Endogenous Preference and Coalition Formation

By Roger A, McCain
Professor, School of Economics, Drexel University
Email: mccainra@drexel.edu
Telephone: office 215 895 2176 cell 610 453 4648

In most of economic theory, and in game theory, preferences are given as data. Thus, on the one hand, preferences are exogenous determinants of equilibria, coalition structures, and similar stable social situations; and on the other hand, preferences provide exogenous criteria for normative evaluations of alternative social situations, both on standards of efficiency and fairness. While this is highly coherent and coherence is to be valued in theory, the assumption that preferences are exogenous is empirically doubtful. With respect to income, in particular, surveys indicate the satisfaction of a higher income decreases as the person becomes accustomed to it (Frey and Stutzer 2002). Similarly, adaptation is observed to disability, marriage (Kahneman and Krueger 2006), and seatbelt regulations (Goudiet et al. 2014). Supposing that self-reported well-being is correlated with preference, these results pose troubling questions about the exogeneity of preferences. This is an important consideration for policy. For example, after a policy transition that makes seat belts mandatory, the agent's preferences may adapt so that the agent is less averse to seat belt use than they were before. It is also important for social status. As Sen observes, adaptation may lead to an underestimate of the economic inequality of a society. As Sen writes (1997 p. 391) "Deprived people … come to terms with their deprivation, and the psychological indicators of pleasure or desire fulfillment may fail to reflect the extent of real deprivation that these people suffer."

The essay begins with a mostly ordinary-language sketch of the theory of cooperative games in effectivity function form, a theory that focuses on the ability of a coalition to influence the range of possible outcomes in an interaction. Some pragmatically important shortcomings of the theory are discussed, but for a best-case analysis, these are set aside. The paper then explores the implications of endogenous preferences, relying mainly on small-scale examples.

1. A Brief Primer of Effectivity Analysis

While some aspects of game theory may be well known, this essay will begin with fundamental ideas for the sake of consistency and contrast. Schelling (1960) and Aumann (2003), who shared the 2005 Nobel Memorial Prize in economics for their work in game theory have both said that the what we call “game theory” is better described as interactive decision theory. Accordingly, a game theory model will begin with a list of decision-makers whose decisions will interact. This list will conventionally be treated as an indexed set that we may denote by ℕ. In general the index $i \leq N$ will denote an individual decision-maker or agent, but in some three-person examples below the members of ℕ may be denoted as A, B, and C. We are assuming that ℕ is finite with just N members but for some models ℕ might be a continuum. For each agent there is a set of courses of action that the agent may select among. These courses of action are conventionally called “strategies.” We may denote the set of strategies available to i as $\Sigma_i = \{\mathrm{S}_{\mathrm{i},1}, \dots, \mathrm{S}_{\mathrm{i},\mathrm{j}}, \dots\}$, so that $\mathrm{S}_{\mathrm{i,j}}$ is one of the courses of action available to i. For some applications such as “mixed strategies,” Σ_i may be a continuum, but for present purposes we will think of Σ_i as finite.

Now, decisions have consequences, and in many game theory applications the consequence is a set of numerical payoffs to the decision-makers. The payoffs may be numbers of money or utility. However, the outcome of the decisions of the N agents may be a more complex object defined in qualitative terms. Indeed in the famous “Prisoner’s Dilemma,” the outcomes are prison sentences or freedom, and the numbers are really incidental to the analysis. For our purposes the set of outcomes will be denoted by Ω and the individual members of the set, the particular outcomes of particular sets of strategies, as ω, θ, ϕ.

Outcomes, then, are determined by the courses of action that the agents decide to make, in such a way that the outcome depends on all decisions of the various agents. This is the meaning of "interactive decision theory." This dependence of the outcome on strategies is expressed by an *outcome function.* Taking the actions selected by the N decision-makers, $\{S_{1j}, \ldots, S_{i,k}, \ldots, S_{N,l}\}$ as the independent variable, we write $\omega = F(S) = F(S_{1j}, \ldots, S_{i,k}, \ldots, S_{N,l})$. To support a "rational" decision in some sense we require, in addition, some evaluation of the outcomes; moreover, we must allow for the possibility that the different decision-makers may disagree on the evaluations. The most general approach borrows from economic theory the concept of preference. If ω and ϕ are outcomes and agent i strictly prefers ω to ϕ, we may write $\omega \mathcal{P}_i \phi$. If the outcomes are vectors of numerical payoffs to the agents, as in most game theory models, we suppose that larger numbers are preferred to smaller (except in the Prisoner's Dilemma).

