DEPARTMENT OF ECONOMICS

Working Paper

FARSIGHTED STABILITY IN HEDONIC GAMES

Effrosyni Diamantoudi
Licun Xue

Working Paper No. 2000-12
Centre for Dynamic Modelling in Economics

ISSN 1396-2426

UNIVERSITY OF AARHUS • DENMARK

CENTRE FOR DYNAMIC MODELLING IN ECONOMICS

DEPARTMENT OF ECONOMICS - UNIVERSITY OF AARHUS - DK - 8000 AARHUS C - DENMARK
☎ +45 89 42 11 33 - TELEFAX +45 86 13 63 34

WORKING PAPER

FARSIGHTED STABILITY IN HEDONIC GAMES

Effrosyni Diamantoudi
Licun Xue

Working Paper No. 2000-12

DEPARTMENT OF ECONOMICS

SCHOOL OF ECONOMICS AND MANAGEMENT - UNIVERSITY OF AARHUS - BUILDING 350
8000 AARHUS C - DENMARK ☎ +45 89 42 11 33 - TELEFAX +45 86 13 63 34

Farsighted Stability in Hedonic Games

Effrosyni Diamantoudi* and Licun Xue*

November 2000

Abstract

We investigate how rational individuals partition themselves into different coalitions in "hedonic games" [see Banerjee, Konishi and Sönmez (1998) and Bogomolnaia and Jackson (2000)], where individuals' preferences depend solely on the composition of the coalition they belong to. We show that the four solution concepts studied in the literature (core, Nash stability, individual stability and contractual individual stability) exhibit myopia on the part of the players. We amend these notions by endowing players with foresight in that they look many steps ahead and consider only credible outcomes. We show the existence and study the properties of the new solutions, as well as their relation to the previous notions.*Journal of Economic Literature Classification Number*: C71, C78. *Keywords*: hedonic games, coalition structures, foresight.

1 Introduction

The significance of coalition formation is manifested through its presence in many facets of our economic, political, and social life as well as by the numerous studies attempting to model it. The plethora of efforts highlights the complexities embedded in the problem in question. In order to circumvent some of these complexities, the study of coalition formation is often limited to specific contexts with more structured preferences, or payoff distributions

*We would like to thank the seminar audience at the University of Copenhagen for useful comments. Both authors are at the Department of Economics, University of Aarhus, Building 350, DK-8000 Aarhus C., Denmark. Email addresses are *faye@econ.au.dk* and *lxue@econ.au.dk*, respectively.

or even constrained coalition formation options. Under such restrictions, it is often possible to analyze more effectively various solution concepts and their properties.

In this spirit, we study coalition formation in "hedonic games", that is, in situations where individuals' preferences depend solely on the composition of the coalition they belong to. The *hedonic aspect* in players' preferences was originally introduced by Dréze and Greenberg (1980) in a context concerning local public goods, where agents' preferences depend on their consumption of the public good as well as the coalition they belonged to. However, the hedonic aspect transcends the local public goods context and spans over a much greater area of economics and sociology. The formation of societies, communities, social clubs and groups are just a few cases where the hedonic aspect is one of the central elements steering their structure.

Recently, several works concentrate on the hedonic notion and its implications in a more general framework. Notably, Banerjee, Konishi and Sönmez (1998) and Bogomolnaia and Jackson (2000) introduced the aforementioned pure hedonic games. Various solution concepts are applied to hedonic games and their rendition is studied. The *Core*, identifies coalition structures that are immune to any beneficial coalitional deviation, addressing situations where coalitional deviations are feasible and costless. However, in many case coalitional deviations are not viable, either because of institutional (legal or physical) constraints or simply because agents are not able to coordinate their actions, especially in situations where the entire set of agents is rather large. As a result, solution concepts that consider only individual deviations are warranted. *Nash Stability* identifies coalition structures where no player wishes to migrate to another coalition in the same structure. Such an analysis is relevant in situations where no permission is required to join a new coalition, a simple example being that of moving from one city to another. *Individual Stability* examines not only the preferences of the migrating player but also the preferences of the coalition this player plans to join. A coalition structure, therefore, is not individually stable if some player wishes to migrate to another coalition whose members permit him to do so, in that they are not hurt by such a migration. Any situation where an individual is hired by a business entity (a coalition), and thus is allowed to join them, serves as an example where individual stability is a more appropriate notion. Lastly, even more restrictive is the notion of *Contractual Individual Stability* that requires, in addition, the members of the original coalition to permit a player to migrate (again in the sense that they are not hurt by the departure of this member). In the spirit of the previous example, consider the case where the newly hired employee has to first break his contract with his previous employer. Unfortunately, most of the afore-mentioned notions do not always provide an answer even to simple problems as we will illustrate later in this section. Therefore, restrictions on players' preferences need to be made to guarantee the existence of different solutions.

In particular, Banerjee, Konishi and Sönmez (1998) study the Core in hedonic games. They define the *top coalition property*, a property that imposes a degree of commonality on players' preferences, while it guarantees the non-emptiness of the Core. The motivation for such a property is served from two perspectives: (i) the multiplicity of economic applications, some of which are presented in their paper, where the property naturally holds, and (ii) the possibility of an empty Core, illustrated through examples, even when strong assumptions on preferences such as additive separability, anonymity or single peakedness are imposed. Bogomolnaia and Jackson (2000) study primarily individually stable coalition structures in hedonic games. They propose the condition of *ordered characteristics* – a condition that builds on single-peaked preferences – and *consistency* to guarantee the existence of individually stable coalition structures and partially their efficiency, while they provide an algorithm to identify weakly Pareto efficient and individually stable coalition structures. Analogously, the motivation for the condition stems from examples satisfying already strong assumptions where individually stable coalition structures do not exist.

This project attempts to build on such endeavors, extending the existing analysis by identifying and rectifying some problems associated with the solution concepts used. While the most analyzed problem concerning the Core is its "frequent" emptiness, the notion has also being criticized for the easiness with which it discards feasible outcomes -which often results to emptiness. In particular, an outcome is excluded from the Core because a coalition can induce an other outcome that benefits all of its members. Whether this outcome survives a further deviation by some other coalition is not examined. Especially, when such a further deviation may render the original deviating coalition worse off than it was in the very beginning, the original deviation may be deterred.

Alternatively, the Core has been criticized for including outcomes that can be dominated if a group of agents does not behave myopically. To illustrate this point consider a group of agents that can induce some outcome which does not necessarily improve upon the original one. It is possible however, that once this new outcome becomes the status quo, further deviations may occur which will render the original deviating coalition better off than it was in the very beginning. Allowing therefore, a potential deviating coalition to foresee many steps ahead, it may proceed with a deviation which ultimately benefits its members, that otherwise may have been halted. Indeed, such an argument applies to Example 5 where a core element is ruled out. However, if players' preferences are strict, then core outcomes exhibit a surprising robustness and survive foresight, as is formally shown in following sections. Put differently, for a core outcome to be excluded from the farsighted solution we propose, it must be the case that some players are indifferent between some coalitions.

In environments where moves can be followed by counter moves and so on, it is natural to allow a deviating coalition to "speculate" on the ultimate result of its action. Obviously, an *ultimate* result must be immune to further deviations. That is, for an outcome to be ultimate, it must be the case that no group of agents has an incentive to deviate, again anticipating the ultimate outcomes of its deviation. Thus, an ultimate result must be in the solution set of the game. Such a reasoning is captured by the notion of *abstract stable set* first introduced by von Neumann & Morgenstern (1944). The notion is characterized by internal and external stability. *Internal stability* guarantees that the solution set is free from inner contradictions in the sense that no outcome in the solution set is dominated by another outcome also in the solution set; *external stability* guarantees that every outcome excluded from the solution set is accounted for in the sense that it is dominated by some outcome inside the solution set. Although the original abstract stable set accommodates the issue of credibility regarding a dominating outcome it does not address foresight. Harsanyi (1974) introduced the concept of indirect dominance to capture foresight. An outcome indirectly dominates another, if there exists a sequence of outcomes starting from the dominated outcome and leading to the dominating one, and at each stage of the sequence the group of players required to enact the inducement prefers the final outcome to its status quo.

