

Where strategic and evolutionary stability depart – a study of minimal diversity games

DIETER BALKENBORG ^{*} DRIES VERMEULEN ^{†‡}

November 14, 2014

Abstract

A minimal diversity game is an n player strategic form game in which each player has m pure strategies at his disposal. The payoff to each player is always 1, unless all players select the same pure strategy, in which case all players receive zero payoff. Such a game has a unique isolated completely mixed Nash equilibrium in which each player plays each strategy with equal probability, and a connected component of Nash equilibria consisting of those strategy profiles in which each player receives payoff 1. The Pareto superior component is shown to be asymptotically stable under a wide class of evolutionary dynamics, while the isolated equilibrium is not. On the other hand, the isolated equilibrium is strategically stable, while the strategic stability of the Pareto efficient component depends on the dimension of the component, and hence on the number of players, and the number of pure strategies.

JEL Codes. C72, D44.

Keywords. Strategic form games, strategic stability, evolutionary stability.

1 Introduction

Consider a cruise ship with n passengers on board. The ship has hit an iceberg, is sinking rapidly, and only has two lifeboats L and R on board. Each passenger has to decide which lifeboat to run to, without any prior knowledge on the decision of the other passengers. On board of each lifeboat there is only room for $n - 1$ passengers, so not all passengers can be saved by the same boat. If at most $n - 1$ passengers choose to run to the same boat, all passengers

^{*}Department of Economics, University of Exeter, Streatham Court, Exeter EX4 4PU, UK. Email: d.g.balkenborg@exeter.ac.uk

[†]Corresponding author. Department of Quantitative Economics, University Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Email: d.vermeulen@maastrichtuniversity.nl

[‡]We thank two anonymous referees for their comments that helped us to improve the presentation of the paper considerably. Also, apart from expressing our gratitude for the lively discussions we had with participants at various conferences when presenting our results, we specifically thank Josef Hofbauer, Jeroen Swinkels, Fabrizio Germano, and Peter Ashwin for many inspiring conversations. We thank Stefano Demichelis for sharing his views on this topic during his stay at Maastricht University.

get saved and the payoff to each passenger is one. When all passengers decide to run to the same boat all passengers drown and the resulting payoff in that case is zero. The question is: How should the passengers choose which boat to run to? Clearly, all that is needed is that not all passengers choose to run to the same boat.

The game thus defined has two types of Nash equilibria. In one there is complete lack of coordination, and each passenger chooses to run to either lifeboat with equal probability. Hence, in this equilibrium there is a positive probability that all passengers run to the same boat, in which case they all drown and have payoff zero. In the second type of equilibrium the coordination problem is solved. Two passengers choose different boats to run to, one to lifeboat L , the other to lifeboat R , while the remaining passengers can, and may, run to whichever boat they want.

The two types of Nash equilibria are very different. Can one convincingly argue that one is more reasonable than the other? In this paper we compare two approaches to answer this question. We show that, from the evolutionary perspective, the equilibria where the coordination problem is solved are most natural. We show that they form a set that is asymptotically stable under a wide class of evolutionary dynamics, while the equilibrium where the coordination problem is not solved is always unstable. We then take a look at the above choice problem from the perspective of strategic stability. For the more demanding notions of strategic stability—essentiality as defined by Wu Wen-Tsün and Jiang Jia-He ([16],[30]), best response stability as defined by Hillas ([11],[12]), and strategic stability in the sense of Mertens ([21],[22])—it turns out that the total number of passengers starts to matter. Surprisingly, the question is not whether the number of passengers is large or small, but whether it is even or odd. When the number of passengers is even, the set of Nash equilibria where the coordination problem is solved contains a strategically stable set, otherwise it does not. The Nash equilibrium where the coordination problem is not solved is always strategically stable, in virtually any sense.

The main purpose of the paper is to substantiate these claims. We think these findings are interesting for at least two reasons. First of all, it provides a whole class of very simple games where the two approaches of evolutionary stability and strategic stability make mutually exclusive predictions, namely whenever the number of players is odd. This contrasts with earlier findings, notably by Swinkels ([25] [26]), and Demichelis and Ritzberger [4], who show that—under a topological restriction on the set of Nash equilibria in question—strategic stability is implied by evolutionary stability. Our findings in case the number of players is odd show that

the topological restriction is indispensable for such results.¹

Secondly, the alternating behavior of strategic stability solutions is puzzling. Although our proofs show that continuity arguments, based on asymmetric payoff perturbations, do provide a basic intuition for our findings, continuity is explicitly ruled out by Kohlberg and Mertens [17] as a reasonable axiom for solution concepts. Thus, our results emphasize the need for game theoretic arguments to justify the alternating behavior we find for the more demanding notions of strategic stability.

Refinement theory became popular when, with the increased use of game theoretic methods in Economics in the 1970's and early 1980's, it became clear that the notion of a Nash equilibrium is too weak as a solution concept to analyze relevant economic models. Refinements such as perfect, proper and sequential equilibrium were introduced and widely used. However, as it turned out, also refinements of Nash equilibrium allowed for many possible solutions, including very implausible ones. Moreover, refinements were often only motivated and introduced on an ad-hoc, example-driven basis. This encouraged game theorists to try to find a more systematic and consistent way to further refine among Nash equilibria. One influential approach in this context is the theory of equilibrium selection, initiated by Harsanyi and Selten [9], which has the ambitious aim to select a unique Nash equilibrium for every game.

A second influential approach to refinement theory is due to Kohlberg and Mertens [17]. In contrast to Harsanyi and Selten, Kohlberg and Mertens [17], henceforth called K&M, argue that any satisfactory selection criterion cannot select a single equilibrium, but forces us to select *sets* of closely related equilibria. Their second major contribution to the ongoing discussion was to produce a list of basic criteria such refined sets of equilibria should satisfy. K&M's point of view was that, while a full-fledged axiomatic approach may be out of reach, the search for a satisfactory solution concept should be guided at least by a list of properties that are desirable from a game theoretical perspective. Their own initial notion of strategically stable sets of equilibria violates some of these criteria. However, Mertens [21] proposes a notion of strategic stability that satisfies all the requirements made by K&M and that moreover passes several additional plausibility tests.

On the other side of the spectrum, earlier experimental research—e.g. by McKelvey and Palfrey [20] on the centipede game—made it clear that classical game theory, in particular Nash equilibrium, is not always a good predictor of human behavior. These results inspired a new

¹In fact Demichelis and Ritzberger [4] already have a remark to this extent in the context of a 4-person game, for the case of strategic stability in the sense of Mertens.

line of research in game theory that focused on models of bounded rationality, in an attempt to avoid the traditional game-theoretic approach to take the assumption that players are rational to its ultimate conclusions. One such attempt uses learning dynamics to model boundedly rational behavior, in line with earlier work by evolutionary biologists on animal behavior.

In their seminal work Maynard Smith and Price [19] showed that under certain conditions such learning dynamics converge to Nash equilibrium. Nash equilibria that are selected via learning dynamics tend to be rather “refined” Nash equilibria. This raises the question whether the outcomes predicted by the learning approach assuming only bounded rationality² might coincide with the outcomes given by the refinement approach, based on the extreme emphasis on rationality.

Indeed, the results of Swinkels [25] and Demichelis and Ritzberger [4] imply that any asymptotically stable Nash equilibrium is automatically strategically stable. However, as K&M emphasized, in general one should look for *sets* of Nash equilibria. Swinkels [26] noted that the analysis for sets of Nash equilibria is complicated by the fact that topological properties of the sets start to matter. He showed for a wide class of dynamics that asymptotically stable sets of the dynamics contain a strategically stable (even a hyper-stable) set as defined by K&M, *provided* that the set is contractible. D&R considerably sharpen this result. They show for a wide class of dynamics that sets that are asymptotically stable under the dynamics contain a stable set of Nash equilibria in the sense of Mertens, provided the set has non-zero Euler characteristic.

While these results show a strong connection between evolutionary and strategic stability under certain topological restrictions on the solution set, it does not rule out the existence of convincing examples where the two types of stability are not related. Indeed Demichelis and Ritzberger [4] give an example of a game where strategic stability in the sense of Mertens does not agree with evolutionary stability. The aim of this paper is to provide a large class of games where this phenomenon occurs, notably for many of the stronger types of strategic stability such as essentiality and best response stability.³

More concretely, we present a natural class of coordination games, called minimal diversity games. Minimal diversity games are coordination games with a slight twist. In classical coordination games, all players have to choose the same action in order to coordinate. Here the

²In the literature on bounded rationality evolutionary models are re-interpreted as learning models.