These elements – often simplified in some ways – are common to the research literatures of game theory. That literature, however, has some major divisions. The most important is between noncooperative and cooperative game theory. For noncooperative game theory, decision-maker i considers only the outcomes $\omega = F\{\overline{S_{1j}}, \ldots, S_{i,k}, \ldots, \overline{S_{N,l}}\}$, taking the strategies of the other agents as given and attempting to choose the "best response" to their strategies in terms of i's own preferences. In that sense there is no coordination of strategies in noncooperative game theory[1]. In cooperative game theory a group of agents $c = \{\iota, \ldots\ \kappa\} \subseteq \mathbb{N}$ may form a coalition and coordinate their strategy decisions to obtain an outcome that is in their mutual interest. This

[1] In many game theory courses, including some at the doctoral level, only noncooperative game theory is taught.

mutual interest may come at the expense of the nonmembers of the coalition, but in some games the grand coalition of all of the agents in the game may be able to do still better.

Within cooperative game theory there is a further division between transferable utility (TU) and nontransferable utility (NTU) games. For TU games it is assumed 1) that the outcome is a set of numerical payoffs, 2) that the payoffs can be costlessly and reliably transferred from one member of the coalition to another, and 3) larger numerical outcomes, including or net of transfers, are preferred to smaller numbers. Within the literature of NTU games some contributions begin by addressing the question, “What outcomes can this coalition bring about? For what outcomes is it *effective*?” This effectivity theory is the starting point of the present essay. Here is a simple example to illustrate the further development of the theory: the Rendezvous Game.

2. The Rendezvous Game

Agents A and B would like to get together. They can choose between two destinations for their rendezvous: North or South. To do that they must, at least, choose the same destination. But this is a cooperative game: they can form a coalition and jointly decide on one destination or the other. Thus, the *outcomes* for this game are *rendezvous* and *no rendezvous*. However, Agent C wants to prevent them from getting together. Perhaps C suspects that they will act jointly to do C some harm, if they get together; but we don’t need to know the answer to that question – all we need to know is that while A and B prefer rendezvous to no rendezvous, C prefers no rendezvous to rendezvous. To prevent the rendezvous C can block either North or South, but not both. Thus, each agent chooses between strategies “North” and “South,” and the relation of outcomes to strategies is shown in Table 1.

Table 1. Strategies and Outcomes in the Rendezvous Game

		C			
		North		South	
		B		B	
		North	South	North	South
A	North	no rendezvous	no rendezvous	rendezvous	no rendezvous
	South	no rendezvous	rendezvous	no rendezvous	no rendezvous

At the next stage we pose the question, for what outcomes is a coalition *effective*? In general, the grand coalition of all agents is effective for any outcome, and in this game there are six ways that the grand coalition can bring about no rendezvous and two ways that it can bring about rendezvous. But neither of these outcomes is in the mutual interest of all agents. We want to ask, what outcomes can coalitions with one or two members bring about? There is a class of games, *simple* games, for which that question has an unambiguous answer. But the Rendezvous Game is not a simple game. For games that are not simple, we may find there are two kinds of effectivity: we say that the coalition M={i, j, …, k}is α-effective for ω if there are strategies that will result in ω regardless of the strategies chosen by players who are not members of the coalition, and that the coalition M is β-effective for ω if, given the strategies chosen by nonmembers of the coalition, there are strategies that the coalition may choose that will result in ω. This is a difference in information. We are saying that, for β-effectivity, the coalition needs to know what strategies the nonmembers will choose, but for α-effectivity, they do not need to know that since they have a strategy combination that will yield ω regardless what strategies nonmembers choose.

The Rendezvous game illustrates the distinction of α-effectivity from β-effectivity. First we ask whether the coalition of A and B is α-effective for rendezvous. The answer is no. Recall, the coalition M is α-effective for outcome ω if there is a strategy that assures ω no matter which strategies are chosen by agents not in c. Whichever destination the coalition of A and B chooses, C can block that destination by choosing the same strategy. But, conversely, {C} is not α-effective for no rendezvous, since whichever destination {C} blocks leaves it open to {A,B} to choose the other destination and get together.

