# An Impossible Trinity for Interactive Epistemology*

Eduardo Zambrano†

First version: November $11^{th}$, 2004

This version: November $9^{th}$, 2005

*I would like to thank John Nachbar, Amanda Friedenberg, Francisco Rodríguez, Muhamet Yildiz and participants at the 2005 Spring Midwest Economic Theory Conference and at a Notre Dame seminar for their comments to a previous version of this paper.

†Department of Finance, Mendoza College of Business, University of Notre Dame, Notre Dame, IN 46556. Phone: 574-631-4597, fax: 574-631-5255, e-mail: ezambran@nd.edu.

**Abstract**

In this paper I devise a novel epistemic model of a normal form game where players have beliefs over epistemic types and over decision rules, that is, over maps from a set of epistemic types to the action space. I show that in any such model of a game, for a large set of games, if beliefs over decision rules allow the possibility of inference and have supports that are sufficiently diverse then they are inconsistent with the opponent employing a rational decision rule in deriving her choice. Hence, as in the learning in games literature, it is true that reasoning about equilibrium in an epistemic model is a more stringent requirement than being in equilibrium.

# 1 Introduction

The purpose of this paper is to study whether an answer exists to a basic question in Economics: can players "deduce their way" to a Nash equilibrium of the game that they are about to play? This is the canonical question that lies at the center of the literature on the foundations of game theory, as well as the counterpart to the central question from the literature on learning in games: can players "learn their way" to a Nash equilibrium of the game?

A number of papers, which prominently include Hart and Mas-Colell (2003), Foster and Young (2001) and Jordan (1993) have shown that it is notoriously difficult to formulate sensible adaptive dynamics that guarantee convergence to Nash equilibrium. These results, it is said, stand in contrast to those obtained in the epistemic literature, where it is now known what assumptions about rationality, knowledge and belief logically imply that the actions or the beliefs of the players constitute a Nash equilibrium (Aumann and Brandenburger, 1995).

In this paper I elaborate from the insights developed by Nachbar (2005, 1997) on learning in repeated games to argue that the epistemic framework suffers from the same kind of difficulties that the learning literature has in justifying Nash equilibrium play if one requires those models to be explicit about whether any kind of *decision process* could have led the players to "do what they do" at a given state of the world.

Key to the development of my results is understanding the formal similarities between a repeated game and an epistemic model of a game. A repeated game calls for each player to make choices over time by selecting a strategy for the repeated game (a map from the set of finite histories to the action set) and to form beliefs over the strategies employed by the other players. In an epistemic model of a game one can define an analogous object at each state of the world: what I call a *decision rule*, that is, a map from the set of types that a player could have had to that player's action set. Hence, one can define an epistemic model of a game by a collection of decision rules for each player together with beliefs by the players over the decision rules employed by the other players. Once this is in place I can ask questions about how belief diversity and the *possibility* of accurate inference interact with players employing rational decision rules in deriving their choices.

I show that in any such enhanced model of a game, for a large set of games, there is an *impossible trinity*: if beliefs over decision rules allow the possibility of inference and have supports that are sufficiently diverse then they are inconsistent with the opponent employing a rational decision rule in deriving her choice.

## 1.1 A preliminary example

The following example comes directly from Nachbar (2005) and sets the stage for the appropriate example in the context of a static game.

Consider an infinitely repeated two-player game with discounting and perfect monitoring based on a stage game where each player has two stage game actions, $L$ and $R$. Suppose that each player's belief over what transpires in the repeated game is a beta distribution over the opponent's i.i.d. strategies. Both players choose a best response to their beliefs over time. This is the model known in the literature on learning in games as *fictitious play*.

The starting point of Nachbar (2005) is that each player in the fictitious play model is certain that her opponent plays a strategy that is *independent of history* even though, if players optimize even approximately, neither player's actual strategy is independent of history. The natural response to this inconsistency is to enrich each player's beliefs by including more complicated strategies along with the i.i.d. strategies. The problem is that, as we include more complicates strategies in a player's beliefs, we also make her best response more complicated, so that the richer theory may still be inconsistent.

Nachbar extends this observation about the i.i.d. model to general Bayesian repeated game models that satisfy natural conditions regarding sequential rationality, learnability, and a belief diversity condition.

## 1.2 Example 1: Cournot

Consider the standard model of Cournot competition between two players A and B. The demand function is given by $P = a - bQ$ and each player has constant marginal costs equal to $c_A$ and $c_B$ for players A and B respectively. Think of player A as choosing their output by means of a decision rule $r_A$ that dictates how much to produce as a function of "type," where a type is simply a probability distribution over player B's rules and player B's types.[1] A type is, therefore, a list $t_A = (t_A^1, t_A^2, \ldots)$ where $t_A^1$ is player A's belief over the rule chosen by B, $t_A^2$ is player A's belief over B's belief about A's rule, and so on. Similarly for player B. Consider the simplest and most common version of this model, namely, one where player A's belief is defined over the set of constant rules $\overline{r}_B \in \mathbb{R}_+$ for B (hence beliefs over rules completely

[1] Weinstein and Yildiz (2005) used a similar example in a different setting to illustrate the importance of properly specifying the higher order uncertainty in a game for the computation of the Bayes Nash equilibria of the game.

determine beliefs over output) and both players choose a best response to their belief.

The starting point of this paper is that each player in this version of the Cournot model is certain that its opponent uses a rule that is *independent of type* even though, if the players attempt to maximize profits even approximately neither player's actual rule is independent of type. Formally we would say that each belief $t_A$ of player A regarding the output choice of player B is defined over rules of the form $r(t_B) = \overline{r}_B$ for all types $t_B$ of B. If player A maximizes profit its best response is

$$r_A(t_A) = \frac{(a - c_A) - E_A\left(\overline{r}_B; t_A^1\right)}{2b} \tag{1}$$

which is a rule for player A that depends on player A's first order type $t_A^1$, that is, on player 1's beliefs over the presumed constant rule chosen by player B.

Just as in the learning in games literature, the natural response to this inconsistency is to enrich each player's beliefs by including more complicated rules along with the rules that are independent of type. The problem is that, as we include more complicated rules in a player's beliefs, we also make her best response rule more complicated, so that the richer theory may still be inconsistent.

