Interdependent Choices*

Bruno Salcedo†

January 1st, 2017

Abstract Interdependent-choice Rationalizability (ICR) and Interdependent-choice Equilibrium (ICE) are simple and tractable solution concepts for strategic environments that allow for the choices of some agents to depend on those of others. They can be interpreted in two ways: (1) as robust predictions that do not depend on the details of the sequential and informational structures, or (2) as a characterization of self-enforceable contracts in environments in which the timing of actions is flexible, and actions are verifiable. I find that choice interdepence allows for cooperation in some but not all prisoners' dilemmas, and derive the optimal deal that a district attorney should offer the prisoners.

Keywords Interdependent choices · Robust Predictions · Implementation · Prisoners' dilemma

JEL classification C72 · D86

*A previous version of this manuscript was circulated under the tile "Implementation without commitment in moral hazard environments" (arXiv:1303.0916).

†Department of Economics, Cornell University, salcedo@cornell.edu. I wrote this paper under the invaluable guidance and supervision of Ed Green. I wish to gratefully acknowledge the comments and suggestions from Kalyan Chaterjee, Nail Kashaev, Vijay Krishna, Ruilin Zhou, Juan Block, Tymofiy Mylovanov, Bulat Gafarov and Bruno Sultanum, as well as Wiroy Shin and the participants of the 1st Prospects in Economic Research Conference at the Pennsylvania State University, the 2013 Midwest Economic Theory Meeting at Michigan State University, the 24th International Game Theory Conference at Stony Brook, the 2013 Asian Meeting of the Econometric Society at National University Singapore, and the 2013 Latin American Meeting of the Econometric Society at Colegio de México. I gratefully acknowledge the Human Capital Foundation (http://hcfoundation.ru/en/), and particularly Andrey P. Vavilov, for research support through the Center for the Study of Auctions, Procurement, and Competition Policy (http://capcp.psu.edu/) at the Pennsylvania State University. All remaining errors are my own.

1. Introduction

When agents make choices independently and simultaneously, standard notions of rationality postulate that each agent chooses a myopic best response to a fixed belief about his opponents' behavior. The story is quite difference when the decisions of some agents may depend on the actual behavior of others. In such settings, it is not sufficient for agents to consider the material consequences of their acts taking the behavior of their opponents as given. Agents must also consider the way that their own choices might affect those of others.

Choice interdependence is particularly relevant in moral hazard environments with no Pareto efficient Nash equilibria. With complete information, the problem of moral hazard can be solved if agents can enforce complete contracts (Coase theorem), or if they interact repeatedly and are patient enough (folk theorems). This is possible because written contracts or publicly observed histories serve as coordination devices allowing interdependence. This paper abstracts the notion of choice interdependence from such settings to investigate the extent to which its power remains in environments without *binding commitment*, *repetition*, or *monetary transfers*.

I consider situations in which each agent has to choose an perform an *irreversible action* at a *time of his choosing*, and such actions are *verifiable* but *not contractible*. Examples of such environment can be, citizens casting a vote in an election, prisoner choosing whether to accept or reject a sentence reduction in exchange for a confession, or competing firms choosing which features to include in their products. Suppose that a researcher knows the actions available to each agent and the agents' preferences over act profiles, and know that this this information is common knowledge among the players. However, she does *not* know any additional details about the environment. In particular, she does not know the specific game tree being played. What predictions could she make about the outcome of the environment? Could anything other than a correlated equilibria emerge? Would a folk-theorem-like result apply?

I propose two tractable and simple solution concepts to help address these questions. *Interdependent-choice Rationalizability* (ICR) is a form of rationalizability that allows for agents to believe that the choices of others depends on their own. *Interdependent-choice equilibrium* (ICE) with respect to a set of credible threats, is a simple notion of equilibrium defined by a finite set of affine inequalities. Define an *extensive form mechanism* (EFM) to be an extensive form game that is compatible with the information the researcher has about the environment. Proposition 3 states that, if an outcome is implementable as a Nash equilibrium of an EFM, then it must be an ICE. Proposition 4 states that, if an outcome is an ICE with respect to the set of ICR actions, then it is implementable as a perfect Bayesian (PB) equilibrium of an EFM. Proposition 5 states that, in generic 2×2 environments, an outcome can arise as a sequential equilibrium of an EFM if and only if it is an ICE with respect to the set of ICR actions.

Even without contracts, ICE can go well beyond the set of correlated equilibria. However, it yields sharper predictions than standard models with full commitment. The set of interdependent belief systems that are consistent with an EFM are restrictive enough to rule out a folk-theorem-like result. In particular, in sections 2 and 5, I show that cooperation possible in some but not all prisoners' dilemmas. Proposition 6 characterizes exactly how and for which prisoners' dilemmas cooperation is possible. I use this characterization to find the optimal deal that a district attorney should offer to the prisoners (Proposition 8).

There are many different EFMs that implement the same outcome as an equilibrium. In section 6, I introduce a class of mechanisms called *mediated mechanisms* that are canonical in the sense of Forges (1986). That is, every outcome that can be implemented in some EFM can *also* be implemented in a mediated mechanism. In a mediated game, a non-strategic mediator manages the play through private recommendations. The salient features are that the mediator can choose the timing of the recommendations, and make her recommendations contingent on the past behavior of the agents.[1] For these mechanisms to be feasible, it is important that the timing of the actions is flexible, and that actions can be directly observed or verified by an impartial third party. The set of mediated mechanisms is not the only canonical canonical class of EFM. However, it is a natural choice in accordance with the principle that implementation is easier when the mediator can observe everything, while the agents have as little information as possible (Myerson, 1986).

Being an ICE with respect to ICR actions is a sufficient but not a necessary

[1]The mediators in my model are remarkably powerful in comparison to the ones in classic studies that only allow mediators to engage simultaneous pre-play recommendations such as Aumann (1987), or simultaneous communication at fixed stages of the game such as Myerson (1986). In contrast, they are less powerful than mediators who can take actions on behalf of the players as in Moulin and Vial (1978), Monderer and Tennenholtz (2009) and Forgó (2010).

condition to be PB implementable. This is because, after observing an unexpected event in an extensive form game, a player might believe that the past choices of his opponent were not rational. Hence, despite the fact that ICR is equivalent to common knowledge of rationality with interdependent beliefs (cf. Halpern and Pass (2012)), it is possible for rational players to choose off-the-equilibrium-path actions which are not ICR. To deal with this issue, section 7 introduces the notions of *forward-looking interdependent-choice rationalizability* (FICR), and *quasi-sequential equilibrium* (QSE). QSE is a notion of equilibrium for extensive form games that lies in-between PB equilibrium and sequential equilibrium. Proposition 10 asserts that an outcome is QS implementable if and only if it is an ICE with respect to FICR actions.

Comparison with the literature

Choice interdependence is a common theme across various literatures. Besides the well established literatures on repeated games and games with contracts, different literatures allow for different forms of implicit repetition, commitment or transfers. The literature on counterfactual variations can be thought of as a reduced representation of repeated games (Kalai and Stanford, 1985). Commitment can be traced back to Moulin and Vial (1978) and Kalai (1981), which allow players to delegate choices to a mediator, or make *binding* preplay announcements. With unrestricted commitment, one obtains folk theorems (Kalai et al., 2010). Recent relevant works in this area include papers on commitment (Bade et al., 2009, Renou, 2009) and delegation (Forgó, 2010). Also related is the recent literature on revision games (Kamada and Kandori, 2009, 2011), which allow for a specific form of pre-play binding communication, see also Calcagno et al. (2014) and Iijima and Kasahara (2015).

Other literatures allow for counterfactual reasoning without being explicit about the mechanism that generates choice interdependence. Seminal examples include Rapoport (1965) and Howard (1971). More recently, Halpern and Rong (2010) and Halpern and Pass (2012), analize equilibrium and rationaliability with counterfactual beliefs. Once again, folk theorems hold if no further restrictions are imposed. There is a question as to when and which forms of counterfactial reasoning are consistent with the idea that players have independent free wills (Gibbard and Harper, 1980, Lewis, 1979). The present work thus considers only those counterfactual beliefs which can arise from different sequential and informational structures of an EFM.

In this regard, the current work is closely related to (Nishihara, 1997, 1999). Nishihara proposes a mechanism that allows for cooperation in some prisoners' dilemma games. An important difference is that my mechanism is exaclty timeable is the sense of Jakobsen et al. (2016). See Section 6.2 for more details. Also, I go beyond cooperation in the prisoners' dilemma and characterizes every outcome which can be implemented in any finite game.

Interdependent choices have also been studied in other contexts. Eisert et al. (1999) shows that cooperation is possible in a prisoners dilemma, when players can condition their choices on certain *quantum* randomization devices with *entangled* states. Tennenholtz (2004)'s program equilibrium generates choice interdependence for games *between computer programs* by allowing them to read each other's code before taking an action. Similarly, Levine and Pesendorfer (2007)'s self-referential equilibrium allows player's to receive a signal about their opponent's intentions before making their own choice. See also Block (2013) and Block and Levine (2015). I rule out this kind of signals as they represent a from of commitment: from the moment of *deciding* which action to play, to the moment of actually *performing* it. ICE can arise in settings where choices are instantaneous, and players can hide their intentions.

2. Cooperation in a prisoners' dilemma

Consider the following variation of the prisoners' dilemma. Two prisoners awaiting trial are offered a sentence reduction in exchange for a confession. As usual, this is a one-shot interaction and there are no contracts nor monetary transfers. However, I do not assume that the prisoners must make their choices independently nor simultaneously. Instead, they can choose the timing of their actions and hire a non-strategic lawyer (she) to help them coordinate. Formally, suppose the following:

(i) Each prisoner $i = 1, 2$, must submit an official signed statement at a time of his choosing $t_i \in [0, 1]$.

(ii) The statement specifies whether the prisoner chooses to to cooperate with his accomplice (C), or defect by accepting the deal (D). Once submitted, the statement cannot be modified or withdrawn.

(iii) The lawyer can make private non-binding recommendations to each prisoner at any moment in time.

(iv) Each prisoner can show the lawyer a certified copy of his submitted statement as *hard proof* that he cooperated or defected.

(v) The prisoners' preferences are summarized in the following payoff matrix, where $g, l > 0$ are fixed parameters.

	C	D
C	$1\ ,\ 1$	$-l\ ,\ 1+g$
D	$1+g\ ,\ -l$	$0\ ,\ 0$

Figure 1 – Payoff matrix for the prisoners' dilemma.

If $g < 1$, it is possible for the prisoners to cooperate in equilibrium. They could instruct the lawyer to proceed as follows. She will randomly and privately choose two times $r_i \in [0, 1]$. At date r_i, she will recommend prisoner i to immediately submit his statement with some recommended action. In equilibrium, both prisoners will follow such recommendations and immediately report back showing the copy of their statements as proof of compliance. The lawyer will always recommend C along the equilibrium path, and D after any detectable deviation.

The recommendation dates should be drawn as $r_i = 1 - 1/n_i$ where $n_1, n_2 \in \mathbb{N}$ are chosen as follows. First, the lawyer will draw $n \in \mathbb{N}$ from a geometric distribution with parameter $\rho \in (0, 1 - g)$. This is possible only because of the assumption $g < 1$.[2] Then, with probability $1/2$, she sets $n_1 = n$ and $n_2 = n + 1$, and with probability $1/2$ she sets $n_2 = n$ and $n_1 = n + 1$.

I will now show that following the lawyer's recommendations constitutes a sub-game-perfect equilibrium of the induced game. There are different types of histories to consider. First, suppose that prisoner i is recommended to cooperate

[2]The condition $g \leq 1$ is both sufficient and necessary to be able to implement cooperation as an ICE. See Corollary 7 in section 5.

at time $t_i = 0$. In this case, i knows for sure that he is the first player to receive a recommendation. Hence, if he cooperates and shows evidence of this to the lawyer, then $-i$ will also cooperate and his payoff will be $u_i(\mathrm{C}, \mathrm{C}) = 1$. Otherwise, if he deviates by defecting, waiting, or refusing to show evidence of his cooperation to the lawyer, then his accomplice will defect and his payoff will be no better than $u_i(\mathrm{D}, \mathrm{D}) = 0$. Hence, in this case, it optimal for i to be obedient and cooperate.