³Which implies the same result for even stronger variants such as hyperstability and strategic stability in the sense of Mertens.

task faced by the players is precisely to *avoid* choosing the same action. Each game within this class has a set of Nash equilibria that consists of one isolated completely mixed Nash equilibrium and one connected component of Pareto dominant Nash equilibria.⁴ For two families of games within the class of minimal diversity games⁵ we show that any sufficiently strong notion of strategic stability exclusively selects the isolated completely mixed equilibrium, while evolutionary stability exclusively selects the component of Pareto dominant Nash equilibria. Thus, minimal diversity games provide such examples of games where the predictions made by evolutionary stability and strategic stability are different, even in the strong sense that they may be mutually exclusive.

Moreover, and more interestingly, we establish an intriguing alternating pattern in the strategic stability of the Pareto dominant Nash component. It was already known in several examples, such as Example 2 in Demichelis and Ritzberger [4] (an example originally presented in Hofbauer and Swinkels [15]), that strategic and evolutionary stability might make different predictions.

However, as is shown in section 3, for any minimal diversity game with n players and m strategies the set of efficient Nash equilibria is a connected component of Nash equilibria of dimension $d = (n-1)(m-1) - 1$. Our results show that for the two classes of games we consider, 2 players and m actions, and n players and 2 actions, this component of efficient Nash equilibria is strategically stable precisely when d is even. We conjecture, that this phenomenon is not merely “accidental”, but true in general: for any minimal diversity game strategic stability and evolutionary stability make mutually exclusive predictions precisely when the dimension d of the set of efficient Nash equilibria is odd. This would be entirely in line with the conjecture of Swinkels that the topology of the component of Nash equilibria in question, specifically its Euler characteristic, is the decisive factor in the connection (and the distinction!) between the two types of stability.

In the light of the efforts of Kohlberg and Mertens to justify strategic stability requirements through purely game theoretic desiderata this is a disconcerting observation. It is far from easy to think of such a purely game theoretic desideratum that would force us to accept dimensionality arguments, or more generally purely geometric characteristics, to be decisive

⁴The only exception are those games where the number of players is 2 and the number of pure strategies for each player also equals 2. In this case the set of Pareto dominant Nash equilibria consists of two isolated points (that is, a sphere of dimension zero).

⁵In one family there are two players with an odd number of strategies, and in the other there are two strategies and an odd number of players.

factors for strategic stability. A concept of strategic stability for which this sort of peculiar behavior based on geometric characteristics does not arise would have to be defined via a radically different approach.

In relation to these last few observations it is also interesting to see that, in those cases where the efficient set is not strategically stable, we can also show that it is not essential in the sense of Wu Wen-Tsun and Jiang Jia-He [16], [30]. This implies that many evolutionary dynamics never converge to any equilibrium.⁶ In this sense strategic stability matters for evolutionary stability even when at first sight it seems to be irrelevant.

K&M already stressed the necessity to discard the focus on generic classes of strategic form games, but rather study natural subclasses within the class of strategic form games. K&M made this observation in the context of extensive form games, where the class of games considered was defined by fixing the game tree of an extensive form game, and by varying the payoffs of players at the terminal nodes of the game tree. In a similar vein our class of minimal diversity games is a natural class of potential games, very much in the spirit of games such as for example Colonel Blotto games.⁷ In the class of potential games minimal diversity games are characterized only by the requirement that players, similar to the task of players in Colonel Blotto games, need to “avoid coordination” in order to maximize their joint payoff.

The paper is organized as follows. After some preliminaries we introduce the class of minimal diversity games. We show that the set of Nash equilibria of each minimal diversity game consists of an isolated completely mixed equilibrium and a connected component G of efficient equilibria of dimension $d = (n - 1) \cdot (m - 1) - 1$, where n is the number of players and m the number of strategies. We then show for a wide class of dynamics (including the replicator dynamics as well as the class of Nash dynamics defined by D&R) that the set G is asymptotically stable, while the completely mixed equilibrium is not. Thereafter we turn to strategic stability where we first observe that all equilibrium components of these games are strategically stable in the sense of Kohlberg and Mertens. We then derive the results indicated above, first for games with two players and an arbitrary number of strategies, and next for games with two strategies and an arbitrary number of players.

⁶This finding strengthens the observation in D&R that these sets of Nash equilibria are not robustly evolutionary stable. Notice, however, that there are still asymptotically stable attractors not containing fixed points in the nearby games.

⁷See among others Hart [10], and Monderer and Shapley [23] for motivation and explanation of these games.

2 Preliminaries

This paper concerns the class of minimal diversity games. However, many concepts applied here are defined for general normal form games. It will hence be useful to have some basic notation and terminology available for these games.

A finite normal form game consists of a finite set of players $N = \{1, \dots, n\}$, and for each player $i \in N$ a finite pure strategy set S_i and a payoff function $u_i: S \rightarrow \mathbb{R}$ on the set $S := \prod_{i \in N} S_i$ of pure strategy profiles. We denote the game by (S, u) , where $u = (u_i)_{i \in N}$ is the vector of payoff functions. A mixed strategy σ_i of player i is a vector $(\sigma_i(s_i))_{s_i \in S_i}$ that assigns a probability $\sigma_i(s_i) \geq 0$ to each pure strategy $s_i \in S_i$. We denote the set of mixed strategies of player i by Σ_i . The set of all profiles $\sigma = (\sigma_i)_{i \in N}$ of mixed strategies is denoted by Σ . The support of a mixed strategy σ_i is the set of all pure strategies s_i with $\sigma_i(s_i) > 0$. The multilinear extension of the payoff function u_i of player i to the set Σ of all strategy profiles is given by the formula

$$u_i(\sigma) = \sum_{s \in S} \prod_{j \in N} \sigma_j(s_j) u_i(s).$$

By $u_i(\sigma \mid s_i)$ we denote the payoff to player i when player i plays pure strategy $s_i \in S_i$ while his opponents adhere to the mixed strategy profile σ . A strategy profile $\sigma \in \Sigma$ is a Nash equilibrium when $u_i(\sigma) \geq u_i(\sigma \mid s_i)$ holds for every player i and every pure strategy s_i of player i .

MINIMAL DIVERSITY GAMES A minimal diversity game is a normal form game (S, u) such that $S_i = M = \{1, \dots, m\}$ for every player $i \in N$, and

$$u_i(s_1, \dots, s_n) = \begin{cases} 0 & \text{if } s_1 = s_2 = \dots = s_n \\ 1 & \text{else.} \end{cases}$$

To simplify notation for minimal diversity games, for player i and pure strategy k , we denote the probability $\sigma_i(k)$ that player i assigns to pure strategy k in strategy profile σ by σ_{ik} .

STRATEGIC STABILITY In this paper we use several different notions of strategic stability, notably strong stability, defined by Kojima, Okada and Shindoh [18], regularity, defined by Harsanyi [8], essentiality, defined by Wu Wen-Tsun and Jiang Jia-He [16], [30], KM stability, hyperstability and full stability, defined by K&M, strategic stability in the sense of Mertens, defined by Mertens [21], and best response stability, defined by Hillas [11]. KM stability was simply called stability in K&M, and the same holds for best response stability in Hillas [11]. We use the terms KM stability and best response stability in this paper to avoid confusion.

We do not define most of these notions, because we rely on the results of D&R and Hillas et

al. [12] for most of our conclusions. We explicitly use the definitions of essentiality and strong stability though, and for that reason we state here their formal definitions. For information on the remaining notions of strategic stability we refer the interested reader to the original papers.

For two games (S, u) and (S, v) with the same player set and the same set S of strategy profiles, we write

$$\|u - v\| = \max \{|u_i(s) - v_i(s)| \mid s \in S, i \in N\}.$$

A closed set $C \subset \Sigma$ of Nash equilibria of the game (S, u) is called essential when for every open set $U \subset \prod_i \mathbb{R}^{S_i}$ containing C there is a $\varepsilon > 0$ such that every game (S, v) with $\|u - v\| < \varepsilon$ has a Nash equilibrium in U .

A Nash equilibrium σ of (S, u) is called strongly stable (Kojima, Okada and Shindoh [18]) if there exist open sets $U \ni u$ and $V \ni \sigma$ and a continuous function $g: U \rightarrow V$ such that for any $v \in U$ and $\tau \in V$ we have $g(v) = \tau$ precisely when τ is a Nash equilibrium of (S, v) .

Notice that the definition automatically implies $g(u) = \sigma$. A strongly stable equilibrium is strategically stable in virtually any sense: it is perfect, proper, essential, best response stable, and even strategically stable in the sense of Mertens.