Instead, we might ask whether {A,B} is β-effective for rendezvous. The answer is yes. Recall, a coalition M is β-effective for ω if, given any strategies non-members of M may choose, there are strategies that M may choose that assure ω. Thus, if {C} chooses North, {A,B} can choose South, and vice versa. But in a similar way, {C} is β-effective for no rendezvous. Here is a qualification: if we interpret β-effectivity as an information condition, the conditions for effectivity of {A,B} is that {A,B} has the advantage of going second, while the condition for effectivity of {C} is that {C} has that advantage, and both of these cannot be true. However, what we can say is that there is no coalition structure *and specific set of strategies* that will be stable in the sense that no coalition or potential new coalition $M' \neq M$ can improve on them by shifting to an outcome that is preferred by all members of the coalition.

What coalitions and outcomes can be thought of as *stable*? As in much of the literature of TU games, the concept of the core is applied to answer this question. Consider first simple games, so that we need not distinguish α and β effectivity. Suppose that a coalition M is effective for the outcome ω. Suppose, however, that there is a coalition M' and M' is effective for outcome θ; and further, for some agent ι a member of M' we have $\theta\mathcal{P}_\iota\omega$ while there is no member γ of M' for whom $\omega\mathcal{P}_\gamma\theta$. That is, we are supposing that for the members of c*, θ is Pareto-preferable to ω. Then it is in the interest of M' to bring about the shift from ω to θ. We say that ω is dominated θ via M'. The set of outcomes for which some coalition is effective, and which are undominated, comprise the *core* of a simple NTU game. The core may be thought of as a concept of stability, in that the members of the core are stable against deviations to benefit other coalitions. The core may have no members, so that we say that the core is null, or has one or many members. To find a unique member of the core is, of course, a particularly satisfactory result.

But, again, for nonsimple games effectivity can be construed in two ways, α and β effectivity. Thus, given M and ω, if there is no coalition M' and θ with M' α-effective for θ and θ Pareto-preferable to ω for the members of c*, then ω is α-undominated and is a member of the α-core of the game. Substituting β for α in this statement we characterize β-undomination and the β-core of the game. Notice that if an outcome is α-dominated it is certainly β-dominated, but may be β-dominated although not α-dominated, so that the α-core may be larger than the β-core.

Now we ask what outcomes are in the α-core of the rendezvous game. We have seen that neither {A,B} nor {C} can assure the outcome they prefer. Neither the singletons {A} and {B} can do so either, and two-agent coalitions {A, C} and {B,C} cannot agree on a preferred outcome. The grand coalition {A, B, C} is (always) effective for any outcome. Neither {A,B} nor {C} is effective to veto the outcome they do not like, so neither rendezvous or no rendezvous can be excluded from the α-core of the rendezvous game. What about the β-core? We have seen that {A,B} is β-effective for rendezvous and both prefer it to no rendezvous so no rendezvloous is β-dominated, excluded from the β–core. But, equally, {C} is effective for no rendezvous and prefers it. Thus, either outcome can be vetoed by a (singleton or two-person) coalition that unanimously prefers the other. The β-core of this game is a null set.

3. Endogenous Preferences, Farsighted and Shortsighted Decisions

Thus far we have followed the conventional assumption of economics and game theory that the outcome of the game and the preferences of agents are two quite different things, and the preferences are treated as data. To say that preferences are endogenous is, however, to say that the preference system is one of the outcomes. That is, among the elements of an outcome ω is a set of preferences over outcomes, $\mathcal{P}_{i,\omega}$, where the preferences are those of agent i realized in outcome ω. As a first step, we may consider the decisions of an individual decision-maker, without considering interactive decisions. Letting the outcomes be ω and θ, there are four cases:

1. $\omega\mathcal{P}_{i,\omega}\theta$, $\omega\mathcal{P}_{i,\theta}\theta$
2. $\theta\mathcal{P}_{i,\omega}\omega$, $\theta\mathcal{P}_{i,\theta}\omega$
3. $\omega\mathcal{P}_{i,\omega}\theta$, $\theta\mathcal{P}_{i,\theta}\omega$
4. $\theta\mathcal{P}_{i,\omega}\omega$, $\omega\mathcal{P}_{i,\theta}\theta$

The first two are unproblematic. The last could give rise to cycles. Case 3 would not give rise to cycles but would seem to extend the set of stable outcomes and might be quite common as a result of adaptation of preferences to existing circumstances. However, 3 and 4 raise a distinct question: can the decision-maker be foresighted about these changes of preferences? We could say that the decision is farsighted if it is made on the basis of the new preference system that is realized in the new outcome the deviating coalition brings about. If, on the one hand, the decision-maker is foresightful in case 4, then it seems the cycle disappears: agent i reflects that, even though θ seems preferable, the switch to θ would not actually improve agent i's situation, and thus continue to decide for ω. In case 3, however, foresight deepens the mystery. Agent i has no motive to change the outcome, but nevertheless anticipates that they would be pleased if they did. It seems that in this case, no rational choice may be made simply on the basis of the preference systems $\mathcal{P}_{i,\omega}$ and $\mathcal{P}_{i,\theta}$. One possibility might be a system of preferences over preference systems (David George 2003). Another possibility could be an appeal to other values than preferences.

The dichotomy of shortsighted and farsighted decisions further complicates the categories of coalitional stability. First, for agent i and outcomes ω, θ and ϕ, we denote by $\theta \mathcal{P}_{i\omega} \phi$ a case in which agent i prefers outcome θ to outcome ϕ when outcome ω is realized. Then given coalition M and outcomes ω and ϕ, with M effective for ω, we may say

1a) if there is a coalition M' that is α-effective for θ, and for some member i of coalition M' $\theta \mathcal{P}_{i\omega} \omega$ but there is no member i of M' for whom $\omega \mathcal{P}_{i\omega} \theta$ then M and ω are shortsighted α-dominated by θ via c*.

1b) The shortsighted α-core of the game comprises all outcomes for which some coalition is effective and which are not shortsighted α-dominated.

2a) if there is a coalition M' that is α-effective for θ, and for some member i of coalition M' $\theta \mathcal{P}_{i\theta} \omega$ but there is no member i of M' for whom $\omega \mathcal{P}_{i\theta} \theta$ then M and ω are farsighted α-dominated by θ via c*.

2b) The farsighted α-core of the game comprises all outcomes for which some coalition is effective and which are not farsighted α-dominated.

3a) if there is a coalition M' that is β-effective for θ, and for some member i of coalition M' $\theta \mathcal{P}_{i\omega} \omega$ but there is no member i of M' for whom $\omega \mathcal{P}_{i\omega} \theta$ then M and ω are shortsighted β-dominated by θ via c*.

3b) The shortsighted β-core of the game comprises all outcomes for which some coalition is effective and which are not shortsighted α-dominated.

4a) if there is a coalition M' that is β-effective for θ, and for some member i of coalition M' $\theta \mathcal{P}_{i\theta} \omega$ but there is no member i of M' for whom $\omega \mathcal{P}_{i\theta} \theta$ then M and ω are farsighted β-dominated by θ via c*.

4b) The farsighted β-core of the game comprises all outcomes for which some coalition is effective and which are not farsighted β-dominated.

4. Examples with Endogenous Preferences

This section will give three simple examples of coalition formation with endogenous preferences.

i

The first example may be thought of it as a game of education in economic development. For simplicity of presentation we consider a three-person game, although a more realistic example would presumably involve much larger numbers. As before the three agents are A, B, and C. Their interaction can give rise to two outcomes denoted as a backward and an advanced state. In the backward state B and C are very poor and A is moderately poor. A's advantage in the backward state derives from skills A has and that the others lack. In the advanced state, all possess skills and the complementarity of their skilled efforts increases the productivity of each. However, one strategy available to A is to teach the skills to B and C. A's strategies are to teach or not to teach, while the strategies of B and C are to study or not to study. If the strategies are teach, study, study, then the outcome is the advanced state; otherwise it is the backward state. Thus, in the advanced state, the material welfare of A is improved, but the material welfare of B and C are enhanced by a greater margin so that A, while better off than B and C, exceeds them by a smaller margin. The outcome function is shown in Table 2 and the preferences in Table 3.

