Математическая Теория Игр и её Приложения, т.2, в.2, с. 79-98

УДК 519.612:632.4
ББК 22.18

МОДЕЛЬ ЭНДОГЕННОГО ФОРМИРОВАНИЯ КОАЛИЦИЙ С ДВУМЯ ТИПАМИ ИГРОКОВ

Денис С. Степанов
Факультет Вычислительной Математики и Кибернетики
Московский Государственный Университет
119991, Москва, Ленинские горы, 2-й уч. корпус
e-mail: dn.step@gmail.com

Рассматривается модель формирования коалиций игроками, характеризуемыми значением некоторого параметра (географическое положение, идеальная точка). Новизна постановки заключается в предположении о *неоднотипности* игроков: к основному множеству добавляется небольшая доля игроков с функцией выигрыша, отличной от функции выигрыша игроков основного типа. Рассмотрены различные концепции коалиционной устойчивости и получены соответствующие необходимые и достаточные условия. Проанализировано соотношение данных условий для игр с одним и двумя типами игроков.

Ключевые слова: коалиционная устойчивость, равновесие Нэша, слабое коалиционное равновесие (СКР).

1. Введение

Одно из современных направлений в теории игр связано с исследованием моделей эндогенного формирования коалиций в больших неоднородных популяциях игроков. В такой игре игроки похожи в смысле вида функции выигрыша и множества стратегий, но различаются по некоторому параметру $x \in X$ (*идеальная точка*), при этом все множество игроков описывается распределением по указанному параметру. Стратегия игрока — выбор некоторой коалиции, стратегия коалиции представляет собой точку из X и определяется по заданному правилу в зависимости от состава коалиции. Размер коалиции пропорционален доле игроков, вошедших в нее. Игроки однородны по функции выигрыша, которая зависит от двух параметров: возрастает по размеру коалиции и убывает по расстоянию между идеальной точкой x и стратегией коалиции.

Подобные модели используются в политологии и экономической географии при изучении вопросов устойчивости разбиения населения по странам [1, 2], а также по юрисдикциям (муниципалитетам или регионам) внутри страны [3, 4]. Они находят также применение при анализе устойчивых разбиений избирателей по политическим партиям [5, 6], социальных сетей (в интернет-сообществах) и коалиций вообще [7, 8]. В указанных исследованиях авторы рассматривают вопросы существования и коалиционной устойчивости равновесий Нэша и изучают их свойства.

Однако, в данной области актуальной и практически неисследованной проблемой является учет в моделях неоднородности игроков не только по значению параметра, но и по характеру зависимости выигрыша от аргументов, то есть, другими словами, во всех упомянутых исследованиях предполагается, что игроки однотипны с точки зрения вида функции выигрыша. В реальной жизни можно наблюдать множество примеров, подтверждающих, что выигрыши агентов с одинаковой идеальной точкой в целом могут сильно отличаться или меняться с течением времени (вследствие каких-либо событий).

В настоящей работе рассматриваются структуры, являющиеся равновесными в модели с однотипными игроками, и исследуются их свойства в игре с двумя типами игроков, отличающихся характером зависимости выигрыша от аргументов. Для таких структур были найдены необходимые и достаточные условия *локальной устойчивости* (то есть устойчивости к образованию новых коалиций путем раскола или объединения существующих) и исследована их связь с аналогами для случая однотипной группы игроков. Также найден критерий существования *слабого коалиционного равновесия* (то есть равновесия, в котором не выгодно образование вообще любых новых коалиций — например, из частей нескольких существующих коалиций).

2. Модель

Сначала приведем, следуя [8], формальное описание игры в случае, когда игроки относятся к одному типу. Пусть множество идеальных точек можно представить как отрезок $[0,1] = X$, а игроки равномерно распределены по идеальным точкам на данном отрезке. Есть достаточно большой набор меток M: «Коалиция 1», «Коалиция 2»,…(например, в случае формирования политических партий это «социалисты», «демократы», «либералы» и т.д.). Пусть $M = \{1, 2 \ldots, \bar{M}\}$. Каждый из игроков, выбирая соответствующую метку, становится членом коалиции или же не вступает ни в одну из коалиций (метка «0»). Стратегия (политика) коалиции определяется как медиана распределения членов коалиции по идеальным точкам (самое распространенное в литературе правило). Размер коалиции пропорционален доле игроков, выбравших соответствующую метку.

В общем случае коалицией i является множество игроков, выбравших метку i. В дальнейшем ограничимся анализом ситуаций, в которых каждая коалиция $i \in M$ характеризуется интегрируемой функцией $\delta_i(x), x \in X$, описывающей плотность распределения по идеальным точкам игроков, выбравших метку i. В этом случае размер r_i коалиции $i \in M$ формально определяется как $r_i = \int_0^1 \delta_i(x)dx$.

Ситуация игры $\mathcal{D}$ определяется как множество коалиций ненулевого размера $I \subset M$ и набор $(\delta_i(x), i \in I)$ интегрируемых функций, показывающих долю игроков с идеальной точкой x, выбравших коалицию $i \in I$: $\mathcal{D} = \left\{\delta_i(x) \geq 0 : \sum_{i \in I} \delta_i(x) \leq 1, x \in [0,1], i \in I\right\}$. Пусть $\delta_0(x) = 1 - \sum_{i \in I} \delta_i(x)$ — доля игроков, не вступивших в коалиции. Стратегия P_i коалиции задается условием $\int_0^{P_i} \delta_i(x)dx = \int_{P_i}^1 \delta_i(x)dx$. Выигрыш игрока с идеальной точкой x, вошедшего в коалицию $i \in I$, определяется как

$$U(x, i, \mathcal{D}) = R\left(r_i\left(\mathcal{D}\right)\right) - L\left(\left|P_i\left(\mathcal{D}\right) - x\right|\right), \tag{2.1}$$

где $R(\cdot), L(\cdot)$ — некоторые положительные монотонно возрастающие функции, причем $L''(\cdot) \geq 0, R''(\cdot) \leq 0$. Выигрыш игрока в случае отказа от вступления в коалиции равен 0. Обозначим данную игру как $\mathcal{G}_1$.

