

Centre interuniversitaire de recherche  
en économie quantitative

CIREQ

**Cahier 01-2013**

*Every Choice Function is  
Backwards-Induction Rationalizable*

Walter BOSSERT and  
Yves SPRUMONT



Le **Centre interuniversitaire de recherche en économie quantitative (CIREQ)** regroupe des chercheurs dans les domaines de l'économétrie, la théorie de la décision, la macroéconomie et les marchés financiers, la microéconomie appliquée et l'économie expérimentale ainsi que l'économie de l'environnement et des ressources naturelles. Ils proviennent principalement des universités de Montréal, McGill et Concordia. Le CIREQ offre un milieu dynamique de recherche en économie quantitative grâce au grand nombre d'activités qu'il organise (séminaires, ateliers, colloques) et de collaborateurs qu'il reçoit chaque année.

*The **Center for Interuniversity Research in Quantitative Economics (CIREQ)** regroups researchers in the fields of econometrics, decision theory, macroeconomics and financial markets, applied microeconomics and experimental economics, and environmental and natural resources economics. They come mainly from the Université de Montréal, McGill University and Concordia University. CIREQ offers a dynamic environment of research in quantitative economics thanks to the large number of activities that it organizes (seminars, workshops, conferences) and to the visitors it receives every year.*

## **Cahier 01-2013**

### *Every Choice Function is Backwards-Induction Rationalizable*

Walter BOSSERT and Yves SPRUMONT

Ce cahier a également été publié par le Département de sciences économiques de l'Université de Montréal sous le numéro (2013-01).

*This working paper was also published by the Department of Economics of the University of Montreal under number (2013-01).*

Dépôt légal - Bibliothèque nationale du Canada, 2012, ISSN 0821-4441

Dépôt légal - Bibliothèque et Archives nationales du Québec, 2013

ISBN-13 : 978-2-89382-642-4

# Every Choice Function is Backwards-Induction Rationalizable\*

WALTER BOSSERT and YVES SPRUMONT

Department of Economics and CIREQ  
University of Montreal  
walter.bossert@umontreal.ca  
yves.sprumont@umontreal.ca

This version: January 31, 2013

**Abstract.** A choice function is backwards-induction rationalizable if there exists a finite perfect-information extensive-form game such that, for each subset of alternatives, the backwards-induction outcome of the restriction of the game to that subset of alternatives coincides with the choice from that subset. We prove that every choice function is backwards-induction rationalizable. *Journal of Economic Literature* Classification Numbers: C72; D70.

\* We thank Sean Horan for useful discussions. Financial support from the Fonds de Recherche sur la Société et la Culture of Québec and the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

# 1 Introduction

This paper studies how the equilibrium outcome of a game varies with the set of feasible alternatives. We focus on finite perfect-information extensive-form games where all players have strict preferences over the alternatives attached to the terminal nodes. In any such game, Kuhn’s (1953) backwards-induction algorithm yields a unique outcome. If only a subset of alternatives remains feasible, pruning the original tree of all branches leading to infeasible alternatives defines a new game. Applying Kuhn’s algorithm to this restricted game again yields a unique outcome. Applying the algorithm to every restricted game thus generates a choice function, that is, a mapping selecting a single alternative from each feasible subset. We are interested in the general properties of this choice function, namely, those that do not depend on the number of players or the particular structure of the game tree.

It should be clear that such a choice function need not be well-behaved. Consider for instance the game  $G^4$  in Figure 1 where we indicate the relevant parts of the players’ preferences next to each decision node and the corresponding outcome next to each terminal node. The backwards-induction outcome of that game is alternative 3. The outcome of the game restricted to the feasible set  $\{1, 2\}$  is 1, the outcome restricted to  $\{1, 3\}$  is 1, and the outcome restricted to  $\{2, 3\}$  is 2. This is the choice function  $f^4$  in Table 1, omitting the trivial choices from singletons. The collective choice behavior described by this choice function is highly irrational since the alternative chosen from the universal set is the only alternative that is never chosen from any pair—the Condorcet loser of the base relation associated with  $f^4$ .

Up to a relabeling of the alternatives, Table 1 depicts all the possible three-alternative choice functions. The reader will easily check that each of the choice functions  $f^1, f^2, f^3, f^4$  is generated by the game bearing the corresponding label in Figure 1. This shows that when there are only three conceivable alternatives, the backwards-induction outcome of a game may change in a totally arbitrary fashion with the set of feasible alternatives.

$B$	$f^1(B)$	$f^2(B)$	$f^3(B)$	$f^4(B)$
$\{1, 2\}$	1	1	1	1
$\{1, 3\}$	1	3	1	1
$\{2, 3\}$	2	2	2	2
$\{1, 2, 3\}$	1	1	2	3

Table 1: The three-alternative case.

