

УДК 519.8
ББК 22.18

МОДЕЛИ РЕФЛЕКСИВНЫХ ИГР В ЗАДАЧАХ УПРАВЛЕНИЯ ЭКОЛОГО- ЭКОНОМИЧЕСКИМИ СИСТЕМАМИ

Новиков Д. А.¹, Чхартишвили А. Г.²
(Институт проблем управления РАН, Москва)

Обосновывается возможность и целесообразность использования аппарата теории рефлексивных игр для описания задач принятия решений и управления эколого-экономическими системами.

Ключевые слова: эколого-экономическая система, рефлексивная игра, структура информированности, информационное равновесие.

1. Введение

Задача управления эколого-экономической системой [14, 15], элементы которой способны к целенаправленному поведению [3], с теоретико-игровой точки зрения состоит в том, чтобы создать для управляемых субъектов (агентов) игру с такими правилами [5, 8], чтобы ее исход был как можно более благоприятным для управляющего органа (центра) [7, 13]. Поэтому необходимым этапом решения задачи управления является теоретико-игровой анализ, позволяющий центру спрогнозировать реакцию управляемой системы на те или иные управляющие воздействия.

¹ Дмитрий Александрович Новиков, доктор технических наук, профессор, член-корреспондент РАН, заместитель директора (novikov@ipu.ru).

² Александр Гедванович Чхартишвили, доктор физико-математических наук, ведущий научный сотрудник (sandro_ch@mail.ru).

Одним из методов теоретико-игрового моделирования являются рефлексивные игры (см. [12]), позволяющие учитывать сложную информированность (в том числе взаимную информированность [1, 2, 6]) агентов. На сегодняшний день рефлексивные игры нашли свое применение для описания информированности, совместного принятия агентами решений и решения соответствующих задач информационного управления (управления информированностью агентов) в самых разных областях: корпоративное управление, экономика, маркетинг, политика и пр. [4, 9, 10, 11, 12].

Информированность агентов в рефлексивной игре задается структурой, которую составляют представления о существенных параметрах ситуации, а также о представлениях оппонентов (других агентов). Решением рефлексивной игры является информационное равновесие – набор действий реальных и фантомных (существующих в сознании реальных) агентов, в рамках которого каждый агент максимизирует целевую функцию исходя из своей информированности.

В данной статье рассмотрено несколько моделей эколого-экономических систем, в которых исследуется зависимость исхода взаимодействия агентов от структуры их информированности (информационное равновесие). Если информированность агента является ложной (т.е. агент заблуждается относительно условий игры), то наблюдаемый им результат может как оказаться для него неожиданным, так и соответствовать ожиданиям. Если второе выполняется для всех агентов, то имеет место стабильное информационное равновесие [12], условия существования которого также исследуются в рассматриваемых далее моделях.

2. «Число агентов на рынке»

Пусть имеются n однородных (одинаковых) агентов (занумеруем их элементами множества $N = \{1, \dots, n\}$), выбирающих неотрицательные объемы производства $x_i \geq 0$ и имеющих целевые функции

$$(1) \quad f_i(x) = x_i - (x_i)^2 / 2 - \frac{\chi}{n} \sum_{j \in N} x_j,$$

где $x = (x_1, x_2, \dots, x_n)$, $\chi \geq 0$ – коэффициент штрафов. Первое слагаемое в выражении (1) соответствует выручке агента от продажи произведенной им продукции по единичной цене, второе слагаемое – затратам, третье – штрафам за загрязнение окружающей среды (будем считать, что штраф за суммарное загрязнение, пропорциональное суммарному объему производства, распределяется поровну между агентами). Пусть выражение (1) является общим знанием среди агентов, а последовательность функционирования следующая: агенты одновременно и независимо выбирают объемы производства, после чего каждому из них сообщается размер наложенного на него штрафа.

Если число агентов является среди них общим знанием, то при такой информированности каждый агент выберет действие

$$(2) \quad x_i^* = 1 - \chi / n,$$

максимизирующее его целевую функцию (1).