Заметим, что выигрыш игрока в коалиции нулевого размера заведомо неположителен, поэтому будем предполагать, что доля игроков, выбравших такие коалиции, равна нулю.

В данной модели ситуация игры $\mathcal{D} = (\delta_i(x), i \in I)$ является *равновесием Нэша (РН)*, если для любого $x \in X$ и $i \in \bar{I}$ из того, что $\delta_i(x) > 0$ следует $i \in \operatorname*{Argmax}_{j \in \bar{I}} U(x, j, \mathcal{D})$, где $\bar{I} = I \cup 0$. То есть в РН каждый игрок выбирает коалицию, максимизирующую его выигрыш. Заметим, что *атомарная структура* (ни один из игроков не вступил в коалицию) — всегда РН.

Ниже исследуются *регулярные РН (РРН)*, то есть равновесия, в которых нет коалиций с одинаковой стратегией. Понятие *РРН* введено в [8], поскольку нерегулярные равновесия заведомо неустойчивы к объединению коалиций.

Рассмотрим следующую модификацию игры $\mathcal{G}_1$ — игру $\mathcal{G}_2$. К исходному множеству игроков (*типа 1* или *старого*, *основного* типа) добавляется некоторое количество игроков (*типа 2* или *нового* типа) с иной функцией выигрыша таким образом, что доля игроков нового типа в общем множестве игроков равна $\lambda > 0$. При этом рассматривается два варианта относительного распределения игроков разных типов:

- D1 Игроки обоих типов равномерно распределены на всем отрезке $[0, 1]$.
- D2 Игроки типа 2 равномерно распределены по идеальным точкам только на некотором сегменте $S \subset [0, 1]$ длины λ, а игроки типа 1 — на дополнении этого сегмента.

Пусть функция $J_t(x) \geq 0$ показывает долю игроков типа $t \in \{1, 2\}$ среди игроков с идеальной точкой $x \in [0, 1]$. Тогда $J_1(x) + J_2(x) = 1$ и в случае D1 функция $J_2(x) \equiv \lambda$, а в случае D2 — выражается как:

$$J_2(x) = \begin{cases} 1 & \text{при } x \in S; \\ 0 & \text{при } x \in [0, 1] \setminus S. \end{cases}$$

В игре $\mathcal{G}_2$ ситуация игры $\mathcal{D}$ определяется как множество коалиций ненулевого размера $I \subset M$ и набор интегрируемых функций $\delta_i^t(x)$, показывающих долю игроков типа $t \in \{1,2\}$ с идеальной точкой x, выбравших коалицию $i \in I$:

$$\mathcal{D}(I)=\left\{\delta_i^t(x) \geq 0 : \sum_{i\in I}\big((1-\lambda)\delta_i^1(x)+\lambda\delta_i^2(x)\big)\leq 1, x\in[0,1], i\in I, t=1,2\right\} \tag{2.2}$$

Таким образом, в игре $\mathcal{G}_2$ каждая коалиция $i \in I$ задается парой функций (δ_i^1, δ_i^2). Пусть $\delta_i(x) = (1-\lambda)\delta_i^1(x) + \lambda\delta_i^2(x)$ для $i \in I$ и $\delta_0(x) = 1 - \sum\limits_{i\in I}\delta_i(x)$. Размер r_i и стратегия P_i коалиции $i \in I$ определяется так же, как в игре $\mathcal{G}_1$. Выигрыш игрока в игре $\mathcal{G}_2$ определяется аналогично формуле (2.1), но каждому типу $t \in \{1,2\}$ соответствуют свои функции $R_t(\cdot)$ и $L_t(\cdot)$:

$$U_t(x,i,\mathcal{D}) = R_t\left(r_i\left(\mathcal{D}\right)\right) - L_t\left(\left|P_i\left(\mathcal{D}\right) - x\right|\right). \tag{2.3}$$

Причем $L_t(\cdot), R_t(\cdot) \geq 0, L_t'(\cdot), R_t'(\cdot) \geq 0, L_t''(\cdot) \geq 0$ и $R_t''(\cdot) \leq 0$. В дальнейшем для обозначения выигрыша игрока x типа $t \in (1,2)$ в коалиции i размера r_i со стратегией P_i будем использовать запись $U_t(x, r_i, P_i)$.

В игре $\mathcal{G}_2$ равновесие Нэша — такая ситуация игры $\mathcal{D} = (\delta_i^t(x), i \in I,\ t = 1,2)$ в которой, для любого типа $t \in \{1,2\}, x \in X$ и $i \in \bar{I}$ из того, что $\delta_i^t(x) > 0$ следует $i \in \operatorname*{Argmax}\limits_{j\in\bar{I}} U_t(x,j,\mathcal{D})$, где $\bar{I} = I \cup 0$. Понятие регулярного РН вводится аналогично игре $\mathcal{G}_1$.

Поиск равновесных структур в данной теоретико-игровой модели, исходя из определения, предполагает решение сложных экстремальных задач на функциональном пространстве: необходимо найти множество $\bar{I} \subset M \bigcup\{0\}$ и такой набор интегрируемых функций $\bar{\mathcal{D}}\left(\bar{I}\right) = \left\{\delta_i^t(x), i \in \bar{I}, t = 1,2\right\}$, удовлетворяющий условию (2.2), что для любого $x \in [0,1]$ из того, что $\delta_i(x) > 0$, следует, что $i \in \operatorname*{Argmax}\limits_{j\in\bar{I}} U(x, r_j, P_j)$, где r_j и P_j в свою очередь зависят от $\bar{\mathcal{D}}$. Применение стандартных оптимизационных методов для решения данной задачи не приводит к содержательным результатам. В настоящей работе развивается подход, предложенный в работе [8], для описания равновесных коалиционных структур и их коалиционной устойчивости в игре $\mathcal{G}_1$.