An example of this would be the following: consider now that, for whatever reason, each belief $t_A$ of player A regarding player B's decisions is defined over rules like (1) above, that is, over rules of the form $r_B(t_B^1) = \alpha - \beta E_B(\overline{r}_A; t_B^1)$ for all first-order types $t_B^1$ of B. If player A maximizes profit subject to a belief of this kind his best response is

$$\begin{aligned} r_A(t_A) &= \frac{(a - c_A) - E_A(r_B(t_B); t_A)}{2b} \\ &= \frac{1}{2b}\left\{\left[(a - c_A) - E_A\left(\alpha; t_A^1\right)\right] + E_A\left[\beta E_B\left(\overline{r}_A; t_B^1\right); t_A^2\right]\right\} \end{aligned}$$

which is a rule for player A that depends on player A's first order belief about B's rules, $t_A^1$, as well as on player A's belief about player B's belief about A's rule, $t_A^2$. Hence, if both players have beliefs of this kind and are maximizing profits then their beliefs over rules will necessarily be mis-specified, namely, both players would be using more complicated decision rules than the ones they consider possible for one another. Moreover, if in addition player A knew that B's beliefs were of this form then player A would be certain and wrong that player B was not using a profit maximizing rule in deriving their choice of output.

A final observation: nothing precludes in either version of the Cournot model that the type held by each player at a particular state of the world be such that they are certain of the profile of output choices chosen by both. Then profit maximization naturally would lead to the Cournot outcome being actually chosen by both players even as the players are certain than the opponent is not employing a rational decision rule in deriving their choice. It is in this sense that, as in the learning in games literature, reasoning about equilibrium in an epistemic model is a more stringent requirement than being in equilibrium.

In this paper I extend these observations about the model of a 'Cournot' game to general epistemic models of non-cooperative games that satisfy natural conditions regarding rationality, the possibility of prediction, and a belief diversity condition.

As it is quite clear, the present paper owes much of its inspiration and methodology to the work of Nachbar (1997, 2005) on repeated games. The exercise I carry out in the present paper is not, however, a straightforward translation of his research into the language of interactive epistemology. There is simply too much that does not translate directly, and modelling choices had to be made throughout the entire exercise. Hence I believe it is important to keep in proper perspective what the similarities and the differences are between Nachbar's research and the present paper. I explain those similarities and differences in detail in Section 4.

It is also important to compare the results presented in this paper with the seminal work by Aumann and Brandenburger (1995), who developed epistemic conditions that lead to Nash equilibria. The results presented here do not contradict in any way those of Aumann and Brandenburger (1995), but they do show that there is a sense in which they justify Nash equilibrium "by accident," namely, as the players themseves are wrong about the decision making rule each player used in making their choice, unless they have sufficiently restricted beliefs over decision rules. I compare the results from the present paper to those by Aumann and Brandenburger (1995) in Section 4.

The structure of the rest of the paper is as follows. In Section 2 I develop the basic theory of interactive epistemology that I use in the paper. In Section 3 I present the main results of the paper. In Section 4 I discuss the interpretation of the results together with the related literature.

# 2 Interactive Epistemology

## 2.1 Preliminaries

Let $\mathcal{M}^{\mathcal{R}}$ be a *decision structure* for a normal form two-player game $G = \langle A_1, A_2, g_1, g_2 \rangle$. Structure $\mathcal{M}^{\mathcal{R}}$ consists of $\langle R_1, R_2, T_1, T_2 \rangle$ where each element of $R_i$ is a map $r_i : T_i \to A_i$ and each element ot $T_i$ is a probability distribution over $R_{-i} \times T_{-i}$. Members of $\Omega = R_1 \times T_1 \times R_2 \times T_2$ are called *states (of the world).* The interpretation is that at state $\omega = (r_1, t_1, r_2, t_2) \in \Omega$ player 1 is of type $t_1$ and chooses action $r_1(t_1)$ with the understanding that, *had player 1 been* of type $t_1'$ (which she is not) she *would have* chosen action $r_1(t_1')$. Elements of $R_i$ are called *decision rules* for player $i$.

Notice that a decision structure encompasses the traditional model of a game used in Aumann and Brandenburger (1995) to deliver epistemic conditions for Nash equilibrium. Such model is, in essence, a structure $\mathcal{M}^{AB} = \langle A_1, A_2, T_1^{AB}, T_2^{AB} \rangle$, where each element of $T_i^{AB}$ is a probability distribution over $A_{-i} \times T_{-i}^{AB}$ and the states of the world are members of $\Omega^{AB} = A_1 \times T_1^{AB} \times A_2 \times T_2^{AB}$. I call $\mathcal{M}^{AB}$ an *AB structure* and $\Omega^{AB}$ an *AB state space.* Since a type for a player in a decision structure has beliefs over the other players' types and their decision rules, this induces beliefs over their actions as well. Both models hence specify at each state the actions chosen by each player together with their beliefs about the actions chosen by others. Therefore, the results by Aumann and Brandenburger (1995) can be formulated and proved in the present setup in exactly the same way. The difference is that, in the present setup, we can also ponder about the nature of the decision rules that led the players to undertake the actions chosen at each state.

In line with the notation in Aumann and Brandenburger (1995) call the conjecture over player 2's actions induced by player 1's type $t_1$ by $\phi_1\left(;t_1\right).$ Similarly for player 2.

## 2.2 Rationality of actions and of decision rules

Given $\varepsilon \geq 0$ *a pair* $(r_1, t_1)$ *is* $\varepsilon-$*rational* if for every $a_1 \in A_1$,

$$\sum_{a_2 \in A_2} \phi_1\left(a_2; t_1\right) g_1\left(r_1(t_1), a_2\right) + \varepsilon \geq \sum_{a_2 \in A_2} \phi_1\left(a_2; t_1\right) g_1\left(a_1, a_2\right)$$

Notice that for $\varepsilon = 0$ this is the standard concept of a players being rational at a given state of the world $(r_1, t_1, r_2, t_2)$, that is, if for each player $i$ the action chosen by player $i$ at that state, $r_i(t_i)$, maximizes the player's expected payoff, and the expectation is taken with respect to the beliefs $t_i$ the player holds at the given state.

Given $\varepsilon \geq 0$ *a decision rule* $r_1$ *is* $\varepsilon-$*rational* if for every $t_1$ in the domain of $r_1$ the pair $(r_1, t_1)$ is $\varepsilon-$rational. Heuristically, a rational decision rule for player 1 is like a reaction function for player 1. In what follows I will say that a player that uses a $\varepsilon-$rational decision rule at $\omega$ is a $\varepsilon-$*rational decision maker at* $\omega$. This terminology allows a player to be rational at $\omega$ without the player being a rational decision maker at $\omega$, a distinction that will become useful in Section 4.