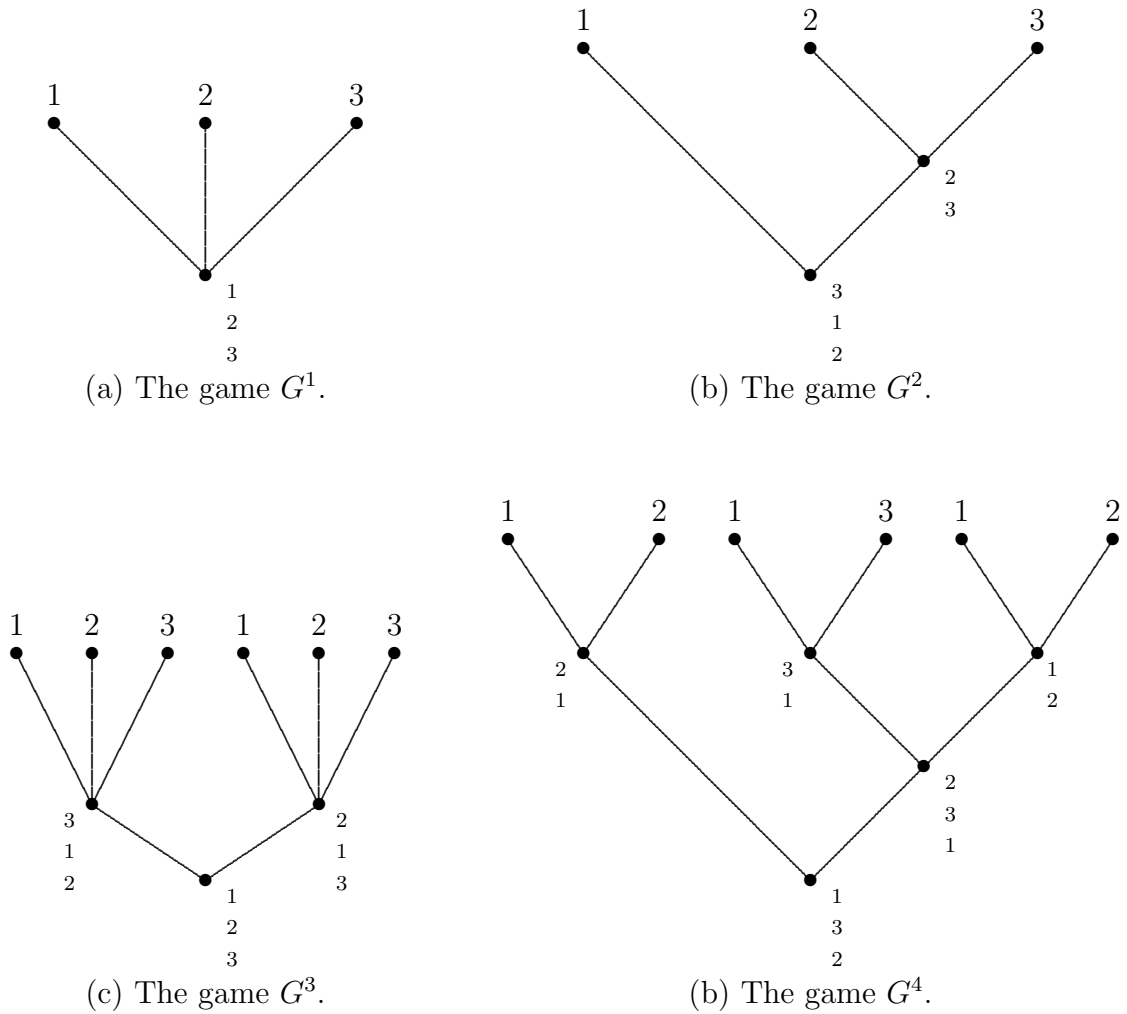


Figure 1: The three-alternative case.

This result generalizes. Say that a choice function is *backwards-induction rationalizable* if there exists a finite perfect-information extensive-form game such that, for each subset of alternatives, the backwards-induction outcome of the restriction of the game to that subset of alternatives coincides with the choice from that subset. We prove that every choice function is backwards-induction rationalizable.

Our paper belongs to an emerging literature applying the revealed preference approach to the study of collective decisions. The general goal is to identify the testable restrictions of the main theories of collective decision-making when individual preferences are not observable. We briefly review here the part of that literature which deals with noncooperative game-theoretic solution concepts.

Yanovskaya (1980), Sprumont (2000) and Galambos (2005) consider choice correspon-

dences defined over Cartesian subproducts of a given  $n$ -player normal game form. They identify necessary and sufficient conditions under which there exist  $n$  preferences over the conceivable joint actions such that the joint actions selected from each subgame form coincide with the Nash equilibria of the corresponding subgame. Lee (2012) characterizes choice behavior that is rationalizable via Nash equilibria of zero-sum games.

Ray and Zhou (2001) fix a finite-length game tree and an assignment of the decision nodes to a given set of  $n$  agents. They study choice functions defined over the restrictions of this game form. They give conditions under which there exist  $n$  (strict) preferences over the terminal nodes such that the alternative chosen in each restricted game form coincides with the backwards-induction outcome of the game generated by these preferences and the restricted game form. See also Ray and Snyder (2003).

In contrast with Ray and Zhou (2001) we do not fix the number of players or the grand game form. The choice functions we consider are therefore defined on subsets of alternatives—as in Arrow’s (1959) classical formulation—rather than on restricted game forms. We do this in order to identify the behavioral implications of backwards induction which do not depend on the particular characteristics of the environment. It turns out that there is none. Xu and Zhou (2007) perform the same exercise as ours under the rather particular restriction that each alternative can be attached to a terminal node of the grand game form once and only once. This restriction does constrain the choice functions that can be rationalized. While Xu and Zhou (2007) note (without a proof) that all three-alternative choice functions are backwards-induction rationalizable in the unrestricted-population case without their additional condition, they explicitly state that the general case represents an open question.

The unrestricted-population approach we follow in this paper is well established in economic theory. Our “anything-goes theorem” is in the spirit of results obtained in completely different contexts such as the Debreu (1974) – Mantel (1974) – Sonnenschein (1973) theorem on the aggregate excess demand of an exchange economy or McGarvey’s (1953) theorem on majority tournaments.

