Non-equivalence between all and canonical elaborations

Satoru Takahashi¹

Received: 30 April 2019 / Accepted: 13 June 2019 / Published online: 2 September 2019

© The Author(s) 2019, corrected publication 2020

Abstract

In a minimal diversity game, I show that the set of potential maximizers is robust to canonical elaborations, but not to all elaborations. This is the first example to demonstrate that the set-valued notion of robustness to canonical elaborations is strictly weaker than the set-valued notion of robustness to all elaborations.

Keywords Robustness · Elaborations · Canonical elaborations · Minimal diversity games · Potential games

1 Introduction

Kajii and Morris (1997a) introduce the notion of robustness to all elaborations. They say that an equilibrium of a complete-information game is robust to all elaborations if every nearby incomplete-information game, called an elaboration, admits a Bayesian Nash equilibrium in a neighborhood of the equilibrium. In their unpublished paper, Kajii and Morris (1997b) propose other notions of robustness based on various subclasses of elaborations, among which is the class of canonical elaborations. A canonical elaboration is an elaboration where each type has either the same payoff function as in the complete-information game (a “normal” type) or a strictly dominant action (a “commitment” type). Canonical elaborations have played an important role in the literature. In fact, to date, whenever the literature establishes the non-robustness of a given equilibrium to all elaborations, it does so by constructing canonical elaborations with no Bayesian Nash equilibrium near the equilibrium. Examples include Rubinstein (1989), Example 3.1 and Lemma 5.5 in Kajii and Morris (1997a), Oyama and Takahashi (2011), and Section 5.1.2 in Haimanko and Kajii (2016).

Subsequently, Ui (2001) shows that if a game admits a potential with a unique maximizer, then the potential maximizer is robust to canonical elaborations. However, he neither proves nor disproves whether the potential maximizer is robust to all elaborations, and leaves open whether robustness to

---

✉ Satoru Takahashi
ecsst@nus.edu.sg

¹ National University of Singapore, Singapore, Singapore

canonical elaborations is a strictly weaker notion than robustness to all elaborations (Ui 2001, footnote 6). Morris and Ui (2005) extend the two notions of robustness—robustness to all elaborations and to canonical elaborations—to the respective set-valued notions, and show that the set of (generalized) potential maximizers is robust to canonical elaborations. Nonetheless, Ui's problem remains open (Morris and Ui 2005, footnote 9). Recently, Pram (2019) replaces the solution concept used in elaborations, from Bayesian Nash equilibrium to agent-normal-form correlated equilibrium, and shows that the two notions of robustness are equivalent.¹

In this paper, I show that in a minimal diversity game introduced by Balkenborg and Vermeulen (2016), the set of potential maximizers is robust to canonical elaborations, but not to all elaborations. To the best of my knowledge, this is the first example to demonstrate that the set-valued notion of robustness to canonical elaborations is strictly weaker than the set-valued notion of robustness to all elaborations. This example also implies that Pram's (2019) equivalence result, at least if it is extended to the respective set-valued robustness notions, relies on the use of correlation devices.

2 Set-valued robustness

A complete-information game consists of a finite set $N = \{1, 2, \dots, n\}$ of players, a finite set $A = \prod_{i \in N} A_i$ of action profiles, and a payoff function profile $\mathbf{g} = (g_i)_{i \in N}$, $g_i : A \rightarrow \mathbb{R}$ for $i \in N$. Suppressing N and A , I denote this game by $\mathbf{g}$. A mixed action profile $\alpha = (\alpha_i)_{i \in N}$, $\alpha_i \in \Delta(A_i)$ for $i \in N$, is an equilibrium of $\mathbf{g}$ if

$$g_i(\alpha) - g_i(a_i, \alpha_{-i}) \geq 0$$

for any $i \in N$ and $a_i \in A_i$.² Here, the domain of g_i is extended to mixed action profiles, and $\alpha_{-i} = (\alpha_j)_{j \neq i}$.

An elaboration of $\mathbf{g}$ is an incomplete-information game consisting of the same sets of players and action profiles as $\mathbf{g}$ as well as a countable set $T = \prod_{i \in N} T_i$ of type profiles, a bounded type-dependent payoff function profile $\mathbf{u} = (u_i)_{i \in N}$, $u_i : A \times T \rightarrow \mathbb{R}$ for $i \in N$, and a common prior $P \in \Delta(T)$. Let $T_{-i} = \prod_{j \neq i} T_j$. Without loss of generality, I assume that $P(\{t_i\} \times T_{-i}) > 0$ for any $i \in N$ and $t_i \in T_i$, and hence the conditional probability $P_i(t_{-i}|t_i) := P(t_i, t_{-i})/P(\{t_i\} \times T_{-i})$ is well defined. I denote this game by $(\mathbf{u}, P)$.

Given elaboration $(\mathbf{u}, P)$, let

$$T_i^{u_i} = \{t_i \in T_i \mid u_i(a, t_i, t_{-i}) = g_i(a) \text{ for all } a \in A \text{ and } t_{-i} \in T_{-i}\}$$

¹ Pram (2019) states his equivalence result for the singleton-valued robustness notions, but it easily extends to the respective set-valued robustness notions for any nonempty, closed, and convex set of action distributions.

² Except for action profiles in games $\mathbf{h}$ and $\tilde{\mathbf{h}}$ (see footnote 7 below), I abuse notations and denote a profile of actions or types by $x = (x_i)_{i \in N}$ and (x_i, x_{-i}) interchangeably.

be the set of all types of player i who know that their payoffs are given by g_i . Let $T^u = \prod_{i \in N} T_i^{u_i}$. I say that $(\mathbf{u}, P)$ is an ε -elaboration of $\mathbf{g}$ if $P(T^u) = 1 - \varepsilon$. I also say that $(\mathbf{u}, P)$ is a canonical elaboration of $\mathbf{g}$ if every type in $T_i \setminus T_i^{u_i}$ has a strictly dominant action.

A behavioral strategy profile $\sigma = (\sigma_i)_{i \in N}$, $\sigma_i : T_i \rightarrow \Delta(A_i)$, is a Bayesian Nash equilibrium of $(\mathbf{u}, P)$ if

$$\sum_{t_{-i} \in T_{-i}} P_i(t_{-i} | t_i) (u_i(\sigma(t), t) - u_i(a_i, \sigma_{-i}(t_{-i}), t)) \geq 0$$

for any $i \in N$, $t_i \in T_i$, and $a_i \in A_i$, where the domain of $u_i(\cdot, t)$ is extended to mixed action profiles, and $\sigma_{-i}(t_{-i}) = (\sigma_j(t_j))_{j \neq i}$. A strategy profile σ induces an action distribution $\sigma_P \in \Delta(A)$ by

$$\sigma_P(a) = \sum_{t \in T} P(t) \prod_{i \in N} \sigma_i(t_i)(a_i)$$

for each $a \in A$. For action distributions $\mu, \nu \in \Delta(A)$ and a nonempty and closed set $\mathcal{E} \subseteq \Delta(A)$, I denote $\|\mu - \nu\| = \max_{a \in A} |\mu(a) - \nu(a)|$ and $d(\mu, \mathcal{E}) = \min_{\nu \in \mathcal{E}} \|\mu - \nu\|$.