Now suppose that i is recommended to cooperate at some time $t_i = 1 - 1/n_i > 0$. As before, if he cooperates and shows evidence of this to the lawyer, then $-i$ will also cooperate and his payoff will be $u_i(\mathrm{C}, \mathrm{C}) = 1$. If he deviates by defecting, his payoff will equal $(1+g)$ times the probability that player $-i$ has already cooperated and won't have an opportunity to react to i's deviation. This probability is given by

$$\Pr(t_i > t_{-i} | t_i) = \frac{1/2 \cdot \Pr(n = n_i - 1)}{1/2 \cdot \Pr(n = n_i) + 1/2 \cdot \Pr(n = n_i - 1)} = \frac{1}{2 - \rho}.$$

Hence, i's expected utility from defecting is bounded above by

$$\frac{1}{2 - \rho}(1 + g) < \frac{1}{2 - (1 - g)}(1 - g) = 1,$$

which implies that cooperating is optimal.

Now consider a history when prisoner i is recommended to defect. Because this recommendation only arises after a detectable deviation, i can believe that his accomplice has already confessed. And, therefore, it is optimal for i to do the same. Finally, it is easy to see that there is no reason why a prisoner would like to submit his statement *before* being instructed to do so by the lawyer. Hence, always following the lawyer's recommendations constitutes a sequential equilibrium that results in full cooperation.

3. A framework for robust predictions

Each player $i \in I = \{1, 2\}$ is to choose and take one and only one irreversible action a_i from a finite set A_i.[3] Let $\mathscr{A} = \{\times_{i\in I} A_i' \,|\, \forall i \in I,\ A_i' \subseteq A_i\}$ denote the set of action subspaces. i's preferences are represented by $u_i : A \to \mathbb{R}$. No form of contract or binding agreement with regard to these actions can be enforced. Each agent must freely choose which action to take. The restriction to two players is for exposition purposes. The model can be extended to n-player environments, but the required notation is cumbersome.

3.1. Extensive form mechanisms

The tuple $E = (I, A, u)$ is only a *partial* characterization of the environment. It says nothing about the order in which choices will be made, nor about the information that each player will have when deciding which action to take.[4] In particular, choices need not be independent nor simultaneous. Instead, players could be playing some extensive form game that generates choice interdependence. For instance, players could condition their choices on correlated random signals. Alternatively, it could be the case that players take their actions sequentially in a fixed order, so that the decisions the later movers depends on the choices of those who moved first. I assume that the agents could be playing *any* extensive form game that is consistent with our partial description of the environment, and with the no-commitment assumption.

Extensive form games are defined as in Osborne and Rubinstein (1994), with some differences in notation. An extensive form game is characterized by a tuple $G = (M, X, j, \mathscr{H}, s_0, v)$. M is the *finite* set of moves and $X \subseteq_{t\in\mathbb{N}} M^t$ is the *countable* set of nodes. $j(x) \in I \cup \{0\}$ is the agent moving at x, where 0 represents Nature (or a mediator). Z and Y_i are the sets of terminal nodes and i's decision

[3] I employ the stabdard notation $-i$ for i's opponent, $a = (a_i, a_{-i}) \in A = \times_i A_i$ for action profiles, $\alpha \in \Delta(A)$ for joint distributions, $\alpha_i \in \Delta(A_i)$ for marginal distributions, and $\alpha(\cdot\,|a_i) \in \Delta(A_{-i})$ for conditional distributions.

[4] Here, the term "information" refers to information about the actions taken by other players. The environment E is assumed to be common knowledge. In contrast, Bergemann and Morris (2013) derive robust predictions to assumptions on the information that the players have about the payoff structure.

nodes, respectively. $\mathscr{H}_i$ partitions Y_i into information sets and satisfies *perfect recall*. Let $M(x)$ and $M(H)$ denote the sets of moves available at a node $x \in X$ or an information set $H \in \mathscr{H}$, respectively. $s_0 : Y_0 \to \Delta(M)$ specifies the players' common prior beliefs about Nature's behavior. Finally, $v_i : Z \to \mathbb{R}$ represents i's preferences over terminal nodes.

The following definition captures the set of extensive form games that are consistent with our partial description of the environment, while ruling out side-payments, enforceable contracts, delegation, and any other form of *binding* commitment regarding their choices about actions a_i.

Definition 3.1 An *extensive form mechanism* (EFM) for E is an extensive form game G such that for every terminal node z there exists an action profile a^z and a tuple of decision nodes $(x_i^z)_{i \in I} \in \times_{i \in I} Y_i$ such that

(i) $v(z) = u(a^z)$.

(ii) For each i, $A_i \subseteq M_i(x_i^z)$, and both x_i^z and (x_i^z, a_i^z) are predecessors of z.

(iii) For each player i, decision node $x_i' \in Y_i$, and action $a_i' \in A_i$, if both x_i' and (x_i', a_i') are predecessors of z, then $x_i' = x_i^z$ and $a_i' = a_i^z$.

The first requirement is that each terminal node z can be associated with an action profile a^z, and utility over terminal nodes is given by the utility from the corresponding action profiles. The second requirement is that each player i actually chooses his own action a_i^z. Also, it rules out partial commitment by requiring that, at the moment of choosing a_i^z, player i could have chosen *any* other action in A_i. The third requirement is that choices are irreversible. Conditions (ii) and (iii) together imply that for every every action profile $a \in A$, there exists at least one terminal node z such that $a^z = a$.

For exposition purposes, Definition 3.1 is more restrictive than necessary. In particular, it requires that there is at most a single way to perform each action in each information set. This precludes features that might seem relevant, such as allowing players to decide on the spot whether to take an action publicly or privately. Appendix A proposes a more general definition of EFMs. As it turns out, all the results of the paper remain true with either definition.

3.2. Robust predictions

Given an EFM, a (behavior) strategy for player i is a function that assigns a distribution $s_i(\cdot|H) \in \Delta(M(H))$ to each information set $H \in \mathscr{H}_i$. Let S_i denote the set of i's strategies, and $S = \times_i S_i$. Let $\zeta(x|s, s_0)$ denote the probability that the game will reach node x given that Nature chooses according to s_0 and players follow s. Each strategy profile s induces a distribution $\alpha^s \in \Delta(A)$ over acts of the environment given by $\alpha^s(a) = \sum_{z \in Z(a)} \zeta(z|s, s_0)$, where $Z(a) = \{z \in Z \mid a^z = a\}$.

Definition 3.2 A distribution over acts $\alpha \in \Delta(A)$ is (Nash, sequentially, PB, ...) *implementable*, if there exists an EFM and a corresponding (Nash, sequential, PB, ...) equilibrium s^* that induces it.

The set of implementable outcomes is exactly the set of robust predictions that a researcher could make if he knew the partial description of the environment E, and he believed that the agents could be playing any equilibrium of any EFM. Alternatively, it can be interpreted as the set of outcomes that the players could implement as self-enforceable agreements without the use of binding commitment devices. The main objective of the paper is to characterize the sets of implementable outcomes under different notions of equilibrium and rationality.

4. Interdependent choice equilibrium and robust predictions

4.1. Interdependent choice equilibrium

This section characterizes the set of outcomes that are implementable in the sense of Definition 3.2. This is done via the notion of interdependent choice equilibrium defined as follows.

Definition 4.1 A distribution over action profiles $\alpha \in \Delta(A)$ is an *interdependent-choice equilibrium* (ICE) with respect to a set of credible threats $B \in \mathscr{A}$, if there exists some $\theta : A \to \Delta(I)$ such that for all $i \in I$ and $a' \in A_i$

$$\sum_{a_{-i}\in A_{-i}} \alpha(a)\Big(u_i(a) - \Big(1-\theta(i|a)\Big)u_i(a'_i, a_{-i}) - \theta(i|a)\underline{w}_i(a'_i|\alpha, B)\Big) \geq 0, \qquad (1)$$

where $\underline{w}_i(a'_i|B^*_{-i}) := \min\Big\{u_i(a'_i, a_{-i}) \mid \quad a_{-i} \in B^*_{-i}\Big\}$ and $B^*_{-i} = \text{supp}(\alpha_{-i}) \cup B_{-i}$. For the case $B = A$, I omit the reference to B and simply sat that α is an ICE.

Some intuition about the definition ICE can be acquired by analyzing two extreme cases. Condition (1) can be thought of as an incentive constraint for player i, requiring that a_i should be more profitable than deviating to a'_i. If $\theta(i|a) = 0$, this condition coincides with the incentive constraints for correlated equilibrium. In this case, player i computes the expected utilities from a_i and a'_i using the same distribution $\alpha_{-i}(\,\cdot\,|a_i)$. In other words, i beliefs that the behavior of $-i$ does not depend on his own choice. In the opposite extreme, if $\theta(i|a) = 1$, player i believes that if he deviates from a_i to a'_i, his opponent will react by choosing the harshest punishment in B^*_{-i}.[5] In this case, if $B^* = A$, condition (1) coincides with the definition of individual rationality.

The definition of ICE requires $\theta(\,\cdot\,|a)$ to be a probability measure. This requirement captures the fact that choice interdependence in my model should be consistent with an equilibrium of some EFM.[6] For that to be the case, $-i$'s action can *only* depend on i's action if $-i$ makes his choice *after* observing a signal about $-i$'s action. Of course, it cannot be the case that i moves before $-i$ and, at the same time, $-i$ moves before i. One can think of $\theta(i|a)$ to be the probability that i is the first player to move, conditional on being on an equilibrium path of play in which the chosen action profile is a.

Proposition 1 (ICE properties) *For any $B \in \mathscr{A}$, the set of ICE with respect to B is*

[5]The definition of ICE does not include any incentive constraints for such punishments, *as if* agents could commit to punish any deviations. This is *not* an assumption, but rather a technical tool to keep the model tractable. As it turns out, ICE can be used to characterize implementation under different notions of perfection, by choosing B adequately.

[6]This is the point where my model differs from models with full commitment, such as Kalai et al. (2010), Halpern and Pass (2012), Tennenholtz (2004) or Block and Levine (2015). This restriction is the reason why the set ICE does not result in a folk-theorem-like result. In particular, setting $\theta(i|a) = \delta \in (0,1)$, condition (1) resembles the recursive characterization of SPNE of repeated games due to Abreu et al. (1990). With that analogy in mind, requiring $\theta(\,\cdot\,|a)$ to be a probability measure is tantamount to assuming that the average discount factor cannot exceed 1/2.

non-empty and contains the set of correlated equilibria and is contained in the set of individually rational outcomes. The set of ICE with respect to A is a non-empty closed and convex polytope.

Example 4.1 Two partners decide whether to work (W) or shirk (S) in a joint-venture, their payoffs are depicted in Figure 2. The figure also shows the sets of payoffs corresponding to individual rationality, Nash equilibrium with public randomization, correlated equilibrium, and ICE. In this example, all the Pareto efficient outcomes correspond to ICE.

(W, W) is not a Nash equilibrium because, whenever an agent is working, his opponent prefers to shirk. It is an ICE because, a player who considers shirking knows that with some probability, his opponent will learn of this defection and react by also shirking. The payoff vector $(1, 1)$ cannot be attained as an ICE, because it requires players to shirk with high probability. Since each player always prefers that his opponent works, this leaves too little room to punish deviations.

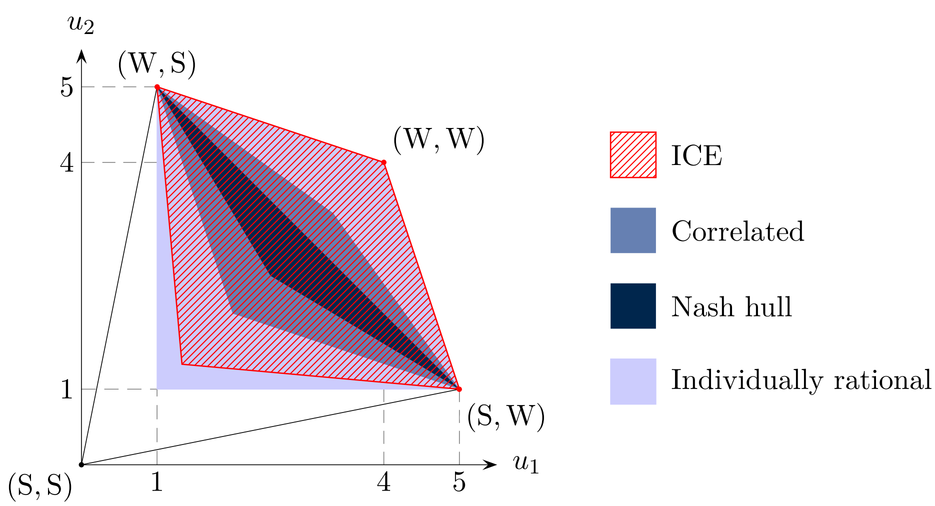

Figure 2 – Equilibrium payoffs for a teamwork game.