## 2 Definitions

Let  $A = \{1, \dots, m\}$  be a finite universal set of alternatives. For any non-empty subset  $B$  of  $A$ , the power set of  $B$  excluding the empty set is denoted by  $\mathcal{P}(B)$ . A *choice function* (on  $B$ ) is a mapping  $f: \mathcal{P}(B) \rightarrow B$  such that  $f(C) \in C$  for all  $C \in \mathcal{P}(B)$ . An *ordering* is a reflexive, complete, transitive and antisymmetric relation. We denote by  $\mathcal{R}_B$  the set of

orderings on  $B$ .

A choice function  $f$  on  $B$  is *best-element rationalizable* if there exists an ordering  $R^*$  on  $B$  such that

$$f(C) = \max(C; R^*)$$

for all  $C \in \mathcal{P}(B)$ , where  $\max(C; R^*) = \{a \in C \mid aR^*b \text{ for all } b \in C\}$ . In this case, we say that  $R^*$  is a *best-element rationalization* of  $f$  or that  $R^*$  *best-element rationalizes*  $f$ . If there are at most two alternatives in  $B$ , any choice function  $f$  on  $B$  is best-element rationalizable: the case in which  $B$  is a singleton is trivial and if  $B$  contains exactly two elements,  $f$  is best-element rationalized by the relation  $R^*$  given by  $f(B)R^*b$  where  $\{b\} = B \setminus \{f(B)\}$ .

The choice function  $f$  is (irrational) *of degree*  $k \geq 1$  if

(i) there exists an ordering  $R^*$  on  $B$  and a collection  $\mathcal{C} = \{C_1, \dots, C_k\}$  of  $k$  distinct subsets of  $B$  such that  $f(C) \neq \max(C; R^*)$  for all  $C \in \mathcal{C}$  and  $f(C) = \max(C; R^*)$  for all  $C \in \mathcal{P}(B) \setminus \mathcal{C}$ ; and

(ii) for all orderings  $R'$  on  $B$  and for all collections  $\mathcal{C}' = \{C'_1, \dots, C'_{k'}\}$  of  $k' < k$  distinct subsets of  $B$ , there exists  $C \in \mathcal{P}(B) \setminus \mathcal{C}'$  such that  $f(C) \neq \max(C; R')$ .

In this case, we say that  $R^*$  *underlies* the choice function  $f$  and we call  $\mathcal{C}$  the collection of *critical sets of  $f$  with respect to  $R^*$* .

Let  $\prec$  be a transitive and asymmetric relation on a non-empty and finite set  $N$ . We say that  $n \in N$  is a *direct predecessor* of  $n' \in N$  if  $n \prec n'$  and there exists no  $n'' \in N$  such that  $n \prec n'' \prec n'$ . Equivalently, we say that  $n'$  is a *direct successor* of  $n$  in this case. The set of direct predecessors of  $n \in N$  is denoted by  $P(n)$ , and  $S(n)$  is the set of direct successors of  $n$ .

A *tree*  $\Gamma$  is given by a quadruple  $(\mathbf{0}, D, T, \prec)$ , where:

(i)  $\mathbf{0}$  is the *root*;

(ii)  $D$  is a finite set of *decision nodes* such that  $\mathbf{0} \in D$ ;

(iii)  $T$  is a non-empty and finite set of *terminal nodes* such that  $D \cap T = \emptyset$ ;

(iv)  $\prec$  is a transitive and asymmetric *precedence relation* on the set of all nodes  $N = D \cup T$  such that:

(iv.a)  $P(\mathbf{0}) = \emptyset$  and  $|S(\mathbf{0})| \geq 1$ ;

(iv.b) for all  $n \in D \setminus \{\mathbf{0}\}$ ,  $|P(n)| = 1$  and  $|S(n)| \geq 1$ ;

(iv.c) for all  $n \in T$ ,  $|P(n)| = 1$  and  $S(n) = \emptyset$ .



When considering two trees  $\Gamma$  and  $\Gamma'$ , we identify the components in the obvious fashion, that is,  $\Gamma = (\mathbf{0}, D, T, \prec)$  and  $\Gamma' = (\mathbf{0}', D', T', \prec')$ . Likewise, the sets of direct predecessors and direct successors of a node  $n \in N$  according to  $\prec$  are  $P(n)$  and  $S(n)$ , whereas these sets for  $n \in N'$  according to  $\prec'$  are denoted by  $P'(n)$  and  $S'(n)$ . This should not create any ambiguity.

A *path* in  $\Gamma$  from a decision node  $n \in D$  to a terminal node  $n' \in T$  (of length  $K \in \mathbb{N}$ ) is an ordered  $(K + 1)$ -tuple  $(n_0, n_1, \dots, n_K) \in N^{|K+1|}$  such that  $n = n_0$ ,  $\{n_{k-1}\} = P(n_k)$  for all  $k \in \{1, \dots, K\}$  and  $n_K = n'$ .

A *game* (on  $B$ ) is a triple  $G = (\Gamma, g, R)$  where  $\Gamma$  is a tree,  $g: T \rightarrow B$  is an *outcome function* that assigns an alternative  $g(n) \in B$  to each terminal node  $n \in T$ , and  $R: D \rightarrow \mathcal{R}_B$  is a *preference assignment* map that assigns an ordering  $R(n)$  on  $B$  to each decision node. This is a simplified version of a perfect-information game because the set of players is implicitly identified with the set of decision nodes and preferences are assumed to be antisymmetric. Again, we use the obvious notation  $(\Gamma', g', R')$  for a game  $G'$  etc.