### 2.2.1 Example 2

Fix $\varepsilon = 0$. Consider the following two player game $G$:

*Bob*

| | | $c$ | $d$ |
|---|---|---|---|
| *Ann* | $C$ | $3,3$ | $0,0$ |
| | $D$ | $0,0$ | $1,1$ |

and the following decision structure for $G$: Ann has only one type, $t_a$, and two decision rules, $r_a^1$ and $r_a^2$ defined on Ann's only type. Bob, in turn has two possible types, $t_b^1$ and $t_b^2$, and two possible decision rules: $r_b^1$ and $r_b^2$. The types and the decision rules are defined as follows:

$t_a$:

| | $r_b^1$ | $r_b^2$ |
|---|---|---|
| $t_b^1$ | $c, \frac{1}{3}$ | $c, \frac{1}{6}$ |
| $t_b^2$ | $d, \frac{1}{3}$ | $c, \frac{1}{6}$ |

$t_b^1$:

| | $r_a^1$ | $r_a^2$ |
|---|---|---|
| $t_a$ | $C, \frac{3}{4}$ | $D, \frac{1}{4}$ |

$t_b^2$:

| | $r_a^1$ | $r_a^2$ |
|---|---|---|
| $t_a$ | $C, \frac{1}{5}$ | $D, \frac{4}{5}$ |

.

The numbers written in the matrix for type $t_a$ of Ann depict Ann's beliefs over the product $\{t_b^1, t_b^2\} \times \{r_b^1, t_b^2\}$. The letters in the matrix depict Bob's choices for different rule-type pairs. For example, type $t_a$ of Ann puts probability $\frac{1}{6}$ to Bob being of type $t_b^2$ and employing rule $r_b^2$. Also, when Bob is of type $t_b^2$ and employs rule $r_b^2$ Bob's choice is $c$. In fact, all types of Bob choose $c$ under decision rule $r_b^2$. Notice that the beliefs over types and decision rules induce beliefs over actions as follows: $\phi_a(\cdot; t_a) = \frac{2}{3}c + \frac{1}{3}d$, $\phi_b(\cdot; t_b^1) = \frac{3}{4}C + \frac{1}{4}D$ and $\phi_b(\cdot; t_b^2) = \frac{1}{5}C + \frac{4}{5}D$.

Consider the state $(r_a^1, t_a, r_b^1, t_b^1)$. It is easy to see that both Ann and Bob are rational at that state because $r_a^1(t_a) = C$ is a best response to $\phi_a(\cdot; t_a)$

and $r_b^1(t_b^1) = c$ is a best response to $\phi_b(\cdot; t_b^1)$. Moreover, both Ann and Bob's decision rules $r_a^1$ and $r_b^1$ are rational at that state. For Ann this is immediate from the fact that $r_a^1$ is defined only over $t_a$. For Bob this follows since $r_b^1(t_b^2) = d$ is a best response to $\phi_b(\cdot; t_b^2)$.

Rule $r_b^2$, on the other hand, is not a rational decision rule for Bob, since $r_b^2(t_b^2) = c$ is not a best response to $\phi_b(\cdot; t_b^2)$. ☒

The set of decision rules $R_i$ for each player $i$ could potentially be very large, as well as the domain of the rules in $R_i$, which would follow if, for example, they were fully defined over a complete type space, namely, a type space such that for every probability distribution over the decision rules and beliefs of the opponents there was a type of the player which would hold that probability distribution as its belief.

In what follows I concentrate my attention on a subset $\widehat{R}_i$ of $R_i$. Think of $\widehat{R}_1$ as the set of decision rules that player 2 will consider possible for player 1 and viceversa. Let $R = R_1 \times R_2$ and $\widehat{R} = \widehat{R}_1 \times \widehat{R}_2$.

Given $\varepsilon \geq 0$ set $\widehat{R}$ is *consistent with* $\varepsilon-$*rationality* iff player 1 has a $\varepsilon-$rational decision rule $r_1$ in $\widehat{R}_1$ and player 2 has a $\varepsilon-$rational decision rule $r_2$ in $\widehat{R}_2$.

## 2.3 The possibility of inference

Fix $p \in (0,1]$. *Sets* $\widehat{R}_1$ *and* $\widehat{R}_2$ *permit p-inference* iff for any $r_1 \in \widehat{R}_1$, any $t_1$ in the domain of $r_1$ and any $r_2 \in \widehat{R}_2$ there is $t_2$ in the domain of $r_2$ such that $\phi_2(r_1(t_1);t_2) \geq p$. Similarly for player 2. The interpretation is that we wish to consider decision rules for each player such that, for any decision rule that player 1 could use and any type that player 1 could have, that such rule-type pair is one whose choice player 2 could anticipate, *in principle,* when player 2 reasons according to a decision rule in $\widehat{R}_2$.

Loosely, when sets $\widehat{R}_1$ and $\widehat{R}_2$ permit $p$-inference then I will say that (accurate) inference is possible in principle, which is to say that, whatever my choice, I am certain that you can at least conceive of it, and even be able to predict its occurrence, were you to have information that perhaps you do not actually have.

### 2.3.1 Example 2 (cont.)

In the example from the previous sub-section sets $\widehat{R}_a = \{r_a^1, r_a^2\}$ and $\widehat{R}_b = \{r_b^1, r_b^2\}$ permit $\frac{1}{3}-$inference: every action chosen for Ann by some rule-type pair for Ann is believed by Bob to occur with probability at least $\frac{1}{3}$ by some type from each rule in $\widehat{R}_b$. To wit: action $C$ for Ann is believed to occur with probability $\frac{3}{4}$ by type $t_b^1$ in the domain of both rules $r_b^1$ and $r_b^2$ and action $D$ for Ann is believed to occur with probability $\frac{4}{5}$ by type $t_b^2$ in the domain of both rules $r_b^1$ and $r_b^2$. Also, every action chosen for Bob by some rule-type pair for Bob is believed by Ann to occur with probability at least $\frac{1}{3}$ by some type from each rule in $\widehat{R}_a$. In this case there is only one type of Ann, and according to $t_a$ (which is clearly in the domain of both $r_a^1$ and $r_a^2$) both $c$ and $d$ occur with probability of at least $\frac{1}{3}$. ☒

That inference is possible in principle says nothing as to whether it takes place in practice. In general, it is silent as to whether accurate inference is implied at any particular state of the world. It seems, therefore, a desirable and rather innocuous assumption.[2]

A much weaker version of accurate inference goes as follows:

Fix $p \in (0,1]$. *Sets* $\widehat{R}_1$ *and* $\widehat{R}_2$ *permit weak p-inference* iff for any $r_1 \in \widehat{R}_1$, any $r_2 \in \widehat{R}_2$ there is $t_1$ in the domain of $r_1$ and $t_2$ in the domain of $r_2$ such that $\phi_2(r_1(t_1); t_2) \geq p$. Similarly for player 2. The interpretation is that for any decision rule that player 1 could use, at least some of its choices could

[2] In the example above, for instance, at state $\left(r_a^1, t_a, r_b^1, t_b^2\right)$ the choice made by Ann is $r_a^1(t_a) = C$. Bob gives such choice at that state a probability of $\frac{1}{4}$. Making the inference that Ann chooses $C$ with probability of at least $\frac{1}{3}$ is *possible* according to Bob's rule $r_b^1$, but not *necessary* at state $\left(r_a^1, t_a, r_b^1, t_b^2\right)$ : Bob would have had to be of type $t_b^1$ for this to be true, which he is not.

be anticipated, *in principle,* when player 2 reasons according to a decision rule in $\widehat{R}_2$.