Definition 1 (Morris and Ui 2005) A set of action distributions, $\mathcal{E} \subseteq \Delta(A)$, is robust to all (resp., canonical) elaborations in $\mathbf{g}$ if it is nonempty and closed, and for any $\delta > 0$, there exists $\bar{\varepsilon} > 0$ such that any (resp., canonical) ε -elaboration $(\mathbf{u}, P)$ of $\mathbf{g}$ with $\varepsilon < \bar{\varepsilon}$ admits a Bayesian Nash equilibrium σ such that $d(\sigma_P, \mathcal{E}) \leq \delta$.

With some abuse of terminology, I say that an action distribution $\mu \in \Delta(A)$ is robust to all (resp., canonical) elaborations when the corresponding singleton set $\{\mu\}$ is robust to all (resp., canonical) elaborations; a pure action profile $a^* \in A$ is robust to all (resp., canonical) elaborations when the Dirac measure on a^* is robust to all (resp., canonical) elaborations.

Neither a full characterization of robustness to all elaborations nor to canonical elaborations has been known so far. Instead, the literature has identified several sufficient conditions as well as necessary conditions for robustness. (1) Most of the known sufficient conditions, except Ui (2001) and some of the conditions in Morris and Ui (2005), are for robustness to all elaborations, such as a unique correlated equilibrium (Kajii and Morris 1997a, Proposition 3.2), a $\mathbf{p}$ -dominant equilibrium with $\sum_{i \in N} p_i < 1$ (Kajii and Morris 1997a, Proposition 5.3), and a monotone potential maximizer with either $\mathbf{g}$ or the monotone potential being supermodular (Morris and Ui 2005, Proposition 2). Obviously, the two robustness notions, i.e., robustness to all elaborations and to canonical elaborations, are equivalent under any of these conditions. (2) Also, having no other strict $\mathbf{p}$ -dominant equilibrium with $\sum_{i \in N} p_i \leq 1$ is a necessary condition for robustness to canonical elaborations.³ Thus, the two robustness notions are equivalent under its negation, i.e., having another strict $\mathbf{p}$ -dominant equilibrium with $\sum_{i \in N} p_i \leq 1$. (3) For almost every binary-action

³ This follows from scrutinizing Lemma 5.5 in Kajii and Morris (1997a) and its proof.

supermodular game, if an action profile is not a monotone potential maximizer, then it is not robust to canonical elaborations (Oyama and Takahashi 2019). Combined with Morris and Ui (2005, Proposition 2), this result implies that the two robustness notions are generically equivalent in binary-action supermodular games. (4) The two robustness notions are equivalent also for the largest or smallest action profile of a supermodular game.⁴

3 A minimal diversity game

Balkenborg and Vermeulen (2016) introduce the class of minimal diversity games with n players and m actions.⁵ Here, I focus on $n = 3$ and $m = 2$. That is, there are three players, $N = \{1, 2, 3\}$, and each player $i \in N$ has two pure actions, $A_i = \{0, 1\}$. In this section, I use player indices in modulo 3, e.g., $2 + 2 = 1$, and identify each mixed action $\alpha_i \in \Delta(\{0, 1\})$ with the probability of taking pure action 1, $x_i = \alpha_i(1) \in [0, 1]$. Players have common payoffs $g_i : A \rightarrow \mathbb{R}$ given by

$$g_i(a) = \begin{cases} 0 & \text{if } a_1 = a_2 = a_3, \\ 1 & \text{otherwise.} \end{cases}$$

This game, denoted by $\mathbf{g} = (g_1, g_2, g_3)$, admits a potential, say g_1 (Monderer and Shapley 1996). The best response functions in $\mathbf{g}$ are summarized as follows:

Lemma 1 *In game $\mathbf{g}$, given the opponents' mixed actions $x_{i+1}, x_{i+2} \in [0, 1]$,*

- *if $x_{i+1} + x_{i+2} < 1$ (resp., > 1), then player i strictly prefers action 1 (resp., 0) to action 0 (resp., 1);*
- *if $x_{i+1} + x_{i+2} = 1$, then player i is indifferent between the two actions.*

It is easy to see that the minimal diversity game $\mathbf{g}$ has two kinds of equilibria: randomization with equal probabilities, $(x_1, x_2, x_3) = (1/2, 1/2, 1/2)$, and mixed action profiles in the form of $(x_1, x_2, x_3) = (0, x, 1)$, $0 \leq x \leq 1$, and their permutations.

Denote $\mathcal{E} := \Delta(A \setminus \{(0, 0, 0), (1, 1, 1)\})$. Note that $\mathcal{E}$ is the set of all distributions over the set of potential maximizers. Note also that the equilibrium of the first kind, $(x_1, x_2, x_3) = (1/2, 1/2, 1/2)$, induces the uniform distribution over A , which does not belong to $\mathcal{E}$. On the other hand, all equilibria of the second kind induce action distributions in $\mathcal{E}$.

⁴ I believe that this is a “folklore” theorem among researchers on this topic. For the sake of completeness, a proof is reported in the Appendix.

⁵ Balkenborg and Vermeulen (2016) use this class of games to illustrate the differences between evolutionary stability and strategic stability. In particular, they show that the set of potential maximizers in mixed action profiles is asymptotically stable under a large class of evolutionary dynamics and KM-stable (Kohlberg and Mertens 1986), but that its essentiality (Wu and Jiang 1962), best response stability (Hillas 1990), and Mertens stability (Mertens 1989, 1991) depend on the parity of m and n .

Proposition 1 $\mathcal{E}$ is robust to canonical elaborations in $\mathbf{g}$.

Proof Follows from an extension of Ui (2001) and Morris and Ui (2005, Theorem 4) to non-product-set potential maximizers. $\square$

Proposition 2 $\mathcal{E}$ is not robust to all elaborations in $\mathbf{g}$.

Proof I use the following game $\mathbf{h} = (h_\alpha, h_\beta, *)$ among players α , β , and γ as a building block of my construction of elaborations:

$$\mathbf{h} = \begin{array}{c} \begin{array}{cc} & 0 & 1 \\ 0 & \begin{array}{|cc|} \hline 1, 0, * & 0, 1, * \\ \hline \end{array} \\ 1 & \begin{array}{|cc|} \hline 0, 2, * & 2, 0, * \\ \hline \end{array} \end{array} & \begin{array}{c} \begin{array}{cc} & 0 & 1 \\ 0 & \begin{array}{|cc|} \hline 2, 0, * & 0, 2, * \\ \hline \end{array} \\ 1 & \begin{array}{|cc|} \hline 0, 1, * & 1, 0, * \\ \hline \end{array} \end{array} \end{array},$$

where player α chooses a row, player β chooses a column, and player γ chooses a matrix. Player γ 's payoffs are irrelevant to my construction, and hence are not specified. I denote by $\mathbf{h}^x$ the induced two-player game between players α and β given player γ 's mixed action $x \in [0, 1]$:

$$\mathbf{h}^x = \begin{array}{c} \begin{array}{cc} & 0 & 1 \\ 0 & \begin{array}{|cc|} \hline 1+x, 0 & 0, 1+x \\ \hline \end{array} \\ 1 & \begin{array}{|cc|} \hline 0, 2-x & 2-x, 0 \\ \hline \end{array} \end{array} \end{array}.$$

(Recall my convention that “mixed action x ” refers to the mixed action that takes pure action 0 with probability $1-x$ and 1 with probability x .)