3. Равновесие Нэша

Напомним, что в игре $\mathcal{G}_1$ под коалицией $i \in I$ мы понимаем соответствующую функцию плотности δ_i. Носитель $\operatorname{supp}\delta_i$ функции δ_i задает множество S_i идеальных точек игроков, выбравших коалицию i. В общем случае $\delta_i(x)$ для данного $x \in [0,1]$ может принимать любое значение из отрезка $[0,1]$ и, следовательно, по множеству S_i нельзя однозначно восстановить исходную коалицию. Но в равновесии Нэша $\delta_i(x) = 1$ при $x \in \operatorname{int} S_i$ и множество S_i однозначным образом (с точностью до множества игроков, расположенных на его границе, мера которого равна нулю) соответствует множеству игроков, выбравших коалицию $i \in I$. Поэтому при поиске равновесий под коалицией $i \in I$ будем понимать более простой объект — множество S_i.

Можно показать, что в игре $\mathcal{G}_1$ в регулярном равновесии отрезок [0,1] разбивается на непересекающиеся интервалы идеальных точек игроков, выбравших одну коалицию (или не вступивших ни в одну из коалиций). Более того в работе [8] показано, что существует всего три типа равновесных разбиений игроков на коалиции (см. теорему 3.1).

Рассмотрим уравнение $R_t(r) - L_t(r/2) = 0$ для некоторого $t \in \{1,2\}$. Отметим, что если данное уравнение имеет положительное решение, то в указанных предположениях оно единственно. Обозначим данное решение через r^*. Рассмотрим некоторое натуральное число m. Для игры $\mathcal{G}_1$ с игроками типа $t \in \{1,2\}$ справедлива следующая теорема.

Теорема 3.1. ([8], стр.1522) *В игре $\mathcal{G}_1$ существуют следующие три типа РРН.*

a) Разбиение отрезка $[0,1]$ на m коалиций одинакового размера $r = 1/m$ является РРН тогда и только тогда, когда $r \leq r^*$. В частности, если $r^* \geq 1$ или $R_t(r) > L(r/2)$ для любого $r > 0$, то такая структура является РРН для любого m; если $L_t'(0) > 2R_t'(0)$, то единственное РРН — атомарная структура.

b) Если $r^* \in (0,1)$, то для любого $u \in \left(0, \max\limits_r \left(R_t(r) - L_t\left(r/2\right)\right)\right)$ существуют два решения $r_1(u), r_2(u) \in (0,1)$ уравнения $u = R_t(r) - L_t\left(r/2\right)$, и если существуют натуральные числа m_1, m_2 такие, что $m_1 r_1(u) + m_2 r_2(u) = 1$, то разбиение отрезка $[0,1]$ на m_1 коалиций размера $r_1(u)$ и m_2 коалиций размера $r_2(u)$ является РРН (коалиции

могут располагаться в любом порядке).
c) Для любого $m < 1/r^*$ любое разбиение отрезка $[0,1]$ на m коалиций размера r^* и $l \leq m+1$ интервалов игроков, воздержавшихся от вступления в коалиции, является РРН.

Не существует других РРН кроме указанных.

Рассмотрим типы равновесий, описанные в пунктах b) и c) утверждения. Каждый из данных типов равновесий оказывается неустойчивым в некотором смысле. Коалиционная структура, соответствующая равновесию типа b) в игре $\mathcal{G}_1$, в ситуации общего положения не будет равновесной в игре $\mathcal{G}_2$, поскольку из того, что $R_t(r_1) - L_t(r_1/2) = R_t(r_2) - L_t(r_2/2)$ для $t = 1$, вообще говоря, не следует выполнение данного равенства для $t = 2$. А равновесие типа c) неустойчиво к образованию новых коалиций (из игроков, выбравших стратегию «0»). Таким образом, при исследовании эффектов, возникающих в равновесиях, при переходе от игры $\mathcal{G}_1$ к игре $\mathcal{G}_2$ будем рассматривать только структуры, описанные в пункте a) утверждения. Обозначим подобную структуру через K_m, а через r — размер коалиции в $K_m : r = 1/m$. Данная структура является равновесной тогда и только тогда, когда выигрыш игрока, находящегося на границе коалиции, неотрицателен:

$$R_t(r) - L_t(r/2) \geq 0. \tag{3.1}$$

В отличие от игры $\mathcal{G}_1$ в игре $\mathcal{G}_2$ коалиции $i \in I$ соответствует пара функций плотности (δ_i^1, δ_i^2). Аналогичным образом носитель функции $\mathrm{supp}\delta_i^t$ задает множество S_i^t идеальных точек игроков типа $t \in \{1,2\}$, выбравших коалицию i, и в равновесии множество S_i^t однозначным образом (с точностью до множества игроков, расположенных на его границе, нулевой меры) соответствует множеству игроков типа $t \in \{1,2\}$, выбравших коалицию $i \in I$. В этом случае под коалицией $i \in I$ будем понимать пару множеств (S_i^1, S_i^2). Это оказывается целесообразно и при исследовании коалиционной устойчивости равновесий, поскольку если $S_i^t = S_j^t, t = 1,2$, и выгодно образование коалиции $i : \delta_i^t(x) \leq 1, x \in S_i^t, t = 1,2$, то выгодно образование коалиции $j : \delta_j^t(x) = 1, x \in S_j^t, t = 1,2$.

Для коалиции $i \in I$ такой, что каждое из множеств $S_i^t, t = 1,2$, представляет собой отрезок, *границей коалиции* i будем называть множество крайних точек соответствующих интервалов. В этом случае *соседними* для коалиции i будут коалиции, пересечение границ которых с границей коалиции i не пусто.

4. Локальная устойчивость

Равновесие Нэша устойчиво в том смысле, что отклонение одного игрока от состояния равновесия не приводит к увеличению его выигрыша. В то же время хорошо известны примеры, когда одновременное отклонение нескольких игроков от равновесия приводит к увеличению выигрыша всех этих игроков. В связи с этим в литературе рассматриваются различные понятия коалиционной устойчивости равновесий. Основная часть данной работы также посвящена исследованию устойчивости коалиционных структур, являющихся регулярными равновесиями Нэша как в игре $\mathcal{G}_1$ (с игроками типа 1), так и в игре $\mathcal{G}_2$ (с игроками обоих типов), к образованию новых коалиций.