Table 2. Education Game

		C			
		Study		Don’t	
		B		B	
		Study	Don’t	Study	Don’t
A	Teach	Advanced	Backward	Backward	Backward
	Don’t	Backward	Backward	Backward	Backward

Table 3. Preferences in the Education Game

	Backward	Advanced
A	Advanced$\mathcal{P}$Backward	Backward$\mathcal{P}$Advanced
B	Advanced$\mathcal{P}$Backward	Advanced$\mathcal{P}$Backward
C	Advanced$\mathcal{P}$Backward	Advanced$\mathcal{P}$Backward

The preferences are explained as follows: in the backward state, because of their material deprivation, the agents prefer an outcome that leads to greater absolute wealth and thus alleviates their poverty. In the advanced state, having secure material subsistence, the agents prefer an outcome that enhances their relative status. For B and C the advanced state is preferable on both grounds, but for A, the backward state is preferable for the greater margin of A’s well-being over that of B and C.

There are seven possible non-null coalitions: the grand coalition and three two-person and three singleton coalitions. Only the grand coalition is effective for advanced, while every non-null coalition is effective for backward. In each case, the effectivity is both α and β effectivity, that is, this is a simple game. Then we see that advanced *shortsighted dominates* backward via the grand coalition, while backward *shortsighted dominates* advanced via {A}. Thus the shortsighted core of this game is empty – there is no stable coalition structure or outcome in the case of shortsighted decisions. This is the case of cycles. However, with farsighted decisions, A will veto a move by the grand coalition to realize advanced, so that backward is not dominated by advanced via the grand coalition. Perhaps this is "conservatism." However, if advanced is realized, A will oppose a movement to backward by any coalition of which A is a member, and no coalition without A will choose the outcome backward in any case. Thus advanced as an outcome realized by the grand coalition will not be dominated by backward via any coalition and is also a member of the farsighted core of the game. With farsight, in this case, both outcomes are stable.

ii

For the next example, once again, a three-person game will be used for expositional simplicity, although once again larger numbers would be more realistic. In outcomes ω and θ, the three agents are subject to a social role such that some economic activities are considered inappropriate for them. This exclusion is not absolute. The two strategies available to them are "seek" (opportunities considered inappropriate) or "don't." If an agent does seek opportunities they will be successful and improve their material conditions but this improvement of the material circumstances may not be enough to offset the disadvantages resulting from taking an "inappropriate" role, so that the outcome of "don't" may be preferred to the outcome of "seek." On the other hand, if a sufficient number of agents seek these opportunities, the norm becomes incredible, and the disadvantages of "inappropriate" roles disappear. For this example, a sufficient number is all three. The game is shown in Table 4 and the preferences in Table 5.

Table 4. Opportunity Game

		C			
		Seek		Seek	
		B		B	
		Seek	Don’t	Seek	Don’t
A	Seek	ϕ	θ	θ	θ
	Don’t	θ	θ	θ	ω

Table 5. Preferences in the Opportunity Game

	ω	θ	ϕ
A	$\omega\mathcal{P}_{A\omega}\theta$, $\omega\mathcal{P}_{A\omega}\phi$, $\theta\mathcal{P}_{A\omega}\phi$	$\omega\mathcal{P}_{A\theta}\theta$, $\phi\mathcal{P}_{A\theta}\theta$, $\phi\mathcal{P}_{A\theta}\omega$	$\phi\mathcal{P}_{A\phi}\theta$, $\phi\mathcal{P}_{A\phi}\omega$, $\theta\mathcal{P}_{A\phi}\omega$
B	$\omega\mathcal{P}_{B\omega}\theta$, $\omega\mathcal{P}_{B\omega}\phi$, $\theta\mathcal{P}_{B\omega}\phi$	$\omega\mathcal{P}_{B\theta}\theta$, $\phi\mathcal{P}_{B\theta}\theta$, $\phi\mathcal{P}_{B\theta}\omega$	$\phi\mathcal{P}_{A\phi}\theta$, $\phi\mathcal{P}_{A\phi}\omega$, $\theta\mathcal{P}_{A\phi}\omega$
C	$\omega\mathcal{P}_{C\omega}\theta$, $\omega\mathcal{P}_{C\omega}\phi$, $\theta\mathcal{P}_{C\omega}\phi$	$\omega\mathcal{P}_{C\theta}\theta$, $\phi\mathcal{P}_{C\theta}\theta$, $\phi\mathcal{P}_{C\theta}\omega$	$\phi\mathcal{P}_{C\phi}\theta$, $\phi\mathcal{P}_{C\phi}\omega$, $\theta\mathcal{P}_{C\phi}\omega$