In criticizing the Core for its myopia and lack of credibility, we suggested that agents should look ahead and speculate on the ultimate result of their initial actions. Baring in mind that different sequences of actions may take place, it is natural to question how the players are to treat situations where the ultimate result may not be unique; moreover, how they are to behave when some ultimate outcomes may make them better off while others make them worse off. The answer to such a question depends solely on the behavioral characteristic of the players making the decision. Two extreme approaches are that of optimism (implicit in the notion of abstract stable set), where a coalition would proceed with the deviation as long as one of the ultimate outcomes makes its members better off, whereas a conservative coalition would not proceed with the deviation unless every possible ultimate outcome makes its members better off. Amalgamations of stability and indirect dominance, are studied by Chwe (1994) and Xue (1998) where both behavioral assumptions are investigated along with the existence of the solution concepts.

The criticism developed thus far concerning the Core applies also to the stability notions that allow for only individual deviations. In the same manner a deviating coalition considers all plausible outcomes that may arise from its initial deviation, an individual migrating from one coalition to another can raise the same concerns and question the possibility of others joining later on, or existing members departing. Moreover, the welcoming coalition (in the case of individual stability) may, for example, wonder whether the admission of a new undesirable member may bring later on far more desirable migrants. Similarly, the remaining coalition (in the case of contractual individual stability) may permit a favorite member to depart believing that it will, later on, be replaced by someone even more desirable, and so on.

To amend existing solution concepts, we introduce notions that allow players to foresee arbitrarily many steps ahead, and consider only the plausible (as opposed to any feasible) outcomes that are likely to result from their actions. Although the next sections define formally and motivate more precisely all proposed solution concepts, we illustrate through the well known problem of roommates how (most) solutions remain silent.

Example 1 (*The roommate problem*) *Three individuals have to agree in sharing a room that can only accommodate two of them. Each individual prefers to share the room with only one other person rather than be alone and of course none of them wishes that all three squeeze in one room. Unfortunately, their preferences are rather contradicting: the first prefers to share the room with the second rather than with the third, while the second prefers to share the room with the third rather than with the first, and lastly the third prefers to share the room with the first rather than with the second. Their preferences are summarized as follows:*

$$\begin{aligned}\{1,2\} &\succ_1 \{1,3\} \succ_1 \{1\} \succ_1 \{1,2,3\} \\ \{2,3\} &\succ_2 \{1,2\} \succ_2 \{2\} \succ_2 \{1,2,3\} \\ \{1,3\} &\succ_3 \{2,3\} \succ_3 \{3\} \succ_3 \{1,2,3\}\end{aligned}$$

It is easy to see that the Core of the game is empty. For example consider the partition of $\{\{1,2\},\{3\}\}$; it is blocked by $\{2,3\}$ since 2 prefers to be with 3 rather than with 1, and 3 prefers to be with 2 rather than be alone. The same argument blocks any partition involving a pair and a singleton.

Unfortunately, the Core is not the only solution encountering problems. No coalition structure is Nash stable either. Consider the same partition of $\{\{1,2\},\{3\}\}$, which is blocked via 2 that wishes to join 3. In fact, since 3 welcomes 2 this coalition structure is not individually stable either. Similar arguments render all partitions Nash and individually unstable.

The only solution that provides an answer to this problem is contractual individual stability. Although the grand coalition and the singletons are excluded from the solution, the partitions involving a pair and a singleton are contractually individually stable. Consider the partition $\{\{1,2\},\{3\}\}$, although 2 would like to join 3 and 3 welcomes 2, 1 will not permit 2 to depart from their coalition. In fact, contractual individual stability offers a very appealing analysis to the problem if 1 and 2 had indeed singed a contract together for their room. In the event, though, where no commitments have taken place the notion is not applicable, and no answer is given to the problem.

In contrast, all four extensions of the solution concepts proposed in this paper suggest the same solution, that is, they support only all the coalition structures involving a pair and a singleton. The fundamental argument embedded in all the notions is as follows: consider again the partition $\{\{1,2\},\{3\}\}$, player 2 will not join 3 fearing that 3 may, later on, join 1 and thus leave 2 alone. If the individuals are conservative and farsighted, they will not dissolve a "mediocre" partnership in pursuance of the "perfect" one.

In Section 2 we introduce formally the model and the solution concepts. We also provide examples where the old notions are not silent to illustrate the distinction and refinement of the new ones. Existence results concerning the solution concepts are presented in Section 3. In the same section we investigate the relation of the new notions to existing ones when different restrictions are imposed on players' preferences. Lastly, in Section 4 we discuss conjectures and open questions of interest to the topic.

2 Definitions

We first introduce the basic notations and the four existing stability notions discussed in the introduction. Then we proceed to introduce the stability notions under foresight.

2.1 Preliminaries

- Let N be a finite set of players.

- A coalition S is a non-empty subset of N..

- Let $\mathcal{S}$ denote the collection of all coalitions and for $i \in N$, let $\mathcal{S}(i)$ denote the collection of coalitions that contain i, that is, $\mathcal{S}(i) = \{S \subset N \mid i \in S\}$.

Hedonic game: A hedonic game G is a pair $(N, \{\succeq_i\}_{i \in N})$, where N is a finite set of players and for all $i \in N$, $\succeq_i$ is a complete, reflexive, and transitive binary relation on $\mathcal{S}(i)$, representing i's preferences over coalitions that contain i. We use $\succ_i$ to denote the asymmetric part of $\succeq_i$ (i.e., strict preferences) and $\sim_i$ the indifference relation.

- A coalition structure, P, is a partition of N, that is, $P = \{S_1, S_2, ..., S_k\}$, $\bigcup_{j=1}^{k} S_j = N$ and for all $i \neq j$, $S_i \bigcap S_j = \emptyset$.

- Let $\mathcal{P}$ be the collection of all coalition structures and $\mathcal{P}(S) = \{P \in \mathcal{P} \mid S \in P\}$ be the collection of all coalition structures that contain coalition S.

- For $P \in \mathcal{P}$ and $i \in N$, let $S_P(i)$ be the coalition $S \in P$ such that $i \in S$.

The four main stability notions in the literature are defined below. Core stability is a notion of coalitional stability.

Core stability: A coalition structure $P \in \mathcal{P}$ is core stable or in the core of G if there does not exist "a blocking coalition" $S \subset N$ such that $S \succ_i S_P(i)$ for all $i \in S$.

The remaining three notions are often referred to as non-cooperative stability notions.

Nash stability: A coalition structure $P \in \mathcal{P}$ is Nash stable if there do not exist $i \in N$ and $S \in P \cup \{\emptyset\}$ such that $S \cup \{i\} \succ_i S_P(i)$.

Individual stability: A coalition structure $P \in \mathcal{P}$ is individually stable if there do not exist $i \in N$ and $S \in P \cup \{\emptyset\}$ such that $S \cup \{i\} \succ_i S_P(i)$ and $S \cup \{i\} \succeq_j S$ for all $j \in S$.

Contractural individual stability: A coalition structure $P \in \mathcal{P}$ is contractually individually stable if there do not exist $i \in N$ and $S \in P \cup \{\emptyset\}$ such that (1) $S \cup \{i\} \succ_i S_P(i)$ and $S \cup \{i\} \succeq_j S$ for all $j \in S$; (2) $S_P(i) \setminus \{i\} \succeq_j S_P(i)$ for all $j \in S_P(i) \setminus \{i\}$.

2.2 Farsighted coalitional stability

According to core stability, if P is under consideration, any coalition $S \subset N$ can form and "object to" P. If S forms and before other players regroup, the resulting coalition structure is

$$P' = \{S\} \cup \{T \setminus S \mid T \in P \text{ and } T \setminus S \neq \emptyset\};$$

in this case, we write $P \xrightarrow{S} P'$.