### 2.3.2 Example 2, again

Sets $\widehat{R}_1$ and $\widehat{R}_2$ as defined for Example 2 permit weak $\frac{2}{3}-$inference. For Ann it is the case that each of the choices she makes under each of her decision rules is believed to occur with probability at least $\frac{2}{3}$ by a type in the domain of both $r_b^1$ and $r_b^2$ (for $r_a^1$ the type for Bob is $t_b^1$ and for $r_a^2$ the type for Bob is $t_b^2$). For Bob it is the case that one of his choices from using rule $r_b^1$ (choice c) is believed to occur with probability $\frac{2}{3}$ by Ann's type $t_a$, and identically from Bob using rule $r_b^2$ (rule $r_b^2$ always chooses $c$ and type $t_a$ of Ann believes $c$ occurs with probability $\frac{2}{3}$).

## 2.4 Belief diversity

Sets $\widehat{R}_1$ and $\widehat{R}_2$ could in principle be very small or, alternatively, contain rules defined over a very small domain. When this is the case the possibility of inference is, in fact, a kind of rational expectations assumption, a very strong assumption which I do not wish to make. For instance, if $\widehat{R}_1$ and $\widehat{R}_2$ each contained rules defined, respectively, over a single point in $T_1$ and $T_2$

then the possibility of inference would amount to both players predicting, with probability greater than $p$, the actual choices made by players using the rules in $\widehat{R}_1$ and $\widehat{R}_2$.

Because we wish to permit the possiblity of inference without imposing a rational expectations assumption we would like sets $\widehat{R}_1$ and $\widehat{R}_2$ be sufficiently diverse, both in terms of the richness of rules in them as well as in terms of the richness of the domain of the rules in them. Assumptions $CS$ and $T$ below guarantee such diversity.

Let $\gamma_{12} : A_1 \to A_2$ be a function that 'translates' actions from the action space of player 1 to the action space of player 2. Let $\pi_{21} : T_2 \to T_1$ be a function that 'translates' types from the type space of player 1 to the type space of player 2.

Rule $r_2$ for player 2 is a *mechanical brother of rule* $r_1$ iff there are translations $\gamma_{12}$ and $\pi_{21}$ such that

$$r_2 (t_2) = \gamma_{12} (r_1 (t_1)) \text{ for } t_1 = \pi_{21} \ (t_2) .$$

Similarly for player 1.

*Sets* $\widehat{R}_1$ *and* $\widehat{R}_2$ *satisfy CS* (for caution and symmetry) iff every mechanical brother $r_2$ of $r_1 \in \widehat{R}_1$ is in $\widehat{R}_2$; similarly for player 1.

Condition $CS$ is inspired on a condition from the literature on learning in repeated games developed by Nachbar (2005), which in turn is a variation on what Nachbar (1997) called *neutrality.* It reflects two more primitive criteria: caution and symmetry. I explain them in turn using Nachbar's (2005) terminology and wording.

By caution I mean that if a decision rule $r_2$ is in $\widehat{R}_2$ then all simple variants of $r_2$ are also in $\widehat{R}_2$. Caution is motivated by the idea that a prudent player would not want to rule out as impossible simple variants of the decision rules to which the player assigns positive probability. By symmetry I mean that for each decision rule $r_2$ in $\widehat{R}_2$ there is some decision rule $r_1$ that is a simple variant of $r_2$. Symmetry is motivated by the idea that, aside from differences in the composition and cardinality of the action sets, the sets $\widehat{R}_1$ and $\widehat{R}_2$ are structurally similar, that is, they contain decision rules of a similar "mechanical" quality.

### 2.4.1 Example 2, one more time

The sets $\widehat{R}_1$ and $\widehat{R}_2$ defined in Example 2 do not satisfy condition $CS$. To see this consider the decision rule $r_a^1$ and the functions $\gamma_{12}$ and $\pi_{12}$ defined by $\gamma_{12}(C) = \gamma_{12}(D) = d$ and $\pi_{12}(t_a) = t_b^1$. The mechanical brother of $r_a^1$ under $\gamma_{12}$ and $\pi_{12}$ would be a decision rule $r_b^*$ that would choose action $d$ when player 2 is of type $t_b^1$. Notice that such rule is not in $\widehat{R}_2$.

### 2.4.2 Example 3

Here is a decision structure for the game used in Example 2 that satisfies condition $CS$: $\langle (r_a^1, r_a^2, r_a^3, r_a^4), (r_b^1, r_b^2, r_b^3, r_b^4), (t_a^1, t_a^2), (t_b^1, t_b^2) \rangle$, where the decision rules for Ann and Bob are defined as follows:

| | $r_a^1$ | $r_a^2$ | $r_a^3$ | $r_a^4$ |
|---|---|---|---|---|
| $t_a^1$ | $c$ | $c$ | $d$ | $d$ |
| $t_a^2$ | $d$ | $c$ | $c$ | $d$ |

| | $r_b^1$ | $r_b^2$ | $r_b^3$ | $r_b^4$ |
|---|---|---|---|---|
| $t_b^1$ | $C$ | $C$ | $D$ | $D$ |
| $t_b^2$ | $D$ | $C$ | $C$ | $D$ |

Verifying that condition $CS$ is satisfied is a straightforward yet tedious exercise. I ommit the details here.

As seen, condition $CS$ ensures diversity of the set of rules contained in $\widehat{R}$. Condition $T$ ensures that rules in $\widehat{R}$ are defined over a sufficiently rich set of types.

*Sets* $\widehat{R}_1$ *and* $\widehat{R}_2$ *satisfy* $T$ iff for every $r_1$ in $\widehat{R}_1$ and given a relation $\mathcal{P} \subseteq T_1 \times \widehat{R}_2 \times T_2$ there is a bijection $h : T_1 \to \widehat{R}_2 \times T_2$ such that $(t_1, h_1(t_1), h_2(t_1)) \in$

$\mathcal{P}$ for every $t_1$ in the domain of $r_1$. Similarly for player 2. The interpretation is that, for each rule $r_1$ in $R_1$ for player 1 we wish to have enough types and rules for player 2 such that there is always a type $t_2$ for player 2 that could uniquely identify a rule $r_2$ and a type $t_1$ in the domain of $r_1$ that satisfies $\mathcal{P}$.