Рассмотрим возможную информированность агентов относительно их количества n . Если каждый из агентов считает, что их общее число составляет \hat{n} и это является общим знанием, то каждый из агентов рассчитывает на следующий размер своего штрафа:

$$(3) \quad \frac{\chi}{\hat{n}} \left(1 - \frac{\chi}{\hat{n}}\right) \hat{n} = \chi \left(1 - \frac{\chi}{\hat{n}}\right).$$

Наблюдая фактическое значение своего штрафа

$$(4) \quad \frac{\chi}{n} \left(1 - \frac{\chi}{n}\right) n = \chi \left(1 - \frac{\chi}{n}\right),$$

ни один из агентов не имеет оснований усомниться в правильности своих представлений (так как правые части выражений (3) и (4) совпадают). Следовательно, информационное равновесие (2) является стабильным при любом (в том числе ложном) представлении агентов \hat{n} (являющимся среди них общим знанием) об их числе n . Это свойство является следствием того, что штраф пропорционален среднему действию агентов и, следовательно, влияние их количества «компенсируется» суммарным действием.

В заключение настоящего раздела отметим, что вывод о стабильности в том числе ложных равновесий не зависит от параметра χ системы штрафов, т.е. в рассмотренной модели переход к истинному информационному равновесию за счет варьирования системы штрафов невозможен – необходимы меры информационного воздействия на агентов, изменяющих их индивидуальную и/или взаимную информированность.

3. «Совместное производство»

Предположим, что в регионе функционируют n предприятий, выпускающих однородную продукцию и имеющих целевые функции

$$(5) f_i(x) = \lambda x_i - \frac{x_i^2}{2(r_i + \beta X_{-i})} - \chi x_i,$$

где $X_{-i} = \sum_{j \neq i} x_j$, $x_i \geq 0$ – возможные действия агентов; $r_i > 0$ –

типы (параметры) агентов; $\lambda > 0$ – цена единицы продукции; $\chi \geq 0$ – коэффициент штрафов; β – неотрицательный параметр. В соответствии с выражением (5) затраты каждого агента зависят от деятельности других агентов (например, имеет место трансфер технологий).

Предположим, что выражения (5) и все значения всех входящих в них параметров являются общим знанием среди агентов. Обозначим

$$X = \sum_{j=1}^n x_j, R = \sum_{j=1}^n r_j.$$

Напомним, что равновесие Нэша – это набор действий агентов, в котором действие каждого агента максимизирует его целевую функцию (при фиксированных действиях остальных агентов). В данном случае для нахождения равновесия Нэша приравняем к нулю производную целевой функции (5) агента по его действию x_i , что после несложных преобразований приводит к следующему выражению:

$$(6) \quad x_i = \frac{(r_i + \beta X)(\lambda - \chi)}{1 + \beta(\lambda - \chi)}.$$

Суммируя выражения (6) по всем агентам, получаем соотношение, из которого найдем суммарное действие:

$$(7) \quad X = \frac{R(\lambda - \chi)}{1 - \beta(\lambda - \chi)(n-1)}.$$

Подставляя (7) в (6), получаем окончательно:

$$(8) \quad x_i^* = \frac{\lambda - \chi}{1 + \beta(\lambda - \chi)} \left(r_i + \frac{\beta R(\lambda - \chi)}{1 - \beta(\lambda - \chi)(n-1)} \right).$$

Видно (см. выражение (7)), что с ростом цены и/или числа агентов увеличивается значение суммарного равновесного объема производства, а с ростом «силы штрафов» χ этот показатель уменьшается.

Исследуем теперь случай, когда представления агентов относительно величины n и значений $r = \{r_i\}$ могут быть ложными. Поскольку i -й агент знает свои тип и действие, он может рассчитать (см. (5)) значение выражения

$$\beta X_{-i} = \beta(X - x_i) = \frac{r_i \gamma}{1 + \gamma} + \frac{\gamma}{1 + \gamma} \cdot \frac{R}{1 - \gamma(n-1)}$$

(здесь использовано обозначение $\gamma = \beta(\lambda - \chi)$ для сокращения записи).