I also construct another game $\tilde{\mathbf{h}} = (\tilde{h}_\alpha, \tilde{h}_\beta, *)$ by relabeling player β 's action 0 as action 1, and action 1 as action 0, i.e.,

$$\tilde{\mathbf{h}} = \begin{array}{c} \begin{array}{cc} & 0 & 1 \\ 0 & \begin{array}{|cc|} \hline 0, 1, * & 1, 0, * \\ \hline \end{array} \\ 1 & \begin{array}{|cc|} \hline 2, 0, * & 0, 2, * \\ \hline \end{array} \end{array} & \begin{array}{c} \begin{array}{cc} & 0 & 1 \\ 0 & \begin{array}{|cc|} \hline 0, 2, * & 2, 0, * \\ \hline \end{array} \\ 1 & \begin{array}{|cc|} \hline 1, 0, * & 0, 1, * \\ \hline \end{array} \end{array} \end{array}.$$

I denote by $\tilde{\mathbf{h}}^x$ the induced two-player game given player γ 's mixed action $x \in [0, 1]$:

$$\tilde{\mathbf{h}}^x = \begin{array}{c} \begin{array}{cc} & 0 & 1 \\ 0 & \begin{array}{|cc|} \hline 0, 1+x & 1+x, 0 \\ \hline \end{array} \\ 1 & \begin{array}{|cc|} \hline 2-x, 0 & 0, 2-x \\ \hline \end{array} \end{array} \end{array}.$$

$\square$

Fig. 1 Interactions among t_i^* , t_{i+1}^* , t_{i+2}^* , $t_{i+1}^{i,f}$, $t_{i+2}^{i,f}$, $t_i^{i,s}$, and $t_{i+2}^{i,s}$

Both $\mathbf{h}^x$ and $\tilde{\mathbf{h}}^x$ are 2×2 games with unique mixed action equilibria as follows:

Lemma 2

1. Game $\mathbf{h}^x$ has a unique equilibrium $((1+x)/3, (1+x)/3)$.
2. Game $\tilde{\mathbf{h}}^x$ has a unique equilibrium $((1+x)/3, (2-x)/3)$.

Note that in both games $\mathbf{h}^x$ and $\tilde{\mathbf{h}}^x$, player α 's equilibrium action $(1+x)/3$ changes monotonically with respect to player γ 's action x , and in particular, player α takes action 0 more (resp., less) likely than action 1 if and only if player γ takes action 0 more (resp., less) likely than action 1. Also note that in game $\tilde{\mathbf{h}}^x$, players α and β take mixed actions with complementary probabilities, i.e., $(1+x)/3 + (2-x)/3 = 1$. These properties will be used later in the proof of Lemma 3.

Inspired by Balkenborg and Vermeulen (2016, Section 7.1), for each ε , $0 < \varepsilon < 1$, I construct the following elaboration $(\mathbf{u}, P)$ of $\mathbf{g}$. Each player has five types, $T_i = \{t_i^*, t_i^{i+1,f}, t_i^{i+2,f}, t_i^{i,s}, t_i^{i+1,s}\}$, where t_i^* is a “normal” type with the same payoff function as g_i , and all other types are divided into the “first generation” (with superscript f) and the “second generation” (with superscript s).⁶

The prior $P \in \Delta(T)$ is given by

$$P(t) = \begin{cases} 1 - \varepsilon & \text{if } (t_1, t_2, t_3) = (t_1^*, t_2^*, t_3^*), \\ \varepsilon/9 & \text{if } (t_i, t_{i+1}, t_{i+2}) = (t_i^*, t_{i+1}^{i,f}, t_{i+2}^{i,f}), (t_i^{i,s}, t_{i+1}^{i,f}, t_{i+2}^{i,s}), \\ & \text{or } (t_i^{i,s}, t_{i+1}^*, t_{i+2}^{i,s}) \text{ with some } i \in N, \\ 0 & \text{otherwise.} \end{cases}$$

Since $0 < \varepsilon < 1$, there are ten type profiles that are assigned with strictly positive probabilities. In words, normal type t_i^* interacts with the other normal types (t_{i+1}^*, t_{i+2}^*) , first-generation types $(t_{i+1}^{i,f}, t_{i+2}^{i,f})$, and second-generation types $(t_{i+1}^{i+2,s}, t_{i+2}^{i+2,s})$; first-generation types $t_{i+1}^{i,f}$ and $t_{i+2}^{i,f}$ interact with normal type t_i^* , and $t_{i+1}^{i,f}$ also interacts with second-generation types $(t_i^{i,s}, t_{i+2}^{i,s})$; second-generation types $t_i^{i,s}$ and $t_{i+2}^{i,s}$ interact with t_{i+1}^* and $t_{i+1}^{i,f}$. See Fig. 1.

⁶ As it will be clear in the example before Lemma 3 and the proofs of Lemmas 3 and 4, I will construct the elaboration $(\mathbf{u}, P)$ in such a way that once normal type t_i^* 's action is given, it uniquely determines the equilibrium actions of first-generation types of superscript i , $t_{i+1}^{i,f}$ and $t_{i+2}^{i,f}$, and in turn, $t_{i+1}^{i,f}$'s action uniquely determines the equilibrium actions of second-generation types of superscript i , $t_i^{i,s}$ and $t_{i+2}^{i,s}$.

Given the prior P , each type has the following posterior:

$$\begin{aligned}
 P_i(t_{-i} | t_i^*) &= \begin{cases} (9 - 9\varepsilon)/(9 - 7\varepsilon) & \text{if } (t_{i+1}, t_{i+2}) = (t_{i+1}^*, t_{i+2}^*), \\ \varepsilon/(9 - 7\varepsilon) & \text{if } (t_{i+1}, t_{i+2}) = (t_{i+1}^{if}, t_{i+2}^{if}) \text{ or } (t_{i+1}^{is}, t_{i+2}^{is}), \\ 0 & \text{otherwise,} \end{cases} \\
 P_{i+1}(t_{-(i+1)} | t_{i+1}^{if}) &= \begin{cases} 1/2 & \text{if } (t_i, t_{i+2}) = (t_i^*, t_{i+2}^{if}) \text{ or } (t_i^{is}, t_{i+2}^{is}), \\ 0 & \text{otherwise,} \end{cases} \\
 P_{i+2}(t_{-(i+2)} | t_{i+2}^{if}) &= \begin{cases} 1 & \text{if } (t_i, t_{i+1}) = (t_i^*, t_{i+1}^{if}), \\ 0 & \text{otherwise,} \end{cases} \\
 P_i(t_{-i} | t_i^{is}) &= \begin{cases} 1/2 & \text{if } (t_{i+1}, t_{i+2}) = (t_{i+1}^*, t_{i+2}^{is}) \text{ or } (t_{i+1}^{if}, t_{i+2}^{is}), \\ 0 & \text{otherwise,} \end{cases} \\
 P_{i+2}(t_{-(i+2)} | t_{i+2}^{is}) &= \begin{cases} 1/2 & \text{if } (t_i, t_{i+1}) = (t_i^{is}, t_{i+1}^*) \text{ or } (t_i^{is}, t_{i+1}^{if}), \\ 0 & \text{otherwise.} \end{cases}
 \end{aligned}$$

Payoffs $\mathbf{u} = (u_1, u_2, u_3)$, $u_i : A \times T \rightarrow \mathbb{R}$, are given as follows:

- Each normal type t_i^* has the complete-information payoff function g_i regardless of the other players' types. That is,

$$u_i(a, t_i^*, t_{-i}) = g_i(a)$$

for any $a \in A$ and $t_{-i} \in T_{-i}$.