For convenience, we shall extend players' preferences over coalitions to coalitional preferences over coalition structures.

Coalitional preferences: A coalition $S \subset N$ strictly prefers P' to P, denoted $P' \succ_S P$, where $P, P' \in \mathcal{P}$, if $S_{P'}(i) \succ_i S_P(i)$ for all $i \in S$; similarly, a coalition $S \subset N$ prefers P' to P, denoted $P' \gtrsim_S P$, where $P, P' \in \mathcal{P}$, if $S_{P'}(i) \gtrsim_i S_P(i)$ for all $i \in S$.

Given a coalition structure $P \in \mathcal{P}$, if some coalition $S \subset N$ has an incentive to form, thereby inducing P', we say P' dominates P.

(Direct) Dominance: For $P, P' \in \mathcal{P}$, P' (directly) dominates P, or $P' > P$, if $P \xrightarrow{S} P'$ and $P' \succ_S P$.

Thus, core can be alternatively defined as follows:

Core: Core of G is the set of coalition structures that are not dominated with respect to $>$, i.e., core of G is the (abstract) core of abstract system[1] $(\mathcal{P}, >)$. Formally,

$$\text{Core}(\mathcal{P}, >) = \{P \in \mathcal{P} \mid \nexists P' \in \mathcal{P} \text{ such that } P' > P\}.$$

[1] An abstract system consists of a set and a binary relation on this set. See von Neumann and Morgenstern (1944).

Myopia is reflected in the fact that a blocking coalition does not consider the possibility that the rest of the players may regroup. That is, given P', another coalition can form to induce a new coalition structure and so on. If players are farsighted, they should consider the ultimate outcomes of their actions. Thus, a coalition may choose to "deviate" to a coalition structure, which does not necessarily make its members better off, as long as its deviation leads to final coalition structures that benefit all its members; similarly, a coalition may choose not to deviate to a coalition structure it prefers if its deviation eventually leads to coalition structures that make its members worse off. The following "indirect dominance" due to Harsanyi (1974) and Chwe (1994) captures foresight.

Indirect dominance: For $P, P' \in \mathcal{P}$, P' indirectly dominates P, or $P' \gg P$, if there exists a sequence of coalition structures $P^1, P^2, \ldots, P^k$, with $P^1 = P$ and $P^k = P'$, and a sequence of coalitions $S^1, S^2, \ldots, S^{k-1}$ such that $P^j \stackrel{S^j}{\rightarrow} P^{j+1}$ and $P' \succ_{S^j} P^j$ for all $j = 1, 2, \ldots, k-1$.

We can now examine different solutions of the abstract system $(\mathcal{P}, \gg)$. First, consider the (abstract) core of $(\mathcal{P}, \gg)$.

The (abstract) core of $(\mathcal{P}, \gg)$:

$$\text{Core}(\mathcal{P}, \gg) = \{P \in \mathcal{P} \mid \nexists P' \in \mathcal{P} \text{ s.t. } P' \gg P\}.$$

It is easy to see that $\text{Core}(\mathcal{P}, \gg) \subset \text{Core}(\mathcal{P}, >)$. However, $\text{Core}(\mathcal{P}, \gg)$, like $\text{Core}(\mathcal{P}, >)$, does not consider the credibility of the dominating alternative. In fact, as we shall see in Section 3, $\text{Core}(\mathcal{P}, \gg)$ is too exclusive and hence is very likely to be empty. The following von Neumann-Morgenstern (vN-M) (abstract) stable set amends the core by insisting that a dominating alternative to be credible.

vN-M stable set of $(\mathcal{P}, \gg)$: $\mathcal{R} \subset \mathcal{P}$ is *vN-M internally stable* if there do not exist $P, P' \in \mathcal{R}$ such that $P' \gg P$; $\mathcal{R}$ is *vN-M externally stable* if for all $P \in \mathcal{P} \setminus \mathcal{R}$, there exists $P' \in \mathcal{R}$ such that $P' \gg P$. $\mathcal{R}$ is *vN-M stable* is it is both internally and externally stable.

As shown in Greenberg (1990), vN-M stable set implicitly assumes "optimistic behavior". To see this, we first introduce the following notation.

Likely outcomes given P **:** For $\mathcal{Q} \subset \mathcal{P}$ and $P \in \mathcal{P}$, let

$$\mathcal{Q}|_{P,\gg} = \{P' \in \mathcal{Q} \mid P' = P \text{ or } P' \gg P\}.$$

Thus, if $\mathcal{Q}$ is a "solution set", then $\mathcal{Q}|_{P,\gg}$ is the set of likely outcomes when P is under consideration.

Now we can rewrite the definition of vN-M stable set.

vN-M stable set redefined: $\mathcal{R}$ is vN-M internally stable if $Q \in \mathcal{R}$ implies that there do not exist $P \in \mathcal{P}$ and $S \subset N$ such that $Q \xrightarrow{S} P$ and $P' \succ_S Q$ for *some* $P' \in \mathcal{R}|_{P,\gg}$. $\mathcal{R}$ is vN-M externally stable if $Q \in \mathcal{P} \setminus \mathcal{R}$ implies that there exist $P \in \mathcal{P}$ and $S \subset N$ such that $Q \xrightarrow{S} P$ and $P' \succ_S Q$ for *some* $P' \in \mathcal{R}|_{P,\gg}$. $\mathcal{R}$ is vN-M stable if it is both internally and externally stable.

Optimism is reflected in the phrase "for *some* $P' \in \mathcal{R}|_{P,\gg}$". In fact, the vN-M stable set for $(\mathcal{P}, \gg)$ entails *over-optimism* on the part of a departing coalition. Indeed, it is easy to see from the above definition and the definition of indirect dominance that coalition S, in contemplating a deviation, expects the other players to act in a way that is in the best interest of S[2]. The following notion of conservative stable set amends this problem.

Conservative stable set: $\mathcal{Q}$ is *conservatively internally stable* if $Q \in \mathcal{Q}$ implies that there do not exist $P \in \mathcal{P}$ and $S \subset N$ such that $Q \xrightarrow{S} P, \mathcal{Q}|_{P,\gg} \neq \emptyset$, and $P' \succ_S Q$ for *all* $P' \in \mathcal{Q}|_{P,\gg}$. $\mathcal{Q}$ is *conservatively externally stable* if $Q \in \mathcal{P} \setminus \mathcal{Q}$ implies that there exist $P \in \mathcal{P}$ and $S \subset N$ such that $Q \xrightarrow{S} P, \mathcal{Q}|_{P,\gg} \neq \emptyset$, and $P' \succ_S Q$ for *all* $P' \in \mathcal{Q}|_{P,\gg}$. $\mathcal{Q}$ is *conservatively stable* if it is both conservatively internally and externally stable.

[2] See Xue (1998) for detailed discussion on this matter in a more general framework.

The conservative stable set is closely related to Greenberg's (1990) "conservative stable standard of behavior" (CSSB) and Chwe's (1994) "consistent set". The following example illustrates the myopia embedded in the core and how it is rectified through the conservative stable set.

Example 2 (*Expanding the Core*) *Consider a game with 4 players and the following preference orderings:*

$$\begin{aligned}
\{1,4\} &\succ_1 \{1,2,3\} \succ_1 \{1,2\} \succ_1 \{1\} \succ_1 \cdots \\
\{1,2,3\} &\succ_2 \{1,2\} \succ_2 \{2,3\} \succ_2 \{2\} \succ_2 \cdots \\
\{1,2,3\} &\succ_3 \{2,3\} \succ_3 \{3,4\} \succ_3 \{3\} \succ_3 \cdots \\
\{3,4\} &\succ_4 \{1,4\} \succ_4 \{4\} \succ_4 \cdots
\end{aligned}$$

The rest of relevant coalitions are ranked strictly below the above coalitions. The core[3] *contains only the following partition:*

$$\{\{1,4\},\{2,3\}\}.$$

Note that partition $\{\{1,2\},\{3,4\}\}$ is not in the core since coalition $\{1,2,3\}$ can strictly improve upon it. However, if players 2 and 3 are farsighted they will realize that once partition $\{\{1,2,3\},\{4\}\}$ is the status quo, coalition $\{1,4\}$ will form and bring about $\{\{1,4\},\{2,3\}\}$ which is not as good for 2 and 3 as the original one. Therefore, $\{\{1,2\},\{3,4\}\}$ should not be ruled out by $\{1,2,3\}$ if 2 and 3 are farsighted.