Из последнего соотношения видно, что если агент считает истинными значениями количества агентов и суммы их типов величины \hat{n} и \hat{R} соответственно, то равновесие является стабильным [12] при выполнении следующего условия:

$$(9) \quad \frac{\hat{R}}{1 - \gamma(\hat{n} - 1)} = \frac{R}{1 - \gamma(n - 1)}.$$

При выполнении условия (9) значение целевой функции совпадает с ожиданиями агента.

4. «Пороговые штрафы»

В настоящем разделе рассматривается модель, которая отражает следующую ситуацию: агенты штрафуются, если резуль-

тат их суммарных усилий $\sum_{i \in N} x_i$ по охране окружающей среды от загрязнения производственными отходами меньше заданного порога. Штраф является настолько крупным, что делает производство нерентабельным, поэтому агентам необходимо избежать его (или отказаться от производственной деятельности). При отсутствии штрафа i -й агент получает доход V_i , $i \in N$.

Результат является возрастающей функцией от суммы усилий, прилагаемых каждым агентом. Будем считать, что агенты могут иметь различные представления о параметрах этой функции и, следовательно, о такой величине порога θ , что штраф не платится при выполнении условия $\sum_{i \in N} x_i \geq \theta$.

Пусть реализация действия $x_i \geq 0$ требует от i -го агента затрат $c_i(x_i, r_i)$, где $r_i > 0$ – его тип (параметр, описывающий индивидуальные характеристики), $i \in N$.

Относительно функций затрат агентов предположим, что $c_i(x_i, r_i)$ – непрерывная возрастающая по x_i и убывающая по r_i функция, причем $c_i(0, r_i) = 0$, $i \in N$.

Обозначим через X' множество всевозможных наборов действий (x_1, \dots, x_n) и определим множество индивидуально рациональных действий агентов:

$$IR = \{x \in X' / \forall i \in N \ V_i \geq c_i(x_i, r_i)\}.$$

Нетрудно видеть, что $IR = \prod_{i \in N} [0; x_i^+]$, где

$$x_i^+ = \max \{x_i \geq 0 / c_i(x_i, r_i) \leq V_i\}, \ i \in N.$$

Обозначим

$$X(\theta) = \{x \in X' / \sum_{i \in N} x_i = \theta\}.$$

Рассмотрим последовательно различные варианты информированности агентов о значении параметра $\theta \in \Theta$.

Вариант I. Значение $\theta \in \Theta$ является общим знанием. Тогда равновесием игры агентов является параметрическое (т.е. зависящее от параметра θ) равновесие Нэша, принадлежащее множеству $E_N(\theta) = IR \cap X(\theta)$.

Вариант II. Представления агентов о величине порога попарно различны, но при этом набор $\{\theta_i\}$ является общим знанием. Иными словами, имеет место асимметричное общее знание.

Не ограничивая общности, занумеруем агентов таким образом, чтобы их представления возрастали: $\theta_1 < \dots < \theta_n$. Структура возможных равновесий в этой ситуации описывается следующим утверждением.

Утверждение 1. Если $\theta_i \neq \theta_j$ при $i \neq j$, то информационным равновесием могут быть (в зависимости от соотношения между параметрами) следующие $n + 1$ исходов: $\{x^* \mid x_i^* = 0, i \in N\}$; $\{x^* \mid x_k^* = \theta_k, x_i^* = 0, i \in N, i \neq k\}, k \in N$. Содержательно это означает следующее: либо никто из агентов не прикладывает усилий, либо усилия прикладывает один k -й агент, выбирая действие θ_k .

В более общем случае $\theta_1 \leq \dots \leq \theta_n$. (представления агентов могут совпадать) может появиться область равновесий аналогично варианту I. Содержательно это означает, что в равновесии усилия прикладывают те агенты, которые одинаково представляют себе величину порога.