- First-generation types t_{i+1}^{if} and t_{i+2}^{if} play game $\tilde{\mathbf{h}}$ in the roles of players α and β , respectively, when they interact with normal type t_i^* in the role of player γ . Also, type t_{i+1}^{if} always receives zero payoff when he interacts with second-generation types (t_i^{is}, t_{i+2}^{is}) . That is,

$$\begin{aligned}
 u_{i+1}(a, t_{i+1}^{if}, t_{-(i+1)}) &= \begin{cases} \tilde{h}_\alpha(a_{i+1}, a_{i+2}, a_i) & \text{if } (t_i, t_{i+2}) = (t_i^*, t_{i+2}^{if}), \\ 0 & \text{if } (t_i, t_{i+2}) = (t_i^{is}, t_{i+2}^{is}), \\ \text{arbitrary} & \text{otherwise,} \end{cases} \\
 u_{i+2}(a, t_{i+2}^{if}, t_{-(i+2)}) &= \begin{cases} \tilde{h}_\beta(a_{i+1}, a_{i+2}, a_i) & \text{if } (t_i, t_{i+1}) = (t_i^*, t_{i+1}^{if}), \\ \text{arbitrary} & \text{otherwise} \end{cases}
 \end{aligned}$$

for any $a \in A$.⁷

⁷ In games $\mathbf{h}$ and $\tilde{\mathbf{h}}$, I fix the order of players α , β , and γ and denote action profiles in the form of (player α 's action, player β 's action, player γ 's action). For example, $\tilde{h}_\alpha(a_{i+1}, a_{i+2}, a_i)$ means the payoff of player $i+1$ in game $\tilde{\mathbf{h}}$ when he takes action a_{i+1} in the role of player α , and players i and $i+2$ take actions a_i and a_{i+2} in the roles of players γ and β , respectively.

- Second-generation types $t_i^{i,s}$ and $t_{i+2}^{i,s}$ always receive zero payoff when they interact with normal type t_{i+1}^* . Also, $t_i^{i,s}$ and $t_{i+2}^{i,s}$ play game **h** in the roles of players α and β , respectively, when they interact with first-generation type $t_{i+1}^{i,f}$ in the role of player γ . That is,

$$u_i(a, t_i^{i,s}, t_{-i}) = \begin{cases} 0 & \text{if } (t_{i+1}, t_{i+2}) = (t_{i+1}^*, t_{i+2}^{i,s}), \\ h_\alpha(a_i, a_{i+2}, a_{i+1}) & \text{if } (t_{i+1}, t_{i+2}) = (t_{i+1}^{i,f}, t_{i+2}^{i,s}), \\ \text{arbitrary} & \text{otherwise,} \end{cases}$$

$$u_{i+2}(a, t_{i+2}^{i,s}, t_{-(i+2)}) = \begin{cases} 0 & \text{if } (t_i, t_{i+1}) = (t_i^{i,s}, t_{i+1}^*), \\ h_\beta(a_i, a_{i+2}, a_{i+1}) & \text{if } (t_i, t_{i+1}) = (t_i^{i,s}, t_{i+1}^{i,f}), \\ \text{arbitrary} & \text{otherwise} \end{cases}$$

for any $a \in A$.

Note that $(\mathbf{u}, P)$ is an ε -elaboration of **g**. Thus, to prove Proposition 2, it suffices to show that for any small $\varepsilon > 0$, $(\mathbf{u}, P)$ has no Bayesian Nash equilibrium that induces an action distribution in a neighborhood of $\mathcal{E}$.

Before proceeding to a formal proof, let me explain why equilibria in the form of $(0, x, 1)$ of the complete-information game **g**, which induce action distributions in $\mathcal{E}$, disappear in the elaboration $(\mathbf{u}, P)$. Consider, for example, the mixed action profile $(x_1, x_2, x_3) = (0, 3/4, 1)$. Recall that $(0, 3/4, 1)$ is an equilibrium of **g** where players 1 and 3 have strict incentives to take actions 0 and 1, respectively, and player 2 is indifferent between the two actions. In the elaboration $(\mathbf{u}, P)$, however, there is no Bayesian Nash equilibrium $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ that satisfies $(\sigma_1(t_1^*), \sigma_2(t_2^*), \sigma_3(t_3^*)) = (0, 3/4, 1)$. To see this, first note that since first-generation types $t_2^{1,f}$ and $t_3^{1,f}$ play game $\tilde{\mathbf{h}}$ with normal type t_1^* , and type $t_2^{1,f}$'s incentive is not affected by the interaction with second-generation types $(t_1^{1,s}, t_3^{1,s})$, by Lemma 2(2), types $t_2^{1,f}$ and $t_3^{1,f}$ take the unique equilibrium of $\tilde{\mathbf{h}}^0$, i.e., $\sigma_2(t_2^{1,f}) = 1/3$ and $\sigma_3(t_3^{1,f}) = 2/3$. Note that $\sigma_2(t_2^{1,f}) = 1/3 < 1/2$ because $\sigma_1(t_1^*) = 0 < 1/2$ and player α 's equilibrium action in $\tilde{\mathbf{h}}^x$ is increasing in x . Similarly, since first-generation types $t_3^{2,f}$ and $t_1^{2,f}$ play game $\tilde{\mathbf{h}}$ with normal type t_2^* , and type $t_3^{2,f}$'s incentive is not affected by the interaction with second-generation types $(t_2^{2,s}, t_1^{2,s})$, by Lemma 2(2), types $t_3^{2,f}$ and $t_1^{2,f}$ take the unique equilibrium of $\tilde{\mathbf{h}}^{3/4}$, i.e., $\sigma_3(t_3^{2,f}) = 7/12$ and $\sigma_1(t_1^{2,f}) = 5/12$. Note that types $t_3^{2,f}$ and $t_1^{2,f}$ take action 1 with complementary probabilities, i.e., $7/12 + 5/12 = 1$, thanks to the relabeling of actions in $\tilde{\mathbf{h}}$. Second, since second-generation types $t_1^{1,s}$ and $t_3^{1,s}$ play game **h** with $t_2^{1,f}$, and types $t_1^{1,s}$'s and $t_3^{1,s}$'s incentives are not affected by the interaction with the normal type t_2^* , by Lemma 2(1), types $t_1^{1,s}$ and $t_3^{1,s}$ take the unique equilibrium of $\mathbf{h}^{1/3}$, i.e., $\sigma_1(t_1^{1,s}) = \sigma_3(t_3^{1,s}) = 4/9$, which is below $1/2$ because $\sigma_2(t_2^{1,f}) = 1/3 < 1/2$ and