However, the reader may have already observed that the following sequence of events may take place:

$$\begin{aligned}
\{\{1,2\},\{3,4\}\} \quad &\xrightarrow{\{1,3\}} \{\{2\},\{1,3\},\{4\}\} \xrightarrow{\{2,3\}} \{\{1\},\{2,3\},\{4\}\} \\
&\xrightarrow{\{1,4\}} \{\{1,4\},\{2,3\}\}.
\end{aligned}$$

Indeed partition $\{\{1,4\},\{2,3\}\}$ indirectly dominates $\{\{1,2\},\{3,4\}\}$. But, $\{\{1,4\},\{2,3\}\}$ is not the only final outcome once $\{1,3\}$ induces $\{\{2\},\{1,3\},\{4\}\}$.

[3] Throughout the paper, we refer the core to $\mathrm{Core}(\mathcal{P},>)$, i.e., the core as originally defined in the literature.

In fact, $\{\{1,2\},\{3,4\}\}$ *indirectly dominates* $\{\{2\},\{1,3\},\{4\}\}$ *through the following sequence:*

$$\{\{2\},\{1,3\},\{4\}\} \xrightarrow{\{1,2\}} \{\{1,2\},\{3\},\{4\}\} \xrightarrow{\{3,4\}} \{\{1,2\},\{3,4\}\}.$$

Thus, $\{\{1,2\},\{3,4\}\}$ *is conservatively stable since* $\{1,3\}$ *cannot deviate and guarantee higher payoffs for 2 and 3. In fact, both partitions* $\{\{1,4\},\{2,3\}\}$ *and* $\{\{1,2\},\{3,4\}\}$ *belong to the maximal conservative stable set. Note also that* $Core(\mathcal{P}, \gg) = \emptyset$ *since* $\{\{1,4\},\{2,3\}\}$ *is indirectly dominated by* $\{\{1,2\},\{3,4\}\}$ *via the following sequence:*

$$\begin{aligned}\{\{1,4\},\{2,3\}\} \quad & \xrightarrow{\{2,4\}} \{\{1\},\{2,4\},\{3\}\} \xrightarrow{\{1,2\}} \{\{1,2\},\{3\},\{4\}\} \\ & \xrightarrow{\{3,4\}} \{\{1,2\},\{3,4\}\}.\end{aligned}$$

The next example illustrates that conservative stable set can refine the core.

Example 3 (*Refining the Core*) *Consider a game with 4 players and the following preference ordering:*

$$\begin{aligned}\{1,2\} &\succ_1 \{1,3\} \succ_1 \{1,4\} \succ_1 \{1\} \succ_1 \cdots \\ \{2,4\} &\succ_2 \{1,2\} \sim_2 \{2,3\} \succ_2 \{2\} \succ_2 \cdots \\ \{1,3\} &\succ_3 \{3,4\} \sim_3 \{2,3\} \sim_3 \{3\} \succ_3 \cdots \\ \{3,4\} &\succ_4 \{2,4\} \succ_4 \{1,4\} \succ_4 \{4\} \succ_4 \cdots\end{aligned}$$

The rest of relevant coalitions are ranked strictly below the above coalitions. The core contains the two partitions

$$\{\{1,2\},\{3,4\}\} \text{ and } \{\{1,3\},\{2,4\}\}.$$

Given $\{\{1,2\},\{3,4\}\}$*, however, players 2 and 3 will form a coalition, expecting the following chain of events*

$$\begin{aligned}\{\{1,2\},\{3,4\}\} \quad & \xrightarrow{\{2,3\}} \{\{1\},\{2,3\},\{4\}\} \xrightarrow{\{1,3\}} \{\{1,3\},\{2\},\{4\}\} \\ & \xrightarrow{\{2,4\}} \{\{1,3\},\{2,4\}\}.\end{aligned}$$

In fact, the unique stable set contains only $\{\{1,3\},\{2,4\}\}$.

2.3 Notions of farsighted non-cooperative stability

Nash stability, individual stability, and contractual individual stability consider only one-step "deviations" initiated by individuals. We can introduce a farsighted stability notion as an alternative for each of these three concepts.

2.3.1 Farsighted Nash stability

To define farsighted Nash stability, we first formalize what each deviating player can achieve. If $i \in N$ leaves a coalition in P to join another coalition also in P or to stay alone, the resulting coalition structure, P', is defined as follows: (1) $S_P(i) \setminus \{i\} \in P'$ if $S_P(i) \setminus \{i\} \neq \emptyset$; (2) $T \cup \{i\} \in P'$ for some $T \in P \cup \{\emptyset\}$; (3) $S \in P'$ for all $S \in P$ such that $S \neq S_P(i)$ and $S \neq T$. We shall write $P \xrightarrow{\{i\}} P'$ in this case. Now, we can adapt the notions of farsighted stability in Section 2.1, observing that only individuals can move according to $\xrightarrow{\{i\}}$.

The following example depicts the myopia embedded in Nash stability.

Example 4 (*An undesired guest*) [4] *Consider a game with 3 players and the following preference orderings:*

$$\begin{aligned} \{1,2\} &\succ_1 \{1\} \succ_1 \{1,2,3\} \succ_1 \{1,3\} \\ \{1,2\} &\succ_2 \{2\} \succ_2 \{1,2,3\} \succ_2 \{2,3\} \\ \{1,2,3\} &\succ_3 \{2,3\} \succ_3 \{1,3\} \succ_3 \{3\} \end{aligned}$$

No partition is this game is Nash stable. For example, $\{\{1,2\},\{3\}\}$[5] *is not Nash stable because 3 prefers to join* $\{1,2\}$ *rather than stay alone. However,* $\{\{1,2,3\}\}$ *is not Nash stable either. In fact, players 1 and 2 will eventually bring back* $\{\{1,2\},\{3\}\}$*. If players are farsighted, then* $\{\{1,2\},\{3\}\}$ *is stable.*

2.3.2 Farsighted individual and contractual stability

Recall that in the previous subsection, player $i \in N$ can change $P \in \mathcal{P}$ to another $P' \in \mathcal{P}$ such that $P \xrightarrow{\{i\}} P'$. However, if player i needs the permission

[4] See Bogomolnaia and Jackson (2000).

[5] This partition is, however, core stable.

of $S_{P'}(i) \setminus \{i\}$ to join as is the case with individual stability, then i alone cannot change P to P' simply because $P \xrightarrow{\{i\}} P'$. We write $P \xrightarrow[S_{P'}(i)]{\{i\}} P'$ to denote that i needs the permission of $S_{P'}(i) \setminus \{i\}$ to change P to P'. Thus, P is individually stable if $P \xrightarrow[S_{P'}(i)]{\{i\}} P'$, $P' \succ_{\{i\}} P$, and, $P' \succsim_{S_{P'}(i)} P$. If players are farsighted, they should consider that P' may be replaced by another coalition structure and it is the "final" coalition structures that matter. This applies to both the migrating player i and the welcoming coalition $S_{P'}(i) \setminus \{i\}$. In particular, a coalition, in deciding whether to admit a new member, considers the final coalition structures as well. To capture the foresight of all the involved players, we define the following indirect dominance relation.

Indirect individual dominance For $P, P' \in \mathcal{P}$, P' indirectly individually dominates P, or $P' \gg P$, if there exists a sequence of coalition structures $P^1, P^2, \ldots, P^k$, with $P^1 = P$ and $P^k = P'$, and a sequence of individuals $i^1, i^2, \ldots, i^{k-1}$ such that $P^j \xrightarrow[S_{P^{j+1}}(i)]{\{i^j\}} P^{j+1}$ and $P' \succ_{i^j} P^j$ and $P' \succsim_{S_{P^{j+1}}(i)} P^j$ for all $j = 1, 2, \ldots, k-1$.