In what follows I refer to $CS$ and $T$ together as condition $CST$.

# 3 The impossible trinity

An action $a_1^*$ is weakly dominant iff, for any $a_2 \in A_2$,

$$g_1\left(a_1^*, a_2\right) \geq \max_{a_2 \in A_2} g_1\left(a_1, a_2\right).$$

The definition for player 2 is similar.

Game $G$ satisfies No Weak Dominance (NWD) iff neither player has a weakly dominant action.

The main result of the paper is an *impossible trinty.*

**Theorem 1** *Suppose that NWD holds in game $G$. Then there is $\overline{\varepsilon} > 0$ and $\overline{p} \in (0, 1]$ such that, for any $\varepsilon \in [0, \overline{\varepsilon})$ and any $p \in (\overline{p}, 1]$ and any $\widehat{R} \subset R$, if $\widehat{R}$ satisfies $CST$ and permits weak $p$-inference then $\widehat{R}$ is inconsistent with $\varepsilon-$rationality.*

The interpretation is that, for a large set of games, if beliefs allow the possibility of inference and have supports that are sufficiently diverse then they are inconsistent with the opponent employing a rational decision rule. Then, to the extent that players employ a rational decision rule at a particular state of the world, then the players beliefs *over decision rules* will necessarily be mis-specified at that state (even though they may not be missspecified *about the actions chosen at that state*).

**Remark 2** *The result breaks down for games with a weakly dominant action for each player because* $CS$ *guarantees that* $\widehat{R}$ *would always be consistent with rationality, regardless of what we assume about beliefs, as* $\widehat{R}$ *would contain the decision rule for each player that always chooses the weakly dominant action for that player.*

The proof of Theorem 1 relies on the notion of an evil twin, which I define below.

## 3.1 Evil twins

Fix $p$ and $\varepsilon$. *Rule* $r_2 \in \widehat{R}_2$ *is an* $(\varepsilon, p)-$*evil twin of* $r_1 \in \widehat{R}_1$ iff whenever $\phi_1 (r_2 (t_2) ; t_1) \geq p$ for some $t_1$ in the domain of $r_1$ and some $t_2$ in the domain of $r_2$ then $(r_1, t_1)$ is not $\varepsilon-$rational. A similar definition holds for player 2.

Fix $p$ and $\varepsilon$. $\widehat{R}$ has the $(\varepsilon, p)$ −*evil twin* property iff for any decision rule $r_1 \in \widehat{R}_1$ there is a decision rule $r_2 \in \widehat{R}_2$ that is an $(\varepsilon, p)$ −evil twin of $r_1$.

#### 3.1.1 Example 2, one more time

It is not hard to see that, in Example 2, $r_a^2$ is an $\left(0, \frac{4}{5}\right)$ −evil twin of $r_b^2$, for there is only one type for Bob in the domain of $r_b^2$ (type $t_b^2$) that predicts, with probability $\frac{4}{5}$, the actual choice made by Ann when of type $t_a$ and she uses rule $r_a^2$ (her choice is action $D$). Bob's choice, in turn, when he is of such type $t_b^2$ and uses rule $r_b^2$ is $c$, which is a poor choice for Bob when Ann chooses D and Bob can predict this. ☒

Combining these definitions yields the following observation.

**Theorem 3** *For any $p$ and $\varepsilon$, if $\widehat{R}$ permits weak $p-$inference and has the $(\varepsilon, p)$ −evil twin property then $\widehat{R}$ does not contain any $\varepsilon-$rational decision rule for either player.*

**Proof.** Pick any arbitrary $r_1 \in \widehat{R}_1$. I will show it is not an $\varepsilon-$rational decision rule for player 1. By the $(\varepsilon, p)$ −evil twin property there is $r_2 \in \widehat{R}_2$ such that whenever $\phi_1 (r_2 (t_2) ; t_1) \geq p$ for some $t_1$ in the domain of $r_1$ and some $t_2$ in the domain of $r_2$ then $(r_1, t_1)$ is not $\varepsilon-$rational. Since $\widehat{R}$ permits weak $p-$inference for such $r_2$ then there is $t_1$ in the domain of $r_1$ and $t_2$ in the domain of $r_2$ such that $\phi_1 (r_2 (t_2) ; t_1) \geq p$. ■

The key technical fact underlying Theorem 1 is that the evil twin property follows from $CST$.

**Theorem 4** *Suppose that no player has a weakly dominant action in game $G$. Then there is $\overline{\varepsilon} > 0$ and $\overline{p} \in (0,1]$ such that for any $\varepsilon \in [0, \overline{\varepsilon})$ and any $p \in (\overline{p}, 1]$ and any $\widehat{R}$, if $\widehat{R}$ satisfies $CST$ then $\widehat{R}$ has the $(\varepsilon, p)-$evil twin property.*

Notice that the proof of Theorem 1 follows immediately from Theorems 3 and 4.

## 3.2 The proof ot Theorem 4

I begin by showing how to construct evil twins. As in Nachbar (1997, 2005) define the function $\widetilde{a}_2 : A_1 \rightarrow A_2$ for any action $a_1 \in A_1$ by

$$\widetilde{a}_2 (a_1) \in \arg\max_{a_2 \in A_2} \left[ \max_{a_1' \in A_1} g_1 (a_1', a_2) - g_1 (a_1, a_2) \right].$$

Loosely, when player 2 chooses $\widetilde{a}_2 (a_1)$ player 1 has maximal incentives not to choose $a_1$.

Pick a decision rule $r_1 \in \widehat{R}_1$. Fix $t_1$. Find all $r_2$, and $t_2$ in the domain of $r_2$, such that $\phi_1(r_2(t_2); t_1) \geq p$. Pick one such pair $(r_2, t_2)$. If the condition cannot be met, pick the pair $(r_2, t_2)$ arbitrarily. For each $t_1$ pick the pairs $(r_2, t_2)$ such that they are distinct, component by component, for each $t_1$ in the domain of $r_1$. By condition $T$, this can be done. This process defines a map $\pi_{21} : T_2 \to T_1$ such that each point $(t_2, \pi_{21}(t_2))$ is associated with a different decision rule $r_2$ in $\widehat{R}_2$. Define $\widetilde{r}_2$, for each $t_2$, by

$$\widetilde{r}_2(t_2) = \widetilde{a}_2(r_1(t_1))$$

for $t_1 = \pi_{21}(t_2)$. Letting $\gamma_{12} = \widetilde{a}_2$ and with the $\pi_{21}$ defined above condition $CS$ implies that $\widetilde{r}_2 \in \widehat{R}_2$.