В работе рассматриваются следующие понятия коалиционной устойчивости. РН устойчиво к *локальному расколу*, если не существует новой коалиции, являющейся подмножеством некоторой коалиции и обеспечивающей бóльшие выигрыши всем своим членам. РН устойчиво к *объединению*, если не существует новой коалиции, являющейся объединением соседних коалиций и обеспечивающей бóльшие выигрыши всем своим членам. Если РН устойчиво к локальному расколу и объединению, то оно *локально устойчиво.*

Теорема 4.1. ([8], стр. 1522) *В игре $\mathcal{G}_1$ с игроками типа $t \in \{1, 2\}$ равновесная структура $K_m, m = 2, 3, \ldots$, локально устойчива тогда и только тогда, когда*

$$R_t(2r) - R_t(r) \le L_t(r) - L_t(r/2). \tag{4.1}$$

Как было указано в предыдущем разделе, в игре $\mathcal{G}_1$ в равновесной структуре K_m выполнено условие (3.1) неотрицательности выигрыша граничного игрока. Оказывается, что указанное условие является критерием устойчивости структуры к локальному расколу. При невыполнении данного условия части агентов, близких к границе коалиции, невыгодно быть в ее составе (им выгодно выйти из состава данной коалиции или образовать новую коалицию меньшего размера). Таким образом, в данной игре равновесная структура

K_m всегда устойчива к локальному расколу. Условие (4.1) гарантирует, что объединение соседних коалиций невыгодно. Это следует из того, что объединение любого числа соседних коалиций невыгодно, когда невыгодно объединение двух соседних коалиций, а последнее определяется величиной выигрыша граничных агентов объединенной коалиции. Покажем, что для игры $\mathcal{G}_2$ справедливы аналогичные результаты (при любом из предположений D1-D2).

Теорема 4.2. *Пусть для игроков типа 1 выполнены условия (3.1)–(4.1). В этом случае структура K_m является равновесием Нэша и локально устойчива в игре $\mathcal{G}_2$ тогда и только тогда, когда для игроков типа 2 выполнено условие (3.1).*

Доказательство. Следуя доказательству утверждения 2 из работы [8], можно показать, что условие (3.1) обеспечивает устойчивость к индивидуальным отклонениям игроков нового типа, а также к образованию коалиций (достаточно малого размера), состоящих из этих игроков (и, возможно, игроков старого типа). При этом не требуется выполнения условия (4.1), поскольку устойчивость к локальному объединению будет обеспечиваться игроками старого типа, для которых выполнены условия локальной устойчивости (3.1)–(4.1). □

5. Слабое коалиционное равновесие

В данном разделе будет рассматриваться усиление понятия локальной устойчивости — понятие слабого коалиционного равновесия.

Равновесие Нэша называется *слабым коалиционным равновесием (СКР)*, если не существует новой коалиции, обеспечивающей бóльшие выигрыши всем своим членам.

Сначала приведем результаты для игры $\mathcal{G}_1$, а в последующих подразделах — аналогичные результаты для игры $\mathcal{G}_2$ для вариантов распределения игроков D1 и D2.

При исследовании коалиционной устойчивости вводится дополнительное предположение относительно функций $R(\cdot)$ и $L(\cdot)$: $L'''(\cdot) \geq 0, R'''(\cdot) \leq 0$.

5.1. *СКР* в игре $\mathcal{G}_1$

В работе [8] было доказано достаточное условие эквивалентности понятий локальной устойчивости и слабого коалиционного равновесия в случае $R(r) \equiv r$ (теорема 2, стр. 1522). Ниже приведем обобщение данного утверждения.

Теорема 5.1. *В игре $\mathcal{G}_1$ с игроками типа $t \in \{1, 2\}$ локально устойчивая структура K_m является СКР тогда и только тогда, когда выполнены следующие условия:*

$$L'_t\,(r/2) \geq R'_t(r) \ \text{ и } \tag{5.1}$$

$$R_t(r^*) - R(r) + L_t(r^*/2 - r/2) - L_t(r^*/2) \leq 0, \tag{5.2}$$

где $r^* = \min\left\{x^*, \frac{3}{2}r\right\}$, если существует решение уравнения $R'_t(x^*) + \frac{1}{2}L'_t(x^*/2 - r/2) - \frac{1}{2}L'_t(x^*/2) = 0$, и $r^* = r$ иначе.

Дополнительные ограничения (5.1)–(5.2) на локально устойчивые структуры возникают по следующей причине. Из доказательства теоремы 2 из [8] видно, что условия локальной устойчивости (3.1) и (4.1) обеспечивают устойчивость к образованию любых коалиций размера меньше r и больше $2r$ соответственно, но не гарантируют невыгодность образования коалиций, размер которых принадлежит интервалу $(r, 2r)$.

Также в работе [8] приведено достаточное условие *СКР*, в котором ограничения на параметры модели формулируются в более простом виде.

Следствие 5.1. ([8], стр. 1522) *Локально устойчивая структура K_m является СКР, если выполнено условие (5.1) и условие*

$$R'_t(2r) + \tfrac{1}{2}L'_t\,(r/2) - \tfrac{1}{2}L'_t(r) \geq 0. \tag{5.3}$$

5.2. Игра $\mathcal{G}_2$ в предположении D1.

Напомним, что вариант D1 предполагает равномерное распределение игроков двух типов на отрезке $[0, 1]$.

В зависимости от свойств функций R_2 и L_2 добавление игроков нового типа к исходному множеству игроков может может либо «разрушить» *СКР*, то есть привести к тому, что структура, являющаяся

СКР в игре $\mathcal{G}_1$, не будет *СКР* в игре $\mathcal{G}_2$, либо наоборот обеспечить устойчивость, если ее не было. Нас будет интересовать первый исход, который может произойти по двум причинам:

A1 Игроки нового типа являются *бóльшими индивидуалистами.* То есть либо более чувствительны к росту удаленности: $L_2'(\cdot) > L_1'(\cdot)$, либо менее чувствительны к росту размера коалиции: $R_2'(\cdot) < R_1'(\cdot)$ (либо и то и другое одновременно).