Individual conservative stable set (under foresight): $\mathcal{Q}$ is *conservatively internally stable* if $Q \in \mathcal{Q}$ implies that there do not exist $P \in \mathcal{P}$ and $i \in N$ such that $Q \xrightarrow[S_P(i)]{\{i\}} P$, $\mathcal{Q}|_{P,\gg} \neq \emptyset$, $P' \succ_i Q$, and $P' \succsim_{S_P(i)} Q$ for *all* $P' \in \mathcal{Q}|_{P,\gg}$. $\mathcal{Q}$ is *conservatively externally stable* if $Q \in \mathcal{P} \setminus \mathcal{Q}$ implies that there exist $P \in \mathcal{P}$ and $i \in N$ such that $Q \xrightarrow[S_P(i)]{\{i\}} P$, $\mathcal{Q}|_{P,\gg} \neq \emptyset$, $P' \succ_i Q$, and $P' \succsim_{S_P(i)} Q$ for *all* $P' \in \mathcal{Q}|_{P,\gg}$ $\mathcal{Q}$ is *conservatively stable* if it is both conservatively internally and externally stable.

Note that for $\mathcal{Q} \subset \mathcal{P}$ and $P \in \mathcal{P}$,

$$\mathcal{Q}|_{P,\gg} = \{P' \in \mathcal{Q} \mid P' = P \text{ or } P' \gg P\}.$$

We can also define a similar notion of individual contractual stable set under foresight by modifying the above definitions to capture the fact that a deviating player also needs the permission of the coalition from which he departs and this coalition is also farsighted. The following example illustrates the implications of foresight in extending individual stability and individual contractual stability.

Example 5 (*Switching partners*) *Consider a game with 4 players. The following is part of the preference ordering of the players. The rest of the relevant coalitions are ranked strictly below these coalitions.*

$$\begin{aligned} \{1,4\} &\succ_1 \{1,2\} \\ \{2,3\} &\succ_2 \{2,3,4\} \succ_2 \{1,2\} \\ \{2,3\} &\succ_3 \{3,4\} \\ \{1,4\} &\succ_4 \{3,4\} \end{aligned}$$

$\{\{1,2\},\{3,4\}\}$ *is individually stable and contractually individually stable: For example, if 2 leaves his current partner 1, then 1 will be worse off; moreover, if 1 joins the coalition of 3 and 4, both 3 and 4 will be worse off. If players are farsighted, however, they will anticipate that once 2 leaves 1 and joints* $\{3,4\}$*, 4 will subsequently leave 2 and 3 to join 4. That is,*

$$\{\{1,2\},\{3,4\}\} \xrightarrow[\{.\}]{\{2\}} \{\{1\},\{2,3,4\}\} \xrightarrow[\{.\}]{\{4\}} \{\{1,4\},\{2,3\}\}.$$

Thus, not only players 2 and 4 have incentive to initiate their moves but also other relevant coalitions have incentive to approve their moves. Note that the final coalition structure $\{\{1,4\},\{2,3\}\}$ *Pareto dominates* $\{\{1,2\},\{3,4\}\}$*.*

3 Results

In this section, we first establish the existence of the notions we propose and then proceed to analyze their properties.

3.1 Existence

The following theorem shows that the conservative stable set always exists; this is in contrast to the fact that the core may be empty and the vN-M stable set may fail to exist.

Theorem 1 *A hedonic game admits a conservative stable set and a maximal conservative stable set with respect to set inclusion.*

Proof. Let $\mathcal{Q}^0 = \mathcal{P}$ and for $k = 0, 1, \ldots$ define recursively,

$$\mathcal{Q}^{k+1} = \left\{ \begin{array}{c} Q \in \mathcal{Q}^k \mid \nexists P \in \mathcal{Q}^k \text{ and } S \subset N \text{ s.t. } Q \xrightarrow{S} P,\ \mathcal{Q}^k|_{P,\gg} \neq \emptyset, \\ \text{and } P' \succ_S Q\ \forall P' \in \mathcal{Q}^k|_{P,\gg} \end{array} \right\}.$$

We shall first show that there exits k^* such that $\mathcal{Q}^{k^*+1} = \mathcal{Q}^{k^*}$. Since $\mathcal{P}$ is finite, it is sufficient to show that for all k and $P \in \mathcal{P}$, there exist $P' \in \mathcal{Q}^k$ such that $\mathcal{Q}^k|_{P,\gg} \supset \mathcal{Q}^k|_{P',\gg} \ni P'$. Obviously, this is true for $k = 0$. Assume the claim is true for k. Let $P^1 \in \mathcal{Q}^k \setminus \mathcal{Q}^{k+1}$, then there exist $Q \in \mathcal{Q}^k$ and $S \subset N$ such that $P^1 \xrightarrow{S} Q$, $\mathcal{Q}^k|_{Q,\gg} \neq \emptyset$, and $P' \succ_S P^1$ for all $P' \in \mathcal{Q}^k|_{Q,\gg}$. Thus, $P' \gg P^1$ for all $P' \in \mathcal{Q}^k|_{Q,\gg}$. Given that $\gg$ is irreflexive, $\mathcal{Q}^k|_{P^1,\gg} \supsetneq \mathcal{Q}^k|_{Q,\gg}$. Since the claim is true for k, then there exists $P^2 \in \mathcal{Q}^k$ such that $\mathcal{Q}^k|_{Q,\gg} \supset \mathcal{Q}^k|_{P^2,\gg} \ni P^2$. Thus, we $\mathcal{Q}^k|_{P^1,\gg} \supsetneq \mathcal{Q}^k|_{P^2,\gg} \ni P^2$. Obviously, $\mathcal{Q}^{k+1}|_{P^1,\gg} \supset \mathcal{Q}^{k+1}|_{P^2,\gg}$. If $P^2 \in \mathcal{Q}^{k+1}|_{P^2,\gg}$, we are done. Otherwise, there is $P^3 \in \mathcal{Q}^k$ such that $\mathcal{Q}^k|_{P^2,\gg} \supsetneq \mathcal{Q}^k|_{P^3,\gg} \ni P^3$. Since $\mathcal{Q}^k$ is finite, there exist ℓ such that $\mathcal{Q}^k|_{P^1,\gg} \supset \mathcal{Q}^k|_{P^2,\gg} \supset \cdots \supset \mathcal{Q}^k|_{P^\ell,\gg} \ni P^\ell$ and $P^\ell \in \mathcal{Q}^{k+1}$.

$\mathcal{Q}^{k^*+1} = \mathcal{Q}^{k^*}$ implies that $\mathcal{Q}^{k^*}$ is conservatively internally stable. To show conservative external stability, suppose $P \in \mathcal{P} \setminus \mathcal{Q}^{k^*}$. Then there exists $k < k^*$ such that $P \in \mathcal{Q}^k \setminus \mathcal{Q}^{k+1}$. Thus, there exist $Q \in \mathcal{Q}^k$ and $S \subset N$ such that $P \xrightarrow{S} Q$, $\mathcal{Q}^k|_{Q,\gg} \neq \emptyset$,and $P' \succ_S P$ for all $P' \in \mathcal{Q}^k|_{Q,\gg}$. Since $\mathcal{Q}^k|_{Q,\gg} \supset \mathcal{Q}^{k^*}|_{Q,\gg} \neq \emptyset$, we have $P' \succ_S P$ for all $P' \in \mathcal{Q}^{k^*}|_{Q,\gg}$. Thus, $\mathcal{Q}^{k^*}$ is conservatively external stable.