The rule $\widetilde{r}_2$ defined above is, in fact, an $(\varepsilon, p)$ −evil twin of $r_1$, for appropriately defined $\varepsilon$ and $p$ which means that $\widehat{R}$ has the $(\varepsilon, p)$ −evil twin property. The rest of the proof is in the Appendix.

# 4 Discussion

*a. Interpretation of the results*

Let $p$ and $\varepsilon$ be as in Theorem 1. Consider a state $\omega$ of $\Omega$ such that the types $t_1$ and $t_2$ held by the players at that state have beliefs over decision rules with supports that satisfy weak $p-$inference and $CST$. Then, by Theorem 1, those supports must not contain an $\varepsilon-$rational decision rule for either player. The implication is that if both players employ rational decision rules at $\omega$ then beliefs over decision rules at $\omega$ will be misspecified as each player would be using a rule that necessarily would be outside of the support of the opponent's belief. Therefore, under the hypotheses of Theorem 1 the players can use rational decision rules or have well-specified beliefs over decision rules at a given state of the world, but not both. *Under belief diversity, if players use rational decision rules, then the possibility of accurate inference fails.*

Moreover, *if players use rational decision rules, correctly know each other's payoffs,*[3] *and have mutual certainty of belief diversity and of the possibility of accurate inference, then they are each certain and wrong that the other is not using a rational decision rule.*

To see this formally say that player 1 is certain of $E \subseteq R_2 \times T_2$ at $\omega = (r_1, t_1, r_2, t_2)$ iff $p_1(E; t_1) = 1$. Similarly for player 2. Let $\widehat{R}_2(t_1)$ be the support of 1's beliefs over 2's decision rules at $\omega = (r_1, t_1, r_2, t_2)$ and, given a

[3] *Notice that "knowledge of each other's payoffs" is automatically satisfied in the present setup as payoffs are constant throughout the entire state space. I state it here for completeness, as it would be necessary in a model that explicitly dealt with uncertainty over payoffs.*

subset $\widehat{R}_1$ of $R_1$ define $\left[\widehat{R}_1\right]$ by

$$\left[\widehat{R}_1\right] := \left\{(r_1, t_1) \in R_1 \times T_1 : \widehat{R}_1 \times \widehat{R}_2(t_1) \text{ satisfy weak } p - \text{inference and } CST\right\}.$$

Define player 2 to be certain that weak $p-$inference and $CST$ is satisfied at $\omega = (r_1, t_1, r_2, t_2)$ iff

$$p_2\left(\left[\widehat{R}_1\right]; t_2\right) = 1,$$

where $\widehat{R}_1$ is the support of 2's belief over 1's decision rules at $\omega$. Similarly for player 1.

I can state the result formally now as follows:

**Theorem 5** *Suppose NWD holds in game G. Let $p$ and $\varepsilon$ be as in Theorem 1. Assume players are certain that weak $p-$inference and CST hold at $\omega$. Then player 1 is certain that 2 is not using a $\varepsilon-$rational decision rule at $\omega$. Similarly for player 2.*

The proof is in the Appendix.

One is then left with a formalism populated by rational decision makers that are certain that their opponents aren't rational decision makers themselves! Not that one would want this, but this is precisely the point: it is what would follow from the seemingly innocuous assumptions of rational decision, mutual knowledge of payoffs, and mutual certainty of belief diversity and the possibility of accurate inference.

*b. Relation to Aumann and Brandenburger (1995)*

In a seminal paper Aumann and Brandenburger (1995) developed epistemic conditions that lead to Nash equilibria. A natural question to ask is what relationship, if any, the results presented in this paper have with the Aumann and Brandenburger (1995) results.

The results presented here do not contradict in any way those of Aumann and Brandenburger (1995), but they do show that there is a sense in which they justify Nash equilibrium "by accident," namely, as the players themseves are wrong about the decision making rule each player used in making their choice, unless they have sufficiently restricted beliefs over decision rules.

Consider being at a state of the world $\omega$ from $\Omega$ where $p$ and $\varepsilon$ are as in Theorem 1 and the types $t_1$ and $t_2$ held by the players at that state have beliefs over decision rules with supports that satisfy weak $p-$inference and $CST$. Assume, further, that the decision rules used at $\omega$ by every player are $\varepsilon-$rational and that there is mutual certainty of action profiles at that state. Then, by the *preliminary observation* in Aumann and Brandenburger (1995), the actions chosen at that state are a Nash equilibrium although the players beliefs over decision rules are mis-specified: the rules used by the players at $\omega$ by Theorem 1 cannot be in the support of each other's beliefs over decision rules.

Consider now being at a state of the world $\omega$ from $\Omega$ where $p$ and $\varepsilon$ are as in Theorem 1. Assume that players are certain that weak $p-$inference and $CST$ hold at $\omega$. Assume, further, that the players correctly know about each other's payoffs, conjectures and rationality at $\omega$. Then, by Theorem A in Aumann and Brandenburger (1995), the conjectures are a Nash equilibrium, even though, by Theorem 4 in the present paper, the players are certain that each other is not using a rational decision rule at $\omega$.

Again, the results do not contradict the seminal work by Aumann and Brandenburger (1995) in any way. The results, however, reveal a difficulty in interpreting Nash equilibrium as the result of a deliberative process through which players understand how each other makes the choices they make when players are modelled as $\varepsilon-$*rational decision makers* whose beliefs over decision rules satisfy a support diversity condition, allow the possibility of inference, and who know about all this about each other. This is so because even if their actions or their beliefs were in equilibrium, the players would be by necessity wrong about the decision making rule each player used in making their choice and even about them being rational decision makers to begin with.

To sum it up: Players in the present setup may understand what they do or believe (and what they do or believe may be a Nash equilibrum) yet they cannot understand how each other reasons their way to the choices they make. It is in this sense, as in the learning in games literature, that reasoning about equilibrium in an epistemic model is a more stringent requirement than being in equilibrium.

*c. Similarities and differences with Nachbar (1997, 2005)*

For the purpose of this discussion I will focus the comparison to Nachbar (2005) and keep the discussion at an informal level.