player α 's equilibrium action in $\mathbf{h}^x$ is increasing in x . Lastly, consider the incentive of normal type t_2^* , who plays game $\mathbf{g}$ with (t_1^*, t_3^*) , (t_1^{2f}, t_3^{2f}) , and (t_1^{1s}, t_3^{1s}) . Since types t_1^* and t_3^* play different pure actions, type t_2^* is indifferent between actions 0 and 1 in the interaction with (t_1^*, t_3^*) . Type t_2^* is also indifferent between the two actions in the interaction with (t_1^{2f}, t_3^{2f}) because $\sigma_1(t_1^{2f}) + \sigma_3(t_3^{2f}) = 1$. Moreover, since t_1^{1s} and t_3^{1s} take action 0 more likely than action 1, type t_2^* strictly prefers action 1 to action 0 in the interaction with (t_1^{1s}, t_3^{1s}) . Overall, t_2^* strictly prefers action 1 to action 0, which contradicts with $\sigma_2(t_2^*) = 3/4$.

The following two lemmas generalize the above argument.

Lemma 3 Let $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ be a Bayesian Nash equilibrium of $(\mathbf{u}, P)$.

1. If $\sigma_i(t_i^*) < 1/2$ (resp., $> 1/2$) and $\sigma_i(t_i^*) + \sigma_{i+2}(t_{i+2}^*) \leq 1$ (resp., ≥ 1), then $\sigma_{i+1}(t_{i+1}^*) = 1$ (resp., 0).
2. In particular, if $\sigma_i(t_i^*) = 0$ (resp., 1), then $\sigma_{i+1}(t_{i+1}^*) = 1$ (resp., 0).

Proof of Lemma 3 Since part (2) follows immediately from part (1), I will show only part (1). By symmetry between actions 0 and 1, I assume without loss of generality that $\sigma_i(t_i^*) < 1/2$ and $\sigma_i(t_i^*) + \sigma_{i+2}(t_{i+2}^*) \leq 1$. Denote $x := \sigma_i(t_i^*)$ and $y := \sigma_{i+1}(t_{i+1}^*)$. First, since first-generation types t_{i+1}^{if} and t_{i+2}^{if} play game $\tilde{\mathbf{h}}$ with t_i^* , and t_{i+1}^{if} 's incentive is not affected by the interaction with second-generation types (t_i^{1s}, t_{i+2}^{1s}) , by Lemma 2(2), types t_{i+1}^{if} and t_{i+2}^{if} take the unique equilibrium of $\tilde{\mathbf{h}}^x$, i.e., $\sigma_{i+1}(t_{i+1}^{if}) = (1+x)/3$ and $\sigma_{i+2}(t_{i+2}^{if}) = (2-x)/3$. Similarly, since first-generation types t_{i+1}^{i+1f} and t_{i+2}^{i+1f} play game $\tilde{\mathbf{h}}$ with normal type t_{i+1}^* , and type t_{i+2}^{i+1f} 's incentive is not affected by the interaction with second-generation types $(t_{i+1}^{1s}, t_i^{i+1s})$, by Lemma 2(2), types t_{i+2}^{i+1f} and t_i^{i+1f} take the unique equilibrium of $\tilde{\mathbf{h}}^y$, i.e., $\sigma_{i+2}(t_{i+2}^{i+1f}) = (1+y)/3$ and $\sigma_i(t_i^{i+1f}) = (2-y)/3$. Second, since second-generation types t_i^{1s} and t_{i+2}^{1s} play game $\mathbf{h}$ with t_{i+1}^{if} , and types t_i^{1s} 's and t_{i+2}^{1s} 's incentives are not affected by the interaction with the normal type t_{i+1}^* , by Lemma 2(1), types t_i^{1s} and t_{i+2}^{1s} take the unique equilibrium of $\mathbf{h}^{(1+x)/3}$, i.e.,

$$\sigma_i(t_i^{1s}) = \sigma_{i+2}(t_{i+2}^{1s}) = \frac{1 + \frac{1+x}{3}}{3} = \frac{4+x}{9} < \frac{1}{2}, \quad (1)$$

where the inequality follows since $x < 1/2$. Lastly, consider the incentive of the next normal type t_{i+1}^* , who plays game $\mathbf{g}$ with (t_i^*, t_{i+2}^*) , $(t_i^{i+1f}, t_{i+2}^{i+1f})$, and (t_i^{1s}, t_{i+2}^{1s}) .

- Since $\sigma_i(t_i^*) + \sigma_{i+2}(t_{i+2}^*) \leq 1$, by Lemma 1, type t_{i+1}^* weakly prefers action 1 to action 0 in the interaction with (t_i^*, t_{i+2}^*) .
- Since $\sigma_i(t_i^{i+1f}) + \sigma_{i+2}(t_{i+2}^{i+1f}) = (2-y)/3 + (1+y)/3 = 1$, by Lemma 1, type t_{i+1}^* is indifferent between actions 0 and 1 in the interaction with $(t_i^{i+1f}, t_{i+2}^{i+1f})$.

- By (1) and Lemma 1, type t_{i+1}^* strictly prefers action 1 to action 0 in the interaction with $(t_i^{i,s}, t_{i+2}^{i,s})$.

Overall, t_{i+1}^* strictly prefers action 1 to action 0, and hence $\sigma_{i+1}(t_{i+1}^*) = 1$. $\square$

Lemma 4 *The elaboration $(\mathbf{u}, P)$ has a unique Bayesian Nash equilibrium $\sigma = (\sigma_1, \sigma_2, \sigma_3)$, where $\sigma_i(t_i) = 1/2$ for all $i \in N$ and $t_i \in T_i$.*

Proof of Lemma 4 First, I will prove by contradiction that $\sigma_i(t_i^*) = 1/2$ for all $i \in N$. Suppose that $\sigma_i(t_i^*) \neq 1/2$ for some $i \in N$. By the symmetry among players, I assume without loss of generality that $i = 1$ maximizes $|\sigma_i(t_i^*) - 1/2|$. By the symmetry between actions 0 and 1, I also assume without loss of generality that $\sigma_1(t_1^*) < 1/2$. Note that

$$\frac{1}{2} - \sigma_1(t_1^*) = \left| \sigma_1(t_1^*) - \frac{1}{2} \right| \geq \left| \sigma_3(t_3^*) - \frac{1}{2} \right| \geq \sigma_3(t_3^*) - \frac{1}{2},$$

and hence $\sigma_1(t_1^*) + \sigma_3(t_3^*) \leq 1$. Therefore, by Lemma 3(1), I have $\sigma_2(t_2^*) = 1$. Applying Lemma 3(2) iteratively, I have $\sigma_3(t_3^*) = 0$ and hence $\sigma_1(t_1^*) = 1$. This is a contradiction.