4.2. Interdependent choice rationalizability

The salient feature of environments with choice interdependence, is that rational agents do not make their choices given a fixed belief. Instead, they compute expected utility with respect to *counterfactual beliefs* $\lambda_i : A_i \to \Delta(A_{-i})$, where $[\lambda_i(a_i)](a_{-i})$ represents i's assessed likelihood that his opponent will choose a_{-i} *if i plays a_i*. This section proposes a notion of rationalizablity that incorporate this idea.

Before doing so, I will introduce some notation. Expected utility with respect to counterfactual beliefs is denote by $U_i(a_i, \lambda_i) = \sum_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i}) \cdot [\lambda_i(a_i)](a_{-i})$. An action a_i is said to be a best response to λ_i if $U_i(a_i, \lambda_i) \geq U_i(a_i', \lambda_i)$ for all $a_i' \in A_i$. For $A' \in \mathscr{A}$, $\Lambda_i(A')$ is the set of counterfactual beliefs such that $[\lambda_i(a_i)](A'_{-i}) = 1$ for every $a_i \in A_i$.

Definition 4.2 (Interdependent choice rationalizability (ICR))

- $a_i^* \in A_i$ is *IC-rationalizable* (ICR) with respect to $A' \in \mathscr{A}$, if and only if it is a best response to some $\lambda_i \in \Lambda_i(A')$.
- $A' \in \mathscr{A}$ is *self-IC-rationalizable* if and only if every action profile in A' consists of actions that are ICR with respect to A'.
- The set of *ICR action profiles*, A^{ICR}, is the largest self-IC-rationalizable set.

ICR is analogous to (correlated) rationalizability (cf. Pearce (1984), Bernheim (1984)), simply replacing correlated beliefs with counterfactual beliefs. It coincides with the notion of *minimax ratonalizability* developed independently and simultaneously by Halpern and Pass (2012). Halpern and Pass additionally show that it is equivalent to rationality and common certainty of rationality in epistemic models with counterfactual reasoning.

Let $\mathrm{ICR}_i(A')$ denote the set of i's actions that are ICR with respect to A'. A^{ICR} is guaranteed to exist because $\mathrm{ICR}(\cdot)$ is $\subseteq$-monotone. Consequently, the union of all self self-IC-rationalizable sets is also self-IC-rationalizable. It is nonempty because it always contains the set of rationalizable action profiles. A^{ICR} can be found in a tractable way using the notion of absolute dominance defined ahead.

Definition 4.3 Given two actions $a_i, a_i' \in A_i$, a_i *absolutely dominates* a_i' with respect to $A' \in \mathcal{A}$ if and only if $\min_{a_{-i} \in A'_{-i}} u_i(a_i, a_{-i}) > \max_{a_{-i} \in A'_{-i}} u_i(a_i', a_{-i})$.

In other words, a_i absolutely dominates a_i', if and only if the best possible payoff from playing a_i' is strictly worse than the worst possible payoff from playing

a_i. Absolute dominance is much simpler than strict dominance in computational terms because a player can conjecture different reactions for each alternative action, and thus mixed actions need not be considered. The following proposition ensures that ICR(A') results from eliminating absolutely dominated actions, and A^{ICR} results from repeating this process iteratively.

Proposition 2 *An action is ICR with respect to A' if and only if it is not absolutely dominated in A', and the iterated removal of absolutely dominated actions is order independent and converges in finite time to A^{ICR}.*

4.3. Robust predictions under choice interdependence

So far, two types of solution concepts have been introduced. First, the set of implementable outcomes is the set of robust predictions that a researcher can make that do not depend on the details of the environment. Second, ICE and ICR are tractable solution concepts that are easy to compute. this section shows that ICE and ICR can help to characterize the sets of implementable outcomes under different standard notions of equilibrium for extensive form games. The first result is that, every outcome that can arise as an equilibrium of an EFM is an ICE.

Proposition 3 *If a distribution over action profiles is Nash implementable then it is an ICE.*

Proposition 3 implies that an outcome which is *not* an ICE requires some form of commitment, side payments, or repetition to be implemented in equilibrium. In situations in which (i) action spaces and preferences are known, (ii) agents cannot use binding commitment devices regarding their choices form A_i, and (iii) agent's choices are expected to be in equilibrium, every equilibrium outcome is an ICE. This prediction is robust even in the researcher does not know the details of the timing and information structures.

The converse of Proposition 3 is also true. Every ICE is Nash implementable. However, because EFM are extensive form games, plausible equilibrium predictions should involve sequential rationality constraints. In particular, the definition of ICE does not involve incentive constraints for additional off-the-equilibrium-path punishments in $B^* \setminus \{\text{supp}(\alpha)\}$. The following proposition provides sufficient conditions for implementation as perfect Bayesian (PB) equilibrium.

Proposition 4 *Every ICE with respect to ICR is PB implementable.*

The condition in Proposition 4 is sufficient bur *not* necessary for PB implementation in general games. Section 7.2 explains why it is not necessary, and provides condition that is both sufficient and necessary for general games. As it turns out, ICE with respect to ICR is necessary and "almost" sufficient for sequential implementation in generic 2×2 environments.

Proposition 5 *In 2×2 environments without repeated payoffs, the set of ICE with respect to ICR is dense in the set of sequentially implementable outcomes.*

5. When is cooperation possible in a prisoners' dilemma

There are different versions of the classic prisoners' dilemma story, and some of them fit my model closely. A signed statement is an irreversible action with flexible timing. Contracts between the prisoners are might not be enforceable by the legal system. The right to have an attorney present provides a mechanisms for actions to be verifiable. So, as long as repetition or transfers can be ruled out, the set of sequentially implementable outcomes provides a reasonable set of robust predictions for behavior in a prisoners' dilemma with rational prisoners. Applying proposition 5 results in the following complete characterization.

Proposition 6 *In the prisoners' dilemma from Figure 1, a joint distribution $\alpha \in \Delta(A)$ is sequentially implementable if and only if:*

$$\left(\frac{1-g}{l}\right)\alpha(\mathrm{C},\mathrm{C}) \geq \alpha(\mathrm{D},\mathrm{C}) + \alpha(\mathrm{C},\mathrm{D}). \tag{2}$$

Let us further analize condition (2). First, notice that there are no constraints on $\alpha(\mathrm{D},\mathrm{D})$, this is because defecting is already a dominant strategy. Second, notice that the right hand side is always non-negative, and can always be made equal to 0 by setting $\alpha(\mathrm{D},\mathrm{C}) = \alpha(\mathrm{C},\mathrm{D}) = 0$. This implies that the cooperative strategy profile (C,C) can be played with positive probability only if $g \leq 1$, and, if $g \leq 1$, then it can be played with full probability.

Corollary 7 *If $g \leq 1$, then (C,C) is an ICE and, if $g > 1$, the only ICE is (D,D).*

Finally, notice that the probabilities of the asymmetric outcomes are bounded above by a linear function of the probability of the cooperative outcome. This is a natural condition, because a player can only benefit from cooperating if his opponent is also cooperating. In particular, when $g = 1$, the left hand side of condition (2) equals 0, which means that players cannot assign any probability to asymmetric outcomes. For $g \in (0, 1)$, this bound is relaxed so that the set of ICE payoffs fans out, see figure 3. However, the expected utility of players who do not confess has to be strictly greater than 1, because otherwise the threat of opponent's defection would not have any bite. This implies that there always exist individually rational payoffs which cannot be achieved by any Nash implementable distribution.

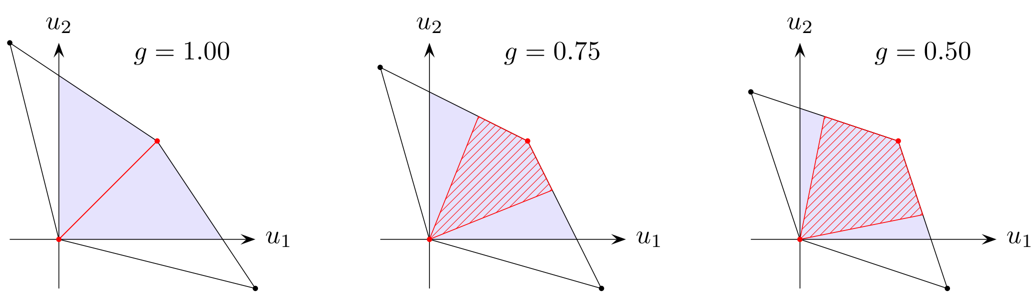

Figure 3 – ICE for the prisoners' dilemma with $l = 0.5$ and different values of g.

5.1. Optimal sentence reduction

The analysis can be taken one step forward. Consider the problem of a DA that must choose which deal to offer the prisoners as to maximize the total amount of time served. Proposition 6 shows that cooperation may or may not be possible depending on the specific payoffs. Hence, the DA's problem is non trivial as she must figure out the worst offer that the prisoners will accept in equilibrium. For simplicity, I assume that the prisoners have linear preferences over the amount of time served.

Suppose that, if n prisoners were to confess, the DA could secure a maximum sentence of at most $\bar{\mu}_n \in \mathbb{N}$ days in prison for each prisoner, where $\bar{\mu}_2 > \bar{\mu}_1 > \bar{\mu}_0 > 1$ are exogenous parameters.[7] The DA can commit to offering an anonymous sentencing policy $\mu = (\mu_0, \mu_1^-, \mu_1^+, \mu_2) \in \mathbb{Z}_+^4$. If nobody confesses, each prisoner will be sentenced to μ_0 days. If both prisoners confess, each will be sentenced to μ_2 days. If only one prisoner confesses, he will be sentenced to μ_1^- days and his accomplice to μ_1^+ days. A policy μ is feasible if $\mu_0 \leq \bar{\mu}_0$, $\mu_1^+, \mu_1^- \leq \bar{\mu}_1$ and $\mu_2 \leq \mu_2$.

The timing is as follows. First, the DA chooses a feasible policy μ. This induces the environment E in figure 4. Then, the prisoners play the equilibrium which minimizes their total time served. Here, I consider two scenarios. In the first scenario, the DA can force agents to make choices independently, so that the set of equilibria is just the set of NE of the environment. In the second scenario, the DA cannot prevent the prisoners from coordinating their choices, so that the relevant set of equilibria is the set of ICE.

	C	D
C	$-\mu_0, -\mu_0$	$-\mu_1^-, -\mu_1^+$
D	$-\mu_1^+, -\mu_1^-$	$-\mu_2, -\mu_2$

Figure 4 – Environment induced by a feasible policy μ.

Proposition 8 *The maximum total sentence time that the DA can guarantee under independent choices is* $2\bar{\mu}_1 - 2$. *Under interdependent choices, the time that the DA can guarantee is* $2\min\{2\bar{\mu}_0, \bar{\mu}_1\} - 2$. *Hence, the cost to the DA of choice interdependence is* $2\max\{0, \bar{\mu}_1 - 2\bar{\mu}_0\}$.

Under independent choices, the optimal mechanism is to set the maximum possible sentence for those people who do not confess, and the minimal sentence reduction in exchange for a confession. This results in a *per capita* sentence close to $\bar{\mu}_1$. The resulting game is a prisoners' dilemma. Using the same mechanism,

[7] Restricting attention to natural numbers guarantees existence of an optimal policy.

the condition $g > 1$ from Corollary 7 is satisfied if $\bar{\mu}_1 \leq 2\bar{\mu}_0$. Which means that mutual defection is also the only ICE and the prisoners would accept the deal, even under choice interdependence. Hence, in that case, there are no interdependence rents for the agents.

In contrast, when $\bar{\mu}_1 > 2\bar{\mu}_0$, using the same mechanism would not be optimal because mutual cooperation would be an ICE that would result in a total sentencing time per capita of at most $\bar{\mu}_0$. In appendix B.4 I show that, in that case, the optimal mechanism is given by $\mu_0^* = \bar{\mu}$, $\mu_1^+ = 0$, $\mu_1^- = \bar{\mu}_1$, and $\mu_2^* = \min\{2\bar{\mu}_0, \bar{\mu}_1\} - 1$. This is the harshest prisoners' dilemma in which mutual cooperation is *not* an ICE. Interestingly, the maximum possible per capita sentence in that case is only twice the amount of time that the DA could convict the prisoners for without a confession.