3 The Nash equilibria of a minimal diversity game

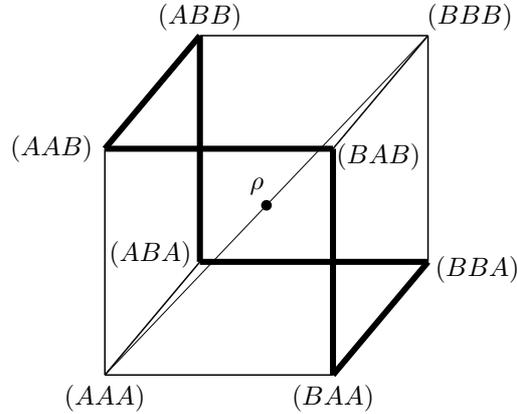
In this section we show that the set of Nash equilibria of a minimal diversity game consists of two components. One component consists of a single isolated completely mixed Nash equilibrium. The other component has dimension $d = (n - 1) \cdot (m - 1) - 1$, and consists of all strategy profiles in which at least two players play a different pure strategy.

It is easy to verify that the completely mixed strategy profile in which each player randomizes with equal probability $\frac{1}{m}$ between all his pure strategies is a Nash equilibrium with an expected payoff of $1 - \left(\frac{1}{m}\right)^{n-1}$ for each player. We denote this Nash equilibrium by $\rho = (\rho_i)_{i \in N}$.

Secondly, every strategy profile where one (and hence every) player gets the maximal expected payoff one is clearly a Nash equilibrium. We denote the set of all such Nash equilibria by G . Note that G is in fact the set of all Pareto-efficient strategy profiles in the game.

To illustrate the set G we consider one of the simplest minimal diversity games, namely the game with three players and two strategies each. In this case each player's mixed strategy space is a line segment, so that the set of strategy profiles can be identified with a cube, as in

the following graph.



The pure strategy profiles are the corners of the cube. The Nash equilibrium ρ in completely mixed strategies is in the center of the cube. The set G of Pareto efficient Nash equilibria is the cycle on the boundary consisting of six line segments. Along each line segment each pure strategy is not used by one of the players. More generally, we show the following two facts.

- (1) ρ and the strategy profiles in G are the only Nash equilibria of a minimal diversity game.
- (2) G is a connected component of Nash equilibria that is homeomorphic to a sphere of dimension $d = (n - 1)(m - 1) - 1$.

Consequently, ρ is an isolated Nash equilibrium and G is a connected component of Nash equilibria.⁸

Proposition 3.1 *The Nash equilibria of a minimal diversity game are precisely the strategy profiles in G , together with the completely mixed strategy profile ρ .*

Proof. Let σ be a Nash equilibrium of the game in completely mixed strategies. For each $k \in M$, write $\sigma_k = \prod_{i \in N} \sigma_{ik}$. Then, when $\sigma_{ik} > 0$, $\prod_{j \neq i} \sigma_{jk} = \frac{\sigma_k}{\sigma_{ik}}$. So, since the game is a minimal diversity game, in case $\sigma_{ik} > 0$ the expected payoff of player i from playing pure strategy k at profile σ is

$$1 - \prod_{j \neq i} \sigma_{jk} = 1 - \frac{\sigma_k}{\sigma_{ik}}.$$

Therefore, since in a completely mixed equilibrium each player is indifferent between all his

⁸There is one exception, namely when $m = n = 2$. In this case G consists of two isolated Nash equilibria, and hence it is homeomorphic to the two endpoints of the unit interval.

strategies in M we have

$$1 - \frac{\sigma_k}{\sigma_{ik}} = 1 - \frac{\sigma_l}{\sigma_{il}}$$

for all players $i \in N$ and pure strategies $k, l \in M$. Therefore, for two given players i and j we have

$$\frac{\sigma_{il}}{\sigma_{ik}} = \frac{\sigma_l}{\sigma_k} = \frac{\sigma_{jl}}{\sigma_{jk}}.$$

Hence $\sigma_{ik} = \frac{\sigma_{il}}{\sigma_{jl}} \sigma_{jk}$ for all $k, l \in M$. Summing over k we obtain $1 = \frac{\sigma_{il}}{\sigma_{jl}}$. Hence $\sigma_{il} = \sigma_{jl}$ for every $l \in M$, which shows that all players use the same strategy. Therefore $(\sigma_{ik})^n = \prod_{j \in N} \sigma_{jk} = \sigma_k$ for every $i \in N$ and $k \in M$. Hence, since $\frac{\sigma_l}{\sigma_{il}} = \frac{\sigma_k}{\sigma_{ik}}$, it follows that $\sigma_{il} = \sigma_{ik}$ for every $i \in N$ and $k \in M$. Thus each player uses each pure strategy in M with equal probability and we conclude that $\sigma = \rho$.

Suppose next that σ is a Nash equilibrium, and that $\sigma_{ik} = 0$ for some $i \in N$ and $k \in M$. Then every player $j \neq i$ can ensure himself the maximal payoff 1 by playing pure strategy k . So, since σ is Nash equilibrium, every player $j \neq i$, and then also i , receives payoff 1 under σ . Hence $\sigma \in G$ by definition. ■

The completely mixed strategy combination ρ is hence an isolated Nash equilibrium. Our proof that G is a topological sphere is based on the following construction. Consider again the previous example and the previous graph. Take a plane that is perpendicular to the line going through the points (A, A, A) , (B, B, B) (and hence ρ). Project the space of all strategy combinations (the cube) onto this plane. The cube is then mapped onto a region in the plane bounded by a regular hexagon. The set G is mapped precisely onto the boundary of the hexagon, which is a topological cycle.

In the proof below we generalize this construction. We project the polyhedron of all strategy combinations onto a linear subspace D . The image E under this projection of the polyhedron of strategy profiles is then itself a polyhedron. We show that G is mapped precisely onto the boundary of the polyhedron E . And the boundary of the compact and convex polyhedron E is, of course, a topological sphere.

It turns out that the linear space D onto which one has to project can be described very easily. Namely, if one thinks of strategy combinations as $m \times n$ matrices with column sums 1, then L is the space of all $m \times n$ matrices with column- and row sums zero. The inequalities that define the polyhedron E can be explicitly computed, which is done in the proof.

Proposition 3.2 *The set G of Pareto efficient Nash equilibria of a minimal diversity game is homeomorphic to a sphere of dimension $d = (n - 1) \cdot (m - 1) - 1$.*

Proof. It suffices to show that G is homeomorphic to the boundary of a compact and convex set of dimension $(n-1) \cdot (m-1)$. A mixed strategy profile $\sigma = (\sigma_{ik})_{i \in N, k \in M}$ of the minimal diversity game is a point in the Euclidean vectorspace \mathbb{R}^{mn} . Let D be the set of vectors $d \in \mathbb{R}^{mn}$ such that $\sum_k d_{ik} = 0$ for every i and $\sum_i d_{ik} = 0$ for every k . This is a linear subspace of \mathbb{R}^{mn} of dimension $mn - n - m + 1 = (n-1)(m-1)$.

Geometrically, the set D is the linear space of vectors that are perpendicular to the affine space generated by the pure strategy combinations where all players make the same choice (that is, those pure strategy combinations not in G).

Let \mathcal{J} be the set of maps $J: M \rightarrow N$ that are not constant (that is, $J^{-1}(i) \neq M$ for every i). Let E be the set of points $d \in D$ that satisfy the inequality

$$\sum_l d_{J(l),l} \geq -1$$

for every $J \in \mathcal{J}$. Note that for $0 \in D$ all these inequalities are strict, so that E also has dimension $(n-1)(m-1)$. Moreover, E is bounded, hence a polytope. So, the (relative) boundary F of E (in D) is homeomorphic to a sphere of dimension $d = (n-1)(m-1) - 1$.

The polytope E is in fact the orthogonal projection of Σ onto D .⁹ Let P denote the orthogonal projection onto D . We show that the orthogonal projection P is indeed a homeomorphism between G and F . For $J \in \mathcal{J}$, write

$$G_J = \{\sigma \mid \sigma_{J(k),k} = 0 \text{ for all } k \in M\}.$$

It is straightforward to check that the set G_J is a face of Σ of dimension d , and that G is the union of these faces G_J . Let r^k denote the strategy profile where each player plays pure strategy k . For any $\sigma \in \Sigma$ we have by a straightforward calculation that

$$P(\sigma) = \sigma - \frac{1}{n} \sum_k \left(\sum_j \sigma_{jk} \right) \cdot r^k. \quad (*)$$

A. We show that $P(G) \subseteq F$. Take $\sigma \in \Sigma$. By definition, $P(\sigma) \in D$. Further, using $(*)$, for $J \in \mathcal{J}$ it follows that

$$\sum_l P(\sigma)_{J(l),l} = \sum_l \sigma_{J(l),l} - 1. \quad (**)$$

So, since $\sum_l \sigma_{J(l),l} \geq 0$, it follows that $P(\sigma) \in E$. Also, if $\sigma \in G_J$, then $\sum_l \sigma_{J(l),l} = 0$, so that $\sum_l P(\sigma)_{J(l),l} = -1$. This shows that $P(G) \subseteq F$.

⁹We do not need this fact in our arguments though, so we do not prove it here.