Second, given the equilibrium action of normal types, I can specify, as in the proof of Lemma 3, the equilibrium actions of first- and second-generation types by applying Lemma 2 iteratively. Namely, for each $i \in N$, types $t_{i+1}^{i,f}$ and $t_{i+2}^{i,f}$ take the unique equilibrium of $\tilde{\mathbf{h}}^{1/2}$, i.e., $\sigma_{i+1}(t_{i+1}^{i,f}) = (1 + 1/2)/3 = 1/2$ and $\sigma_{i+2}(t_{i+2}^{i,f}) = (2 - 1/2)/3 = 1/2$. Also, types $t_i^{i,s}$ and $t_{i+2}^{i,s}$ take the unique equilibrium of $\mathbf{h}^{1/2}$, i.e., $\sigma_i(t_i^{i,s}) = \sigma_{i+2}(t_{i+2}^{i,s}) = (1 + 1/2)/3 = 1/2$. $\square$

Note that for any ε , $0 < \varepsilon < 1$, the unique Bayesian Nash equilibrium specified in Lemma 4 induces the uniform distribution over A , which does not belong to $\mathcal{E}$. Thus the proof of Proposition 2 concludes. $\square$

Remark 1 Kajii and Morris (1997a, Remark 4) discuss a more permissive notion of ε -elaborations by not requiring normal types to *know* that their own payoffs are given by g_i , but requiring them to *believe* so with high probability. This change would make the corresponding robustness notions more stringent, and non-robustness results (such as Proposition 2) easier to hold. In fact, under this permissive notion of ε -elaborations, the set $\mathcal{E}$ would not be robust to canonical elaborations in the minimal diversity game $\mathbf{g}$. To see this, note that the complete-information game $\mathbf{v}(\varepsilon) = \mathbf{g} + \varepsilon \mathbf{v}$, where $\varepsilon > 0$ and $\mathbf{v} = (v_1, v_2, v_3)$, $v_i : A \rightarrow \mathbb{R}$, is given by

$$v_i(a) = \begin{cases} 1 & \text{if } a_i \neq a_{i-1}, \\ 0 & \text{otherwise,} \end{cases}$$

has a unique equilibrium $(1/2, 1/2, 1/2)$ (Balkenborg and Vermeulen 2016, Section 7.1). For $0 < \varepsilon < 1$, consider the following canonical elaboration $(\mathbf{u}', P')$: $T_i' = \{t_i^*, t_i^v\}$,

$$u'_i(a, t) = \begin{cases} g_i(a) & \text{if } (t_1, t_2, t_3) = (t_1^*, t_2^*, t_3^*), \\ v_i(a) & \text{if } (t_i, t_{i+1}, t_{i+2}) = (t_i^*, t_{i+1}^v, t_{i+2}^*), \\ a_i & \text{if } t_i = t_i^v, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$P'(t) = \begin{cases} 1 - \varepsilon & \text{if } (t_1, t_2, t_3) = (t_1^*, t_2^*, t_3^*), \\ \varepsilon/3 & \text{if } (t_i, t_{i+1}, t_{i+2}) = (t_i^*, t_{i+1}^v, t_{i+2}^*) \text{ with some } i \in N, \\ 0 & \text{otherwise.} \end{cases}$$

Note that the state (t_1^*, t_2^*, t_3^*) occurs with probability $1 - \varepsilon$, and at this state, each player i believes with probability $(3 - 3\varepsilon)/(3 - \varepsilon)$, which converges to 1 as $\varepsilon \rightarrow 0$, that his payoff is the same as g_i . Noting that $(\mathbf{u}', P')$ is essentially the same as Balkenborg–Vermeulen’s game $\mathbf{v}(\varepsilon/(3 - 3\varepsilon))$, it is easy to show that $(\mathbf{u}', P')$ has a unique Bayesian Nash equilibrium $\sigma_i(t_i^*) = 1/2$ and $\sigma_i(t_i^v) = 1$ for $i \in N$.

Remark 2 In the definition of elaborations, non-normal types may not know their own payoff functions, i.e., their payoffs may depend on other players’ types. But this is not essential. I can modify the minimal diversity game and the elaboration in the proof of Proposition 2, and show the non-equivalence between the notions of robustness to known-own-payoff elaborations and to canonical elaborations. A trick is to incorporate “nature” as an explicit player, where nature knows the type profile of the other players and replicates it as her action.⁸

Remark 3 If a game admits a potential that is concave in the domain of mixed action profiles, then the set of action distributions over potential maximizers is precisely the set of correlated equilibria (Neyman 1997, Theorem 1), and hence robust to all elaborations. Also, if a game admits a (monotone) potential that is supermodular, then the set of action distributions over potential maximizers is robust to all elaborations (Morris and Ui 2005, Proposition 2). The potential of the minimal diversity game is, however, neither concave nor supermodular (with respect to any total order on each action set).

Remark 4 For any canonical elaboration of a potential game, Ui defines the elaboration potential, and shows that the elaboration potential is indeed a potential of (the normal form of) the elaboration (Ui 2001, Lemma 2). This is a key step toward establishing the robustness of the potential maximizer to canonical elaborations. He

⁸ More precisely, let $\mathbf{g}$ and $(\mathbf{u}, P)$ be the minimal diversity game and the elaboration in the proof of Proposition 2, respectively. Then consider another game $\tilde{\mathbf{g}}$ with $\tilde{N} = N \cup \{0\}$, $\tilde{A}_i = A_i$ for $i \in N$, $\tilde{A}_0 = T$, $\tilde{g}_i(a, t) = u_i(a, t)$ for $i \in N$, and $\tilde{g}_0(a, t) = 1_{t=t^*}$. Let $\tilde{\mathcal{E}} = \Delta((A \setminus \{(0, 0, 0), (1, 1, 1)\}) \times \{t^*\})$. Then $\tilde{\mathcal{E}}$ is robust to canonical elaborations in $\tilde{\mathbf{g}}$, but not robust to the elaboration $(\tilde{\mathbf{u}}, \tilde{P})$ with $\tilde{T}_i = T_i$ for $i \in N$, $\tilde{T}_0 = T$, $\tilde{P}(t, t') = P(t)1_{t=t'}$, and $\tilde{u}_i : \tilde{A} \times \tilde{T}_i \rightarrow \mathbb{R}$ with $\tilde{u}_i((a, t), t'_i) = u_i(a, t)$ for $i \in N$ and $\tilde{u}_0((a, t), t') = 1_{t=t^*}$. (Note that “ t ” in $\tilde{u}_i((a, t), t'_i)$ is not a type profile in $(\tilde{\mathbf{u}}, \tilde{P})$, but an action of player 0.)

also notes that his argument can be used to show the robustness to non-canonical elaborations as long as they admit Bayesian potentials in the sense of van Heumen et al. (1996) (Ui 2001, Remark 2). However, the elaboration $(\mathbf{u}, P)$ constructed in the proof of Proposition 2 is based on non-potential games $\mathbf{h}$ and $\tilde{\mathbf{h}}$, and hence does not admit a Bayesian potential.