Let  $G$  be a game on  $B$ . For any  $C \in \mathcal{P}(B)$  such that  $C \subseteq g(T)$ , the *restriction* of  $G$  to  $C$  is the game  $G|C = G' = (\Gamma', g', R')$  on  $C$  given by:

- (i)  $\mathbf{0}' = \mathbf{0}$ ;
- (ii)  $D' = \{n \in D \mid \text{there exist } n' \in g^{-1}(C) \text{ and a path in } \Gamma \text{ from } n \text{ to } n'\}$ ;
- (iii)  $T' = g^{-1}(C)$ ;
- (iv)  $\prec'$  is the restriction of  $\prec$  to  $N' = D' \cup T'$ ;
- (v)  $g'$  is the restriction of  $g$  to  $T'$ ;
- (vi) for all  $n \in D'$ ,  $R'(n)$  is the restriction of  $R(n)$  to  $C$ .

See Figure 2 for an illustration of a game  $G$  on a set  $B = \{1, 2, 3, 4, 5\}$  and its restriction  $G' = G|C$  to  $C = \{1, 4\}$ . The arrows emanating from the terminal nodes point to the respective outcomes according to the outcome functions of the games.

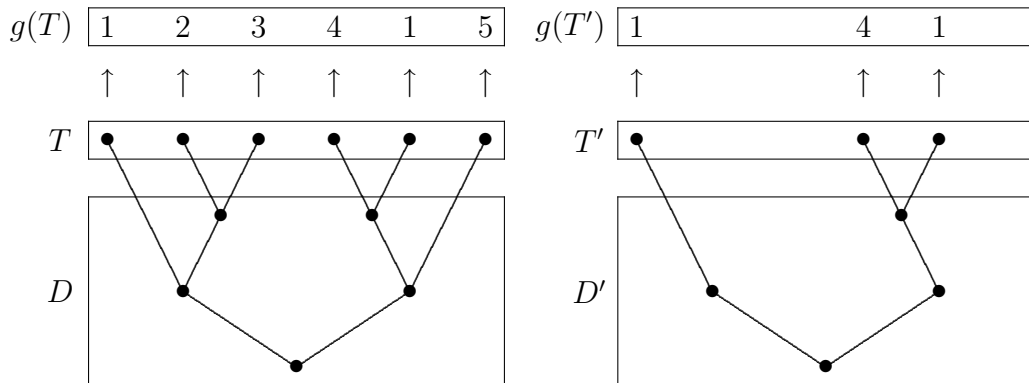


Figure 2: A game on  $B = \{1, 2, 3, 4, 5\}$  and its restriction to  $C = \{1, 4\}$ .

For each decision node  $n \in D$ , we denote by  $e_n(G)$  the backwards-induction outcome of the subgame of  $G$  rooted at  $n$ . This outcome is defined in the usual way: we first set  $e_n(G) = g(n)$  for all  $n \in T$ , then define recursively  $e_n(G) = \max(\{e_{n'}(G) \mid n' \in S(n)\}; R(n))$  for all  $n \in D$ . To simplify notation we write  $e_{\mathbf{0}}(G) = e(G)$ . Because  $G$  is a finite perfect-information game and all preferences are antisymmetric, the backwards-induction outcome of any subgame of  $G$  (including  $G$  itself) exists and is unique. For every  $C \in \mathcal{P}(B)$ , the backwards-induction outcome of  $G|C$  is also well defined.

A choice function  $f$  on  $B$  is *backwards-induction rationalizable* if there exists a game  $G = (\Gamma, g, R)$  on  $B$  such that  $f(C) = e(G|C)$  for all  $C \in \mathcal{P}(B)$ . In this case, we say that  $G$  is a *backwards-induction rationalization* of  $f$  or that  $G$  *backwards-induction rationalizes*  $f$ .

### 3 Results

The following two lemmas provide important steps in the proof of our main result. We begin with choice functions of degree 1 whose only critical set is the set of all alternatives.

**Lemma 1** *Let  $B \in \mathcal{P}(A)$ , let  $R^* \in \mathcal{R}_B$  and let  $b_0 \in B \setminus \max(B; R^*)$ . The choice function  $f$  on  $B$  defined by*

$$f(C) = \begin{cases} \max(C; R^*) & \text{if } C \neq B, \\ b_0 & \text{if } C = B \end{cases}$$

*is backwards-induction rationalizable.*

**Proof.** Let  $B$ ,  $R^*$ ,  $b_0$  and  $f$  be as in the statement of the lemma. Without loss of generality, assume that  $B = \{1, \dots, p\}$  for some  $p$  and that  $R^*$  is the natural ordering  $\leq$

on  $\{1, \dots, p\}$ . Thus, the best element in a non-empty subset  $C$  of  $B$  according to  $R^*$  can be written as  $\min(C)$  without any danger of ambiguity. Moreover,  $b_0 \neq 1$ . Let  $G$  be the game on  $B$  depicted in Figure 3. For simplicity of presentation, we do not include the labels of the terminal nodes but, instead, attach the outcomes according to the outcome function. The preferences assigned to each decision node (that is, to each player) are displayed to the right of the respective node. We simplify our exposition by restriction attention to the part of the preferences that is relevant for the argument. For instance, all that we need to know regarding the preferences associated with nodes  $p + 1, \dots, 2p - 1$  is that one of the two alternatives that can feature as a possible choice at the respective node is preferred to the other. Analogously, there is no need to specify where the alternative  $b_0$  is ranked in the preference assigned to node  $p - 1$  because it can never appear as a possible choice at this node.