B. We show that $F \subseteq P(G)$. Take a point $d \in F$. Since F is the boundary of E , there is a $J \in \mathcal{J}$ with $\sum_l d_{J(l),l} = -1$. Define

$$\sigma = d - \sum_k d_{J(k),k} \cdot r^k.$$

Then for every i ,

$$\sum_l \sigma_{il} = \sum_l d_{il} - \sum_l \sum_k d_{J(k),k} \cdot r_{il}^k = 0 - \sum_l d_{J(l),l} = +1.$$

Further, for for every i, l ,

$$\sigma_{il} = d_{il} - \sum_k d_{J(k),k} \cdot r_{il}^k = d_{il} - d_{J(l),l} = +1 + d_{il} + \sum_{k \neq l} d_{J(k),k} \geq 1 - 1 = 0.$$

So, $\sigma \in \Sigma$. It is straightforward to check that $P(\sigma) = d$. Hence, $F \subseteq P(\Sigma)$.

C. We show that P is one-to-one on G . Take two strategy profiles $\sigma, \tau \in G_J$ with $P(\sigma) = P(\tau)$. Take $l \in M$. Write $i = J(l)$. Then, since $\sigma_{il} = 0 = \tau_{il}$, from (*) it follows that

$$\sum_j (\sigma_{jl} - \tau_{jl}) = \sum_{k \in M} \left(\sum_j (\sigma_{jk} - \tau_{jk}) \right) \cdot (r^k)_{il} = n \cdot (\sigma_{il} - \tau_{il}) = 0.$$

Hence, $\sum_j (\sigma - \tau)_{jl} = 0$ for every l . So, $\sigma - \tau \in D$. Then, since $P(d) = d$ for all $d \in D$, $\sigma - \tau = P(\sigma - \tau) = 0$. It follows that $\sigma = \tau$. Consequently, P is one-to-one on each face G_J .

Now take strategy profiles $\tau, \kappa \in \Sigma$ with $\tau \in G_J$ and $\kappa \notin G_J$. Then, by (**), $\sum_l P(\tau)_{J(l),l} = -1$ while $\sum_l P(\kappa)_{J(l),l} > -1$. Hence, $P(\tau) \neq P(\kappa)$. Thus, P is one-to-one on G . ■

4 Evolutionary dynamics

In this section we show for the large class of evolutionary dynamics, called strongly payoff consistent selection dynamics, that G is asymptotically stable, but ρ is not. The analysis is straightforward, following well-known arguments. However, the dynamics can have fixed points other than Nash equilibria, and some caution is needed to deal with those.

As in D&R we define a *payoff consistent selection dynamics* to be a Lipschitz continuous vector field $f = (f_i)_{i \in N}$ on Σ that does not point outwards on Σ and that satisfies for all players i and all $x \in \Sigma$

$$\langle f_i(\sigma), \nabla_{\sigma_i} U_i(\sigma) \rangle \geq 0. \quad (1)$$

Here, $\langle x, y \rangle$ denotes the inner product between vectors x and y . We call the selection dynamics *strongly payoff consistent* if, in addition, all Nash equilibria are fixed points and if the above

inequality is strict for every strategy profile σ and every player i who has a pure strategy s_i in the support of σ_i with $u_i(\sigma) < u_i(\sigma | s_i)$. The notion of a strongly payoff consistent selection dynamics weakens the notion of a Nash dynamic in D&R just enough so that the replicator dynamics is captured as well as a special case.

Proposition 4.1 *Suppose we have a strongly payoff consistent selection dynamics in a minimal diversity game. Then the only asymptotically stable set of restpoints is the set G of efficient Nash equilibria. The completely mixed Nash equilibrium ρ is not asymptotically stable.*

Proof. Since all players have the same utility function and the dynamic is strongly payoff consistent, the chain rule implies that utility is non-decreasing along a trajectory and strictly increasing in any point where there is a better reply in the support. Now consider a strategy profile σ with $u(\sigma) < 1$ in which no player i has a better reply in the support of σ_i . Then all players must play mixed strategies which have the same support, which we denote by T .

To see this, note that otherwise one player would have a pure strategy s_i in the support of his strategy σ_i which is not in the support of the strategy σ_j of another player. By using only s_i player i can then increase everybody's payoff to 1, a contradiction.

Then the same argument as in the proof of Proposition 3.1, which showed that in a Nash equilibrium in completely mixed strategies every player must play each pure strategy with equal probability $1/m$, establishes here that in the strategy profile σ each player must use each strategy in T with equal probability. It follows that the dynamics can only have a finite number of points outside G for which (1) holds with equality for all i (namely those points where each player plays strategies in T with probability $\frac{1}{T}$, and those outside T with zero probability).

To see that no rest point σ with $u(\sigma) < 1$ is asymptotically stable under the dynamics, notice that we can find strategy profiles τ with $u(\sigma) < u(\tau)$ arbitrarily near to σ . Because utility is increasing on the trajectory starting in τ , the trajectory cannot move towards σ .

Now let c be such that $u(\sigma) < c$ for all the rest points not in G and consider the neighborhood U of G consisting of all τ with $u(\tau) \geq c$. Since utility is strictly increasing along any trajectory starting in $\tau \in U \setminus G$, Theorem 2.6.1 of Hofbauer and Sigmund [14] implies that any ω -limit point σ of the trajectory must satisfy $\dot{u}(\sigma) = 0$. So, inequality (1) holds for all i . Therefore, σ is an element of G . Hence, G is asymptotically stable. ■

Somewhat lengthy calculations (available from the authors on request) show that the finitely many rest points of the replicator dynamics outside G are unstable hyperbolic rest points. The

stable manifold of such a rest point σ with its common support T can be shown to consist of the strategy profiles where all players use identical mixed strategies with support T . (It is straightforward to verify that σ is the unique strategy profile maximizing utility within this set, which is clearly forward invariant under the replicator dynamics. Dimensional arguments then imply that it is indeed the stable manifold.) Thus the stable manifold of each of these restpoints is of lower dimension than Σ . Hence there is an open and dense subset of Σ such that all trajectories starting in it converge to G . In this sense G is “almost” globally asymptotically stable. Since G is a strict equilibrium set in the sense of Balkenborg and Schlag [1], Theorem 6 of their paper implies that G is an asymptotically stable set of stable rest points of the replicator dynamics.

5 Strategic stability

Strategic stability is concerned with sets of Nash equilibria that satisfy necessary conditions for a solution to be acceptable for rational players. Most definitions of strategic stability require robustness of a set of Nash equilibria with respect to certain perturbations of the original game. In this section we are primarily concerned with essentiality (Wu Wen-Tsun and Jiang Jia-He [16], [30]), KM stability and full stability (K&M) and strategic stability in the sense of Mertens [21] and Hillas [11].

We study which of the Nash equilibrium components of a minimal diversity game, the set of Pareto efficient Nash equilibria G or the set $\{\rho\}$ consisting of the Nash equilibrium in completely mixed strategies, contains a strategically stable set of Nash equilibria according to any of the notions mentioned.

From the perspective of strategic stability, the analysis for the completely mixed Nash equilibrium ρ is straightforward. According to the next Theorem, the Nash equilibrium ρ is strongly stable, and hence essential and stable in the sense of Mertens. Then it follows from K&M and Hillas et al. [12] that ρ is also strategically stable according to any of the other notions mentioned above.

Theorem 5.1 *The equilibrium ρ is strongly stable in the sense Kojima, Okada and Shindoh [18]. Consequently ρ is also essential, and the set $\{\rho\}$ is stable in the sense of Mertens. Further, ρ is regular in the sense of Harsanyi.*

Proof. Let (S, u) be a minimal diversity game, and let ρ be its completely mixed equilibrium. We first show that ρ is strongly stable in the sense of Kojima, Okada and Shindoh [18]. For

player i , define the function $f_{i1}: \prod_i \mathbb{R}^{S_i} \rightarrow \mathbb{R}$ by

$$f_{i1}(\sigma) = \sum_k \sigma_{ik} - 1.$$

For player i and pure strategy $k \geq 2$, define the function $f_{ik}: \prod_i \mathbb{R}^{S_i} \rightarrow \mathbb{R}$ by

$$f_{ik}(\sigma) = u(\sigma | k) - u(\sigma | 1).$$

Note that, for $\sigma \in \prod_i \mathbb{R}^{S_i}$ with $\sigma > 0$, the system of equations $f(\sigma) = 0$ yields (1) $\sigma \in \Sigma$, and (2) for each player i all pure strategies of i yield the same payoff against profile σ . Then, for $\sigma \in \prod_i \mathbb{R}^{S_i}$ with $\sigma > 0$, $f(\sigma) = 0$ precisely when σ is a Nash equilibrium of the minimal diversity game.