Considering shortsighted decisions, both ω and ϕ are α-stable. Note that while any coalition is β-effective for θ, θ is never preferred to ω, so θ never dominates ω via any coalition. The grand coalition is effective for ω, so from ω both other outcomes are dominated via the grand coalition. However, from ϕ, ϕ dominates all other outcomes. From θ, both ϕ and ω are preferred to θ, but ϕ is preferred by all agents to ω, so ϕ dominates θ and ω via the grand coalition. Thus θ is unstable.

Considering farsighted decisions (as we have defined farsight above) we find that ϕ and ω dominate one another via the grand coalition. With farsight, in this case, we have a cycle! But this is a limited sort of farsight. (In what follows we will call this preference foresight.) Some farsight is required to anticipate what one’s preferences would be in another outcome, but is it indeed reasonable to choose according to them? In this case a farsighted agent might (1) anticipate the cycle and thus not choose an outcome that would lead to a cycle, but apply some other criterion, or indeed (2) from ω, the farsighted person might anticipate that at θ, $\phi\mathcal{P}_{i\theta}\theta$, $\phi\mathcal{P}_{i\theta}\omega$, so that a transition to θ would be followed by a transition to ω. (We will call this systematic farsight.) Let us modify the game by substituting only $\phi\mathcal{P}_{C\omega}\omega$; so that C is an exception who prefers a transition from ω to ϕ. Since the singleton coalition is β-effective for θ, with systematic farsight, C might choose "seek" to bring about an indirect transition to ϕ. This strategy might be called "creating a role model."

iii

An issue for which endogenous preferences are particularly plausible is interregional mobility. It is plausible that people prefer to remain where they are, but if they move, once they become accustomed to the new home prefer it. This suggests the following example. Again it will be a three-person game, for simplicity. Strategies are to live in a village or the city. For those who live in a city, efforts are complementary, leading to increased productivity. The game is shown in Table 6, with outcomes described in Table 7. Note that outcomes differ not only in productivity but also in the distribution of the population: from Table 6, for example, in ω_2 A and C reside in the city but B in a village.

Table 6. Mobility Game

		C			
		City		Village	
		B		B	
		City	Village	City	Village
A	City	ω_1	ω_2	ω_5	ω_6
	Village	ω_3	ω_4	ω_7	ω_8

Table 7. Productivity in Various Outcomes

ω_1	High productivity in city.
ω_2, ω_3, ω_5	Moderate productivity in city, low productivity in village.
ω_4, ω_6, ω_7, ω_8	Low productivity everywhere.

In general, we may suppose that people prefer to remain where they are. The difference in productivity may offset this preference, if material well-being is great enough in a high-productivity state to compensate for the sacrifice of changing one’s location. Suppose in particular that the high productivity of the 3-person city yields productivity enough to compensate for moving from a village, but the moderate productivity of the two-person city is not; nor is the margin of productivity of the three-person city over the two-person city sufficient to compensate for moving. Then the first-preferences for the three agents and all outcomes are shown in Table 8.