Remark 5 Pram (2019) establishes the equivalence between robustness to all elaborations and to canonical elaborations if agent-normal-form correlated equilibrium is used as the solution concept in elaborations. It implies, in particular, that the elaboration $(\mathbf{u}, P)$ constructed in the proof of Proposition 2 has an agent-normal-form correlated equilibrium in a neighborhood of $\mathcal{E}$. Indeed, the following is one such example: the three normal types play the uniform distribution over $A \setminus \{(0, 0, 0), (1, 1, 1)\}$, and independently, all other types randomize between actions 0 and 1 with equal probabilities. More formally, it is $\lambda \in \Delta(A_1^{T_1} \times A_2^{T_2} \times A_3^{T_3})$ given by

$$\lambda((a_\tau)_{\tau \in T_1 \cup T_2 \cup T_3}) = \begin{cases} 0 & \text{if } a_{t_1} = a_{t_2} = a_{t_3}, \\ 1/(6 \times 2^{12}) & \text{otherwise.} \end{cases}$$

Remark 6 The uniform distribution over A , induced by the completely mixed equilibrium $(1/2, 1/2, 1/2)$ of the minimal diversity game $\mathbf{g}$, is not robust to canonical elaborations. To see this, consider the following elaboration $(\mathbf{u}'', P'')$, a modification of Rubinstein's (1989) email game. For any small $\varepsilon > 0$, type spaces are given by $T_1'' = T_2'' = T_3'' = \mathbb{N}$, type $t_i = 0$ is committed to playing action 0, all other types $t_i \geq 1$ have the payoff function g_i , and the prior $P'' \in \Delta(T'')$ by

$$P''(t) \propto \begin{cases} (1 - \varepsilon)^{k-1} & \text{if } (t_1, t_2, t_3) = (k, k, k) \text{ or } (k, k, k-1) \text{ for some } k \geq 1, \\ \varepsilon(1 - \varepsilon)^{k-1} & \text{if } (t_1, t_2, t_3) = (k, 0, 0) \text{ or } (0, k, 0) \text{ for some } k \geq 1, \\ 0 & \text{otherwise} \end{cases}$$

with normalization factor $\varepsilon/(2 + 2\varepsilon)$ (so that the total probability is equal to 1). This is a $3\varepsilon/(2 + 2\varepsilon)$ -elaboration of $\mathbf{g}$. Following the standard contagion argument, one can show that $(\mathbf{u}'', P'')$ has a unique rationalizable action $\sigma_1(0) = \sigma_2(0) = \sigma_3(k) = 0$ for $k \geq 0$ and $\sigma_1(k) = \sigma_2(k) = 1$ for $k \geq 1$, which approximately induces the Dirac measure on $(1, 1, 0)$. Therefore, the uniform distribution over A is not robust to canonical elaborations in $\mathbf{g}$. In fact, the same argument implies, through symmetry, that every robust-to-canonical-elaboration set in $\mathbf{g}$ contains $\mathcal{E}$. Combined with Proposition 1, it implies that $\mathcal{E}$ is the smallest such set. Also, it follows from the proof of Proposition 2 that every robust-to-all-elaboration set in $\mathbf{g}$ contains the convex hull of the union of $\mathcal{E}$ and the uniform distribution over A .⁹

⁹ Obviously, every minimal robust-to-all-elaboration set is contained in the set of all correlated equilibria of $\mathbf{g}$, i.e., the convex hull of the union of $\mathcal{E}$, the uniform distribution over $\{(0, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0)\}$, and the uniform distribution over $\{(0, 0, 1), (0, 1, 0), (1, 0, 0), (1, 1, 1)\}$.

4 Conclusion

One of the open questions in the literature is whether robustness to canonical elaborations is a strictly weaker notion than robustness to all elaborations. In this paper, I solved this question for set-valued robustness notions affirmatively by showing that in the three-player two-action minimal diversity game, the set of potential maximizers is robust to canonical elaborations, but not to all elaborations. Of course, many questions remain open. For example, can I show the non-equivalence between the two robustness notions in “generic” games? Can I show the non-equivalence between the singleton-valued robustness notions?

Let me elaborate more. Admittedly, the minimal diversity game $\mathbf{g}$ is “doubly” non-generic; potential games are non-generic in the class of all games, and the minimal diversity game is non-generic even in the class of potential games. In particular, the minimal diversity game has six pure action profiles that tie for potential maximizers. Instead of the minimal diversity game, one can slightly perturb its payoffs and obtain a potential game $\mathbf{g}'$ with a unique potential maximizer a^* . By Ui (2001), a^* is robust to canonical elaborations in $\mathbf{g}'$. Then, can I show that a^* is not robust to all elaborations in $\mathbf{g}'$? Notice that a^* is a strict equilibrium of $\mathbf{g}'$. Thus, along any sequence of ε_k -elaborations $(\mathbf{u}_k, P_k)$ of $\mathbf{g}'$ with a bounded number of normal types and $\varepsilon_k \rightarrow 0$ as $k \rightarrow \infty$, for large k , there exists a Bayesian Nash equilibrium where all normal types play a^* (Monderer and Samet 1989; Fudenberg and Tirole 1991, Section 14.4). In particular, a^* is robust to elaborations with a unique normal type per player, such as the elaboration in the proof of Proposition 2. I have not been able to prove (or disprove) the non-robustness of a^* to elaborations with unboundedly many normal types.

Another issue is the use of small incentives in elaborations. At the end of the proof of the first statement of Lemma 3, I concluded that “Overall, t_{i+1}^* strictly prefers action 1 to action 0, and hence $\sigma_{i+1}(t_{i+1}^*) = 1$.” However, note that type t_{i+1}^* believes that he interacts with $(t_i^{i,s}, t_{i+2}^{i,s})$ with probability $\varepsilon/(9 - 7\varepsilon)$, and hence the incentive margin is of the order of ε . To make this argument non-sensitive to small incentives, one would like to construct another elaboration that has no approximate Bayesian Nash equilibrium in a neighborhood of $\mathcal{E}$. In fact, Haimanko and Kajii (2016) propose the notion of “approximate robustness” along this line.¹⁰ Once again, I do not know whether $\mathcal{E}$ is approximately robust to all elaborations or not in the minimal diversity game.¹¹ Note that the sensitivity to small incentives is closely related to the non-genericity of the minimal diversity game discussed in the previous paragraph. For example, if one were able to strengthen Ui’s (2001) result and show that the unique potential maximizer of a generic potential game is robust (or at least approximately robust) to all elaborations, then by the upper hemicontinuity of approximate robustness (Haimanko and Kajii 2016, Theorem 3), it would imply that $\mathcal{E}$ is approximately robust to all elaborations in the minimal diversity game. Conversely, if one

¹⁰ See also Chen et al. (2014) for a similar requirement in the context of the selection of rationalizable actions.

¹¹ Similarly, the argument in Remark 6 does not extend to approximate Bayesian Nash equilibria. I do not know whether the uniform distribution over A is approximately robust to all (or canonical) elaborations or not.

were able to show that $\mathcal{E}$ is not approximately robust to all elaborations in the minimal diversity game, then it would imply that Ui's (2001) result fails without the restriction to canonical elaborations, i.e., even if a game admits a potential with a unique potential maximizer, it may not be robust to all elaborations.