To show that $\mathcal{Q}^{k^*}$ is the largest conservative stable set, assume to the contrary that there exists another conservative stable set $\mathcal{Q}'$ such that there exists $P \in \mathcal{Q}' \setminus \mathcal{Q}^{k^*}$. There exists k such that $\mathcal{Q}^k \supset \mathcal{Q}'$. Since $\mathcal{Q}'$ is stable, $\mathcal{Q}^\ell \supset \mathcal{Q}'$ for all $\ell \geq k$. Thus, $P \in \mathcal{Q}^{k^*}$, a contradiction. ■

Existence of other farsighted notions can be established in the same fashion.

3.2 Farsighted Coalitional Stability and the Core

As illustrated through Example 3, our notion of conservative stability refines the core. That is, some core partition may be ruled out from the maximal conservative stable set if a farsighted coalition can deviate is such a way that its deviation will lead to final partitions that benefit all its members.

However, as the following theorem indicates, when preferences are strict, core partitions are always contained in the maximal conservative stable set; therefore, the Core of hedonic games with strict preferences exhibits a surprising robustness. Intuitively, given strict preferences, whenever a coalition deviates from a core partition, at least one of the final partition that may result will not yield every member of this coalition a strictly higher payoff.

Theorem 2 *Consider a game G with strict preferences and let $\mathcal{Q}^*$ be the maximal conservative stable set. Then Core$(\mathcal{P}, >) \subset \mathcal{Q}^*$.*

Proof. Consider $\bar{\mathcal{Q}} = \{P\}$ such that $P \in$ Core$(\mathcal{P}, >)$. We will show that $\bar{\mathcal{Q}}$ is conservatively stable.

Internal stability is trivial. External stability can also be established with the direct use of Lemma 3 below, which asserts that any partition $P' \notin \bar{\mathcal{Q}}$ is indirectly dominated by P.

Finally, according to Theorem 1 the maximal conservative stable set $\mathcal{Q}^* \supset \bar{\mathcal{Q}}$ and therefore, $P \in \mathcal{Q}^*$. ■

Lemma 3 *Consider a game G with strict preferences and assume that the Core$(\mathcal{P}, >) \neq \emptyset$. If partition $P^* \in$ Core$(\mathcal{P}, >)$, then P^* indirectly dominates any partition $P \in \mathcal{P} \setminus \{P^*\}$.*

Proof. Let $P \in \mathcal{P} \setminus \{P^*\}$. We will show that $P^* \gg P$ by constructing a sequence of moves, starting with P and ending at P^*, where every acting coalition prefers P^* to its status quo. The sequence consists of two parts: in the first part P is decomposed into singletons and in the second part P^* is constructed from the singletons. We start with the first part. For convenience let S_P denote an arbitrary coalition in P.

Since $P^* \in$ Core$(\mathcal{P}, >)$, $\nexists\, S \in P$ such that $P \succ_S P^*$. Given strict preferences, for any S_P such that $|S_P| > 1$, there exists $i \in S_P$ such that $P^* \succ_i P$. Let i initiate the sequence by breaking away from S_P. That is, $T_1 = \{i\}$ induces partition P_1 such that $T_1 \in P_1$ and $(S_P \setminus T_1) \in P_1$.

Again, since $P^* \in$ Core$(\mathcal{P}, >)$, $\nexists\, S \in P_1$ such that $P_1 \succ_S P^*$. Any S_{P_1} with $|S_{P_1}| > 1$ contains j such that $P^* \succ_j P_1$. Let $T_2 = \{j\}$ induce partition

P_2 such that $T_1, T_2, (S_{P_1} \backslash T_2) \in P_2$. We can continue in this manner until we reach partition P_t such that every $|S_{P_t}| = 1$. Obviously, since $P^* \in$ Core$(\mathcal{P}, >)$, for every $\{i\} \in P_t \backslash P^*$, we have $P^* \succ_{\{i\}} P_t$.

Next we proceed with the second part of the sequence which involves constructing P^*. If P^* consists of singletons only, then $P_t = P^*$. Otherwise, at least one coalition in P^* has more than 1 member. Let $P^* = \{S_1, S_2, ..., S_m\}$. Assume, without of loss of generality, that $|S_j| > 1$ for all $j \leq k \leq m$ and $|S_j| = 1$ for all j such that $k < j \leq m$.

Let the next step of the sequence be the formation of S_1, since all of its members would rather be together (which is the case under P^*) than alone (which is the case under P_t). Denote the new partition by P_{t+1}.

Let the next step of the sequence be the formation of S_2. Again all of S_2's members would rather be together (which is the case under P^*) than alone, which is the case under P_{t+1}. Denote the new partition by P_{t+2}. Continue in this manner, until P^* is constructed. Note that along the way, when S_j forms from P_{t+j-1}, all of its members would rather be together (under P^*) than alone (under P_{t+j-1}).

To summarize,

$$\begin{aligned} \text{1st part} &: \quad P \overset{T_1}{\to} P_1 \overset{T_2}{\to} P_2 \overset{T_3}{\to} ... \overset{T_t}{\to} P_t \\ \text{2nd part} &: \quad P_t \overset{S_1}{\to} P_{t+1} \overset{S_2}{\to} P_{t+2} \overset{S_3}{\to} ... \overset{S_{k-1}}{\to} P_{t+k-1} \overset{S_k}{\to} P^*. \end{aligned}$$

While $P^* \succ_{T_i} P_{i-1}$ for $i = 1, ..., t$ where $P = P_0$ and $P^* \succ_{S_j} P_{t+j-1}$ for $j = 1, ..., k$. Thus, $P^* \gg P$. ■

As Example 3 indicates, strict preferences are indispensable for Theorem 2 to hold.

Remark 1 *It is easy to see that if* $|Core(\mathcal{P}, >)| > 1$ *then* $Core(\mathcal{P}, \gg)$ *is always empty, since any one core outcome indirectly dominates all the others. The converse, however, is not true. As is demonstrated by example 2,* $|Core(\mathcal{P}, >)| = 1$, *yet* $Core(\mathcal{P}, \gg) = \emptyset$ *since the unique core outcome is indirectly dominated by another (non-core) outcome and vice versa, ruling each other out.*

Banerjee, Konishi and Sönmez (1998) first introduce the *top-coalition property* as a relaxation of the *common ranking property* which is due to Farell and Scotchmer (1988). It requires that there exists a coalition that is preferred the most by its members compared to any other coalition each member could possibly join. Moreover, the condition holds for any subset of players.

Top-coalition property Given a non-empty set of players $V \subset N$, a coalition $S \subset V$ is a *top-coalition* of V if for every $i \in S$ and any $T(i) \subset V$ we have $S \succeq_i T(i)$. A hedonic game G satisfies the *top-coalition property* if for any non-empty set of players $V \subset N$, there exists a top-coalition of V.

The authors show that under the top-coalition property and strict preferences the Core always contains the unique top-coalition partition defined as follows.

Top-coalition partition Given a game G that satisfies the top-coalition property, a partition $P^* \in \mathcal{P}$ is a *top-coalition partition* if P^* is such that $P^* = \{S_1, S_2, ..., S_j, ...S_k\}$ where S_1 is a top-coalition of N, S_2 is a top-coalition of $N \backslash S_1$, S_3 is a top-coalition of $N \backslash \{S_1 \cup S_2\}$,..., S_k is a top-coalition of $N \backslash \{\cup_{j<k} S_j\}$.

We offer a result, tangent to the one by Banerjee, Konishi and Sönmez (1998), that when a game satisfies the top-coalition property and preferences are strict there exists a unique conservative stable set that contains only the top-coalition partition, yielding, thus, the exact same results as the Core does. Note that the unique conservative stable set is also the unique optimistic stable set[6].

[6] The authors also define a weaker notion of the top-coalition property, appropriately named *weak-top-coalition property* and they show that the Core is always non-empty under the weak-top-coalition property, yet they do not offer any characterization of the solution apart from the fact that it contains at least all the weak-top-coalition partitions. It is easy to construct examples where the Core along with the conservative stable set may contain partitions that are not weak-top-coalition.