Table 8. First Preferences

	A	B	C
ω_1	$\omega_1 \mathcal{P}_A \omega_j$	$\omega_1 \mathcal{P}_B \omega_j$	$\omega_1 \mathcal{P}_C \omega_j$
ω_2	$\omega_1 \mathcal{P}_A \omega_j$	$\omega_2 \mathcal{P}_B \omega_k$	$\omega_1 \mathcal{P}_C \omega_j$
ω_3	$\omega_3 \mathcal{P}_A \omega_j$	$\omega_1 \mathcal{P}_B \omega_j$	$\omega_1 \mathcal{P}_C \omega_j$
ω_4	$\omega_4 \mathcal{P}_A \omega_k$	$\omega_4 \mathcal{P}_B \omega_k$	$\omega_1 \mathcal{P}_C \omega_j$
ω_5	$\omega_1 \mathcal{P}_A \omega_j$	$\omega_1 \mathcal{P}_B \omega_j$	$\omega_5 \mathcal{P}_C \omega_k$
ω_6	$\omega_1 \mathcal{P}_A \omega_j$	$\omega_6 \mathcal{P}_B \omega_k$	$\omega_6 \mathcal{P}_C \omega_k$
ω_7	$\omega_7 \mathcal{P}_A \omega_k$	$\omega_1 \mathcal{P}_B \omega_j$	$\omega_7 \mathcal{P}_C \omega_k$
ω_8	$\omega_1 \mathcal{P}_A \omega_j$	$\omega_1 \mathcal{P}_B \omega_j$	$\omega_1 \mathcal{P}_C \omega_j$

Notice that the preferences here are based on the supposition that agents are aware of the differences in productivity depending on the size of the city and balance the advantages of higher productivity against their aversion to relocation. We may then observe that ω_1 dominates ω_8 via the grand coalition, and that, if ω_1 is realized, ω_1 dominates all other outcomes. However, if, for example, ω_2 is realized, then 1) no coalition that excludes B is effective for ω_1, and 2) B prefers the *status quo* to ω_1. Further, 3) any other outcome requires some agent to move, without a productivity increase sufficient to offset the aversion to mobility. Thus, by similar arguments, all outcomes with partial urbanization are in the core of the game. With farsighted decisions, in the sense that decisions are made according to preferences in the newly realized outcome, ω_1 dominates all other outcomes via the grand coalition.

4. Reflections on Foresight

Can any normative case be made for foresight in the context of endogenous preferences? It does not seem that any such case can be made. We have considered three examples. In the first, farsighted decisions based on the preferences in the newly realized state stabilize what would be a cycle in the case of shortsighted preferences. However, foresight stabilizes the backward state of lower overall wellbeing and more inequality as well as the advanced state. In the second example, foresight leads to cycles whereas the shortsighted case is stable, although, again, the case of restricted opportunity is stabilized along with the case of unrestricted opportunity. This case raises yet another possibility for a foresighted decision-maker: if an agent or a coalition could bring about a transition to a state that is itself unstable and dominated by the case of unrestricted opportunity, a farsighted decision-maker might choose that strategy. In the third example, farsighted decisions lead to the high-productivity outcome of urbanization, but it is not clear that they are rational. If an agent moves from a village to the city, the agent experiences the subjective and objective costs of relocation. The preference to remain in the city reflects the fact that those negative consequences would be experienced a second time. To decide on the basis of the destination preferences ignores some experience in a way that might be seen as irrational.

It might be argued that foresight and rationality contradict endogenous preferences. An agent with foresight would anticipate all relevant aspects of all outcomes, such as a reduced margin of relative well-being for agent A in the Education Game, changing norms and social roles in the Opportunity Game, and disadvantages of relocation in the Mobility Game, and form a relative evaluation of the alternative outcomes based on their different characteristics[2]. If this relative evaluation has certain properties associated with rationality, then it could be expressed as a preference system. Such a preference system would be independent of the realized state. On the one hand, it is plausible that real human beings do not predictably decide with that degree of rationality and foresight. On the other hand, such a farsighted evaluation, independent of the status quo, is necessary for a normative discrimination among the outcomes. In the case of the Education Game, for example, the Advanced state is preferable to the Backward state on the grounds of Rawls' difference principle. What is demanded here seems, however, to go beyond Rawls' (1971) veil of ignorance. It requires not only that the agents choose among political-economic systems as if they were ignorant of the role they would play in the realized state, but also in ways that are independent of the existing status quo. For example, a person brought up in a society with a traditional class system might prefer the predictability of such a system to the unpredictability of a less stratified system, but for valid normative judgments, would be required to ignore those preferences.

However, if our objective is to explore the obstacles and possibilities of forming coalitions to bring about some transformations of the existing status quo, effectivity theory with endogenous preferences with shortsighted decisions seem to be a reasonable starting point.

