1 is not a minmax strategy, then $(s_1, \sigma_2) \in f_1(\sigma)$ is a state as in Case 1.1, so for every $\varepsilon > 0$, $f^3(\sigma)$ contains a state which is in a ε -neighborhood of a mixed strategy Nash equilibrium of G , and therefore $f^\infty(\sigma)$ contains a mixed strategy Nash equilibrium of G .

Case 2: σ_1 is a minmax strategy and σ_2 is not, or σ_1 is not a minmax strategy and σ_2 is. Without loss of generality, assume σ_1 is a minmax strategy.

2.1 If $v_1(\sigma) > v$, then player 2 can profitably switch to a minmax strategy σ_2^* and we are done.

2.2 If $v_1(\sigma) = v$, then since σ_2 is not a minmax strategy, there exists a deviation to a pure strategy $s_1 \in \Delta_1$ such that $v_1(s_1, \sigma_2) > v$. If s_1 is a minmax strategy, then $(s_1, \sigma_2) \in f_1(\sigma)$ is a state as in Case 2.1, so $f_2(\sigma)$ contains a mixed strategy Nash equilibrium of G . If s_1 is not a minmax strategy, then $(s_1, \sigma_2) \in f_1(\sigma)$ is a state as in Case 1.1, and for every $\varepsilon > 0$ it holds that $f^3(\sigma)$ contains a state which is in an ε -neighborhood of a mixed strategy Nash equilibrium of G , so $f^\infty(\sigma)$ contains a mixed strategy Nash equilibrium of G .

□

As a final result, we show the equivalence between the set of mixed strategy Nash equilibria of G and the MSS of the social environment $\tilde{\Gamma}$ for finite two-player games where one of the two players has two pure strategies.

Theorem A.15. Let \tilde{G} be the mixed-extension of a finite two-player game G and let $\tilde{\Gamma}$ be the social environment corresponding to \tilde{G} . Assume that one player has two pure strategies in G . Then the MSS of $\tilde{\Gamma}$ coincides with the set of mixed strategy Nash equilibria of G .

Proof. Assume without loss of generality that player 1 has two pure strategies. Let the set of pure strategies of player 1 be $\{U, D\}$ with generic element $A \in \{U, D\}$ and let the set of pure strategies of player 2 be given by $\{s^1, \dots, s^\ell\}$ with generic element s^j . We also use the notation U and D for the mixed strategy that puts probability 1 on pure strategy U and D , respectively, and similarly for s^j .

Let some $\sigma \in \Delta$ be given. By Corollary A.13, it suffices to show the weak improvement property of $\tilde{\Gamma}$, i.e., $f^\infty(\sigma)$ contains a mixed strategy Nash equilibrium of G . We distinguish between two cases.

Case 1: G has a pure strategy Nash equilibrium, without loss of generality, (U, s^*) .

If σ is a mixed strategy Nash equilibrium of G , we are done, so assume σ is not a mixed strategy Nash equilibrium of G . If player 2 has a profitable deviation from

σ , then there is a pure strategy best response $s^j \in \Delta_2$ such that $(\sigma_1, s^j) \in f(\sigma)$. If (σ_1, s^j) is a mixed strategy Nash equilibrium of G , we are done. If not, then player 1 must have a pure strategy best response to (σ_1, s^j) , say A . Thus, $f^2(\sigma)$ contains a mixed strategy Nash equilibrium of G or a pure strategy profile (A, s^j) . The same conclusion holds if player 1 has a profitable deviation from σ . If the pure strategy profile (A, s^j) is a Nash equilibrium of G , we are done. If not, at least one player has a profitable deviation from it. We distinguish between two cases.

1.1 $A = D$.

1.1.a Assume player 1 can profitably deviate from (D, s^j) . Then it holds that $(U, s^j) \in f(D, s^j)$. If (U, s^j) is a Nash equilibrium of G , we are done. If not, then player 2 can profitably deviate to the Nash equilibrium (U, s^*) of G and we are done.

1.1.b Assume player 2 can profitably deviate from (D, s^j) . Let s^h be a best response for player 2, so $(D, s^h) \in f(D, s^j)$. If this is a Nash equilibrium of G , we are done. Otherwise, player 1 can profitably deviate to (D, s^h) , which brings us back to Case 1.1.a.

1.2 $A = U$.

1.2.a Assume player 2 can profitably deviate from (U, s^j) . It holds that the Nash equilibrium (U, s^*) of G belongs to $f(U, s^j)$, so we are done.

1.2.b Assume player 1 can profitably deviate from (U, s^j) . Then it holds that $(D, s^j) \in f(U, s^j)$. If (D, s^j) is a Nash equilibrium of G , then we are done. Else, player 2 must have a profitable deviation from (D, s^j) , which brings us back to Case 1.1.b.

Case 2: G has no pure strategy Nash equilibrium.

We first show that in every mixed strategy Nash equilibrium of G , player 1 plays both U and D with strictly positive probability. Towards a contradiction, suppose there is a mixed strategy Nash equilibrium (A, σ_2^*) of G such that player 1 plays a pure strategy, without loss of generality, strategy $A = U$. It holds that any

pure strategy of player 2 in the support of σ_2^* is a best response against U . Since G has no pure strategy Nash equilibrium, it must hold that playing D against any pure strategy in the support of σ_2^* gives player 1 a strictly higher payoff than playing U . It follows that D is a profitable deviation for player 1 from (U, σ_2^*) . This contradicts (U, σ_2^*) being a mixed strategy Nash equilibrium of G .