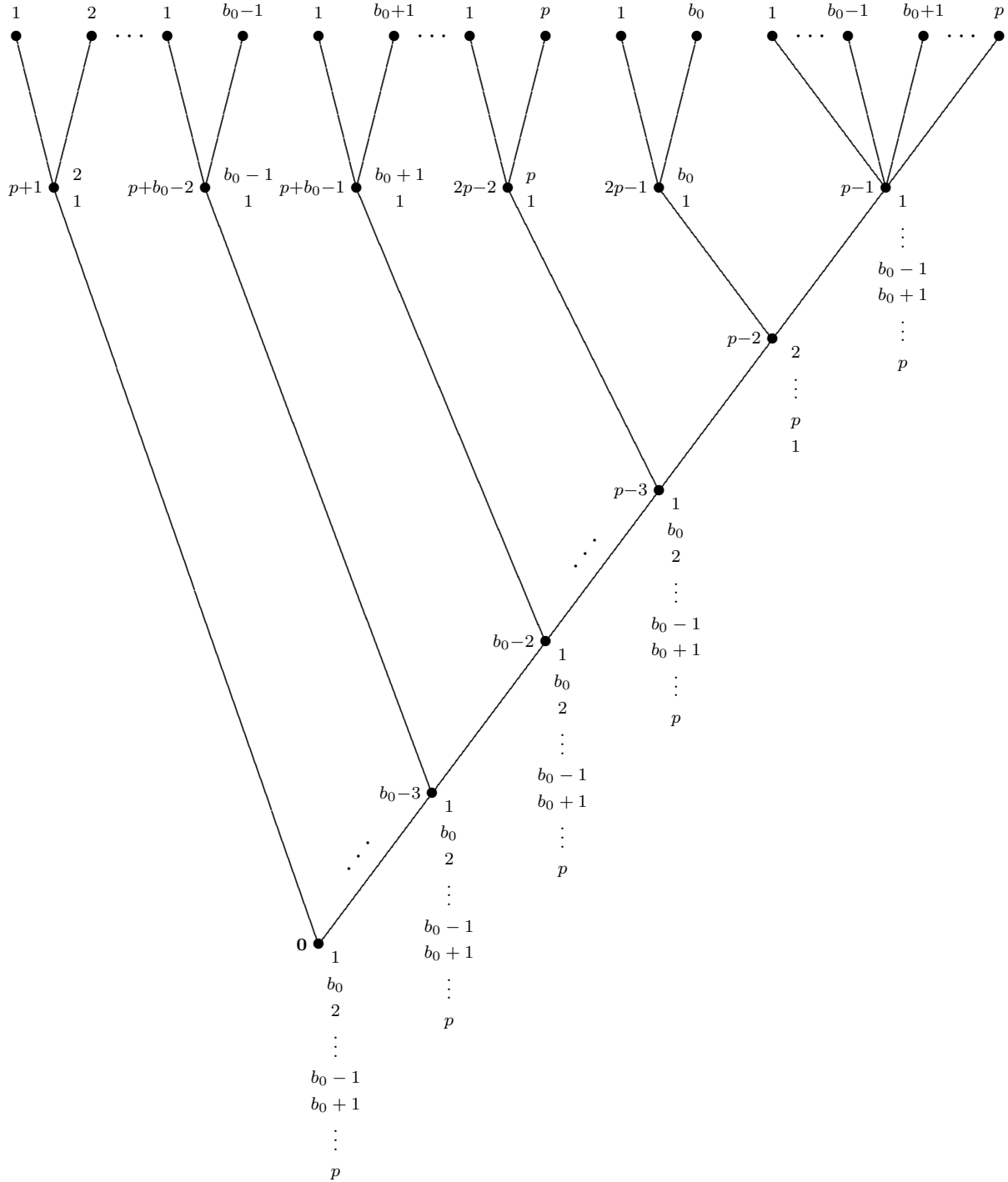


Figure 3: The game  $G$  on  $B$  in the proof of Lemma 1.

We claim that  $f(C) = e(G|C)$  for all  $C \in \mathcal{P}(B)$ . Let  $C \in \mathcal{P}(B)$ . We distinguish four cases.

**Case 1.**  $C = B$ . By definition of  $R(p-1)$  and  $R(2p-1)$ , we have  $e_{p-1}(G|B) = e_{p-1}(G|C) = 1$  and  $e_{2p-1}(G|C) = b_0$ , hence

$$e_{p-2}(G|C) = b_0.$$

Since  $e_{p+q}(G|C) \neq 1$  for all  $q \in \{1, \dots, p-2\}$ , we obtain from the definition of the orderings  $R(p-3), \dots, R(\mathbf{0})$  that

$$e_{p-2}(G|C) = b_0 \Rightarrow e_{p-3}(G|C) = b_0 \Rightarrow \dots \Rightarrow e_{\mathbf{0}}(G|C) = e(G|C) = b_0 = f(C).$$

**Case 2.**  $C \neq B$  and  $\min(C) = 1$ . Since  $C \neq B$ , there exists  $q \in \{1, \dots, p-1\}$  such that  $e_{p+q}(G|C) = 1$ . This in turn implies that

$$e_{q-1}(G|C) = 1. \tag{1}$$

Indeed, if  $q = p-1$ , then  $e_{p+q}(G|C) = 1$  implies  $e_{2p-1}(G|C) = 1$ . Since  $e_{p-1}(G|C) = 1$ , it follows that  $e_{p-2}(G|C) = 1$ , establishing (1). On the other hand, if  $q < p-1$ , then (1) follows from the fact that  $e_{p+q}(G|C) = 1$  and  $\max(C; R(q-1)) = 1$ .