So, since ρ is completely mixed, it is by the Implicit Function Theorem sufficient to show that the Jacobian matrix $\frac{\partial f}{\partial \sigma}(\rho)$ has non-zero determinant. Straightforward calculations show that, for $k \geq 2$,

$$f_{ik}(\sigma) = \prod_{j \neq i} \sigma_{j1} - \prod_{j \neq i} \sigma_{jk}.$$

So, it is easy to check that

$$\frac{\partial f_{ik}}{\partial \sigma_{jl}}(\rho) = \begin{cases} 0 & \text{if } k = 1 \text{ and } j \neq i \\ 1 & \text{if } k = 1 \text{ and } j = i \\ 0 & \text{if } k \geq 2 \text{ and } j = i \\ \left(\frac{1}{m}\right)^{n-2} & \text{if } k \geq 2, j \neq i \text{ and } l = 1 \\ -\left(\frac{1}{m}\right)^{n-2} & \text{if } k \geq 2, j \neq i \text{ and } l \neq 1. \end{cases}$$

It is now straightforward to show that the resulting Jacobian matrix $\frac{\partial f}{\partial \sigma}(\rho)$ has full rank.

We showed that ρ is strongly stable. Therefore ρ is also essential and stable in the sense of Mertens. Further, ρ is completely mixed, and therefore quasi-strict. Hence, by Theorem 2.5.6 of van Damme [29], ρ is regular. \blacksquare

Note that strategic stability in the sense of Mertens also implies many other types of strategic stability such as full stability, KM stability, and best response stability.

Next we study the strategic stability of the set G of efficient Nash equilibria. Consider a minimal diversity game, and let G be its set of efficient Nash equilibria. Any game equivalent to this game¹⁰ and any perturbation defined by restriction of the strategy space of it is a game with identical interests, that is, all players have the same payoff function. Then there are Nash equilibria nearby to strategy profiles equivalent to G in every perturbation close to

¹⁰Two games are equivalent when deletion of all pure strategies that are payoff equivalent to some (possibly mixed) strategy yields the same reduced form.

an equivalent game. These can be found among the strategy profiles maximizing the utility function. The definitions of strategic stability in K&M hence imply the following.¹¹

Proposition 5.2 *The set G of efficient Nash equilibria of a minimal diversity game contains a fully stable set, and hence a KM-stable set.*

Thus, G is strategically stable under the milder notions of strategic stability defined in K&M. We show next that when both the number of players and the number of pure strategies is even, G is strategically stable even under the more demanding notions. However, the remainder of this paper, starting in the next section, will be devoted to the proof that in other cases this no longer holds.

Theorem 5.3 *Suppose both the number of players and the number of pure strategies is even. Then G is essential. Moreover, G contains a strategically stable set in the sense of Mertens.*

Proof. By Proposition 3.2, the set G is homeomorphic to a sphere of dimension $d = (n - 1)(m - 1) - 1$. Thus d is even, which implies that the Euler characteristic of G is $+2$. By Proposition 4.1, G is an asymptotically stable set of rest points under the replicator dynamics. So, by Theorem 1 of Demichelis [2], the index of G is equal to the Euler characteristic, and hence not zero. Theorem 4 in Ritzberger [24] then implies that G is essential,¹² and Theorem 2 of D&R [4] implies that G contains a stable set in the sense of Mertens.¹³ ■

We now come to the main results of the paper. We conjecture that for all minimal diversity games the following result holds. Suppose that the set G of efficient Nash equilibria has odd dimension. Then it does *not* contain a strategically stable set in the sense of Mertens or Hillas, and it also does not contain an essential set.¹⁴

A fairly intuitive proof,¹⁵ using a generalization of the rock-scissors-paper game, is given for bimatrix games. For binary minimal diversity games (that is, minimal diversity games in which players only have two pure actions) we have an elementary proof showing that G does not contain an essential set. Further, specifically for binary minimal diversity games we develop

¹¹The result extends to the set of strategy profiles maximizing the potential in any weighted potential game as defined in Monderer and Shapley [23].

¹²In Ritzberger [24], computation of the index is based on the replicator dynamics. The results of Demichelis and Germano [3] however imply that the computation does not depend on the specific dynamics used.

¹³Alternatively, one can also directly compute the index of ρ , and use the fact that the sum of indices over all Nash components equals $+1$. This however would not result in a considerably shorter proof.

¹⁴We know that G has index zero in this case. The results by Govindan and Wilson [5] then imply that G is not uniformly hyperstable. However, although this yields a partial resolution, it does not (yet) imply that G is not essential in the sense used in this paper.

¹⁵Nevertheless, the proof relies on rather sophisticated tools from Hillas et al. [12].

a technique to linearize the Nash equilibrium correspondence on the class of KM-perturbed games. This then allows us to prove the above conjecture in that case.

Whether any of these techniques can be adapted to more general types of minimal diversity games is not known to us. At least we have proofs for examples with an arbitrarily large number of players, and an arbitrarily large number of pure strategies. These results give us some confidence that the conjecture might be true in its full generality.

6 Bimatrix games

In this section we give the proof of the conjecture for two-player games. Notice that if the number of pure strategies is m in a two-player minimal diversity game, the dimension of the component G is $m-2$. So we obtain one example for each possible dimension of the component. Here we are interested in the case where m is odd.

Consider then a two-player minimal diversity games with an odd number of strategies $m \geq 3$. We show in two steps that G does not contain a best response stable set. It does therefore also not contain a strategically stable set in the sense of Mertens. Moreover, we show that G does not contain an essential set.

First, we introduce a weak notion of strategic stability called independent t -stability. We show that every best response stable set as well as every essential set must contain a t -stable set. Secondly we show that G does not contain an independent t -stable set.

Definition An independent t -perturbation of size ε of a game (S, u) is a collection $(t_i)_{i \in N}$ of maps $t_i: S_i \rightarrow \Sigma_i$ such that $\|t_i(s_i) - s_i\| < \varepsilon$ for every pure strategy $s_i \in S_i$.¹⁶ Each independent t -perturbation of a game defines a t -perturbed game (S, u^t) in normal form, with the same strategy sets S_i as the original game, but new utility functions $u_i^t: S \rightarrow \mathbb{R}$ given by

$$u_i^t(s) = u_i(t(s) \mid s_i),$$

where

$$t(s) = (t_i(s_i))_{i \in N}$$

and $(t(s) \mid s_i)$ is the strategy profile where player i chooses pure strategy s_i , while each other player j adheres to mixed strategy $t(s)_j = t_j(s_j)$. A t -perturbed game obviously defines a particular payoff-perturbed game. The t -perturbations are, however, more general than

¹⁶We slightly abuse notation, and identify a pure strategy s_i with the mixed strategy that puts weight one on pure strategy s_i .

trembling hand perturbations because the trembles of a player are correlated with his intended pure strategy choice. Because of this correlation it matters now for the choice of an optimal strategy (and hence for the Nash equilibria of the perturbed game) whether a player ignores his own trembles or not. In our definition of a t -perturbed game a player ignores his own trembles. As a consequence, a t -perturbation of a game with identical payoff does not have to have identical interests and therefore arguments as in Proposition 5.2 do not apply.

Definition A non-empty connected closed set C of Nash equilibria is an independent t -set if there is a Nash equilibrium close to C in every sufficiently small t -perturbation of the game.

Proposition 6.1 *Let $m \geq 3$ be odd. Then the set G of the two player minimal diversity game with m pure strategies is not an independent t -set.*

Proof. Let $m \geq 3$ be odd. Let $\varepsilon > 0$. For the two-player minimal diversity game with m strategies we construct an independent t -perturbation t of size ε such that the unique Nash equilibrium of the t -perturbed game is ρ . Set

$$t_{ik} = \sum_{l=1}^m \frac{\varepsilon^{(l-k) \bmod m}}{\sum_{l=0}^{m-1} \varepsilon^l} \cdot s_{il}$$

where we denote by s_{il} player i 's l -th pure strategy. If we multiply the payoffs by the common factor $f = \sum_{l=0}^{m-1} \varepsilon^l$, the rescaled payoffs in the t -perturbed game are given by

$$f \cdot u_1^t(l, k) = f \cdot u_2^t(k, l) = f - \varepsilon^{(l-k) \bmod m}.$$

Consider a Nash equilibrium σ of the t -perturbed game. Assume that $\sigma \neq \rho$. We derive a contradiction. Assume without loss of generality that $\sigma_2 \neq \rho_2$. We first show the following claim.

Claim. Let $1 \leq k \leq m$ be such that

$$\sigma_{2k} = \min_{1 \leq l \leq m} \sigma_{2l}.$$

Then $\sigma_{1, k-1} = 0$ (where $\sigma_{1, k-1} = \sigma_{1, m}$ if $k = 1$).