A2 Игроки нового типа являются *бóльшими конформистами.* В этом случае предполагается выполнение неравенств, обратных неравенствам в предположении A1.

Рассмотрим последовательно данные случаи и для каждого из них приведем утверждения, характеризующие *СКР* в игре $\mathcal{G}_2$.

Теорема 5.2. *Пусть выполнено предположение A1. Структура K_m является СКР в игре $\mathcal{G}_2$ тогда и только тогда, когда она является равновесной в игре $\mathcal{G}_2$, то есть для игроков обоих типов выполнено условие (3.1), а также СКР в игре $\mathcal{G}_1$ с игроками типа 1, то есть для игроков типа 1 выполнены условия (4.1)–(5.2).*

Доказательство. Из выполнения предположения A1, а также неравенств (3.1)–(5.2) для игроков типа 1, следует, что возможен единственный случай, при котором возникновение игроков нового типа может привести к тому, что структура K_m, являющаяся *СКР* в игре $\mathcal{G}_1$ (с игроками типа 1), не является *СКР* в игре $\mathcal{G}_2$. Это возможно только, когда игрокам типа 2 выгодно выйти из состава существующей коалиции, посредством образования коалиции достаточно малого размера. Но при выполнении неравенства (3.1) игрокам типа 2 невыгодно создание такой коалиции. □

Заметим, что предположение A2 не ограничивает «степень» конформизма игроков нового типа, которая a priori может быть достаточно велика для того, чтобы разрушить любое *СКР*. Например, если значение $L_2'(\cdot)$ на отрезке [0,1] ограничено достаточно малой величиной, то игрокам типа 2 может быть выгодно всем вместе присоединиться к какой-либо существующей коалиции и тем самым образовать новую коалицию большего размера. В связи с этим необходимо ввести дополнительное ограничение на выигрыш игроков нового типа и самое естественное в данном случае — потребовать выполнение условия устойчивости к локальному объединению для игроков типа 2.

Теорема 5.3. *Пусть выполнено предположение A2 и условия локальной устойчивости (3.1)–(4.1) структуры K_m для каждого из типов. Структура K_m является СКР тогда и только тогда, когда для типа 1 выполнены условия (5.1) и*

$$R_1((1-\lambda)r^*_\lambda + 2r\lambda) - R_1(r) + L_1(r^*_\lambda/2 - r/2) - L_1(r^*_\lambda/2) \leq 0, \quad (5.4)$$

где $r^*_\lambda = \min\left\{x^*, \frac{3}{2}r\right\}$, *если существует решение уравнения* $(1-\lambda+\lambda 2r/x^*)R_1'((1-\lambda)x^*+2r\lambda)+\frac{1}{2}L_1'(x^*/2-r/2)-\frac{1}{2}L_1'(x^*/2)=0$, *и* $r^*_\lambda = r$ *иначе, а для типа 2 — неравенство*

$$L_2'(r/2) \geq \lambda R_2'(r). \quad (5.5)$$

Доказательство. Пусть π_1 — новая (потенциальная) коалиция со стратегией P_1 и размером r_1. Рассмотрим сначала случай $r_1 \geq 2r$. Тогда π_1 формируется из членов как минимум двух коалиций из K_m, и, следовательно, существует игрок x типа $t \in \{1,2\}$, такой что $|P_1 - x| \geq r_1/2$. Справедливо неравенство $R_t(r_1) - L_t(|P_1 - x|) \leq R_t(r_1) - L_t(r_1/2) = U_t(r_1)$. Обозначим $\Delta U_t(r_1) = U_t(r_1) - U_t(r)$. Тогда из устойчивости к локальному объединению вытекает, что $\Delta U_t(2r) \leq 0$. С учетом того, что $\Delta U_t(r) = 0$ и $\Delta U_t(\cdot)$ — вогнутая функция получаем, что для любого размера $r_1 \geq 2r$ прирост выигрыша игрока x при смене коалиции $\Delta U_t(r_1) < 0$.

Пусть $r_1 \leq r$ и коалиция π_1 представляет собой связное множество. Предположим, что она несимметрична относительно своей стратегии. Пусть x — игрок типа $t \in \{1,2\}$, чья идеальная точка расположена ближе всего к стратегии его исходной коалиции P. С учетом предположений относительно знака второй производной функций $R_t(\cdot)$ и $L_t(\cdot)$ прирост выигрыша этого игрока будет минимальным (среди игроков его типа). Пусть ρ обозначает расстояние от P_1 до ближайшей границы исходной коалиции из K_m, которой принадлежит P_1. Фиксируем значение r_1 и подберем параметр $\varepsilon = const$ таким образом, что $|x - P_1| = \varepsilon r_1$, причем $\varepsilon > \frac{1}{2}$ (так как x — ближайший к P игрок). Прирост выигрыша игрока x при переходе в коалицию π_1 равен $\Delta U_t(x, r_1, P_1) = R_t(r_1) - R_t(r) + L_t(z_1) - L_t(z_2)$, где $z_1 = r/2 - (\varepsilon r_1 + \rho), z_2 = |x - P_1| = \varepsilon r_1$ (далее по тексту индекс t опускаем). Заметим стразу, что ΔU убывает по ρ (то есть оптимальное положение коалиции — когда P_1 находится на границе существующих коалиций). Пусть $V(r_1) = \Delta U(r_1) + R(r) - R(r - r_1)$ и, значит, $\Delta U(r_1) \leq V(r_1)$ и функция $V(r_1)$ вогнута.

Существует значение r_1^* такое, что $z_1 = z_2$. При этом $r_1^* \leq \frac{r}{2}$ поскольку при $r_1^* > \frac{r}{2}$ значение $z_1 = \frac{r}{2} - \varepsilon \frac{r_1^*}{2} < \frac{r_1^*}{4} < z_2$ и $\Delta U < 0$. Значение $V(0) = -R(r) + L(r/2) \leq 0$ и $V(r_1^*) < 0$, следовательно, $V(r_1) \leq 0$ для всех $r_1 \leq r/2$.