To finish the proof, we show that $f^\infty(\sigma)$ contains a mixed strategy Nash equilibrium of G . As in the first part of Case 1, we can show that $f^2(\sigma)$ contains a mixed strategy Nash equilibrium of G and we are done, or a pure strategy profile which is not a Nash equilibrium of G . Player 1 or player 2 has a profitable deviation from this pure strategy profile. In the latter case, player 2 can choose a pure strategy best response and in the next step, player 1 can profitably deviate to a pure strategy. In both cases it holds that there is $k \in \mathbb{N}$ such that $f^k(\sigma)$ contains a pure strategy profile (A, s^j) from which player 1 has a profitable deviation. Without loss of generality, let $A = U$.

Observe that for player 1 any completely mixed strategy is a profitable deviation from (U, s^j) . Let σ^* be a mixed strategy Nash equilibrium of G and let $p \in (0, 1)$ denote the probability that σ_1^* puts on U . We distinguish 3 cases.

$$2.1 \quad v_2(D, \sigma_2^*) - v_2(U, \sigma_2^*) > v_2(D, s^j) - v_2(U, s^j).$$

For $\varepsilon \in (0, p)$, let σ_1' be the strategy where player 1 plays U with probability $p - \varepsilon/2$. Since any completely mixed strategy of player 1 is a profitable deviation from (U, s^j) , it holds that $(\sigma_1', s^j) \in f(U, s^j)$. We have that

$$\begin{aligned} v_2(\sigma_1', s^j) &= v_2(\sigma_1^*, s^j) + \frac{\varepsilon}{2}(v_2(D, s^j) - v_2(U, s^j)) \\ &< v_2(\sigma_1^*) + \frac{\varepsilon}{2}(v_2(D, \sigma_2^*) - v_2(U, \sigma_2^*)) \\ &= v_2(\sigma_1', \sigma_2^*), \end{aligned}$$

where the strict inequality uses that σ_2^* is a best response against σ_1^* and the assumption of Case 2.1. It follows that $(\sigma_1', \sigma_2^*) \in f(\sigma_1', s^j)$. Since $\varepsilon > 0$ can be chosen arbitrarily small, this shows that $\sigma^* \in f^\infty(\sigma)$.

$$2.2 \quad v_2(D, \sigma_2^*) - v_2(U, \sigma_2^*) < v_2(D, s^j) - v_2(U, s^j).$$

For $\varepsilon \in (0, 1 - p)$, let σ_1' be the strategy where player 1 plays U with probability $p + \varepsilon/2$. The proof now follows as in Case 2.1.

$$2.3 \quad v_2(D, \sigma_2^*) - v_2(U, \sigma_2^*) = v_2(D, s^j) - v_2(U, s^j).$$

It holds that $(D, s^j) \in f(U, s^j)$.

Let s^h be a best response of player 2 against D and, for $\varepsilon \in (0, 1)$, let σ_2' be the strategy that puts weight $(1 - \varepsilon)$ on σ_2^* and weight ε on s^h . We have that

$$\begin{aligned} v_2(D, \sigma_2') &= v_2(\sigma_2^*) + \varepsilon v_2(D, s^h) - \varepsilon v_2(U, \sigma_2^*) \\ &\geq v_2(\sigma_2^*) + \varepsilon v_2(D, s^j) - \varepsilon v_2(U, \sigma_2^*) \\ &= v_2(D, s^j), \end{aligned} \tag{A.1}$$

where the inequality uses that σ_2^* is a mixed strategy Nash equilibrium of G and the assumption of Case 2.3. Since (D, s^j) is not a Nash equilibrium of G , it holds that $v_2(D, s^h) > v_2(D, s^j)$. By (A.1) and the definition of σ_2' , it now follows that $v_2(D, \sigma_2') > v_2(D, s^j)$, so $(D, \sigma_2') \in f(D, s^j)$. Since (D, s^h) is not a Nash equilibrium of G and s^h is a best response against D , we have that $v_1(\sigma_1^*, s^h) > v_1(D, s^h)$. It follows that

$$v_1(\sigma_1^*, \sigma_2') = (1 - \varepsilon)v_1(\sigma_1^*) + \varepsilon v_1(\sigma_1^*, s^h) > (1 - \varepsilon)v_1(D, \sigma_2^*) + \varepsilon v_1(D, s^h) = v_1(D, \sigma_2'),$$

so $(\sigma_1^*, \sigma_2') \in f(D, \sigma_2')$. Since $\varepsilon > 0$ can be chosen arbitrarily small, we have that $\sigma^* \in f^\infty(\sigma)$, which concludes the proof. □

We analyzed the game of matching pennies in Example A.4 and concluded that better-response dynamics did not single out any strategy profile. The game of matching pennies satisfies the assumptions of both Theorems A.14 and A.15. The MSS of this game therefore consists of the unique mixed strategy Nash equilibrium where each pure strategy is played with probability $1/2$.

References

KNUTH, D. E. (1976): *Marriages Stables*, Les Presses de l'Université de Montréal, Montréal.

MEYN, S. P., AND R. L. TWEEDIE (1993): *Markov Chains and Stochastic Stability*, Springer-Verlag, Berlin.

ROTH, A. E., AND A. POSTLEWAITE (1977): “Weak Versus Strong Domination in a Market with Indivisible Goods”, *Journal of Mathematical Economics*, 4, 131–137.

SHAPLEY, L. S. AND H. SCARF (1974): “On Cores and Indivisibility,” *Journal of Mathematical Economics* 1, 23–37.

TAMURA, A. (1993): “Transformation from Arbitrary Matchings to Stable Matchings,” *Journal of Combinatorial Theory, Series A*, 62, 310–323.