Proof of claim. We assume without loss of generality that $k = m$. We compute that

$$f \cdot u_1^t(m, \sigma_2) = f - \sum_{l=1}^m \varepsilon^{m-l} \sigma_{2l} = f - \sum_{l=1}^{m-1} \varepsilon^{m-l} \sigma_{2l} - \sigma_{2m},$$

while

$$f \cdot u_1^t(m-1, \sigma_2) = f - \sum_{l=1}^m \varepsilon^{(m-1-l) \bmod m} \sigma_{2l} = f - \sum_{l=1}^{m-1} \varepsilon^{m-1-l} \sigma_{2l} - \varepsilon^{m-1} \sigma_{2m}.$$

We obtain

$$\begin{aligned}
f \cdot u_1^t(m, \sigma_2) - f \cdot u_1^t(m-1, \sigma_2) &= \sum_{l=1}^{m-1} \varepsilon^{m-1-l} (1-\varepsilon) \sigma_{2l} - (1-\varepsilon^{m-1}) \sigma_{2m} \\
&= \sum_{l=1}^{m-1} \varepsilon^{m-1-l} (1-\varepsilon) \sigma_{2l} - \sum_{l=1}^{m-1} \varepsilon^{m-1-l} (1-\varepsilon) \sigma_{2m} \\
&= \sum_{l=1}^{m-1} \varepsilon^{m-1-l} (1-\varepsilon) (\sigma_{2l} - \sigma_{2m}) > 0.
\end{aligned}$$

Strict inequality holds because $\sigma_{2m} \leq \sigma_{2l}$ for all l and $\sigma_{2m} < \sigma_{2l}$ for at least one l . This concludes the proof of the claim.

Now we can proceed as follows. From the assumptions that $\sigma_2 \neq \rho_2$ and $\sigma_{2m} = \min_{1 \leq l \leq m} \sigma_{2l}$ we concluded that $\sigma_{1,m-1} = 0$. So, $\sigma_1 \neq \rho_1$ and $\sigma_{1,m-1} = \min_{1 \leq l \leq m} \sigma_{1l}$. Thus the claim implies that $\sigma_{2,m-2} = 0$. Iterating the argument yields $\sigma_{1,m-3} = 0$, then $\sigma_{2,m-4} = 0$ and so on, with the player index alternating between 2 and 1.

Thus, since m is odd, we see that $\sigma_{2l} = 0$ for l odd and $\sigma_{1l} = 0$ for l even. We obtain in particular after $m-1$ steps that $\sigma_{21} = 0$. Now the claim yields $\sigma_{1m} = 0$. Iteration of the argument yields $\sigma_{1l} = 0$ for l odd and $\sigma_{2l} = 0$ for l even. We have shown that $\sigma_{1l} = \sigma_{2l} = 0$ for all l , a contradiction. This completes the proof. \blacksquare

Now we can show the main result of this section.

Theorem 6.2 *Let $m \geq 3$ be odd. Then the set G of the two player minimal diversity game with m pure strategies is not essential. Also it does not contain a best response stable set.*

Proof. Let (S, u) be any normal form game. We argue that a closed set C of Nash equilibria of the game (S, u) that is essential or best response stable necessarily contains an independent t -set. The result then immediately follows from Proposition 6.1.

For essentiality this follows from the observation that, when the size of a t -perturbation is small, also $\|u - u^t\|$ is small by definition of u^t and the continuity of the payoff function u .

Assume that C is best response stable. Then, according to Hillas et al. [12], the set C must contain a so-called CT set.¹⁷ This means in particular that, for every sufficiently small independent t -perturbation, the correspondence BR^t defined by

$$BR^t(\sigma) = \text{convex hull } \{t(s) \mid s \in PB(\sigma)\},$$

¹⁷See Subsection 5.1 of their paper for the definition.

where PB is the pure best reply correspondence, has a fixed point close to C . Let $\sigma \in BR^t(\sigma)$ be such a fixed point. Then, because t is an independent t -perturbation, each σ_i is in the convex hull of the strategies $t_i(s_i)$ with $s_i \in PB_i(\sigma)$. Hence, for each player i there is a vector $\phi(\sigma, t)_i = (\phi(\sigma, t)_i(s_i))_{s_i \in S_i}$ in Σ_i such that

$$\sigma_i = \sum_{s_i \in S_i} \phi(\sigma, t)_i(s_i) \cdot t_i(s_i),$$

while moreover $\phi(\sigma, t)_i(s_i) > 0$ implies that $s_i \in PB_i(\sigma)$. Write $\phi(\sigma, t) = (\phi(\sigma, t)_i)_{i \in N}$. Using the definition of the payoff function u_i^t and the multilinearity of the payoff function u_i it is straightforward to check that

$$u_i^t(\phi(\sigma, t) \mid s_i) = u_i(\sigma \mid s_i)$$

for all pure strategies s_i of player i . Now suppose that $\phi(\sigma, t)_i(s_i) > 0$. Then, as noted before, $s_i \in PB_i(\sigma)$. Thus, according to the above displayed equality, s_i is a pure best reply to $\phi(\sigma, t)$ in the game (S, u^t) . Therefore ϕ_i is a best reply to ϕ in the t -perturbed game (S, u^t) and hence $\phi(\sigma, t)$ is a Nash equilibrium of the t -perturbed game (S, u^t) .

It remains to show that, for sufficiently small t , $\phi(\sigma, t)$ is close to C whenever σ is close to C . This follows readily once we observe that, for small t , $t_i(s_i)$ is close to s_i for all s_i . Hence, $\phi(\sigma, t)_i(s_i)$ is close to $\sigma_i(s_i)$ and $\phi(\sigma, t)$ is close to σ . This concludes the proof. \blacksquare

7 Binary minimal diversity games

The second class of minimal diversity games for which we can prove the conjecture that evolutionary and strategic stability make mutually exclusive choices is the class of binary minimal diversity games. A binary minimal diversity game is a game in strategic form with player set $N = \{1, \dots, n\}$ in which each player has two pure strategies A and B at his disposal. We assume that n is odd and $n \geq 3$. Let S denote the set of pure strategy profiles $s = (s_i)_{i \in N}$ where $s_i \in \{A, B\}$ for all $i \in N$. Each player has the same payoff function $u_i = u$, where $u: S \rightarrow \mathbb{R}$ is defined by

$$u(s) := \begin{cases} 0 & \text{when } s_1 = \dots = s_n \\ 1 & \text{else.} \end{cases}$$

A typical mixed strategy is denoted by $\sigma = (\sigma_{iA}, \sigma_{iB})_{i \in N}$, where σ_{iA} (σ_{iB}) denotes the probability with which player i plays pure strategy A (B). Obviously $\sigma_{iA} \geq 0$, $\sigma_{iB} \geq 0$, and $\sigma_{iA} + \sigma_{iB} = 1$. Alternatively we write $\sigma = (\sigma_i, 1 - \sigma_i)_{i \in N}$ for a generic strategy profile. The space of mixed strategy profiles is denoted by Σ .

7.1 Essentiality

In this section we show, for odd n , that G is not essential in the sense of Wu Wen-Tsun and Jiang Jia-He [16], [30]. Let (S, u) denote the binary minimal diversity game with player set N . It suffices to construct a small perturbation of the game (S, u) that does not have Nash equilibria close to G . Let (S, v) denote the strategic form game with player set N and payoff functions $v = (v_i)_{i \in N}$, where v_i is defined by

$$v_i(s) = \begin{cases} 1 & \text{if } s_i \neq s_{i-1} \\ 0 & \text{else,} \end{cases}$$

with the convention that player 0 equals player n . Take $\varepsilon > 0$. Consider the game $(S, v(\varepsilon))$ with payoff function $v(\varepsilon) = u + \varepsilon v$. We show that the strategy profile in which each player plays both his pure strategies with weight $\frac{1}{2}$ is the unique Nash equilibrium of the game $(S, v(\varepsilon))$.

Theorem 7.1 *The unique Nash equilibrium of the game $(S, v(\varepsilon))$ is the strategy profile $(\sigma_i, 1 - \sigma_i)_{i \in N}$ with $\sigma_i = \frac{1}{2}$ for all $i \in N$.*

Proof. The proof is in three steps.

A. First notice that the strategy profile in which each player plays both his pure strategies with weight $\frac{1}{2}$ is a Nash equilibrium of both the game (S, u) and the game (S, v) . Therefore it is also a Nash equilibrium of the game $(S, v(\varepsilon))$. We show that $(S, v(\varepsilon))$ does not have any other Nash equilibria.

B. Let $(\sigma_i, 1 - \sigma_i)_{i \in N}$ be a Nash equilibrium of the game $(S, v(\varepsilon))$. Suppose that it is an element of the boundary of the space of strategy profiles. Suppose w.l.o.g. that $\sigma_1 = 0$. Then necessarily $\sigma_2 = 1$. Because n is odd, iterating this argument leads to $\sigma_1 = 1$, a contradiction. Hence the game $(S, v(\varepsilon))$ only has completely mixed Nash equilibria.