6. Implementation via mediated mechanisms

There are many different EFMs that implement the same outcome as an equilibrium. One possible way to implement the set of ICE is via a simple class of extensive form games in which a non-strategic mediator manages the players through private recommendations. A *mediated mechanism* is characterized by a triplet (α, θ, B). $\alpha \in \Delta(A)$ is a distribution over action profiles to be implemented. $\theta : A \to \Delta(I)$ specifies a distribution over the order in which players will move, conditional on the action profile to be implemented. $B = \times_i B_i$ specifies actions that can be recommended as *additional* credible threats. The *effective* set of credible threats, $B_i^* = B_i \cup \operatorname{supp} \alpha_i$, also includes the actions played along the equilibrium path.

Mediated mechanisms represent the EFMs described as follows. The game begins with the mediator privately choosing the action profile a^* that she wants to implement (according to α), and the player i^* to move first (according to $\theta(\,\cdot\,|a^*)$). She then "visits" each of the players one by one, visiting i^* first and $-i^*$ second. When visiting each player i, the mediator recommends an action a_i^r, and observes the action actually taken a_i^p. At the moment of making their choices, the players do not possess any information other than the recommendation they receive. The mediator always recommends the intended action to the first player, i.e. $a^r_{i^*} = a^*_{i^*}$. She recommends the intended action to the second player if the first player complied, and one of the worst available punishments in $B^*_{-i^*}$ otherwise, i.e.:

$$\begin{aligned} a^r_{-i^*} &= a^*_{-i^*} && \text{if} \quad a^p_{i^*} = a^*_{i^*}, \\ a^r_{-i^*} &\in \underset{a_{-i^*} \in B^*_{-i^*}}{\arg\min}\ u_{i^*}\left(a^p_{i^*}, a_{-i^*}\right) && \text{if} \quad a^p_{i^*} \neq a^*_{i^*}. \end{aligned}$$

Note that following the mediator's recommendations constitutes a Nash equilibrium if and only if condition (1) is satisfied. Hence every ICE is Nash implementable via a mediated mechanism. This means that mediated mechanisms are a canonical class for Nash implementation in the sense of Forges (1986).

6.1. Mediated mechanism for the prisoners' dilemma

To see an example of a mediated mechanism, consider once again the the prisoners' dilemma from section 2 with a few differences. Now suppose that the lawyer can directly control the timing of meetings, and observe the actions of the players, and the prisoners have no notion way of measuring the passing of time. Then, the prisoners could instruct the lawyer as follows:

> "You must uniformly randomize the order of our meetings. If the DA offers us a (prisoners' dilemma) deal you must always recommend that we do not confess, unless one of us has already confessed, in which case you must recommend that we do confess. Other than those recommendations, you must not provide us with any additional information."

The resulting situation would be the mediated mechanism corresponding to the EFM in Figure 5. The red arrows represent the actions recommended by the lawyer at each information set. Indeed, following such recommendations constitutes a sequential equilibrium if and only if $g \leq 1$. This mechanism is equivalent to the one analyzed in Nishihara (1997, 1999).

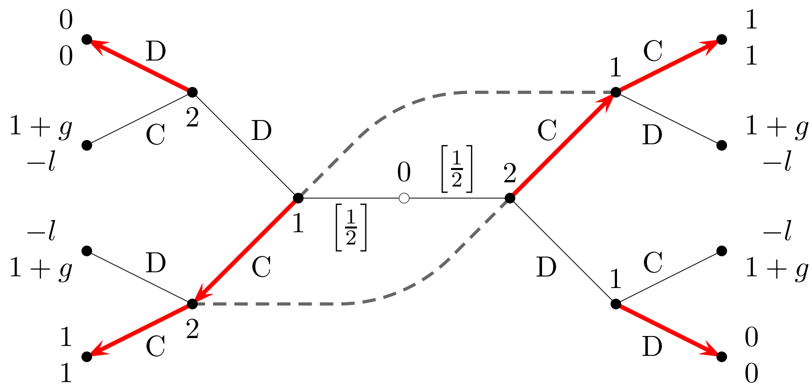

Figure 5 – A mediated mechanism for the prisoners' dilemma.

6.2. Timeability

An crucial feature of the EFM in Figure 5 is that, along the equilibrium path, the prisoners are completely uniformed about the order of play. Each of the prisoners cannot distinguish between the node where he is the first one to move, and the node where he is the second one and his accomplice cooperated before him. And he assign the same probability to each of these two events. For this to be possible, it is crucial that the prisoners cannot measure the passing of time, which might be an implausible assumption. Using the language of Jakobsen et al. (2016), this the mechanism is not exactly timeable and "if the players have a sense of time… [it] cannot be implemented in actual time in a way that respects the information sets." This criticism applies to mediated mechanisms in general.

One way around this issue is to use exactly timeable EFMs that *approximate* the relevant features of mediated mechanisms. For example, the mechanism from Figure 5 can be approximated by the one in Section 2 in which the lawyer made recommendations at randomly selected *dated* moments in time. Note that, in the timeable mechanism, the beliefs of each prisoner about being the first mover at the time of receiving a recommendation converge to $1/2$ uniformly as $\rho \to 0$. Hence, as long as $g < 1$—so that the mechanism in figure 5 is *strictly* incentive compatible—the same outcome can be implemented. The same is true about general ICE and mediated mechanisms. Say that an ICE is *strict*, if condition (1) can be satisfied with inequality whenever $a_i' \neq a_i$.

Proposition 9 *Every strict ICE can be Nash implemented using an exactly timeable EFM.*

6.3. Weaker mediators

Mediators in mediated games are remarkably powerful, beyond what might be available in some real-life situations. It is thus important to keep in mind that mediated mechanisms are one possible way to implement an ICE, but there are be many others. For example, cooperation in a prisoners' dilemma could be implemented via the mediated mechanism from Figure 5. And it could also be implemented via the mechanism from Section 2, in which prisoners are free to choose the timing of their actions and whether they want to keep their actions public or show them to the lawyer. While this mechanism also involves a lawyer, it is possible to consider environments in which similar mechanisms arise naturally without the intervention of a mediator.

Moreover, different outcomes can still be implemented via mechanisms with weaker mediators. For example, suppose that any deviation from the equilibrium path is publicly observed by everyone and not just the mediator. The outcomes that could be implemented by mediated mechanism under such conditions can be characterized by adjusting the worst punishment functions $\underline{w}$ in the definition of ICE. In this case, the relevant punishment function for player i would be $\underline{w}'_i(a'_i) = \min_{a_{-i}\mathrm{BR}_{-i}(a'_i)} u_i(a'_i, a_{-i})$, where BR_{-i} is $-i$'s best response correspondence.

Alternatively, instead of assuming that the mediator controls the order of choices, suppose that she can control the order of her recommendations but players can choose to act before or after they encounter her. In such cases, the mediator could not recommend action-specific punishments. A player who intended to deviate would make his choice after the mediator has left, and thus the mediator could no longer observe the specific deviation. The set of implementable outcomes under these conditions could be characterized by replacing $\underline{w}_i$ with the constant minimax punishment $\underline{w}'_i(a'_i) = \min_{\alpha_{-i}\in\Delta(B^*_{-i})} \max_{a_i\in A_i} U_i(a_i, \alpha_{-i})$.

7. Quasi-sequential implementation

7.1. Quasi-sequential equilibrium

Sequential equilibrium is defined in terms of *sequential rationality* and *belief consistency*. Sequential rationality requires choices to be optimal at the interim stage for every information set in the game. Off the equilibrium path, belief consistency requires players to update their beliefs in accordance with some prior assessment of the relative likelihoods of different trembles or mistakes. Furthermore, it requires that these prior assessments should be common to all players. *Quasi-sequential equilibrium* (QSE) imposes sequential rationality and requires beliefs to be consistent with trembles, but allows players to disagree about which deviations are more likely.

For two player environments, it is useful to allow Nature to assign zero probability to some of its available moves. This is because, when faced with a null event, a player can believe that it was Nature who made a mistake instead of necessarily believing that an opponent deviated from the equilibrium.[8] In order to define consistent beliefs, it is necessary to introduce new notation to denote players' beliefs about Nature's choices, other than s_0. Let S_0 and ${S_0}^+$ denote the sets of mixed strategies and strictly mixed strategies for Nature.

A conditional belief system for i in an extensive form game G, is a function ψ_i mapping i's information sets to distributions over nodes. $\psi_i(y|H)$ is the probability that i assigns in information set H to being in node y. Let Ψ_i denote the set of i's conditional belief systems. An assessment is a tuple $(\psi, s) \in \Psi \times \Sigma$ that specifies both players interim and prior beliefs (or strategies). An extended assessment is a tuple $(\psi, s, s_0') \in \Psi \times \Sigma \times \Sigma_0$ that also specifies prior beliefs on Nature's choices. Given an assessment (ψ, s), an information set H, and an available move m, $V_i(m|H)$ denotes i's expected payoff from choosing m at H. The expectation is taken given his interim beliefs $\psi_i(H)$ regarding the current state of the game, and assuming that future choices will be made according to s.

[8] It is often assumed that Nature assigns positive probability to all of its available moves, but I am unaware of any good arguments to maintain this assumption. Consider for instance the following quote from (Kreps and Wilson, 1982, pp. 868): "To keep matters simple, we henceforth assume that the players initial assessments [on Nature's choices] are strictly positive".

Definition 7.1 (Quasi-sequential equilibrium) An assessment $(\psi^*, s^*) \in \Psi \times \Sigma$ is:

- *Weakly consistent* if and only if for every player there exists a sequence of strictly mixed extended assessments $(\psi^n, s^n, {s'_0}^n)$ satisfying Bayes' rule, and such that $(\psi_i^n, s^n, {s'_0}^n)$ converges to (ψ_i^*, s^*, s_0).

- *Sequentially rational* if and only if $V_i(m|H) \geq V_i(m'|H)$ for all players i, information sets $H \in \mathscr{H}_i$ and moves $m, m' \in M(H)$ with $[s_i^*(H)](m_i) > 0$.

- A *quasi-sequential equilibrium* (QSE) if it is both weakly consistent and sequentially rational.

Sequential rationality requires that the choices that occur off the equilibrium path should be optimal. This implies that players must always believe that the *future* choices of their opponents will be rational, and this fact is common knowledge. However, off the equilibrium path, QSE imposes no restrictions on beliefs about *past* choices, nor agreement of beliefs across different players. In that sense, the difference between QSE and Nash equilibrium can be thought of as a form of *future-looking* rationalizability off the equilibrium path.[9]

The only difference between QSE and sequential equilibrium, is that the former imposes a stronger notion of consistency. Namely, *the same* sequence of strictly mixed assessments should work for all players. Loosely speaking, sequential equilibrium requires choices and beliefs to be in equilibrium, not only along the equilibrium path, but also in every 'subgame'. In contrast, QSE requires equilibrium along the equilibrium path, but only imposes a form of rationalizability in null 'subgames'.

The focus on QSE is partially motivated by the fact that it is the finer refinement for which I can provide a complete characterization. However, there may be situations for which it is more appealing than sequential equilibrium. In general, equilibrium is not a straightforward consequence of rational behavior. In order to guarantee equilibrium one must assume mutual or common knowledge of conjectures (Aumann and Brandenburger, 1995), which may be hard to justify off the equilibrium path.

In this respect, focal point arguments may be questioned because of the complexity of determining whether an equilibrium is sequential. Communication can

[9]This idea closely resembles the notion of common belief in future rationality from Perea (2013).

be questioned along similar lines, because planning for all possible contingencies or agreeing on their likelihood may be too complex. Finally, precedence may provide a justification for equilibrium, but repetition provides no experience about events which only happen off the equilibrium path (Fudenberg and Levine, 1993). Hence there might be situations in which (i) it makes sense to assume agreement *exclusively* along the equilibrium path; and yet (ii) rationality and common certainty of rationality may also be defended in every subgame.

7.2. Forward-looking interdependent-choice rationalizability and QS implementation

The preceding discussion helps to clarify why it is that being an ICE with respect to A^{ICR} is *not* a necessary condition for PB implementation, despite the fact that ICR is equivalent to common knowledge of rationality with interdependent beliefs (cf. Halpern and Pass (2012)). After observing an unexpected event in an extensive form game, a player might believe that the past choices of his opponent were not rational. Hence, even in a sequential equilibrium of an EFM, it is possible for player choose actions that are rational but not ICR off the equilibrum path.

In fact, two kind of actions can *always* be enforced as credible punishments for QS implementation. ICR punishments are admissible because QSE implementation does not require agreement off the equilibrium path. Hence, the player performing the punishment may very well have counterfactual beliefs which rationalize it. Moreover, since QSE only imposes belief of rationality for future choices, beliefs about *past* can be chosen freely. Best responses to arbitrary *degenerate* conjectures are thus also admissible. These two ideas are embodied in the notion of *future-looking interdependent-choice rationalizablity* (FICR).