[2] For a treatment of mobility in these terms see McCain, 2021, Ch. 8.

5. Concluding Summary

This essay has put forward two contributions. The first is the observation that endogenous preferences can be modeled in effectivity cooperative game theory by making the preference system one of the components of the outcome of the interaction. Other models of endogenous preferences (Stigler and Becker 1977, McCain 1979, McCain 1981 e.g.) focus on market equilibria and use concepts derived from human capital theory, but effectivity theory is more general and focused instead on the formation of coalitions, with the core of the game substituting for market equilibrium as a concept of stable social situations. The second contribution is more an instance of raising questions than of answering them: what are the implications for stable social situations if people are farsighted about the changes in their preferences. We have considered three examples: 1) an example in which preferences in one outcome depend on material well-being in one outcome and on relative status in another, more prosperous outcome, in which shortsighted decisions lead to cycles but farsighted decisions stabilize either outcome. 2) An example in which shortsighted decisions stabilize extreme outcomes, but in which a more detailed farsight might lead to strategies of destabilization of a status quo to bring about a different but stable outcome. 3) An example in which (if decisions are shortsighted) a low-productivity status quo will be unstable and dominated by the high-productivity outcome, but intermediate outcomes with moderate productivity may be stable; but with farsighted decisions only the high-productivity outcome is stable. All in all it seems that the answer to the question "what are the implications of foresight" is: it all depends. Specialized models will be needed for specialized cases.

References

Aumann, Robert J. (2003), “Presidential address,” Games and Economic Behavior v. 45, pp. 2-14.

Frey, Bruno S. and Alois Stutzer (2002), “What Can Economists Learn from Happiness Research?” Journal of Economic Literature v. 40, no. 2 (Jun) pp. 402-435.

George, David (2004), Preference Pollution: How Markets Create the Desires We Dislike (University of Michigan Press).

Goudie, Robert J. B and Sach Mukherjee, Jan-Emmanuel de Neve, Andrew J. Oswald and Stephen Wu (2014), “Happiness as a Driver of Risk-avoiding Behaviour: Theory and an Empirical Study of Seatbelt Wearing and Automobile Accidents,” *Economica* v. 81, no. 324 (Oct) pp. 674-697.

Ichiishi, Tatsuro (1987), □-Stable Extensive Game Forms, Mathematics of Operations Research v. 12, no. 4 (Nov) pp. 626-633.

Kahneman, D., & Krueger, A. B. 2006. Developments in the measurement well-being. Journal of Economic Perspectives, Vol. 20, No. 1 (Winter), pp. 3-24

McCain, Roger A. (1979), “Reflections on the Cultivation of Taste,” Journal of Cultural Economics v. 3, no. 1 (Jun) pp. 30-52.

McCain, Roger A. (1981), “Cultivation of Taste, Catastrophe Theory, and the Demand for Works of Art,” American Economic Review v. 71, no. 2 (May) pp. 332-334.

McCain, Roger A. (2021), Comparing Fairness: Relative Criteria of Economic Fairness with Applications (Cheltenham, U.K., Elgar).

McCain, Roger A. (2022), “Cooperative Game Theory and the Theory of the State With a Reconsideration of the Kaldor-Hicks Test,” available at:

https://www.researchgate.net/publication/359064766_Cooperative_Game_Theory_and_the_Theory_of_the_State_With_a_Reconsideration_of_the_Kaldor-Hicks_Test/stats, as of August 8, 2022, originally posted on 01/03/2022.

Moulin, H and Peleg, B (1982), "Cores of Effectivity Functions and Implementation Theory," Journal of Mathematical Economics v. 10, no. 1 (June) pp. 115-145.

Poundstone, William (1992), Prisoner's Dilemma (New York: Doubleday).

Rawls, John (1971), A Theory of Justice (Cambridge: Belknap Press).

Schelling, Thomas (1960), The Strategy of Conflict (Cambridge: Harvard University Press)

Stigler, George and Gary Becker (1977), "De Gustibus non est Disputandum," American Economic Review v. 67, no. 2 (Mar) pp. 67-90.

Sen, Amartya K (1997), "From Income Inequality to Economic Inequality," *Southern Economic Journal* v. 64, no. 2 (Oct.) pp. 383-401.