Acknowledgements I thank Yi-Chun Chen, Atsushi Kajii, Daisuke Oyama, Olivier Tercieux, Takashi Ui, Nicolas Vieille, and Junjie Zhou for their helpful comments and suggestions.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix: supermodular games

I say that a complete-information game $\mathbf{g}$ is *supermodular* if for each $i \in N$,

- A_i is a finite lattice with partial order $\geq_i$ and lattice operators $(\vee_i, \wedge_i)$;
- $g_i(\cdot, a_{-i})$ is supermodular for any $a_{-i} \in A_{-i}$, i.e., $g_i(a_i \vee_i a'_i, a_{-i}) + g_i(a_i \wedge_i a'_i, a_{-i}) \geq g_i(a_i, a_{-i}) + g_i(a'_i, a_{-i})$ for any $a_i, a'_i \in A_i$ and $a_{-i} \in A_{-i}$;
- g_i has increasing differences, i.e., $g_i(a_i, a_{-i}) - g_i(a'_i, a_{-i}) \geq g_i(a_i, a'_{-i}) - g_i(a'_i, a'_{-i})$ for any $a_i, a'_i \in A_i$ and $a_{-i}, a'_{-i} \in A_{-i}$ with $a_i \geq_i a'_i$ and $a_{-i} \geq_{-i} a'_{-i}$.¹²

Let $\bar{a} = (\bar{a}_i)_{i \in N}$ and $\underline{a} = (\underline{a}_i)_{i \in N}$ denote the largest and smallest action profiles, respectively

Proposition 3 Suppose that $\mathbf{g}$ is supermodular and $a^* = \bar{a}$ or $\underline{a}$. Then a^* is robust to all elaborations in $\mathbf{g}$ if and only if it is robust to canonical elaborations in $\mathbf{g}$.

Proof I assume $a^* = \bar{a}$ without loss of generality. Suppose that $\bar{a}$ is robust to canonical elaborations in $\mathbf{g}$. Fix any $\delta > 0$, and let $\bar{\varepsilon} > 0$ be such that any canonical ε -elaboration of $\mathbf{g}$ with $\varepsilon < \bar{\varepsilon}$ has a Bayesian Nash equilibrium that plays $\bar{a}$ with ex-ante probability at least $1 - \delta$.

Fix any $\varepsilon < \bar{\varepsilon}$ and any ε -elaboration $(\mathbf{u}, P)$ of $\mathbf{g}$. Let $(\underline{\mathbf{u}}, P)$ be another elaboration of $\mathbf{g}$ such that it has the same type space as $(\mathbf{u}, P)$, but that each non-normal type of player i has $\underline{a}_i$ as a strictly dominant action. Since $(\underline{\mathbf{u}}, P)$ is a canonical ε -elaboration of $\mathbf{g}$, it has a Bayesian Nash equilibrium $\underline{\sigma} = (\underline{\sigma}_i)_{i \in N}$ that plays $\bar{a}$ with ex-ante probability at least $1 - \delta$. Let $\underline{\Sigma}$ be the set of all behavioral strategy profiles σ such that $\sigma_i(t_i)$ first-order stochastically dominates $\underline{\sigma}_i(t_i)$ (in $\Delta(A_i)$) for any $i \in N$ and

¹² I write $a_{-i} \geq_{-i} a'_{-i}$ if $a_j \geq_j a'_j$ for any $j \neq i$.

$t_i \in T_i$.¹³ By the supermodularity of $\mathbf{g}$ and the Kakutani–Fan–Glicksberg fixed point theorem, $(\mathbf{u}, P)$ has a Bayesian Nash equilibrium in $\underline{\Sigma}$, which plays $\bar{a}$ with ex-ante probability at least $1 - \delta$. $\square$

Proposition 3 easily extends to the respective set-valued robustness notions if the set $\mathcal{E} \subseteq \Delta(A)$ is an upper (resp. lower) set, i.e., if $\mu \in \mathcal{E}$ whenever μ first-order stochastically dominates μ' (resp. μ' first-order stochastically dominates μ) and $\mu' \in \mathcal{E}$.

References

- Balkenborg, D., & Vermeulen, D. (2016). Where strategic and evolutionary stability depart—A study of minimal diversity games. *Mathematics of Operations Research*, 41, 278–292.
- Chen, Y.-C., Takahashi, S., & Xiong, S. (2014). The robust selection of rationalizability. *Journal of Economic Theory*, 151, 448–475.
- Fudenberg, D., & Tirole, J. (1991). *Game Theory*. Cambridge: MIT Press.
- Haimanko, O., & Kajii, A. (2016). Approximate robustness of equilibrium to incomplete information. *International Journal of Game Theory*, 45, 839–857.
- Hillas, J. (1990). On the definition of the strategic stability of equilibria. *Econometrica*, 58, 1365–1390.
- Kajii, A., & Morris, S. (1997a). The robustness of equilibria to incomplete information. *Econometrica*, 65, 1283–1309.
- Kajii, A., & Morris, S. (1997b). *Refinements and Higher Order Beliefs: A Unified Survey*. New York: Mimeo.
- Kohlberg, E., & Mertens, J.-F. (1986). On the strategic stability of equilibria. *Econometrica*, 54, 1003–1037.
- Mertens, J.-F. (1989). Stable equilibria—A reformulation, part I. Definition and basic properties. *Mathematics of Operations Research*, 14, 575–625.
- Mertens, J.-F. (1991). Stable equilibria—A reformulation, part II. Discussion of the definition, and further results. *Mathematics of Operations Research*, 16, 694–753.
- Monderer, D., & Samet, D. (1989). Approximating common knowledge with common beliefs. *Games and Economic Behavior*, 1, 170–190.
- Monderer, D., & Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, 14, 124–143.
- Morris, S., & Ui, T. (2005). Generalized potentials and robust sets of equilibria. *Journal of Economic Theory*, 124, 45–78.
- Neyman, A. (1997). Correlated equilibrium and potential games. *International Journal of Game Theory*, 26, 223–227.
- Oyama, D., & Takahashi, S. (2011). On the relationship between robustness to incomplete information and noise-independent selection in global games. *Journal of Mathematical Economics*, 47, 683–688.
- Oyama, D., & Takahashi, S. (2019). *Generalized Belief Operator and the Impact of Small Probability Events on Higher Order Beliefs*. New York: Mimeo.
- Pram, K. (2019). On the equivalence of robustness to canonical and general elaborations. *Journal of Economic Theory*, 180, 1–10.
- Rubinstein, A. (1989). The electronic mail game: Strategic behavior under “Almost Common Knowledge”. *American Economic Review*, 79, 385–391.
- Ui, T. (2001). Robust equilibria of potential games. *Econometrica*, 69, 1373–1380.
- van Heumen, R., Peleg, B., Tijs, S., & Borm, P. (1996). Axiomatic characterizations of solutions for Bayesian games. *Theory and Decision*, 40, 103–129.
- Wu, W.-T., & Jiang, J.-H. (1962). Essential equilibrium points of n -person non-cooperative games. *Scientia Sinica*, 11, 1307–1322.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

¹³ This condition has a bite only for $t_i \in T_i^{d_i}$; if $t_i \in T_i \setminus T_i^{d_i}$, then any mixed action first-order stochastically dominates $\underline{\sigma}_i(t_i) = \underline{a}_i$.