Definition 7.2 (Future-looking interdependent-choice rationalizablity)

- $a_i^* \in A_i$ is *FICR* with respect to $A' \in \mathscr{A}$ if and only if there exists a belief $\lambda_i^0 \in \Delta(A_{-i})$, a counterfactual belief $\lambda_i^1 \in \Lambda(A')$, and some $\mu \in [0,1]$ such that a_i^* maximizes expected utility with respect to the counterfactual belief $\lambda_i = \mu\lambda_i^0 + (1-\mu)\lambda_i^1 \in \Lambda_i(A)$. Let $\mathrm{FICR}_i(A')$ denote the set of profiles consisting of FICR actions with respect to A'.

- $A' \in \mathscr{A}$ is *self-FICR* if and only if $A' \subseteq \text{FICR}(A')$.

- The set of *FICR* action profiles $A^{\mathsf{FICR}} \in \mathscr{A}$ is the largest self-FICR set.

As before, A^{FICR} is guaranteed to exist because $\text{FICR}(\,\cdot\,)$ is $\subseteq$-monotone, and thus the union of all self-FC-rationalizable sets is self-FC-rationalizable. Also, it is non-empty because it always contains the set of ICR action profiles.

Intuitively, one can think of λ_i^0 as the arbitrary beliefs (degenerate conjectures) over past deviations, and think of λ_i^1 as the conjectures about future FC-rationalizable choices. With this interpretation, an action a_i is FICR with respect to an action space A' if it is a best response to some conjecture λ_i that assigns full probability to actions in A'_{-i}, *only for choices that occur in the future.* λ_i can assign positive probability to any action, provided that this probability is independent from i's choice. The set of FICR actions is exactly the set of credible threats that characterizes QS implementation.

Proposition 10 *A distribution over action profiles is QS implementable if and only if it is an ICE with respect to* A^{FICR}*.*

There are two interesting corollaries of this result. First, since sequential implementability implies QS implementability, Proposition 10 provides as a necessary condition for sequential implementation in arbitrary environments. Second, since ICR actions are FICR, in games with no absolute dominance a distribution is QS implementable if and only if it is an ICE. This means that requiring QSE instead of Nash equilibrium has a small impact, because most games of interest have no absolutely dominated actions.

Corollary 11 *All sequentially implementable outcomes are ICE with respect to* A^{FICR}*.*

Corollary 12 *When there are no absolutely dominated actions, a distribution is quasi-sequentially implementable if and only if it is an interdependent-choice equilibrium.*

This section concludes with a characterization of the FICR operator. Loosely speaking, the following proposition shows that it is equivalent to the elimination of strictly dominated actions in an auxiliary game. Hence, computing A^{FICR} is no more complicated than finding the set of rationalizable actions of a finite game.

Proposition 13 *An action $a_i \in A_i$ is FC-rationalizable with respect to an action subspace $A' \in \mathscr{A}$ if and only if there is no $\alpha_i \in \Delta(A_i)$ such that:*

(i) $\max\left\{u_i(a_i, a_{-i}) \,\middle|\, a_{-i} \in A'_{-i}\right\} < \min\left\{U_i(\alpha_i, a_{-i}) \,\middle|\, a_{-i} \in A'_{-i}\right\}$

(ii) $u_i(a_i, a_{-i}) < U_i(\alpha_i, a_{-i})$ *for all* $a_{-i} \in A_{-i} \backslash A'_{-i}$

References

Abreu, D., Pearce, D., and Stacchetti, E. (1990). Towards a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58(5):1041–1063.

Aumann, R. and Brandenburger, A. (1995). Epistemic conditions for nash equilibrium. *Econometrica*, 63(5):1161–1180.

Aumann, R. J. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, 55(12):1–18.

Bade, S., Guillaume, H., and Renou, L. (2009). Bilateral commitment. *Journal of Economic Theory*, 144(4):1817–1831.

Bergemann, D. and Morris, S. (2013). Robust predictions in games with incomplete information. *Econometrica*, 81(4):1251–1308.

Bernheim, B. D. (1984). Rationalizable strategic behavior. *Econometrica*, 52(4):1007–1028.

Block, J. I. (2013). Timing and codes of conduct. Mimeo, Washington University in St Louis.

Block, J. I. and Levine, D. K. (2015). Codes of conduct, private information, and repeated games. *International Journal of Game Theory*. Forthcoming.

Calcagno, R., Kamada, Y., Lovo, S., and Sugaya, T. (2014). Asynchronicity and coordination in common and opposing interest games. *Theoretical Economics*, 9(2):409–434.

Eisert, J., Wilkens, M., and Lewenstein, M. (1999). Quantum games and quantum strategies. *Physical Review Letters*, 83(15):3077–3080.

Forges, F. (1986). An approach to communication equilibria. *Econometrica*, 54(6):1375–1385.

Forgó, F. (2010). A generalization of correlated equilibrium: a new protocol. *Mathematical Social Sciences*, 60(1):186–190.

Fudenberg, D. and Levine, D. K. (1993). Self-confirming equilibrium. *Econometrica*, 61(3):523–545.

Gibbard, A. and Harper, W. L. (1980). Counterfactuals and two kinds of expected utility. In Harper, W. L., Stalnaker, R., and Pearce, G., editors, *Ifs, conditionals, beliefs, decision, chance and time*. Springer.

Halpern, J. Y. and Pass, R. (2012). Game theory with translucent players. Cornell Univeristy, mimeo.

Halpern, J. Y. and Rong, N. (2010). Cooperative equilibrium. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*.

Howard, N. (1971). *Paradoxes of rationality: theory of metagames and political behavior*. MIT University Press.

Iijima, R. and Kasahara, A. (2015). Gradual adjustment and equilibrium uniqueness under noisy monitoring. Mimeo.

Jakobsen, S. K., Sørensen, T. B., and Conitzer, V. (2016). Timeability of extensive-form games. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, pages 191–199. ACM.

Kalai, A. T., Kalai, E., Lehrer, E., and Samet, D. (2010). A commitment folk theorem. *Games and Economic Behavior*, 69(1):123–137.

Kalai, E. (1981). Preplay negotiations and the prisoner's dilemma. *Mathematical Social Sciences*, 1(4):375–379.

Kalai, E. and Stanford, W. G. (1985). Conjectural variations strategies in accelerated cournot games. *International Journal of Industrial Organization*, 3(2):133–152.

Kamada, Y. and Kandori, M. (2009). Revision games. Mimeo.

Kamada, Y. and Kandori, M. (2011). Asynchronous revision games. Mimeo.

Kreps, D. M. and Wilson, R. (1982). Sequential equilibria. *Econometrica*, 50(4):863–894.

Levine, D. K. and Pesendorfer, W. (2007). The evolution of cooperation through imitation. *Games and Economic Behavior*, 58(2):293–315.

Lewis, D. (1979). Prisoner's dilemma is a Newcomb problem. *Philosophy and Public Affairs*, 8(3):235–240.

Monderer, D. and Tennenholtz, M. (2009). Strong mediated equilibrium. *Artificial Intelligence*, 173(1):180–195.

Moulin, H. and Vial, J. (1978). Strategically zero-sum games: the class of games whose completely mixed equilibria cannot be improved upon. *International Journal of Game Theory*, 7(3-4):201–221.

Myerson, R. B. (1986). Multistage games with communication. *Econometrica*, 54(2):323–358.

Nishihara, K. (1997). A resolution of n-person prisoners' dilemma. *Economic Theory*, 10(3):531–540.

Nishihara, K. (1999). Stability of the cooperative equilibrium in n-person prisoners' dilemma with sequential moves. *Economic Theory*, 13(2):483–494.

Osborne, M. J. and Rubinstein, A. (1994). *A course in game theory*. MIT University Press.

Pearce, D. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52(4):1029–1050.

Perea, A. (2013). Belief in the opponents' future rationality. Maastricht University, mimeo.

Rapoport, A. (1965). *Prisoner's dilemma.* University of Michigan Press.

Renou, L. (2009). Commitment games. *Games and Economic Behavior*, 66(1):488–505.

Tennenholtz, M. (2004). Program equilibrium. *Games and Economic Behavior*, 49(2):363–373.

A. Alternative definition of extensive form mechanisms

This section provides alternative definitions that slightly generalize the notions of EFM (Definition 3.1) and implementation (Definition 3.2). I will use the term EFM' to distinguish the alternative class of mechanisms defined ahead from the one in the main text. The class of EFM' is s is a strict superset of the class of EFM. However, all the results in the paper remain true under either definition.

The first requirement for an extensive form game to be an EFM' is that it must preserve the outcome and preference structure of the environment. That is, there must be a preference-preserving map from terminal nodes (outcomes of the game) to action profiles (outcomes of the environment).

Definition A.1 An *outcome homomorphism* is a function τ from terminal nodes *onto* action profiles preserving preferences, i.e. such that $v(z) = u(\tau(z))$ for every terminal node z. G is *outcome equivalent* to E if it admits an outcome homomorphism.

The next requirement is that each player should *freely* choose his own action at some point in the game. Formalizing this idea requires a form of identifying moves (choices in the game) with actions (choices in the environment). For the remainder of this section, let G be outcome equivalent to E and fix an outcome homeomorphism τ. For every player i and every corresponding decision node y, τ induces a *representation* relationship $\approx_y$ from the set of moves available at y in the game to the set of i's actions in the environment. A move m represents action a_i at y, if and only if choosing m at y in the game has the same effect in (payoff-relevant) outcomes as choosing a_i in the environment. This idea is formalized by the following definition.

Definition A.2 Given a player $i \in I$ and a decision node $y \in Y_i$, a move $m \in M(y)$ *represents* an action $a_i \in A_i$ at y if and only if:

(i) $\tau_i(z) = a_i$ for every $z \in Z(y, m)$

(ii) There exist $m' \in M(y)$ and $z \in Z(y, m')$ such that $\tau_i(z) \neq a_i$

The representation relationship is denoted by $m \approx_y a_i$, and $M^{a_i}(y)$ denotes the set of moves that represent a_i at y. A move is *pivotal* at y if and only if it represents some action.

The first requirement for $m \approx_y a_i$ is that, if i chooses m at y, then the game will end at a terminal node which is equivalent to a_i according to τ_i. This is regardless of any previous or future moves by either i or his opponents. The second requirement is that, after the game reaches y, i could still choose a different move m' after which the game remains open to the possibility of ending at a terminal node that is *not* equivalent to a_i.

Definition A.3 A decision node $y \in Y_i$ is *pivotal* for player $i \in I$ if and only if $M^{a_i}(y) \neq \emptyset$ for every $a_i \in A_i$. $D_i \subseteq Y$ denotes the set of pivotal nodes for i.

In words, a decision node y is pivotal for player i if for every action $a_i \in A_i$ there exists a pivotal move which represents it at y. The central property of EFGs is that *every player makes a pivotal move at a pivotal node* along every possible play of the game. A final technical condition is that a player should always know when he is making a pivotal move representing some action.

Definition A.4 (G, τ) satisfies *full disclosure of consequences* if and only if $\approx_y = \approx_{y'}$ whenever y and y' belong to the same information set.

Finally, the alternative versions of definitions 3.1 and 3.2 are as follows:

Definition 3.1' A *extensive form mechanism* is a tuple $\left(G, \tau\right)$ satisfying full disclosure of consequences and such that for every terminal node z and every player

i, there exists a pivotal node $y \in D_i$ and a pivotal move $m \in M^{\tau_i(z)}$ such that $z \in Z(y, m)$.

Definition 3.2' $\alpha \in \Delta(A)$ is (Nash, sequentially, ...) *implementable* if and only if it there exist a mechanism $\big(G, \tau\big)$ and a (Nash, sequential, ...) equilibrium $s^* \in S$ such that for every $a \in A$:

$$\alpha(a) = \zeta^* \left(\tau^{-i}(a)\right) = \sum_{z \in Z} \zeta(z|s^*, s_0) \cdot \mathbb{1}\Big(\tau(z) = a\Big).$$

B. Omitted proofs

B.1. Necessary conditions for implementation

Proof of Proposition 3. I will proof the Proposition using the alternative definitions from section A. Because every EFM according to Definition 3.1 is also an EFM' according to the alternative definition, the proof implies that the result holds true using either of the two definitions. Consider an EMF' mechanism (G, τ), a NE s^* and let α be the induced distribution. I will show that α is an ICE.