C. Suppose there exists a completely mixed Nash equilibrium $(\sigma_i, 1 - \sigma_i)_{i \in N}$ of the game $(S, v(\varepsilon))$ for which $\sigma_i \neq \frac{1}{2}$ for at least one $i \in N$. Assume w.l.o.g. that $\sigma_i > \frac{1}{2}$. From the equilibrium condition

$$1 - \prod_{j \neq i+1} \sigma_j + \varepsilon(1 - \sigma_i) = 1 - \prod_{j \neq i+1} (1 - \sigma_j) + \varepsilon \sigma_i$$

for player $i + 1$ we can deduce that

$$\frac{\sigma_i}{1 - \sigma_i} = \frac{\prod_{j \neq i, i+1} (1 - \sigma_j) + \varepsilon}{\prod_{j \neq i, i+1} \sigma_j + \varepsilon}.$$

Define $\alpha_j = \frac{\sigma_j}{1-\sigma_j}$ for all $j \in N$. Note that $\alpha_i > 1$, because we assumed that $\sigma_i > \frac{1}{2}$. From $\alpha_i > 1$ we obtain the strict inequality

$$\alpha_i = \frac{\sigma_i}{1-\sigma_i} = \frac{\prod_{j \neq i, i+1} (1-\sigma_j) + \varepsilon}{\prod_{j \neq i, i+1} \sigma_j + \varepsilon} < \frac{\prod_{j \neq i, i+1} (1-\sigma_j)}{\prod_{j \neq i, i+1} \sigma_j} = \prod_{j \neq i, i+1} \frac{1}{\alpha_j}.$$

Then we know that $\alpha_j < 1$ for at least one $j \in N$. So, there must be a player k with $\alpha_k > 1$ and $\alpha_{k-1} \leq 1$. Write $\alpha = \prod_{j \in N} \alpha_j$. From the equilibrium equality for player k we deduce that

$$\alpha_{k-1} \geq \prod_{j \neq k-1, k} \frac{1}{\alpha_j}$$

which can be rewritten as $\alpha \geq \alpha_k$. Hence, $\alpha \geq \alpha_k > 1$. However, since there also exists a player j for which $\alpha_j < 1$, the same line of reasoning yields $\alpha < 1$, a contradiction. ■

In fact the perturbed games are generalized Shapley games as in Hofbauer and Swinkels [15]. This means that in these games any (sufficiently small) open neighborhood of G contains an attractor of *any* strongly payoff consistent dynamics, while it (in the perturbed game!) does not contain any Nash equilibria.

7.2 Strategic stability, definitions of Mertens and Hillas

In this section we argue that for odd n , G is not best response stable in the sense of Hillas [11]. However, since the formal proof is quite elaborate, we only phrase our result, and sketch the outline of the proof.¹⁸

Theorem 7.2 *Let $n \geq 3$ be odd. Then the set G of the binary minimal diversity game with n players does not contain a best response stable set. Consequently, G is also neither homotopy stable, nor stable in the sense of Mertens.*

Sketch of proof. The proof of the theorem consists of several steps, each of which we briefly discuss. Let Γ be a binary minimal diversity game, and let G be its component of Pareto efficient Nash equilibria. In the formal proof we argue that, when n is odd, G is not a CKM set¹⁹ as it is defined in Hillas et al. [12]. The above claims are then implied by the results of that paper.

In order to prove that G is not a CKM set, we first derive a full description of the graph \mathcal{E} of the equilibrium correspondence over KM perturbations. A KM perturbation is a vector $\eta = (\eta_{iA}, \eta_{iB})_{i \in N}$ of non-negative numbers η_{iA} and η_{iB} . In the η -perturbed game player i is

¹⁸The complete proof can be obtained from the authors upon request.

¹⁹CKM stands for: continuous version of Kohlberg-Mertens stability.

forced to play pure strategy A (B) with at least probability of η_{iA} (η_{iB}). The η -perturbed game is denoted by $\Gamma(\eta)$. \mathcal{K} is the set of KM perturbations, and \mathcal{E} is the set of pairs (η, σ) in $\mathcal{K} \times \Sigma$ for which the mixed strategy profile σ is a Nash equilibrium $\Gamma(\eta)$.

In the proof we do not work directly with the graph \mathcal{E} , but rather apply a transformation to the space $\mathcal{K} \times \Sigma$ to obtain a completely linear description of \mathcal{E} . Within this linearized context we show that, for $n \geq 3$ odd, the set G is not a CKM set. This again requires several steps.

A CKM perturbation is a continuous function $\varepsilon: \Sigma \rightarrow \mathcal{K}$. The graph $\text{graph}[\varepsilon]$ of ε is the set of pairs $(\varepsilon(\sigma), \sigma)$ in $\mathcal{K} \times \Sigma$. A closed set $C \subset \Sigma$ is a CKM set if for every sufficiently small CKM perturbation ε there exists a point $(\eta, \sigma) \in \text{graph}[\varepsilon] \cap \mathcal{E}$ that is close to C .

Given the component G , we construct a CKM perturbation ε of small size such that no point in $\text{graph}[\varepsilon] \cap \mathcal{E}$ is close to G . First we construct a KM perturbation, called the initial perturbation, having exactly $2n - 1$ equilibria. One is the completely mixed equilibrium ρ . Then there are $2(n - 2)$ equilibria in which the first k players play the same pure strategy, while the others play the same mixed strategy.²⁰ The exact probabilities used by the mixing players is precisely determined by the perturbation. Finally, there are 2 equilibria where the first $n - 1$ players all play the same pure strategy, while the last player plays the other pure strategy. The completely mixed equilibrium does not concern us, we focus on the $2n - 2$ remaining equilibria.

For example, for $n = 5$, we have the following equilibria. First there is the completely mixed equilibrium (Z, Z, Z, Z, Z) . Then there are the 6 equilibria

$$\begin{array}{ccc} (A, Z, Z, Z, Z) & & (B, Z, Z, Z, Z) \\ (A, A, Z, Z, Z) & \text{and} & (B, B, Z, Z, Z) \\ (A, A, A, Z, Z) & & (B, B, B, Z, Z) \end{array}$$

and finally the two equilibria (A, A, A, B) and (B, B, B, B, A) . In this notation an A (or a B) in position i stands for “player i plays B (or A) with minimum probability”, while a Z in position i means that player i plays a mixed strategy (in the sense that he plays both strategies with strictly more than minimum probability). Only the latter 8 equilibria are of interest to us because these are all close to the sphere G , while (Z, Z, Z, Z, Z) is not.

The next step in the construction is to show that these equilibria are pairwise linked via paths, parametrized by a parameter $\lambda \geq 1$, in the space of perturbations. In the above example, (A, Z, Z, Z, Z) gets linked to (A, A, Z, Z, Z) as shown below. The equilibrium correspondence

²⁰In a perturbed game it is not really possible to play a pure strategy due to the restrictions imposed by the perturbations. When we write that a player plays a pure strategy, in this context we simply mean that he plays the other strategy with minimum weight.

over the path of perturbations for $\lambda \geq 1$, when restricted to the equilibria (A, Z, Z, Z, Z) , (A, Z, B, Z, Z) , (A, A, B, Z, Z) , and (A, A, Z, Z, Z) , looks as follows.

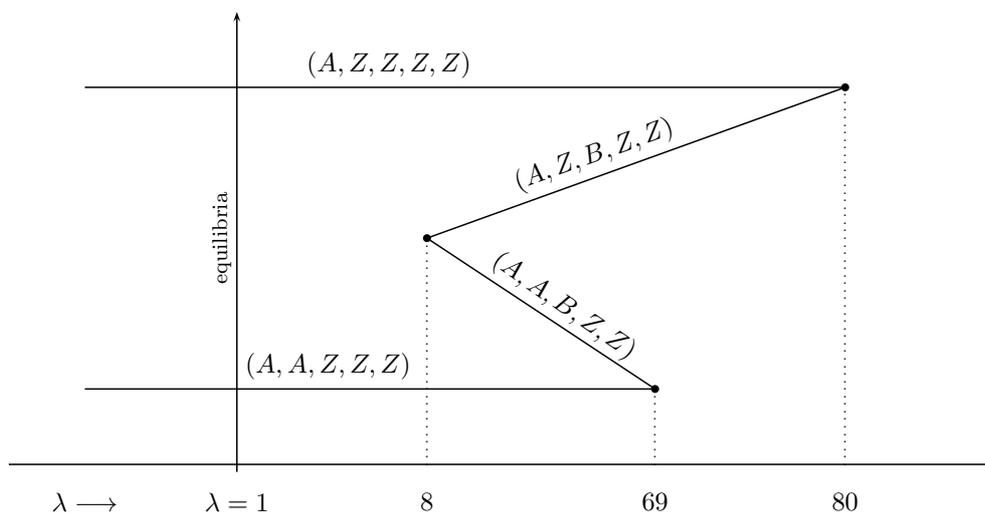


FIGURE: the graph of the equilibrium correspondence for $n = 5$ and $k = 1$.