From (1) and the definition of the orderings  $R(q-2), \dots, R(\mathbf{0})$ , we obtain

$$e_{q-1}(G|C) = 1 \Rightarrow e_{q-2}(G|C) = 1 \Rightarrow \dots \Rightarrow e_{\mathbf{0}}(G|C) = e(G|C) = 1 = f(C).$$

**Case 3.**  $C \neq B$  and  $\min(C) = b_0$ . Recall that  $b_0 \neq 1$ . Therefore  $1 \notin C$  and we obtain

$$e_{p-2}(G|C) = b_0 \Rightarrow e_{p-3}(G|C) = b_0 \Rightarrow \dots \Rightarrow e_{\mathbf{0}}(G|C) = e(G|C) = b_0 = \min(C) = f(C).$$

**Case 4.**  $C \neq B$  and  $\min(C) \notin \{1, b_0\}$ . In this case again,  $1 \notin C$ .

**Case 4.1.**  $b_0 \notin C$ . Since  $1 \notin C$  and  $b_0 \notin C$ , there exists  $q \in \{1, \dots, p-2\}$  such that  $e_{p+q}(G|C) = \min(C)$ . By definition of  $R(q-1), \dots, R(\mathbf{0})$ ,

$$e_{p+q}(G|C) = \min(C) \Rightarrow e_{q-1}(G|C) = \min(C) \Rightarrow \dots \Rightarrow e_{\mathbf{0}}(G|C) = e(G|C) = \min(C) = f(C).$$

**Case 4.2.**  $b_0 \in C$  and  $\min(C) < b_0$ . Then  $e_{p-1}(G|C) = \min(C) \notin \{1, b_0\}$ , hence  $e_{p-2}(G|C) = \min(C)$ . By definition of  $R(p-3), \dots, R(\mathbf{0})$  and because  $e_{p+q}(G|C) \notin \{1, b_0\}$  for  $q \in \{1, \dots, p-2\}$ , we obtain

$$e_{p-2}(G|C) = \min(C) \Rightarrow \dots \Rightarrow e_{\mathbf{0}}(G|C) = \min(C) = f(C),$$

which completes the proof. ■

Our next lemma extends the result of Lemma 1 to all choice functions of degree 1.

**Lemma 2** *Every choice function of degree 1 is backwards-induction rationalizable.*

**Proof.** Suppose  $f$  is a choice function of degree 1 on  $A = \{1, \dots, m\}$ . Without loss of generality, we assume that the underlying ordering  $R^*$  is the natural ordering  $\leq$  on  $A$ . As in the proof of Lemma 1, the best element in a non-empty subset  $B$  of  $A$  according to  $R^*$  can thus be written as  $\min(B)$ . Let  $B_0$  be the unique critical set of  $f$  with respect to  $\leq$ . Thus,  $f(B_0) > \min(B_0)$  and  $f(B) = \min(B)$  for all  $B \in \mathcal{P}(A) \setminus \{B_0\}$ . If  $B_0 = A$ , then  $f$  is backwards-induction rationalizable by Lemma 1. From now on, assume  $B_0 \neq A$ .

Consider the restriction of  $f$  to  $\mathcal{P}(B_0)$ . By Lemma 1, there exists a game  $G_0$  on  $B_0$  that backwards-induction rationalizes this restriction. In particular, this implies that

$$e(G_0|B_0) = f(B_0). \quad (2)$$

Now define a game  $G$  on  $A$  as depicted in Figure 4. The root of the tree is  $\mathbf{0}$ . The set of decision nodes  $D$  is given by the union of  $\{\mathbf{0}, 1, 2\}$  and the set of decision nodes  $D_0$  of  $G_0$ . Label all decision nodes of  $G_0$  so that no ambiguities can arise. The set of terminal nodes  $T$  is the union of the set of terminal nodes  $T_0$  of  $G_0$  and  $2 + |A \setminus B_0|$  additional distinct nodes. We only depict the outcomes assigned to these terminal nodes but not the labels attached to them.

Define  $\prec$  on  $D \cup T$  as follows:  $\mathbf{0}$  is the direct predecessor of 1 and of one of the additional terminal nodes, 1 is the direct predecessor of 2 and of another of the new terminal nodes, 2 is the direct predecessor of the remaining  $|A \setminus B_0|$  new nodes and of the root of the tree  $G_0$ , and the restriction of  $\prec$  to  $D_0$  coincides with  $\prec_0$ .

Let  $g$  be such that  $g(n) = g_0(n)$  for all terminal nodes  $n \in T_0$  and as displayed in the figure for all other nodes. Note that  $g$  assigns a different alternative in  $A \setminus B_0$  to each node in the subset labeled  $A \setminus B_0$ .