(i). *The object of study.* Nachbar studies sets $\widehat{\Sigma}_i$ of *behavior strategies* for player $i$, that is, maps from the set of finite histories to probability distributions over the action space. I study sets $\widehat{R}_i$ of *decision rules at a state of the world*, that is, maps from a set of (epistemic) types to the action space. The interpretation is that at a given state of the world the rule together with a type specify the action chosen by a player, but the rule also specifies what would have the player chosen according to that rule if the player had an epistemic type different from the one the player actually has.

(ii) *Learnability and $p-$inference.* Nachbar takes the belief $\beta_i$ of the player $i$ over the repeated game strategy of the opponent as given and assumes that $\widehat{\Sigma}_i$ satisfies learnability iff for every pair of strategies from $\widehat{\Sigma} := \widehat{\Sigma}_1 \times \widehat{\Sigma}_2$ player $i$ (weakly) learns to predict the path of play. I say that $\widehat{R}_i$ satisfies (weak) $p-$inference, for a given probability level $p$, iff for every pair of rules $r_1$ and $r_2$ from $\widehat{R}$ there is an epistemic type of player $i$ in the domain of $r_i$ that can predict with probability of at least $p$ the choice made by at least one of the types in the domain of $r_{-i}$. Notice that Nachbar's learnability assumption requires learning to take place *necessarily* for every pair of strategies in $\widehat{\Sigma}$, whereas $p-$inference requires inference to be *a possibility* for every pair of decision rules in $\widehat{R}$. Notice also that learnability (and rationality in what follows) is determined with respect to a fixed belief $\beta_i$ for all strategy pairs in $\widehat{\Sigma}$ whereas $p-$inference is determined with respect to epistemic types (hence beliefs) that vary depending on which pair of rules in $\widehat{R}$ is being evaluated.

(iii) *Consistency and $\varepsilon-$rationality.* Nachbar assumes $\widehat{\Sigma}$ satisfies $\varepsilon-$consistency for every player $i$ given beliefs $\beta_i$ iff each player $i$ has a uniform $\varepsilon$ best response in $\widehat{\Sigma}_i$. I assume $\widehat{R}$ is consistent with $\varepsilon-$rationality iff for each player i there is a rule $r_i$ in $\widehat{R}_i$ such that, for every type $t_i$ in the domain of $r_i$, choice

$r_i(t_i)$ maximizes expected utility for player $i$, where the expectation is taken with respect to the beliefs induced by $t_i$. Theorem 1 in Nachbar (2005) finds conditions that lead $\widehat{\Sigma}$ not to be $\varepsilon-$consistent. Theorem 1 in the present paper finds conditions that lead $\widehat{R}_i$ not to be consistent with $\varepsilon-$rationality. Hence, the structure of both theorems regarding rationality of strategies and rules is similar. What varies is the meaning of rationality in either case.

(iv) *Belief diversity: CSP and CST.* $CSP$ in Nachbar (2005) is composed of two conditions, $CS$ (caution and symmetry) and $P$ (pure strategies), both related to the richness of $\widehat{\Sigma}$. $CS$ says that $\widehat{S}_i$ (the set of "pure" behavior strategies in $\widehat{\Sigma}_i$) contains all pure strategies that are simple 'translations' of the pure strategies in $\widehat{S}_{-i}$ and viceversa. Condition $P$ says that if a behavior strategy $\sigma_i$ is in $\widehat{\Sigma}_i$ then at least one pure stragegy that coarsely approximates $\sigma_i$ is contained in $\widehat{\Sigma}_i$ as well.

$CST$ in the present paper is composed of two conditions, $CS$ (caution and symmetry) and $T$ (type diversity), both related to the richness of $\widehat{R}$ and of the domain of the rules in $\widehat{R}$. Condition $CS$ says that, given some 'translation' map $\gamma$ between $A_i$ and $A_{-i}$, and $\pi$ between $T_i$ and $T_{-i}$, if $\widehat{R}_i$ contains $r_i$, then $\widehat{R}_{-i}$ contains $r_{-i}$ such that $r_{-i}$ chooses $\gamma\left(a_i\right)$ at $\pi\left(t_i\right)$ whenever $r_i$ chooses $a_i$ at $t_i$. Notice that I attempt here a "double translation" of actions and types that is not performed by Nachbar in the repeated games setup. This is so because in the repeated game setup it is only necessary to translate the actions chosen given each finite history as behavior strategies for either player are defined over the same set (the set of finite histories).

A condition like Condition $P$ in Nachbar (2005) is not necessary in the present paper as I only deal with "pure" decision rules here. Condition $T$ in the present paper is a type diversity condition (although it is important to notice that $CS$ also requires a certain amount of type diversity). Loosely, $T$, will allow me to do the following: if I have a family $\mathcal{F}$ of rules $r_i$ from $\widehat{R}_i$ that satisfy a certain property $\mathcal{P}$, I can find a rule $r_i'$ in $R_i$ that satisfies property $\mathcal{P}$ and that is defined over only one of the types in the domain of each rule in $\mathcal{F}$. One needs sufficiently many types $t_i$ of player $i$ in the domain of the rules in $\widehat{R}_i$ to do this.

Condition $T$ has no analogues in Nachbar (2005).

Because of the nature of $CS$, it is not surprising that $\widehat{R}_i$ contains irrational decision rules. What is surprising is that $CS$, when combined with $T$ and *weak p−inference*, makes $\widehat{R}_i$ contain *only* irrational decision rules.

(v). *Results.* Nachbar's main theorem is Theorem 1, a theorem that applies to impatient players in games that satisfy a No Weak Dominance condition and to players of any patience in games such that the pure action maxmin payoff is strictly less than the minmax payoff. Because I only deal with one shot games, in this paper discounting is not an issue and therefore I only need to assume that games satisfy the No Weak Dominance condition. It is a very weak condition.

*d. Other related literature*

Binmore (1987) and Canning (1992) explored the consequences of modelling rational agents by means of assuming that players use computable decision procedures in making their choices (procedures that take the decision procedure of the opponent as given) and use *the unsolvability of the halting problem* to show that there is no decision procedure that leads to rational choices against all possible decision procedures of the opponent for all games. The results presented in this paper have the same flavor, except that no computability assumptions are made here. Another difference is that, contrary to Binmore (1987) and Canning (1992), I develop the results in the language of interactive epistemology, that is, in the context of state space models where each state contains information about payoff relevant variables, as well as about the knowledge and belief by each player about each other's knowledge and beliefs.

A typical reaction to results such as those by Binmore (1987) and Canning (1992) is that they are of limited interest due to the computability assumptions made in those papers. It is argued that real decision makers would be more sophisticated that the algorithms considered by Binmore and Canning allow and that their negative results on the impossibility of players rationally understanding each other in equilibrium were not very relevant. The research presented in the present paper shows that the problems they identified are robust to eliminating the computability assumptions and can be readily compared with what is known about the epistemic conditions that lead to equilibrium in state space models.