Теперь предположим, что π_1 представляет собой произвольную (не обязательно связную) коалицию размера $r_1 \leq r$. Фиксируем игрока x, чья идеальная точка наиболее удалена от P_1. Существует связная коалиция размера r_1 со стратегией P_1', в которой x также является граничным игроком, и причем $|P_1' - x| \leq |P_1 - x|$. При этом $\Delta U(x, r_1, P_1) \leq \Delta U(x, r_1, P_1')$. Но выше было показано, что $\Delta U(x, r_1, P_1')$ ≤ 0.

Пусть, наконец, $r \leq r_1 \leq 2r$. Сначала покажем, что создание несимметричной коалиции (в смысле взаимного расположения множеств идеальных точек игроков типов 1 и 2) невыгодно. Существуют два варианта оптимального расположения π_1 относительно исходной структуры: когда π_1 содержит (a) одну или (b) две точки, соответствующие стратегиям некоторых коалиций из K_m.

Покажем для варианта (a) невыгодность создания несимметричных коалиций (в случае (b) доказательство проводится аналогичным образом). Обозначим: x_1, x_2 – граничные агенты типа T_1, z – расстояние от указанных граничных агентов до стратегии P_1 в случае симметричной коалиции, ρ – расстояние от агента x_i, до стратегии его исходной коалиции в случае симметричной коалиции, $\delta > 0$ – смещение стратегии P_1 относительно симметричного случая, d – смещение агента x_1 относительно его положения в симметричной конфигурации (см. рис. 1). Для того, чтобы «компенсировать» игроку x_1 потерю выигрыша из-за смещения политики P_1 в сторону от него, необходимо подобрать достаточно большое смещение d. Покажем, что это невозможно без того, чтобы выигрыш агента x_2 не снизился по сравнению с его выигрышем в симметричной конфигурации. Прирост выигры-

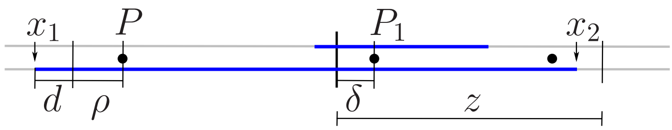

Рисунок 1. Несимметричная коалиция в случае (а).

ша каждого из агентов в этом случае равен (индексы при функциях $R(\cdot)$ и $L(\cdot)$ опускаем):

$$\begin{cases} \Delta_1 U(d,\delta) = R(r_1) - R(r) + L(\rho + d) - L(z + d + \delta), \text{для } x_1, \\ \Delta_2 U(d,\delta) = R(r_1) - R(r) + L(\rho - d) - L(z - d - \delta), \text{для } x_2. \end{cases}$$

Предположим, что существует функция $d(\delta)$ (дифференцируемая при $\delta \geq 0$) такая, что существует значение $\delta_0 > 0$ такое, что $\Delta U_i(d(\delta_0), \delta_0) > 0, i = 1, 2$. В силу непрерывности рассматриваемых функций данные неравенства выполняются в некоторой окрестности точки δ_0 и $d'(\delta) > 0$ на некотором интервале из этой окрестности. Продифференцировав по δ каждый из приростов и сложив полученные выражения, получим:

$$\Delta_1 U' + \Delta_2 U' = d' \left[L'(\rho + d) - L'(\rho - d)\right] - d' \left[L'(z + d + \delta) - \right.$$
$$\left. - L'(z - d - \delta)\right] - \left[L'(z + d + \delta) - L'(z - d - \delta)\right].$$

Следовательно, так как интервал $(z-d-\delta, z+d+\delta)$ больше интервала $(\rho-d, \rho+d)$, значение $z+d+\delta$ больше $\rho+d$, а функция $L(\cdot)$ выпукла, то $\Delta_1 U' + \Delta_2 U' < 0$ для любого δ и, значит, $\Delta_1 U + \Delta_2 U < 0$. Таким образом, пришли к противоречию и $d = \delta = 0$.

С учетом доказанного выше, далее рассматриваются только симметричные коалиции, при этом также сохраняется деление на два случая — (a) и (b), описанное выше в доказательстве. Рассмотрим случай (a). Поместим начало координат в точку $P_1 = P$. Пусть $\frac{x}{2}$ — наиболее удаленный от стратегии коалиции агент нового типа из коалиции π_1, $\frac{z}{2}$ — соответствующий агент старого типа. Для любой конфигурации π_1 существует $\tau > 0$ такое, что $z = \tau x$. Заметим, что конфигурация π_1 такая, что x или z больше $2r$ или в которой x (или z) меньше r невыгодна. Таким образом, в случае (a) x и z могут принимать значения только из диапазона: $x, z \in [r, 2r]$.

Прирост выигрыша крайнего агента типа 1 при переходе в коалицию π_1 равен $\Delta U_1(x) = \Delta R_1 + L_1\left(r - \frac{x}{2}\right) - L_1\left(\frac{x}{2}\right)$, где $\Delta R_1 = R_1\left((1-\lambda)x + \lambda\tau x\right) - R_1(r)$ (прирост для игрока нового типа выписывается аналогичным образом). Заметим, что $\Delta U_1(r) = R_1((1-\lambda)r + \lambda\tau r) - R_1(r) = 0$ при $\tau = 1$ ($z = r$) и $\Delta U_1(r) > 0$ при $\tau > 1$. При этом $\Delta U_1''(x) \leq 0$ и при $\tau = 1$ значение $\Delta U_1'(r) \leq 0$ (следует из условия CKP) и, значит, $\Delta U_1'(x) \leq 0$ для любого x. Таким образом, для существования устойчивости к образованию подобных коалиций необходимо и достаточно чтобы для типа 2 конфигурация π_1 с $\tau > 1$ была невыгодна (для любого x). Значение ΔU_2 при $x = r$ равно $\Delta U_2 \mid_{x=r} = R_2\left((1-\lambda)r + \lambda\tau r\right) - R_2(r) + L_2\left(r - \frac{\tau r}{2}\right) - L_2\left(\frac{\tau r}{2}\right)$ и равно нулю при $\tau = 1$. При этом вторая производная по τ прироста выигрыша отрицательна при $\tau \geq 1$ для любого x. Следовательно, для того чтобы при $x = r$ значение ΔU_2 было меньше нуля необходимо и достаточно, чтобы производная $(\Delta U_2)'_\tau$ при $\tau = 1$ была отрицательной, что выполнено при выполнении условия (5.5).