The preference assignment  $R$  is such that  $R(\mathbf{0}) = R^*$ ,  $R(1)$  has  $f(B_0)$  as its best element, followed by the remaining alternatives in  $A$  in their natural order, and  $R(2)$  first ranks the alternatives in  $A$  other than  $f(B_0)$  in their natural order and has  $f(B_0)$  as its worst element. Each ordering  $R(n)$  corresponding to a decision node  $n$  of  $G_0$  is such that it ranks all alternatives in  $B_0$  in the same way they are ranked by  $R_0(n)$ .



definition of  $R(1)$  that  $e_1(G|B) = \min(B_0)$ . Because  $\min(B_0) < f(B_0)$  and  $R(\mathbf{0})$  is equal to  $\leq$ , we obtain  $e(G|B) = \min(B_0) = \min(B)$ .

**Case 2.2.2.**  $B \cap B_0 \neq B_0$ . Since  $G_0$  backwards-induction rationalizes the restriction of  $f$  to  $\mathcal{P}(B_0)$ ,

$$e(G_0|(B \cap B_0)) = \min(B \cap B_0) = \min(B) = \min(B_0)$$

because both  $B$  and  $B_0$  must contain  $\min(B) = \min(B_0)$ . Since  $\min(B) = \min(B_0) \neq f(B_0)$ , we have  $\min(B)R(2)b$  for all  $b \in B \cap (A \setminus B_0)$ . Therefore  $e_2(G|B) = \min(B)$ . Next,  $e(G|B) = \min(B)$  because  $\min(B) = \min(B_0)R(\mathbf{0})f(B_0)$ .

**Case 3.**  $\min(B_0) < \min(B)$ . In this case  $\min(B_0) \notin B$  and hence  $B \cap B_0 \neq B_0$ . This implies

$$e(G_0|(B \cap B_0)) = \min(B \cap B_0).$$

Again, we consider two subcases.

**Case 3.1.**  $\min(B) = f(B_0)$ . Then we obtain  $e(G|B) = f(B_0) = \min(B)$  immediately from the definition of  $R(\mathbf{0})$ .

**Case 3.2.**  $\min(B) \neq f(B_0)$ . Two further subcases can be distinguished.

**Case 3.2.1.**  $\min(B) \in B_0$ . In this case,  $e(G_0|(B \cap B_0)) = \min(B \cap B_0) = \min(B) \neq f(B_0)$ . By definition of  $R(2)$ ,

$$\begin{aligned} e_2(G|B) &= \min\{\min(B \cap (A \setminus B_0)), e(G_0|(B \cap B_0))\} \\ &= \min\{\min(B \cap (A \setminus B_0)), \min(B \cap B_0)\} \\ &= \min(B). \end{aligned}$$

Because  $\min(B_0) \notin B$ , it follows that

$$e_1(G|B) = \min(B). \tag{3}$$

Since  $\min(B) \neq f(B_0)$ , either  $\min(B) < f(B_0)$  or  $\min(B) > f(B_0)$ . In the first case, (3) and the definition of  $R(\mathbf{0})$  imply that  $e(G|B) = \min(B)$ . In the second case,  $f(B_0) \notin B$  and (3) therefore implies  $e(G|B) = \min(B)$  by definition of the game  $G$ .

**Case 3.2.2.**  $\min(B) \in A \setminus B_0$ . Now  $\min(B \cap (A \setminus B_0)) = \min(B) \neq f(B_0)$ . If  $e(G_0|(B \cap B_0)) \neq f(B_0)$ , the definition of  $R(2)$  implies

$$\begin{aligned} e_2(G|B) &= \min\{\min(B \cap (A \setminus B_0)), e(G_0|(B \cap B_0))\} \\ &= \min\{\min(B), e(G_0|(B \cap B_0))\} \\ &= \min(B). \end{aligned}$$



If  $e(G_0|(B \cap B_0)) = f(B_0)$ , the definition of  $R(2)$  implies

$$e_2(G|B) = \min(B \cap (A \setminus B_0)) = \min(B).$$

In both cases  $e_2(G|B) = \min(B)$ . Since  $\min(B_0) \notin B$ , it follows that  $e_1(G|B) = \min(B)$ . The conclusion that  $e(G|B) = \min(B)$  now follows by the same argument as in case 3.2.1.

■

Using the result of the above lemma, our theorem establishes that *all* choice functions are backwards-induction rationalizable.

**Theorem 1** *Every choice function is backwards-induction rationalizable.*

**Proof.** Every best-element rationalizable choice function is trivially backwards-induction rationalizable and Lemma 2 shows that every choice function of degree 1 is backwards-induction rationalizable. We prove the theorem by induction on the degree of irrationality of a choice function. Let  $k > 1$  and, as our induction hypothesis, suppose that every choice function of irrationality degree  $1, \dots, k-1$  is backwards-induction rationalizable. Consider a choice function  $f$  on  $A = \{1, \dots, m\}$ . Suppose that  $f$  is of degree  $k$  and let the ordering  $R^*$  underlie  $f$ . Without loss of generality, let  $R^*$  be the natural ordering  $\leq$  on  $A$ . Let  $\mathcal{C} = \{C_1, \dots, C_k\}$  be the collection of critical sets of  $f$  with respect to  $R^*$ . Thus, by definition,  $f(C) \neq \min(C)$  for all  $C \in \mathcal{C}$  and  $f(C) = \min(C)$  for all  $C \in \mathcal{P}(A) \setminus \mathcal{C}$ . Pick  $C_1, C_2 \in \mathcal{C}$ . We have

$$f(C_1) \neq \min(C_1) \quad \text{and} \quad f(C_2) \neq \min(C_2). \tag{4}$$

For  $i \in \{1, 2\}$ , define the choice function

$$f_i(C) = \begin{cases} \min(C_i) & \text{if } C = C_i, \\ f(C) & \text{if } C \in \mathcal{P}(A) \setminus \{C_i\}. \end{cases}$$

By construction, each  $f_i$  is of irrationality degree less than  $k$  and, by the induction hypothesis,  $f_i$  is backwards-induction rationalizable by a game  $G_i$ .