*e. Extensions*

I believe that results like the ones presented in the present paper may have eluded discovery because they are "projected away" in in the conventional framework, where the primitive is an action as opposed to a decision rule at each state of the world. One may, however, wonder if there is a version of the results that can be expressed in a more conventional framework, such as the framework developed by Battigalli and Siniscalchi (2002), suitably adapted to the study of normal form games. This research is left for future work.

# References


[1] Battigalli, P. and Marciano Siniscalchi, "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory 106*, 356-391, 2002.

[2] Binmore, K., "Modelling Rational Players, Part I," *Economics and Philosophy 3*, 179-214, 1987.

[3] Canning, D., "Rationality, Computability, and Nash Equilibrium," *Econometrica 60*, 877-888, 1992.

[4] Foster, D, and Peyton Young, "On the Impossibility of Predicting the Behavior of Rational Agents, *Proceedings of the National Academy of Sciences 98*, 12848-12853, 2001.

[5] Jordan, J., "Three problems in Learning Mixed-Strategy Nash Equilibria," *Games and Economic Behavior 5*, 368-386.

[6] Hart, O., and Andreu Mas-Colell, "Uncoupled Dynamics Do Not Lead to Nash Equilibrium," *American Economic Review 93*, 1830-1836, 2003.

[7] Nachbar, J., "Prediction, Optimization, and Learning in Repeated Games," *Econometrica 65*, 275-309, 1997.

[8] Nachbar, J.,"Beliefs in Repeated Games, *Econometrica 73*, 459-480, 2005.

[9] Weinstein, J., and Muhamet Yildiz, "Impact of Higher-Order Uncertainty," *mimeo*, 2003.

# 5 Appendix

## 5.1 Decision rule $\widetilde{r}_2$ : an evil twin

Define $w_i(a_i) := \max_{a_i' \in A_i} g_i(a_i', \tilde{a}_{-i}(a_i)) - g_i(a_i, \tilde{a}_{-i}(a_i))$, $\underline{w}_i := \min_{a_i \in A_i} w_i(a_i)$, and $\overline{\varepsilon} := \min\{\underline{w}_1, \underline{w}_2\}$. Note that $w_i(a_i) \geq 0$ with strict inequality when $a_i$ is not weakly dominant. Let $\underline{w}_i := \min_{a_i \in A_i} w_i(a_i) > 0$, $\overline{g}_i := \max_{a \in A} g_i(a)$ and $\underline{g}_i := \min_{a \in A} g_i(a)$. No strategy is weakly dominant, so $\left[\overline{g}_i - \underline{g}_i\right] > 0$.

For the moment fix $p \in (0,1)$ and $\varepsilon > 0$. Pick an arbitrary $\overline{t_1}$ in the domain of $r_1$ and let $a_1^* = r_1(\overline{t_1})$. If there is no $t_2$ in the domain of $\widetilde{r}_2$ such that $\phi_1\left(\widetilde{r}_2(t_2); \overline{t_1}\right) \geq p$ there is nothing to prove. If $\phi_1\left(\widetilde{r}_2\left(\overline{t_2}\right); \overline{t_1}\right) \geq p$ for some $\overline{t_2}$ in the domain of $\widetilde{r}_2$ it is the case that $a_2^* := \widetilde{r}_2\left(\overline{t_2}\right) = \widetilde{a}_2\left(r_1\left(\overline{t_1}\right)\right)$. To see this notice that $\widetilde{r}_2\left(\overline{t_2}\right) = \widetilde{a}_2(r_1(t_1))$ for $t_1 = \pi_{21}\left(\overline{t_2}\right)$. But by condition $T$ there is only one type $t_1$ such that $t_1 = \pi_{21}\left(\overline{t_2}\right)$, and it is precisely $\overline{t_1}$. Hence, $a_2^* := \widetilde{r}_2\left(\overline{t_2}\right) = \widetilde{a}_2\left(r_1\left(\overline{t_1}\right)\right)$.

Given such beliefs, player 1's expected payoff from choosing $a_1^*$ is at most

$$p \cdot g_1(a_1^*, a_2^*) + (1-p)\,\overline{g}_1.$$

If player 1's choice $a_1^*$ is a best response to the player's beliefs her expected payoff from choosing $a_1^*$ would be at least

$$p \cdot \max_{a_1 \in A_1} g_1(a_1, a_2^*) + (1-p)\,\underline{g}_1.$$

Therefore, $\varepsilon-$rationality of the pair $(r_1, t_1)$ requires that

$$\varepsilon + p \cdot g_1(a_1^*, a_2^*) + (1-p)\,\overline{g}_1 \geq p \cdot \max_{a_1 \in A_1} g_1(a_1, a_2^*) + (1-p)\,\underline{g}_1,$$

or

$$\varepsilon + (1-p)\left[\overline{g}_1 - \underline{g}_1\right] \geq p\left[\max_{a_1 \in A_1} g_1(a_1, a_2^*) - g_1(a_1^*, a_2^*)\right]$$

but $a_2^* = \widetilde{a}_2\left(a_1^*\right)$, so $\max_{a_1 \in A_1} g_1(a_1, a_2^*) - g_1(a_1^*, a_2^*) = w_1\left(a_1^*\right) \geq \underline{w}_1$. Therefore,

$$\varepsilon + (1-p)\left[\overline{g}_1 - \underline{g}_1\right] \geq p\underline{w}_1,$$

which cannot hold, for example, for $\varepsilon < \underline{w}_1$ and $p > \frac{\epsilon + \left[\overline{g}_i - \underline{g}_i\right]}{\underline{w}_i + \left[\overline{g}_i - \underline{g}_i\right]}$. This guarantess the existence of the desired $\overline{\varepsilon}$ and $\overline{p}$ which means, in the end, that for sufficiently low $\varepsilon$ and sufficiently high $p$ the rule $\widetilde{r}_2$ is an $(\varepsilon, p)-$evil twin of $r_1$. $\boxtimes$

# 6 The proof of Theorem 4

Let $\widehat{R}_2$ be the support of 1's belief over 2's decision rules at $\omega$. By definition, player 1 is certain that player 2 is using a decision rule in $\widehat{R}_2$. By assumption, player 1 is also certain that $\widehat{R}_2$ and player 2's supports of beliefs over player 1's decision rules satisfy weak $p-$inference and $CST$. Then, by Theorem 1, $\widehat{R}_2$ does not contain an $\varepsilon-$rational decision rule for player 2 and hence player 1 is certain that player 2 does not use a $\varepsilon-$rational decision rule at $\omega$.