Рассмотрим вариант расположения (b). Данная конфигурация является для игроков старого типа более выгодной, когда $\Delta R_1 + L_1(|x/2 - r/2|) - L_1(x/2) \geq \Delta R_1 + L_1(r - x/2) - L_1(x/2)$, что эквивалентно $x \geq \frac{3r}{2}$ и x в данном случае находится в пределах между $\frac{3r}{2}$ и $2r$ (аналогичное справедливо для типа 2). Условие, обеспечивающее устойчивость к образованию коалиций указанного типа, имеет вид: $\Delta U_1(x^*_\tau) \leq 0$, где $x^*_\tau \in \left(\frac{3r}{2}, 2r\right)$ — точка максимума прироста выигрыша на отрезке. Можно показать, что производная x^*_τ по τ положительна. Производная прироста выигрыша игрока старого типа по τ в точке x^*_τ равна $(\Delta U_1(x^*_\tau))'_\tau = \lambda x^*_\tau R_1'(r^*_1)$ и, следовательно, функция $\Delta U_1(x^*_\tau)$ возрастает по τ. Таким образом, необходимо найти максимально возможное значение τ и обеспечить невыгодность образования такой конфигурации. Из локальной устойчивости следует, что $z = \tau x \leq 2r$. Значит, с учетом того, что x^*_τ возрастает по τ, максимальное значение τ^* равно $\frac{2r}{x^*_{\tau^*}}$, где $x^*_{\tau^*}$ определяется из условия (5.4) при $\tau = \frac{2r}{x^*_{\tau^*}}$ $\square$

Также приведем два следствия из теоремы, формулирующие условия устойчивости в более удобном виде и позволяющие упростить анализ устойчивости коалиционных структур.

Следствие 5.2. *Структура K_m является СКРв игре $\mathcal{G}_2$, если для типа 1 выполнено достаточное условие СКР для игры $\mathcal{G}_1$, то есть условия (3.1)–(4.1) и неравенства (5.1) и (5.3), а для типа 2 — неравенство (5.5).*

Следствие 5.3. *Пусть структура K_m является СКР в игре $\mathcal{G}_1$ с игроками типа 1. Тогда она является СКР и в игре $\mathcal{G}_2$, если для игроков типа 2 выполнено условие (5.5) и неравенство*

$$\lambda R_2'(3r/2) + \tfrac{1}{2}L_2'(r/4) - \tfrac{1}{2}L_2'(3r/4) \leq 0.$$

Отметим, что основное отличие следствия 5.3 от теоремы и следствия 5.2 в том, что в этом варианте утверждения не требуется дополнительных ограничений на тип 1.

В качестве иллюстрации рассмотрим следующий пример. Пусть выигрыш игрока типа $t \in \{1, 2\}$ равен

$$U_t(x, r, P) = 2\alpha_t r - |x - P|^{k_t}, \tag{5.6}$$

где $\alpha_t \geq 0, k_t \geq 2$. Оказывается, что соотношение на параметры выигрыша игроков каждого типа в *СКР* можно выразить через две переменные: $\beta_t = \frac{2\alpha_t}{(r/2)^{k_t-1}}$ и k_t. С учетом следствия 5.2 ограничения на параметры в *СКР* имеют следующий вид:

$$k_1(2^{k_1-1} - 1) \leq \beta_1 \leq \min\left\{2k_1, 2^{k_1} - 1\right\}, \tag{5.7}$$

$$1 \leq \beta_2 \leq \min\left\{\tfrac{1}{\lambda}2k_2, 2^{k_2} - 1\right\}. \tag{5.8}$$

На рис. 2 изображены кривые, входящие в неравенства (5.7), и заштрихована область, в которой ограничения (5.7) выполняются. Из рисунка в частности видно, что, поскольку m возрастает с ростом β_1, то с ростом эластичности функции потерь k_1 число коалиций в устойчивой структуре возрастает, а чем больше значение параметра α, тем больше должен быть размер коалиций в K_m. Также из (5.8) следует, что с ростом доли λ игроков нового типа значение m уменьшается.

5.3. Игра $\mathcal{G}_2$ в предположении D2

Напомним, что в варианте D2 относительного распределения игроков предполагается распределение игроков нового типа на некотором небольшом сегменте $S \subset [0, 1]$ длины $\lambda < r$. В разделе 3 было

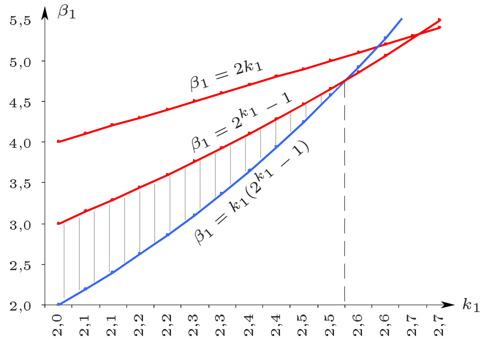

Рисунок 2. Область допустимых значений параметров в *СКР*

показано, что *СКР* возможно только в коалиционной структуре K_m, которая, согласно теореме 4.2, в игре $\mathcal{G}_2$ остается равновесной и локально устойчивой при выполнении для игроков типа 2 условия неотрицательности выигрыша граничного игрока (3.1).

Оказывается, что в предположении A1 (игроки типа 2 являются бо́льшими индивидуалистами) условия *СКР* для варианта относительного распределения игроков D2 полностью аналогичны соответствующим условиям для варианта D1.

Теорема 5.4. *Пусть выполнено предположение A1. Структура K_m является СКР при варианте относительного распределения игроков D2 тогда и только тогда, когда она является СКР при варианте распределения D1.*

В предположении A2 (игроки типа 2 являются бо́льшими конформистами) *СКР* может быть разрушено из-за выгодности создания новых коалиций, расположенных несимметрично относительно исходной равновесной структуры K_m (поскольку граничными игроками могут являться игроки разных типов). Более того, можно показать, что наибольшая «угроза» *СКР* возникает, когда игроки нового типа находятся на границе одной из существующих коалиций.