Theorem 4 *Let a game G with strict preferences satisfy the top-coalition property. Then, it admits a unique conservative stable set $\mathcal{Q}^* = \{P^*\}$ where P^* is the top-coalition partition.*

Proof. We will first show that $\mathcal{Q}^*$ is indeed a stable set and then we will proceed to show uniqueness.

Stability

Internal stability is obvious since $\mathcal{Q}^*$ contains only one element, therefore there cannot be any contradictions.

External stability is satisfied with the direct use of lemma 3. Indeed every outcome $Q_0 \notin \mathcal{Q}^*$ is indirectly dominated by P^* since $P^* \in \text{Core}(\mathcal{P}, >)$.

Alternatively, the construction of the following sequence of coalition structures leading from any arbitrary coalition structure Q_0 to P^* is more intuitive due to the direct use of the top coalitions. In particular, the sequence proposed here omits the 1st part of the sequence proposed in Lemma 3, i.e., the dissolution of the starting partition into singletons. It starts immediately with the 2nd part, that is, the formation of the terminal sequence, P^*.

Step 1: Consider an arbitrary partition $Q_0 \neq P^* = \{S_1, \ldots, S_k\}$. If $S_1 \in Q_0$ proceed to step 2; otherwise let the first step of the sequence be the formation of S_1 since, by definition, $Q_0 \prec_{S_1} P^*$ no matter what Q_0 is, as long as $S_1 \notin Q_0$. Let the new coalition structure be denoted by Q_1, and note that $S_1 \in Q_1$.

Step 2: If $S_2 \in Q_1$ proceed to step 3, otherwise let the second step of the sequence be the formation of S_2, since by definition $Q_1 \prec_{S_2} P^*$. More specifically, we know that all the members of S_2 prefer to be in S_2 rather than in any other coalition involving members of $N \backslash S_1$, which is exactly their situation under Q_1, since S_1 has already formed. Let the new coalition structure be denoted by Q_2, and note that $S_1, S_2 \in Q_2$.

$$\vdots$$

Step k: If $S_k \in Q_{k-1}$, then Q_{k-1} is already P^*, otherwise let the last step of the sequence be the formation of S_k, since be definition $Q_{k-1} \prec_{S_k} P^*$. Note that if S_k is not formed then it must partitioned in smaller coalitions, whose members would rather merge since S_k is a top-coalition of $N \backslash \{\cup_{j<k} S_j\}$. Obviously, the coalition structure S_k induces from Q_{k-1} is exactly P^*.

To conclude, the sequence that leads from any arbitrary Q_0 to P^* is as follows:

$$Q_0 \overset{S_1}{\to} Q_1 \overset{S_2}{\to} Q_2 \overset{S_3}{\to} ...... \overset{S_{k-1}}{\to} Q_{k-1} \overset{S_k}{\to} P^*$$

where for $j = 1, 2, ..., k$ we have $Q_{j-1} \prec_{S_j} P^*$ as argued earlier.

Uniqueness

To prove that $\mathcal{Q}^*$ is the only stable set we will first establish that any stable set $\mathcal{Q} \neq \mathcal{Q}^*$ must contain P^* and then we will proceed by showing that no other partition can be in the stable set $\mathcal{Q}$.

Assume in negation that $P^* \notin \mathcal{Q}$, then according to external stability (at least) $\exists P \in \mathcal{Q}$ such that $P \gg P^*$. Let T be the first coalition that departs from P^* to indirectly induce P. Note that $T \cap S_1 = \emptyset$, since every member of S_1 prefers S_1 the most. Moreover, for every $P' \in \mathcal{P}(S_1)$, $\nexists P \in \mathcal{P}$ and $i \in S_1$ such that $P' \prec_i P$, which implies that members of S_1 will neither participate in T nor in any other coalition involved in inducing P from P^*. This in turn implies that $S_1 \in P$.

Next we shall argue that $T \cap S_2 = \emptyset$ as well. Suppose to the contrary that there exists $i \in T \cap S_2$ and thus there exists $S \in P$ such that $S_2 \prec_i S$. Then this contradicts the original assumption that S_2 is a top-coalition of $N \backslash S_1$ since $S_1 \in P$ and thus $S \subset N \backslash S_1$. Therefore, not only $T \cap S_2 = \emptyset$ but also members of S_2 will not participate, at any stage, of the inducement of P;hence $S_2 \in P$ as well.

In a similar manner we can show that S_k will not participate in the inducement process since it cannot prefer P to P^*, when P already contains $S_1, \ldots, S_{k-1}$; hence $T = \emptyset$, $P = P^*$, thereby contradicting that $P \gg P^*$.

Next we will show that any $Q \in \mathcal{P}$, $Q \neq P^*$ is dominated given $\mathcal{Q}$ and therefore cannot belong to $\mathcal{Q}$ regardless of the rest of the composition of $\mathcal{Q}$. Consider some arbitrary $Q \neq P^*$ such that $S_1 \notin Q$. If S_1 forms from Q it will induce some partition $Q' \in \mathcal{P}(S_1)$ such that $Q' \succ_{S_1} Q$. Moreover, no sequence departing from Q' will involve, at any stage, any members of S_1 as argued earlier. This implies that any partition that may arise from Q' will contain S_1. Recall that we have already shown that $\mathcal{Q}$ contains at least one such partition, P^*. Thus, any $Q \notin \mathcal{P}(S_1)$ is dominated since S_1 can induce partition Q', and whatever outcome arises from Q' is preferred by S_1 to Q.

We have shown that $\mathcal{Q} \subset \mathcal{P}(S_1)$. Now consider some arbitrary partition $Q' \neq P^*$ but $Q' \in \mathcal{P}(S_1)$ and $Q' \notin \mathcal{P}(S_2)$. If S_2 forms from within Q' it will induce some partition $Q'' \in \{\mathcal{P}(S_1) \cap \mathcal{P}(S_2)\}$. Once Q'' is the new status quo, no sequence departing from it and ending up in some element in $\mathcal{Q}$ can involve members of S_2 since no partition in $\mathcal{Q}$ is preferred by some members of S_2 to any partition that already contains S_2. Thus, Q' is dominated since S_2 can induce partition $Q'' \in \{\mathcal{P}(S_1) \cap \mathcal{P}(S_2)\}$, and whatever outcome arises from Q'' is preferred by S_2 to Q'.

Thus, $\mathcal{Q} \subset \{\mathcal{P}(S_1) \cap \mathcal{P}(S_2)\}$. We can proceed in the same manner to show that $\mathcal{Q} \subset \{\cap_{j \leq k} \mathcal{P}(S_j)\} = \mathcal{Q}^* = \{P^*\}$. ■

3.3 Examples on Non-Cooperative Farsighted Stability notions

Example 6 (*Anonymity and single-peakedness*) *Consider a 8-player game with anonymous and single-peaked preferences over the size of coalitions [due to Bogomolnaia and Jackson (2000)]. Players 1 through 4 have peaks at 2 and players 4 through 8 have peaks at 4. A strongly Pareto optimal and individually stable partition is such that coalition* $\{5, 6, 7, 8\}$ *forms and players 1 through 4 partition themselves into pairs. However, partition* $P = \{\{1, 5\}, \{2, 6\}, \{3, 7\}, \{4, 8\}\}$ *is also individually stable. Note that* P *is not strongly Pareto optimal (it is weakly Pareto dominated by any strongly Pareto optimal partition). However,* P *is not individually conservative stable under foresight*[7]*. The unique farsighted optimistic and conservative stable set is the set of strongly Pareto optimal partitions.*

The above result can be generalized as follows. Consider a game with anonymous and single-peaked preferences over the size of coalitions. Assume

[7] P is dominated indirectly by $\{\{1, 3\}, \{2, 4\}, \{5, 6, 7, 8\}\}$ through, for example, the following sequence of moves: 5 joins 3 and 7; 3 joins 1; 6 joins 5 and 7; 8 joins 5, 6, and 7; finally, 4 joins 2.

that for each size k, $\frac{m}{k}$ is an integer, where m is the number of players who have peaks at size k. Then, there exist coalition structures that reach every player's peak. This set of coalitions structures constitutes the unique farsighted stable set (under optimism or conservatism).