Fix any two of actions $a_i^*, a_i' \in A_i$ with $\alpha_i(a_i^*) > 0$ and $a_i' \neq a_i^*$. For each information set $H \in \mathscr{H}_i$, let $M^*(H)$ be the set of moves that represent a_i^* at H *and* are chosen with positive probability. Also, let $\in \mathscr{H}_i^*$ be the set of information sets *along the equilibrium path* in which i chooses a move representing a_i^* with positive probability according to s^*. Finally, let ζ^* be distribution over nodes induced by s^*. All the expectations and conditional distributions in this proof are with respect to ζ^*.

Every $H \in \mathscr{H}_i^*$ must be pivotal, and thus admits a move $m' \in M^{a_i'}(H)$ representing a_i'. Since s^* is a NE, and H is along the equilibrium path, for each $m^* \in M^*(H)$:

$$\mathbb{E}\Big[\, u_i(a_i^*, a_{-i}) \,\Big|\, H, m^* \Big] \geq \mathbb{E}\Big[\, u_i(a_i', a_{-i}) \,\Big|\, H, m' \Big], \tag{3}$$

where H, m denotes the set of nodes $H \times \{m\}$ for $m \in \{m^*, m'\}$.

Let $\Phi^H \subseteq H$ denote the event that τ_{-i} is already determined at H, i.e.:

$$\Phi^H = \left\{ y \in H \,\middle|\, \Big(\forall z, z' \in Z(y)\Big)\Big(\tau_{-i}(z) = \tau_{-i}(z')\Big) \right\}, \tag{4}$$

and let $\bar{\Phi}^H = H \backslash \Phi^H$ be its complement. Notice that the probability of Φ^H and the distribution of $\tau_{-i}^{-1}(a_{-i})$ conditional on Φ^H, are independent from i's choice at H. Hence, by Bayes' rule:

$$\begin{aligned}
\mathbb{E}\Big[u_i(a_i', a_{-i}) \,\Big|\, H, m' \Big] &= \zeta^* \left(\Phi^H \,\middle|\, H, m' \right) \mathbb{E}\Big[u_i(a_i', a_{-i}) \,\Big|\, H, m', \Phi^H \Big] \\
&\qquad + \zeta^* \left(\bar{\Phi}^H \,\middle|\, H, m' \right) \mathbb{E}\Big[u_i(a_i', a_{-i}) \,\Big|\, H, m', \bar{\Phi}^H \Big] \\
&= \zeta^* \left(\Phi^H \,\middle|\, H, m^* \right) \mathbb{E}\Big[u_i(a_i', a_{-i}) \,\Big|\, H, m^* \Big] \\
&\qquad + \zeta^* \left(\bar{\Phi}^H \,\middle|\, H, m^* \right) \mathbb{E}\Big[u_i(a_i', a_{-i}) \,\Big|\, H, m', \bar{\Phi}^H \Big] \\
&\geq \zeta^* \left(\Phi^H \,\middle|\, H, m^* \right) \mathbb{E}\Big[u_i(a_i', a_{-i}) \,\Big|\, H, m^* \Big] + \zeta^* \left(\bar{\Phi}^H \,\middle|\, H, m^* \right) \underline{w}_i(a_i').
\end{aligned}$$

Together with (3), this yields the following inequality which *does not depend on* m':

$$\mathbb{E}\Big[u_i(a_i^*, a_{-i}) \,\Big|\, H, m^* \Big] \geq \zeta^* \left(\Phi^H \,\middle|\, H, m^* \right) \mathbb{E}\Big[u_i(a_i', a_{-i}) \,\Big|\, H, m^* \Big] + \zeta^* \left(\bar{\Phi}^H \,\middle|\, H, m^* \right) \underline{w}_i(a_i').$$

The last inequality holds for for each point in the game where i chooses a_i^* with positive probability. Integrating over them yields:

$$\sum_{a_{-i} \in A_{-i}} \zeta^*(a_i^*, a_{-i}) u_i(a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}} \Big[\zeta^*(-i, a_i^*, a_{-i}) u_i(a_i', a_{-i}) + \zeta^*(i, a_i^*, a_{-i}) \underline{w}_i(a_i') \Big].$$

After rearranging terms, using Bayes to write $\zeta^*(i, a_i^*, a_{-i}) = \zeta^*(i|a_i^*, a_{-i})\zeta^*(a_i^*, a_{-i})$, and factorizing $\zeta^*(a_i^*, a_{-i})$, this equation corresponds to condition (1). Since i, a_i^* and a_i' were arbitrary, it follows that α is an ICE. ■

B.2. Sufficient conditions for implementation

Since every EFM according to Definition 3.1 is also a EFM' according to Definition 3.1', I prove the sufficiency results using the former definition. Proposition 4 is a Corollary of Proposition 10 proven in Section B.5.

Proof of Proposition 5. First, consider the case in which some player i has an absolutely dominated actions, say a_i, and let a_{-i} be $-i$'s *unique* best response to a_i. In this case (a_i, a_{-i}) is the *unique* ICE with respect to A^{ICR}, and it is a (sequential) equilibrium of the simultaneous move game.

Now, suppose that there are no absolutely dominated actions and let α^* be any *strict* ICE. By Lemma 14, there exists a distribution α^0 with full support that can arise from a pure-strategy sequential equilibrium s^0 of an EFM G^0. For each $\mu \in (0,1)$ let $\alpha^\mu = \mu\alpha^0 + (1-\mu)\alpha^*$, and let G^μ be the EFM constructed as follows. At the initial node, chance chooses G^0 with probability μ, and the mediated mechanism G^* (see section 6) corresponding to α^* with probability $(1-\mu)$. Then, a new information partition is created, by combining all the information sets in which the mediator recommends some a_i in G^*, with all the information sets in which i plays a_i in G^0 according to s_i^0. Finally, let s^μ be the strategy profile that follows recommendations in G^*, and mimics s^0 in G^0.

Since A is finite (which means there are finitely many incentive constraints), α^* is a strict ICE and s^0 is a sequential equilibrium of G^0, s^μ is a Nash equilibrium for μ sufficiently close to 0. Also, since α^0 has full support, every information set in G^μ is reached with positive probability, and thus s^μ is a sequential equilibrium. Moreover, $\alpha^\mu \to \alpha^*$ as $\mu \to 0$. Hence, α^* can be approximated by sequentially implementable distributions. ■

Lemma 14 *If a 2×2 environment has no absolutely dominated actions, then there is an sequentially pure-strategy implementable distribution α such that $\alpha(a) > 0$ for every $a \in A$.*

Proof. Let $A_i = \{a_i, b_i\}$ for $i = 1, 2$, and suppose that there are no repeated payoffs nor absolutely dominated actions. The result is straightforward if there are no strictly dominated strategies, because then there exists a completely mixed (sequential) equilibrium. As usual, the randomization can be delegated to chance, so that players use pure strategies in the implementation. The interesting cases are when there are no absolutely dominated strategies, but at least one player has a strictly dominated strategy.

Let $\lambda_i, \lambda_i' \in \Lambda_i$ denote the counterfactual beliefs:

$$\lambda_i(a_{-i}|a_i) = 1 \;\wedge\; \lambda_i(b_{-i}|b_{-i}) = 1, \quad \text{and} \quad \lambda_i'(b_{-i}|a_i) = 1 \;\wedge\; \lambda_i'(a_{-i}|b_{-i}) = 1. \tag{5}$$

If b_i is not absolutely dominated but it is strictly dominated by a_i, then it must be a best response to either λ_i or λ'_i. Furthermore, since there are no repeated payoffs, it must be a *strict* best response. There are two cases to consider depending on whether one or two players have dominated strategies.

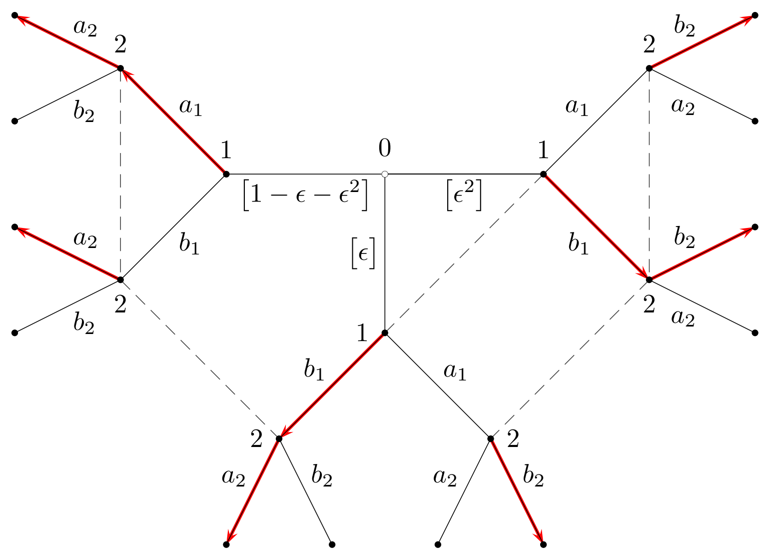

Figure 6 – Implementation of b_1 when it is the only dominated action.

Fist suppose that player 2 has no dominated strategies but b_1 is dominated by a_1. Further assume (without loss of generality) that a_2 is a best response to a_1. This implies that b_2 is the unique best response to b_1, and that (a_1, a_2) is a *strict* NE of the simultaneous move game. If b_1 is a best response to λ_1, then it suffices to have player 1 move first and make his choice public. By backward induction, in the unique SPNE, player 2 will choose a_2 if he chooses a_1 and b_2 if he chooses b_1. Hence, 1's counterfactual beliefs are λ_1 and b_1 is the unique best response.

Otherwise, if a'_1 is a best response to λ'_1, then it can be implemented as an equilibrium of the mechanism in Figure (6), with $\epsilon > 0$ small enough. The equilibrium strategies are represented with arrows. Player's are willing to choose a_i because (a_1, a_2) is a strict Nash equilibrium. Player 2 is willing to choose b_2 because it is a best response to b_1. Player 1 is willing to choose b_1 because his conjectures at that moment are close enough to λ'_1. Since all the information sets are on the equilibrium path, the equilibrium is sequential.

Finally, suppose that both players have strictly dominated strategies, say b_1 and b_2. In this case (a_1, a_2) is a *strict* NE. If b_i is a best response to λ_i, then it can

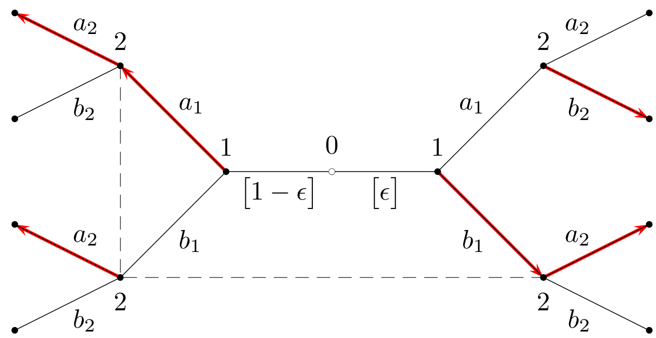

Figure 7 – Implementation of b_1 when b_2 is also dominated.

be implemented as a NE of the mechanism where i moves first and $-i$ chooses b_{-i} along the equilibrium path and punishes deviations with a_{-i}. Otherwise, if b_i is a best response to λ_i', then it can be played with positive probability in a NE of the mechanism depicted in figure 7, with $\epsilon > 0$ small enough. Hence there always exists EFMs G^1 and G^2 with NE in which b_1 and b_2 are played with positive probability.

The proof is not complete because the equilibria are not subgame perfect. For that purpose, one can construct a third mechanism in which nature randomizes between G^1 and G^2 and the simultaneous move game, and every action is played with positive probability along the equilibrium path. Information sets can be connected so that, whenever a player is supposed to choose b_i he believes that he is in G^i. Doing so guarantees that the equilibrium is sequential. ■

Proof of Proposition 9. Suppose that $\alpha \in \Delta(A)$ is an ICE and let θ be such that (α, θ) satisfy (1) strictly for all i and a with $B = A$. For every $\epsilon > 0$, just as in section 2, it is possible to construct random variables t_1^a and t_2^a with the property that

$$|\Pr(t_i^a < t_{-i}^a | t_i^a) - \theta(i|a)| < \epsilon, \tag{6}$$

for $i = 1, 2$. Now, consider the EFM in which Nature chooses an intended action profile a^* according to α, and two moments in time t_1 and t_2 according to a distribution satisfying (6). At moment t_i, player i is recommended to play a_i^* along the equilibrium path, and the harshest punishment for $-i$ if a deviation has already taken place. Since α is a *strict* ICE, for ϵ small enough, following these recommendations constitutes a NE. ■

B.3. Properties of ICE, ICR, and FICR

Proof of Proposition 1. Note that $u_i(a'_i, a_{-i}) \geq \underline{w}_i(a'_i|B^*)$ by definition. Hence, condition (1) becomes tighter for higher values of $\theta(i|a)$. Setting $\theta(i|a) \equiv 0$ in condition (1) yields the definition of correlated equilibrium. Hence, α is a correlated equilibrium, then condition (1) is satisfied for any θ, which implies that α is also an ICE with respect to any B.