Теорема 5.5. *Пусть выполнено предположение A2, структура K_m является СКР в игре $\mathcal{G}_1$ с игроками типа 1 и равновесной в игре $\mathcal{G}_2$. В этих условиях структура K_m будет СКР в игре $\mathcal{G}_2$ тогда и толь-*

ко тогда, когда выполнено одно из двух неравенств:

$$\left[\begin{array}{l} L_1'\,(r/2) \geq 2R_1'(r), \\ L_2'\,(r/2) \leq \frac{2}{3}R_2'(r). \end{array}\right. \tag{5.9}$$

Доказательство. Пусть π_1 — новая (потенциальная) коалиция со стратегией P_1 и размером r_1, и аналогично доказательству теоремы 5.3 рассмотрим сначала случай $r_1 \geq 2r$. Здесь рассуждения аналогичны рассуждениям в соответствующем пункте теоремы 5.3. Единственное отличие заключается в том, что поскольку $|S| = \lambda < r_1$ и в коалицию π входят игроки обоих типов, то устойчивость будет обеспечиваться игроками старого типа.

В случае $r_1 \leq r$ самая «опасная» с точки зрения устойчивости ситуация — когда игроки нового типа находятся на границе существующих коалиций. Но, аналогично доказательству теоремы 5.3, в этом случае также можно показать, что устойчивость будет обеспечиваться условием равновесия (3.1).

При $r \leq r_1 \leq 2r$ нужно также рассматривать два варианта расположения π_1 относительно K_m (см. доказательство теоремы 5.3) и аналогичным образом можно показать, что и в этом случае прирост выигрыша граничного игрока новой коалиции будет минимальным (среди его типа). Обозначим данного игрока через x. Заметим, далее, что в обоих вариантах расположения π_1 прирост граничного игрока старого типа будет тем больше, чем ближе граница π_1 расположена к границе коалиции из K_m и максимален, когда этот игрок является граничным агентом как π_1, так и одной из коалиций из K_m. Прирост его выигрыша в этом случае равен $\Delta U_1(x, r_1, P_1) = R_1(r_1) - R_1(r) + L_1(r/2) - L_1(r_1/2)$, где $r \leq r_1 \leq 2r$. Заметим, что $\Delta U_1(r) = 0$ и $\Delta U_1'' < 0$, и если при этом минимальный прирост выигрыша игроков нового типа будет положителен, то для устойчивости необходимо и достаточно, чтобы $\Delta U_1'(r_1 = r) \leq 0$, и, таким образом, приходим к условиям (5.9) □

Таким образом, только при выполнении более сильного ограничения на выигрыш игроков типа 1 (см. условие (5.1)) структура K_m остается *СКР* при возникновении игроков нового типа.

Вернемся теперь к примеру с функцией выигрыша вида (5.6). С учетом (5.9) ограничения на параметры функций выигрыша игроков старого типа примут вид:

$$k_1(2^{k_1-1}-1) \le \beta_1 \le \min\left\{k_1, 2^{k_1}-1\right\}.$$

Данное ограничение является более сильным, чем условие (5.7) и выполняется только при $k_1 = 2$. Таким образом, только за счет ограничений на свойства функции выигрыша игроков старого типа не удается гарантировать устойчивость. Но при этом, если удается гарантировать выполнение ограничений

$$1 \le \beta_2 \le \frac{1}{\lambda} 2k_2 \left(\left(\frac{3}{2}\right)^{k_2-1} - \left(\frac{1}{2}\right)^{k_2-1}\right),$$

то коалиционная структура останется устойчивой при возникновении группы игроков нового типа, причем при $k_2 \ge 3$ данное условие гарантированно выполняется при выполнении неравенства (5.8), но при $2 \le k_2 \le 3$ наоборот является более сильным.

СПИСОК ЛИТЕРАТУРЫ

1. Alesina A. Spolaore E. *On the Number and Size of Nations* // Quarterly Journal of Economics. 1997. V. 112. N 4. P. 1027–1056.

2. Alesina A. M., Spolaore E. *The size of Nations. Cambridge, MA.* // MIT Press 2003.

3. Bogomolnaia A. M., Le Breton M., Savvateev A., et al. *Stability of Jurisdiction Structures under the Equal Share and Median Rules* // IDEI Working Papers. 2005. N 362.

4. Bogomolnaia A. M., Le Breton M., Savvateev A., et al. *The Egalitarian Sharing Rule in Provision of Public Goods* // Economic Bulletin. 2005. V. 8. N 11. P. 1–5.

5. Gomberg A. M., Marhuenda F., Ortuño-Ortín I. *Endogenous Platforms: The Case of Many Parties* // International Journal of Game Theory. 2005. V. 35. N 2. P. 223–249.

6. Ortuño-Ortín I., Roemer J. E. *Endogenous Party Formation and the Effect of Income Distribution on Policy* // Ivie. 2000. V. AD. N 6. P. 921–935.

7. Sosina Yu. V. *Endogenous Formation of Political Structures and Their Stability* // ORM2004: Труды. М.: МАКС Пресс, 2004 С. 215–216.

8. Vasin A. A., Stepanov D. S. *Endogenous Formation of Political Parties* // Mathematical and Computer Modelling. 2008. V. 48. N 9-10. P. 1519–1526.

TWO TYPES OF PLAYERS IN THE ENDOGENOUS COALITION FORMATION MODEL

Denis S. Stepanov, Moskow State University (dn.step@gmail.com).

Abstract: A model of coalition formation by players whose payoff depends on the value of the parameter (e.g., geographical location, bliss point) is considered. In this model a small portion of the new players with a different payoff function is injected into the main population. This paper considers different types of coalition stability and for each describes corresponding stability criteria. The derived conditions are then compared with the similar criteria in the game with a single type of players.

Keywords: coalition stability, Nash equilibrium, weak coalitional equilibrium (WCE).