Nash stability is demanding: For a coalition partition to be Nash stable, it must be the case that no player has an incentive to join another coalition. It is easy to show that if preferences are strict, a Nash stable coalition structure is also conservatively stable under foresight. To show this, let P be a Nash stable partition. If some player $i \in N$ leaves $S_P(i)$ and joins another coalition, this new partition makes i strictly worse off. Thus, i has an incentive to come back to the $S_P(i)$, thereby inducing P. That is, i fears his own behavior in the future. To rectify this, we can either look at smaller conservative stable sets or modify the behavior of the players.

4 Conclusion

In this paper we proposed notions to amend the four existing solution concepts in the context of hedonic games. Our notions endow players with foresight, allowing them to contemplate over the ultimate results of their actions. Moreover, they capture credibility or consistency, in the sense that any "ultimate" outcome a player hopes for, or fears is indeed a likely outcome, i.e., it already belongs to the solution. We showed that all proposed concepts always exists, without any additional condition, apart from those implies by the model itself. We have also provided formal relations between the new and the old solution concepts under certain restrictions.

What we would like to address in this section is the possibility for further extensions of the solution concepts and the significance of such extensions. The four different approaches, core, Nash stability, individual stability and contractual individual stability are not directly compared to each other (at least in this work) in terms of performance, but are considered distinctly for diverse institutional settings. In particular, the core and contractual individual stability are addressing two different institutional settings, where the remaining coalition may or may not have the power to prevent its members from leaving. Moreover, in the former a coalition can form and orchestrate a deviation whereas in the latter only individuals can deviate. One may consider of analyzing the situation where coalitions can form but in order to do so they require permission of the coalitions they depart from. Analogously, one may extend coalitional deviations within the Nash stability context, that a group of players may form from several coalitions and join some other coalition, even if this other coalition does not enjoy the new group of immigrants. Lastly, coalitional deviations can be incorporated in the individual stability notion (although the name is not appropriate anymore), by allowing a coalition to form and join some other coalition if the forming coalition is strictly better off while the welcoming coalition is not hurt.

In this paper we adopted certain implicit assumption in the possibilities of coalition in terms of what it can induce. In particular, we assumed that once the coalition forms the rest of the players are going to stay put, i.e., the coalitions that were not affected by the formation of the new coalition remain intact, and the ones that were affected remain as new residual coalitions. In addition, once the coalition forms it stays formed until the next move. There are two elements in this process that can be augmented. Firstly, the corrupted coalitions that are assumed to remain as residual coalitions may, as suggested in Hart and Kurz (1983), completely dissolve into singletons. Secondly, a forming coalition S may not stay together but simply restructure its members in an even better partitioning of S, treating thus the concept of formation as coordination among members of S.

Finally, properties of fairness, envy-freeness and population monotonicity are of great importance as is already discussed in Bogomolnaia and Jackson (2000), and naturally the viability of these properties extends to the solution concepts we propose in this work.

5 References

1. Banerjee, S., Konishi, H. & Sonmez, T. (1998). "Core in a Simple Coalition Formation Game," forthcoming in *Social Choice and Welfare.*

2. Bogomolnaia, A. & Jackson M.O. (2000). "The Stability of Hedonic Coalition Structures," mimeo.

3. Chwe, M. S.-Y. (1994). "Farsighted Coalitional Stability," *Journal of Economic Theory* **63**, 299-325.

4. Demange, G. (1982). "Single-Peaked Orders on a Tree," *Mathematical Social Sciences* **3**, 389-396.

5. Demange, G. (1994). "Intermediate Preferences and Stable Coalition Structures," *Journal of Mathematical Economics* **23**, 45-58.

6. Dreze, J. & Greenberg, J. (1980). "Hedonic Coalitions: Optimality and Stability," *Econometrica* **48**, 987-1003.

7. Farrell, J. & Scotchmer, S. (1988). "Partnerships," *Quarterly Journal of Economics*, 279-297.

8. Greenberg, J. (1990). *The Theory of Social Situations: An Alternative Game-Theoretic Approach.* Cambridge University Press.

9. Greenberg, J. (1994). "Coalition Structures," in *Handbook of Game Theory,* (R. Aumann and S. Hart, Eds.) Amsterdam: Elsevier.

10. Greenberg, J. & Weber, S. (1993). "Stable Coalition Structures with Unidimensional Set of Alternatives," *Journal of Economic Theory* **6**, 62-82.

11. Harsanyi, J.C. (1974). "Interpretation of Stable Sets and a Proposed Alternative Definition," *Management Science* **20**, 1472-1495.

12. Hart, S. & Kurz, M. (1983). "Endogenous Formation of Coalitions," *Econometrica* **51**, 1047-1064.

13. Xue, L. (1998). "Coalitional Stability under Perfect Foresight," *Economic Theory* ***11*** *603-627.*

14. von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior.* Princeton University Press.

Working Paper

1999-28	Torben M. Andersen and Niels C. Beier: Noisy Financial Signals and Persistent Effects of Nominal Shocks in Open Economies.
1999-29	Michael Jansson: Consistent Covariance Matrix Estimation for Linear Processes.
2000-1	Niels Haldrup and Peter Lildholdt: On the Robustness of Unit Root Tests in the Presence of Double Unit Roots.
2000-2	Niels Haldrup and Peter Lildholdt: Local Power Functions of Tests for Double Unit Roots.
2000-3	Jamsheed Shorish: Quasi-Static Macroeconomic Systems.
2000-4	Licun Xue: A Notion of Consistent Rationalizability - Between Weak and Pearce's Extensive Form Rationalizability.
2000-5	Ebbe Yndgaard: Labour, An Equivocal Concept for Economic Analyses.
2000-6	Graham Elliott and Michael Jansson: Testing for Unit Roots with Stationary Covariates.
2000-7	Nikolaj Malchow-Møller and Bo Jellesmark Thorsen: A Dynamic Agricultural Household Model with Uncertain Income and Irreversible and Indivisible Investments under Credit Constraints.
2000-8	Niels Haldrup, Antonio Montanés and Andreu Sanso: Measurement Errors and Outliers in Seasonal Unit Root Testing.
2000-9	Erik Harsaae: En kritisk vurdering af den generelle ligevægtsmodel.
2000-10	Rasmus Højbjerg Jacobsen: Why the ECB Should be Ultra-Liberal.
2000-11	Michael Rosholm and Michael Svarer: Structurally Dependent Competing Risks.
2000-12	Effrosyni Diamantoudi and Licun Xue: Farsighted Stability in Hedonic Games.

CENTRE FOR DYNAMIC MODELLING IN ECONOMICS

DEPARTMENT OF ECONOMICS - UNIVERSITY OF AARHUS - DK - 8000 AARHUS C - DENMARK
☎ +45 89 42 11 33 - TELEFAX +45 86 13 63 34

Working papers, issued by the Centre for Dynamic Modelling in Economics:

1999-11 Martin Paldam: The Big Pattern of Corruption. Economics, Culture and the Seesaw Dynamics.

1999-21 Martin Paldam: Corruption and Religion. Adding to the Economic Model?

2000-1 Niels Haldrup and Peter Lildholdt: On the Robustness of Unit Root Tests in the Presence of Double Unit Roots.

2000-2 Niels Haldrup and Peter Lildholdt: Local Power Functions of Tests for Double Unit Roots.

2000-3 Jamsheed Shorish: Quasi-Static Macroeconomic Systems.

2000-4 Licun Xue: A Notion of Consistent Rationalizability - Between Weak and Pearce's Extensive Form Rationalizability.

2000-6 Graham Elliott and Michael Jansson: Testing for Unit Roots with Stationary Covariates.

2000-8 Niels Haldrup, Antonio Montanés and Andreu Sanso: Measurement Errors and Outliers in Seasonal Unit Root Testing.

2000-12 Effrosyni Diamantoudi and Licun Xue: Farsighted Stability in Hedonic Games.