Similarly, if α is an ICE with respect to some B, then condition (1) is satisfied for some θ. Consequently, it is also satisfied setting $\theta(i|a) \equiv 1$. This implies that for every $a'_i \in A_i$ we have

$$\sum_{a_{-i}\in A_{-i}} \alpha(a)u_i(a) \geq \underline{w}_i(a'_i|B^*) = \min_{a'_{-i}\in B^*} u_i(a') \geq \min_{a'_{-i}\in A_{-i}} u_i(a')$$

In particular, this is true for whichever a'_i achieves i's minimax. Therefore,

$$\sum_{a_{-i}\in A_{-i}} \alpha(a)u_i(a) \geq \max_{a'_i\in A_i} \min_{a'_{-i}\in A_{-i}} u_i(a').$$

That is, α is individually rational.

Let $\Gamma \subseteq \Delta(A \times I)$ be the set of distributions satisfying

$$\sum_{a_{-i}\in A_{-i}} \gamma(a)u_i(a) - \gamma(a, -i)u_i(a'_i, a_{-i}) - \gamma(a, i)\underline{w}_i(a'_i|A) \geq 0.$$

Since Γ is defined by a finite set of affine inequalities, it is a closed and convex polytope. The set of ICE with respect to A is the projection of Γ on $\Delta(A)$. ■

Proof of Proposition 2. ICR actions are clearly not absolutely dominated. For the opposite direction, fix an action $a^*_i \in A_i$ that is not absolutely dominated in A'. Let $a^*_{-i} \in \arg\max_{a_{-i}\in A_{-i}} u_i(a^*_i, a_{-i})$. Since a^*_i is not dominated in A', for every $a'_i \in A_i$ there exists some $a_{-i}(a'_i) \in A'_{-i}$ such that $u_i(a^*) \geq u_i\Big(a'_i, a_{-i}(a'_i)\Big)$. Hence a^*_i is a best response to $\lambda_i \in \Lambda_i(A')$, with $\lambda_i(a^*_{-i}|a^*_i) = 1$ and $\lambda_i\Big(a_{-i}(a'_i)\Big) = 1$ for $a'_i \neq a^*_i$.

An elimination procedure can be described by a function $K : \mathscr{A} \to \mathscr{A}$, describing kept actions, such that for $A' \in \mathscr{A}$: (i) never adds new actions, i.e. $K(A') \subseteq A'$; (ii) never eliminates undominated actions, i.e. $\mathrm{CR}(A') \subseteq K(A')$; and (iii) if there are dominated actions then it always eliminates at least one, i.e. $\mathrm{CR}(A') \neq A'$ implies $K(A') \neq A'$.

Now consider the corresponding sequence of surviving actions $(A^n) \in \mathscr{A}^{\mathbb{N}}$ defined recursively by $A^1 = A$ and $A^{n+1} = K(A^n)$. For $n \in \mathbb{N}$ with $A^n \neq \emptyset$, there exists some action profile $a^0 \in A^n$. Ans, since the game is finite, for each player i there exists a best response a_i^* to a_{-i}^0. By (ii) this implies that $a^* \in \mathrm{CR}(A^n) \subseteq A^{n+1}$. Thus, by induction, (A^n) is weakly decreasing sequence of *nonempty* sets. Therefore, since $\mathscr{A}$ is finite, (A^n) converges in finite iterations to a nonempty limit A^*. Since $A^{\mathsf{ICR}} \subseteq \mathrm{CR}(A^{\mathsf{ICR}})$ and $\mathrm{CR}(\cdot)$ is $\subseteq$-monotone, (ii) implies that $A^{\mathsf{ICR}} \subseteq A^n$ for all $n \in \mathbb{N}$, and thus $A^{\mathsf{ICR}} \subseteq A^*$. Finally, (iii) implies that $A^* \subseteq \mathrm{CR}(A^*)$ and thus $A^* \subseteq A^{\mathsf{ICR}}$. ■

Proof of Proposition 13. $a_i^* \in \mathrm{FICR}_i(A')$ if and only if it is a best response to *some* $\lambda_i = \mu\lambda_i^0 + (1-\mu)\lambda_i^1$, with $\lambda_i^0 \in \Delta(A_{-i} \backslash A'_{-i})$, $\lambda_i^1 \in \Lambda(A'_{-i})$ and $\mu \in [0,1]$. Which holds if and only if it is a best response to those beliefs which are more favorable for a_i^*, i.e. beliefs with:

$$\lambda_i^1\left(\operatorname*{arg\,max}_{a_{-i}\in A_{-i}}\left\{u_i(a_i^*,a_{-i})\right\}\,\middle|\,a_i^*\right) = 1, \quad \text{and} \quad \lambda_i^1\left(\operatorname*{arg\,min}_{a_{-i}\in A_{-i}}\left\{u_i(a_i,a_{-i})\right\}\,\middle|\,a_i \neq a_i^*\right) = 1.$$

Hence, after some simple algebra, $a_i^* \in \mathrm{FICR}_i(A')$ if and only if for every $a_i' \in A_i$:

$$(1-\mu)\Big[\bar{w}_i(a_i^*) - \underline{w}_i(a_i')\Big] + \sum_{a_{-i}\notin A'_{-i}} \mu\lambda_i^0(a_{-i})\Big[u_i(a_i^*,a_{-i}) - u_i(a_i',a_{-i})\Big] \geq 0,$$

where $\bar{w}_i(a_i^*, A') \equiv \max_{a_{-i}\in A'_{-i}}\left\{u_i(a_i^*,a_{-i})\right\}$ That is, if and only if it is a best response to some (non-counterfactual) belief in the auxiliary strategic form game $(I, \tilde{A}, \tilde{u})$ with $\tilde{A}_i = A_i$, $\tilde{A}_{-i} = \Big(A_{-i} \backslash A_{-i}\Big) \cup \{a_{-i}^0\}$, and $\tilde{u}_i : \tilde{A} \to \mathbb{R}$ given by:

$$\tilde{u}_i(a_i,a_{-i}) = \begin{cases} u_i(a_i,a_{-i}) & \text{if} \quad a_{-i} \notin A'_{-i} \\ \bar{w}_i(a_i^*, A') & \text{if} \quad a_{-i} \in A'_{-i} \wedge a_i = a_i^* \\ \underline{w}_i(a_i', A') & \text{if} \quad a_{-i} \in A'_{-i} \wedge a_i \neq a_i^* \end{cases}.$$

The result then follows from the well known equivalence between never best responses and dominated actions, cf. Lemma 3 in Pearce (1984). ■

B.4. Prisoners' dilemma

Proof of Proposition 6. Suppose that $\alpha \in \Delta(A)$ is supported as an ICE by θ. Since D is a dominant strategy, the incentive constraints for it are automatically satisfied. After some simple algebra, the incentive constraints when player 1 is asked to play C can be written as:

$$\theta(1|\mathrm{C},\mathrm{C}) \geq \frac{l\alpha(\mathrm{C},\mathrm{D}) + g\alpha(\mathrm{C},\mathrm{C})}{(1+g)\alpha(\mathrm{C},\mathrm{C})}, \tag{7}$$

and the corresponding constraint for player 2 can be written as:

$$\theta(1|\mathrm{C},\mathrm{C}) = 1 - \theta(2|\mathrm{C},\mathrm{C}) \leq \frac{\alpha(\mathrm{C},\mathrm{C}) - l\alpha(\mathrm{D},\mathrm{C})}{(1+g)\alpha(\mathrm{C},\mathrm{C})}. \tag{8}$$

Hence α is an ICE if and only if there exists a number $\theta_0 \in [0,1]$ such that $\theta(1|\mathrm{C},\mathrm{C}) = \theta_0$ satisfies both (7) and (8). This happens if and only if:

$$\frac{\alpha(\mathrm{C},\mathrm{C}) - l\alpha(\mathrm{D},\mathrm{C})}{(1+g)\alpha(\mathrm{C},\mathrm{C})} \geq \frac{l\alpha(\mathrm{C},\mathrm{D}) + g\alpha(\mathrm{C},\mathrm{C})}{(1+g)\alpha(\mathrm{C},\mathrm{C})},$$

which is equivalent to condition (2). Note that $\theta(1|\mathrm{C},\mathrm{C}) + \theta(2|\mathrm{C},\mathrm{C}) = 1$ is the constraint that restricts the set of ICE. This constraint implies that in order to increase $\theta(1|\mathrm{C},\mathrm{C})$ (as to relax (7)), one has to decrease $\theta(2|\mathrm{C},\mathrm{C})$ (which tightens (8)). ■

Proof of proposition 8. The proof is divided into four cases, depending on the properties of the environment induced by μ. First, if (C,C) is an ICE, the total per capita sentence in equilibrium cannot be greater than $-u_i(\mathrm{C},\mathrm{C}) = \mu_0$. Second, suppose that (D,D) is a dominant strategy but, $u_i(\mathrm{D},\mathrm{D}) \geq u_i(\mathrm{C},\mathrm{C})$. In this case, once again, the per capita sentence cannot be greater than $-u_i(\mathrm{D},\mathrm{D}) \leq -u_i(\mathrm{C},\mathrm{C}) = \mu_0$. Third, suppose that (C,C) is *not* and ICE, but (D,D) is *not* dominant. Since (C,C) is not and ICE, it must be the case that

$$u_i(\mathrm{C},\mathrm{C}) < \frac{1}{2}u_i(\mathrm{D},\mathrm{C}) + \frac{1}{2}u_i(\mathrm{D},\mathrm{D}) \quad \Rightarrow \quad \mu_2 < 2\mu_0 - \mu_1^+ \tag{9}$$

Since (C,C) is not a Nash equilibrium and (D,D) is not dominant, it follows that

$$u_i(\mathrm{D},\mathrm{D}) \leq u_i(\mathrm{C},\mathrm{D}). \tag{10}$$

This implies that (D,C) is a Nash equilibrium and thus an ICE. Consequently, in this case, the per capita sentence cannot be greater than

$$-\frac{1}{2}u_i(\mathrm{D},\mathrm{C})-\frac{1}{2}u_i(\mathrm{C},\mathrm{D})\leq-\frac{1}{2}u_i(\mathrm{D},\mathrm{C})-\frac{1}{2}u_i(\mathrm{D},\mathrm{D})<\frac{1}{2}\mu_i^{+}+\mu_0-\frac{1}{2}\mu_2=\mu_0,$$

where the first inequality follows from (10) and the second one from (9). In all these cases, the maximum possible per capita sentence is bounded above by $\mu_0 \leq \bar{\mu}_0 < \min\{\bar{\mu}_1, 2\bar{\mu}_0\}$.

The actual optimal policy belongs to the last remaining case, with (D,D) being strictly dominant, $u_i(\mathrm{C},\mathrm{C}) > u_i(\mathrm{D},\mathrm{D})$ and (C,C) *not* being an ICE. This case corresponds to a prisoner's dilemma in which cooperation is not an ICE. Hence, by corollary 7, $g > 1$ and the only ICE is (D,D). Which means that the optimal policy is obtained by maximizing μ_2 subject to the following constraints

$$0\leq\mu_0\leq\bar{\mu}_0 \quad\wedge\quad 0\leq\mu_1^{+},\mu_1^{-}\leq\bar{\mu}_1 \quad\wedge\quad 0\leq\mu_2\leq\bar{\mu}_2 \tag{11}$$

$$\mu_1^{+}<\mu_0 \quad\wedge\quad \mu_2<\mu_1^{-} \quad\wedge\quad \mu_0<\mu_2 \tag{12}$$

$$2\mu_0>\mu_2+\mu_1^{+} \tag{13}$$

Condition (11) requires the policy to be feasible. Condition (12) requires the induced environment to be a prisoners' dilemma. Condition (13) is equivalent to $g > 1$.