Via another path, although also parametrized by $\lambda \geq 1$, (A, A, A, Z, Z) gets linked to (A, A, A, A, B) . The equilibrium correspondence over the path of perturbations for $\lambda \geq 1$, when restricted to the equilibria (A, A, A, Z, Z) , (A, A, A, Z, B) , and (A, A, A, A, B) , looks as follows.

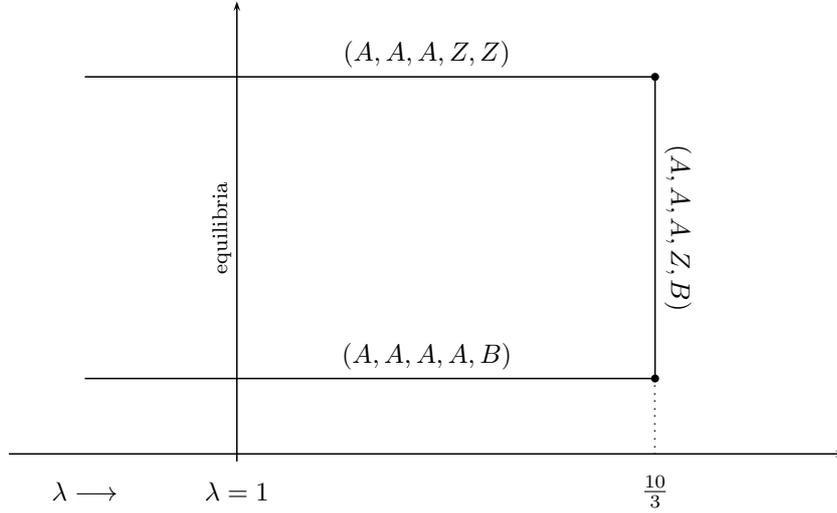


FIGURE: the graph of the equilibrium correspondence for $n = 5$ and $k = 3$.

In the same way (B, Z, Z, Z, Z) gets linked to (B, B, Z, Z, Z) , and (B, B, B, Z, Z) gets linked to (B, B, B, B, A) . These four (in general $n - 1$) paths are then used to construct a CKM perturbation ε that avoids any intersection with the paths connecting the equilibria of the initial perturbation. Hence ε has no points of intersection with \mathcal{E} close to G . ■

8 Discussion and conclusion

In this paper we presented the class of minimal diversity games, a subclass of the class of potential games. The set of Nash equilibria of each minimal diversity game is shown to consist of a symmetric isolated completely mixed equilibrium ρ and a connected component G of strategy profiles that maximize the common payoff function.

The completely mixed equilibrium ρ is strategically stable in virtually any sense: it is perfect, proper, regular, essential, best response stable, and stable in the sense of Mertens. However, ρ is not asymptotically stable.

The connected component G of common payoff maximizers is shown to be asymptotically stable. It also contains a stable set in the sense of Kohlberg and Mertens. The dimension of G is given by $d = (n - 1) \cdot (m - 1) - 1$, where n is the number of players, and m is the number of pure strategies. When d is even, G is essential and stable in the sense of Mertens.

However, when d is odd, we show for the case where $n = 2$ and m is odd as well as the case where $m = 2$ and n is odd, that G is not essential, not best best response stable, and hence also not stable in the sense of Mertens. Thus, these cases provide examples of games where evolutionary stability and the more demanding notions of strategic stability make mutually exclusive predictions. And more interestingly, since the difference in predictions was already observed in several examples, we establish an alternating behavior where, at least for the more demanding types of strategic stability, the stability of G depends on the dimensionality of G . This is in line with the conjecture of Swinkels that the topology of a Nash component, specifically its Euler characteristic, is a decisive factor for strategic stability.

One may wonder whether there are game theoretic considerations that conclusively explain why G ought to be stable, or whether on the contrary there are arguments that explain why G should not be stable. However, no matter what one's stance is in these matters, the alternating behavior of the more demanding notions of strategic stability seems difficult to justify.²¹ In our opinion our results, if nothing else, highlight a distinctive feature of these stronger versions of strategic stability, and any conclusive purely game theoretic foundation of strategic stability will have to explain this phenomenon. However, the correct interpretation of our results is still far from settled, and further research is definitely called for.

References

- [1] Balkenborg D, and Schlag KH (2007) On the evolutionary selection of sets of Nash equilibria. *Journal of Economic Theory* 133: 295–315
- [2] Demichelis S (2000) On the index and asymptotic stability of dynamics. CORE DP 2000/18
- [3] Demichelis S, and Germano F (2000) On the indices of zeros of Nash fields. *Journal of Economic Theory* 94: 192–218
- [4] Demichelis S, and Ritzberger K (2003) From evolutionary to strategic stability. *Journal of Economic Theory* 113: 51–75
- [5] Govindan S, and Wilson R (2005) Essential equilibria. *PNAS* 102: 15706–15711

²¹Nevertheless, one of the anonymous referees communicated to us a remarkable observation in this context. The referee showed that any solution that satisfies Backward Induction together with the Small Worlds axiom (Govindan and Wilson [6]) cannot select the component G in the 3×3 minimal diversity game.

-
- [6] Govindan S, and Wilson R (2012) Axiomatic equilibrium selection for generic two-player games. *Econometrica* 80: 1639–1699
 - [7] Güth W, Schmittberger R, and Schwarze B (1982) An experimental analysis of ultimatum bargaining *Journal of Economic Behavior and Organization* 3: 367–388
 - [8] Harsanyi JC (1973) Oddness of the number of equilibrium points: a new proof. *International Journal of Game Theory* 2: 235–250
 - [9] Harsanyi JC, and Selten R (1988) A general theory of equilibrium selection in games. M.I.T. Press, Cambridge Mass.
 - [10] Hart S (2008) Discrete Colonel Blotto and General Lotto games. *International Journal of Game Theory* 36: 441–460.
 - [11] Hillas J (1990) On the definition of the strategic stability of equilibria. *Econometrica* 58: 1365–1390
 - [12] Hillas J, Jansen M, Potters J, and Vermeulen D (2001) On the relation among some definitions of strategic stability. *Mathematics of Operations Research* 26: 611–635
 - [13] Hofbauer J, and Sigmund K (1988) The theory of evolution and dynamical systems. Cambridge University Press, Cambridge
 - [14] Hofbauer J, and Sigmund K (1998) Evolutionary games and population dynamics. Cambridge University Press, Cambridge
 - [15] Hofbauer J, and Swinkels J (1995) A Universal Shapley-Example. Mimeo
 - [16] Jiang Jia-He (1962) Essential fixed points of the multivalued mappings. *Scientia Sinica* 11: 293–298
 - [17] Kohlberg E, and Mertens J-F (1986) On the strategic stability of equilibria. *Econometrica* 54: 1003–1037
 - [18] Kojima M, Okada A, and Shindoh S (1985) Strongly stable equilibrium points of n -person non-cooperative games. *Mathematics of Operations Research* 10: 650–663
 - [19] Maynard Smith J, and Price GR (1973) The logic of animal conflict. *Nature* 246: 15–18
 - [20] McKelvey RD, and Palfrey TR (1992) An experimental study of the centipede game. *Econometrica* 60: 803–836

-
- [21] Mertens J-F (1989) Stable equilibria – a reformulation, part I. *Mathematics of Operations Research* 14: 575–624
 - [22] Mertens J-F (1991) Stable equilibria – a reformulation, part II: Discussion of the definition and further results. *Mathematics of Operations Research* 16: 694–753
 - [23] Monderer D, and Shapley L (1996) Potential games. *Games and Economic Behavior* 14: 124–143
 - [24] Ritzberger K (1994) The theory of normal form games from the differentiable view point. *International Journal of Game Theory* 23: 207–236
 - [25] Swinkels J (1992) Evolution and strategic stability: from Maynard Smith to Kohlberg and Mertens *Journal of Economic Theory* 57: 333–342
 - [26] Swinkels J 1993 Adjustment Dynamics and Rational Play in Games. *Games and Economic Behavior* 5: 455–484
 - [27] Taylor PD (1979) Evolutionary stable strategies with two types of players. *Journal of Applied Probability* 16: 76–83
 - [28] Taylor PD, and Jonker LB (1978) Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences* 40: 145–156
 - [29] van Damme E (1996) *Stability and perfection of Nash equilibria*. Springer-Verlag Publishers
 - [30] Wu Wen-Tsün and Jiang Jia-He (1962) Essential equilibrium points of n-person non-cooperative games. *Scientia Sinica* 11: 1307–1322