Let  $R_0$  be an ordering on  $A$  such that  $f(C_i) R_0 \min(C_i)$  for  $i \in \{1, 2\}$ . Observe that such an ordering exists if and only if

$$f(C_1) \neq \min(C_2) \quad \text{or} \quad f(C_2) \neq \min(C_1). \tag{5}$$

To see that (5) is implied by (4), suppose that, by way of contradiction,

$$f(C_1) = \min(C_2) \quad \text{and} \quad f(C_2) = \min(C_1).$$

Since  $f(C_i) \in C_i$  for  $i \in \{1, 2\}$ , we must have  $f(C_1) \leq f(C_2)$  and  $f(C_2) \leq f(C_1)$ , hence  $f(C_1) = f(C_2)$  because  $\leq$  is antisymmetric. But then we obtain

$$f(C_1) = f(C_2) = \min(C_1) = \min(C_2),$$

contradicting (4).

Now construct a game  $G$  on  $A$  by letting the root  $\mathbf{0}$  be such that it is associated with the ordering  $R(\mathbf{0}) = R_0$  and has two direct successors, namely, the roots of  $G_1$  and  $G_2$ , and all other characteristics of the game  $G$  are inherited from  $G_1$  and  $G_2$  in the obvious fashion. See Figure 5.

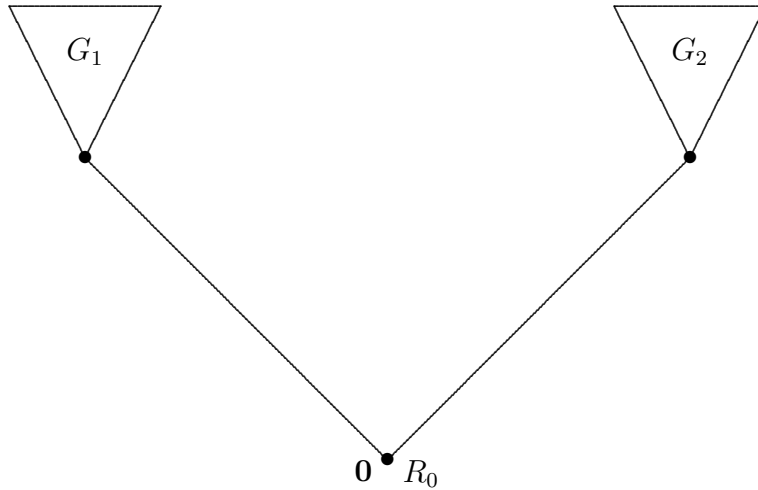


Figure 5: The game  $G$  on  $A$  in the proof of Theorem 1.

By construction, we have

$$e(G|C) = \max(\{e(G_1|C), e(G_2|C)\}; R_0)$$

for all  $C \in \mathcal{P}(A)$ .

If  $C \notin \{C_1, C_2\}$ , it follows that  $e(G_1|C) = e(G_2|C) = f(C)$  and hence  $e(G|C) = f(C)$ .

If  $C = C_1$ , we obtain

$$\begin{aligned} e(G|C) &= e(G|C_1) \\ &= \max(\{e(G_1|C_1), e(G_2|C_1)\}; R_0) \\ &= \max(\{f_1(C_1), f_2(C_1)\}; R_0) \\ &= \max(\{\min(C_1), f(C_1)\}; R_0) \\ &= f(C_1) \\ &= f(C). \end{aligned}$$

If  $C = C_2$ , the above argument can be repeated with the roles of the labels 1 and 2 interchanged. ■

## References

- Arrow, K.J. (1959). Rational choice functions and orderings, *Economica* **26**, 121–127.
- Debreu, G. (1974). Excess demand functions, *Journal of Mathematical Economics* **1**, 15–23.
- Galambos, A. (2005). Revealed preference in game theory, mimeo, Northwestern University.
- Kuhn, H.W. (1953). Extensive games and problems of information, in *Contributions to the theory of games II*, edited by H.W. Kuhn and A.W. Tucker, Princeton: Princeton University Press, 193–216.
- Lee, S. (2012). The testable implications of zero-sum games, *Journal of Mathematical Economics* **48**, 39–46.
- Mantel, R. (1974). On the characterization of aggregate excess demand, *Journal of Economic Theory* **7**, 348–353.
- McGarvey, D. (1953). A theorem on the construction of voting paradoxes, *Econometrica* **21**, 608–610.
- Ray, I. and Snyder, S. (2003). Observable implications of Nash and subgame-perfect behavior in extensive games, mimeo, Brown University.
- Ray, I. and Zhou, L. (2001). Game theory via revealed preferences, *Games and Economic Behavior* **37**, 415–424.
- Sonnenschein, H. (1973). Do Walras’s identity and continuity characterize the class of community excess demand functions? *Journal of Economic Theory* **6**, 345–354.
- Sprumont, Y. (2000). On the testable implications of collective choice theories, *Journal of Economic Theory* **93**, 205–232.
- Xu, Y. and Zhou, L. (2007). Rationalizability of choice functions by game trees, *Journal of Economic Theory* **134**, 548–556.
- Yanovskaya, E. (1980). Revealed preference in noncooperative games, *Mathematical Methods in the Social Sciences* **24**, 73–81 [In Russian].