Note that μ_1^+ and μ_1^- only appear as lower and upper bounds of μ_2, respectively. Hence, it is optimal to set $\mu_1^+ = 0$ and $\mu_1^- = \bar{\mu}_1$. The program thus reduces to maximizing μ_2 subject to

$$2\mu_0>\mu_2>\mu_0 \qquad\wedge\qquad 0\leq\mu_2\leq\bar{\mu}_1 \qquad\wedge\qquad 0\leq\mu_0\leq\bar{\mu}_0.$$

Since $\bar{\mu}_0 < \bar{\mu}_1$, the constraint $\mu_0 \leq \bar{\mu}_0$ is binding, which means that the program reduces to

$$\max\left\{\mu_2 \,\middle|\, \mu_2\leq 2\bar{\mu}_0 \,\wedge\, 0\leq\mu_2\leq\bar{\mu}_1\right\}=\max\{\bar{\mu}_1,2\bar{\mu}_0\}.$$

■

B.5. Quasi-sequential equilibrium

The proof of Proposition 10 is divided in two parts regarding necessity and sufficiency. To establish necessity it suffices to show that given a QSE of an EFM, every action played with positive probability (on or off the equilibrium path) is in A^{FICR}. Then the proof of Proposition 3 applies simply replacing $\underline{w}_i(a_i', A)$ with $\underline{w}_i(a_i, A^{\mathsf{FICR}})$. This fact is established in Lemma 15. Given an EFM and an QSE (s^*, ψ^*), let $A_i^* \subseteq A_i$ denote the set of actions that i plays with positive probability is some information set, i.e.:

$$A_i^* = \left\{ a_i \in A_i \,\middle|\, \Big(\exists H \in \mathscr{H}\Big)\Big(\exists m \in M^{a_i}(H)\Big)\Big([s_i^*(H)](m) > 0\Big) \right\}.$$

Lemma 15 *Every quasi-sequential equilibrium s^* of an extensive form mechanism satisfies $A^* \subseteq A^{\mathsf{FICR}}$.*

Proof. Fix some $a_i^* \in A_i^*$ chosen with positive probability in some $H \in \mathscr{H}_i$, and a move $m^{a_i^*} \in M^{a_i^*}(H)$ that represents a_i^* and is chosen with positive probability. For each other action $a_i' \neq a_i^*$, pick a move $m^{a_i'} \in M^{a_i}(H)$ representing a_i' at H. Now let $\mu = \psi_i^*\Big(\Phi^H \,\Big|\, H\Big) \in [0,1]$, where Φ^H is the event that τ_{-i} is already determined at H, as defined in (4). Finally, let $\lambda_i^0 \in \Delta(A_{-i})$ and $\lambda_i^1 \in \Lambda_i(A^*)$ be the given by:

$$\lambda_i^0(a_{-i}) = \zeta_i^*\left(\tau_{-i}^{-1}(a_{-i}) \,\middle|\, H, \Phi^H\right), \quad \text{and} \quad \lambda_i^1(a_{-i}|a_i) = \zeta_i^*\left(\tau_{-i}^{-1}(a_{-i}) \,\middle|\, H, m^{a_i}, \bar{\Phi}^H\right),$$

and let $\lambda_i = \mu\lambda_i^0 + (1-\mu)\lambda_i^1$.

Being that $\zeta_i^*\Big(\Phi^H \,\Big|\, H, m\Big)$ and $\zeta_i^*\Big(\tau_{-i}^{-1}(a_{-i})|H, m, \Phi^H\Big)$ are independent from m, sequential rationality implies that for every deviation a_i':

$$\begin{aligned}
\sum_{a_{-i}\in A_{-i}} \lambda_i(a_{-i}\,|\,a_i^*)u_i(a_i^*, a_{-i}) &= \sum_{a_{-i}\in A_{-i}} \zeta_i^*\left(\tau_{-i}^{-1}(a_{-i})\,|\,H, m^{a_i^*}\right)u_i(a_i^*, a_{-i}) \\
&\geq \sum_{a_{-i}\in A_{-i}} \zeta_i^*\left(\tau_{-i}^{-1}(a_{-i})\,|\,H, m^{a_i'}\right)u_i(a_i', a_{-i}) \\
&= \sum_{a_{-i}\in A_{-i}} \lambda_i(a_{-i}\,|\,a_i')u_i(a_i', a_{-i}).
\end{aligned}$$

Hence a_i^* is a best response to $\lambda_i^* \in \Lambda_i(A^*)$, and thus $a_i^* \in \mathrm{FICR}(A^*)$. This holds for all i and $a_i^* \in A_i^*$. Hence, $A^* \subseteq \mathrm{FICR}(A^*)$ and thus $A^* \subseteq A^{\mathsf{FICR}}$. ■

The sufficiency proof is constructive, and the mechanics behind the construction are as follows. Every action $a_i^0 \in A_i^{\mathsf{FICR}}$ can be rationalized by some beliefs about future choices in A_{-i}^{FICR} and about arbitrary equilibrium or arbitrary past choices. Off path beliefs are assigned in such a way that, whenever i is asked to choose a_i^0, he naively believes that doing so is in his best interest. Since weak consistency does not imply any consistency requirements *across* players, this can always be done even if it implies that *i must be certain that his opponent is or will be mistaken.*

Proof of sufficiency for Proposition 10. Fix an ICE α with respect to A^{FICR}. I will construct an extensive form mechanism (G, τ) and a QSE (s^*, ψ) implementing it. As an intermediate step, let G^0 denote the mediated game which implements α as an ICE with respect to A^{FICR}. I will add additional off-path histories to guarantee that the equilibrium becomes QS. Since equilibrium path remains unchaged, it is sufficient to ensure that sequential rationality off the equilibrium path, and that the off-path beliefs are weakly consistent.

In the construction, all the players' information sets are pivotal and have a unique pivotal move representing each action, and all the moves in each pivotal information set are pivotal, i.e. $M(H) = \cup_{a_i \in A_i} M^{a_i}(H)$ and $\#M(H) = \#A_i$ for $H \in \mathscr{H}_i$. Furthermore, the only information that a player has at the moment of making his choice is the action that he is supposed to choose. Hence, it is possible to specify equilibrium strategies by labelling each information set with the distribution of actions that the corresponding player is supposed to follow. For instance H^{a_i} represents a pivotal information set in which, according to s^*, i chooses the only move which represents a_i in H^{a_i}.

Fix a player i and some action $a_i^0 \in A_i^{\mathsf{FICR}} \backslash \operatorname{supp}(\alpha_i)$. Since A^{FICR} is self-FC-rationalizable, a_i^0 is a best response to some counterfactual belief $\lambda_i = (1-\mu)\lambda_i^0 + \mu\lambda_i^3$, with $\mu \in [0,1]$, $\lambda_i^0 \in \Delta(A_{-i})$ and $\lambda_i^3 \in \Lambda_i(A^{\mathsf{FICR}})$. $(1-\mu)\lambda_i^0$ can be further decomposed as $(1-\mu)\lambda_i^0 = \gamma\lambda_i^1 + \eta\lambda_i^2$ with $\gamma, \eta \in [0,1]$, $\lambda_i^1 \in \Delta(A_{-i} \backslash A_{-i}^{\mathsf{FICR}})$ and $\lambda_i^2 \in \Delta(A_{-i}^{\mathsf{FICR}})$. Assume without loss of generality that $\lambda_i^3(\bar{a}_{-i}|a_i^0) = 1$ and $\lambda_i^3(\underline{a}_{-i}(a_i')|a_i') = 1$ for every $a_i' \neq a_i^0$, where $\bar{a}_{-i} \in \arg\max_{a_{-i} \in A_{-i}^{\mathsf{FICR}}}\{u_i(a_i^0, a_{-i})\}$ and $\underline{a}_{-i}(a_{-i}) \in \arg\min_{a_{-i} \in A_{-i}^{\mathsf{FICR}}}\{u_i(a_i', a_{-i})\}$.

The entire mechanism starts from an initial node where Nature chooses between G^0 and other additional paths. For each action $a_i^0 \in \mathrm{FICR}_i^{\infty} \backslash \operatorname{supp}(\alpha_i)$, $G^{a_i^0}$ denotes a set of paths on which player i is willing to choose a_i and believe that the future choices of his opponents will be restricted to $\mathrm{FICR}_{-i}^{\infty}$. The set of paths $G^{a_i^0}$ is depicted in Figure 8. The nodes are labelled with circled numbers, and the player moving at each node can be inferred from the subindexes of the information sets.

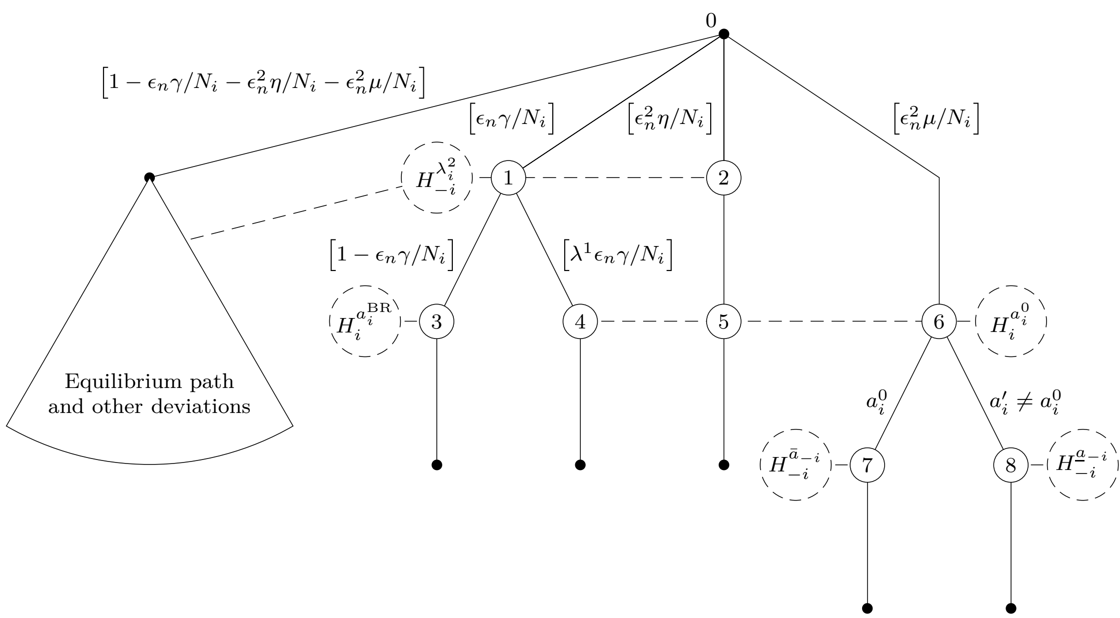

Figure 8 – Incentives for $a_i^0 \in A_i^{\mathsf{FICR}} \backslash A_i^*$.

The numbers within brackets, specify the sequence of mixed strategies that converges to the equilibrium assessment. (ϵ_n) denotes an arbitrary sequence of sufficiently small positive numbers converging to 0, and $N_i = \#A_i^{\mathsf{FICR}}$ is the number of FC-rationalizable actions. The sequence is not strictly mixed, but reach all the relevant information sets with positive probability.[10] The limit of this sequence generates weakly consistent beliefs. Hence, it only remains to verify the incentive constraints:

- At nodes (1) and (2), player $-i$ is willing to make choices according to λ_i^2 because he believes that he is on the equilibrium path.
- At nodes (7) and (8), $\bar{a}_{-i}$ and $\underline{a}_{-i}$ may not be best responses to a_i^0 or a_i'. However, they are in $\mathrm{FICR}_{-i}^{\infty}$ and thus $-i$ is willing to play them either along the equilibrium path, or on $G^{\bar{a}_{-i}}$ and $G^{\underline{a}_{-i}}$. Since $-i$ will consider the

[10]One could use a strictly mixed sequences by assigning probabilities or order ϵ_n^3 or less to other strategies, but this would only complicate the exposition unnecessarily.

deviations to and in $G^{a_i^0}$ to be unlikely (of order at most ϵ^3), the incentives for these actions are independent from what happens in this figure.

- First suppose that the information sets for i are fully contained in the figure:
  - At (3) player i is supposed to choose an action which is a best response to λ_i^2. And therefore his choice is trivially incentive compatible.
  - It is straightforward to see that equilibrium beliefs for player i would generate a conjecture λ_i at nodes (4)–(6), and thus he would be willing to choose a_i^0.
- Now suppose that either $H^{a_i^0}$ or $H^{a_i^{\text{BR}}}$ appear in other parts of the game. There are only two possibilities:
  - They could appear as punishments in the position analogous to (7) or (8) in some $G^{a_{-i}^0}$. From i's perspective, this has probability of order ϵ^3 or lower, and hence it is irrelevant for i.
  - They could appear in the equilibrium path, or in some $G^{a_{-i}^0}$ in the positions of (3) - (8). In such cases, it will also be a best response to the conditional beliefs and thus to the average beliefs.

■