Вариант III. Представления агентов о величине порога различны, но каждый агент считает, что играет в игру с асимметричным общим знанием (вообще говоря, будучи ложно информированным о представлениях оппонентов). В этом случае множество возможных равновесных ситуаций становится максимально возможным: $\prod_{i \in N} [0; x_i^+]$. Более того, справедливо следующее утверждение.

Утверждение 2. Для любого вектора действий $x^* \in \prod_{i \in N} [0; x_i^+]$ существует такая структура информированности

(при которой каждый агент субъективно играет в игру с асимметричным общим знанием), что вектор x^* является единственным равновесием.

Доказательства утверждений 1 и 2 аналогичны доказательствам утверждений в [11, раздел 4.10]

5. «Согласование интересов управляющих органов»

Рассмотрим эколого-экономическую систему, состоящую из одного предприятия (агента) и двух управляющих органов - центров. Стратегией агента является выбор объема производства $x \geq 0$ и уровня безопасности $y \geq 0$, что требует от него затрат $x^2 / 2r$ и $y^2 / 2w$ соответственно ($r > 0, w > 0$). Каждый центр получает от деятельности агента «доход», описываемый функцией $H_i(x, y)$, и выплачивает агенту «стимулирование» $\sigma_i(x, y)$, $i = 1, 2$. Таким образом, целевая функция i -го центра имеет вид

$$(10) \Phi_i(\sigma_i(\cdot), x, y) = H_i(x, y) - \sigma_i(x, y),$$

а целевая функция агента:

$$(11) f(\{\sigma_i(\cdot)\}, x, y) = \lambda x - x^2 / 2r - y^2 / 2w + \sigma_1(x, y) + \sigma_2(x, y).$$

Порядок функционирования следующий: центры одновременно и независимо выбирают функции стимулирования и сообщают их агенту, которое затем выбирает свое действие. Ограничимся рассмотрением множества Парето-эффективных равновесий Нэша игры центров, в которых, как показано в [13], их стратегии имеют вид

$$(12) \sigma_i(x', x, y', y) = \begin{cases} V_i, & x = x', y = y', \\ 0, & \text{иначе;} \end{cases} \quad i = 1, 2.$$

Содержательно центры договариваются о том, что будут побуждать агента выбирать объем производства x' и достигать уровня безопасности y' , и осуществлять совместное стимулирование. Такой режим взаимодействия центров называется режимом сотрудничества [13].

Из вида целевой функции (11) следует, что в отсутствие стимулирования агент будет выбирать нулевой уровень безопасности. Найдем оптимальный для агента объем производства $x^* = \arg \max_{x \geq 0} [\lambda x - x^2 / 2r] = \lambda r$. Из условий оптимальности по

Парето следует, что сумма вознаграждений, получаемых агентом от центров в случае выполнения их рекомендаций, равна

$$(13) V_1 + V_2 = \lambda(x^* - x') - [(x^*)^2 - (x')^2] / 2r + y'^2 / 2w.$$

Условие выгоды сотрудничества для каждого из центров можно сформулировать следующим образом: в режиме

сотрудничества каждый центр должен получить полезность не меньшую, чем он мог бы получить, осуществляя стимулирование агента в одиночку. Полезность i -го центра от «самостоятельного» взаимодействия с агентом равна

$$(14) \Phi_i^* = \max_{x, y \geq 0} [H_i(x, y) - \lambda(x^* - x) + [(x^*)^2 - (x)^2] / 2r - y^2 / 2w].$$

Обозначим через

$$(15) S = \{x \geq 0, y \geq 0 \mid \exists (V_1; V_2) \in \mathfrak{R}_+^2: H_i(x, y) - V_i \geq \Phi_i^*, i = 1, 2;$$

$$V_1 + V_2 = \lambda(x^* - x) - [(x^*)^2 - (x)^2] / 2r + y^2 / 2w\}$$

область компромисса – множество таких действий агента, для реализации которых сотрудничество выгодно для центров.

Обозначим

$$(16) \Phi_0^* = \max_{x, y \geq 0} \{H_1(x, y) + H_2(x, y) - \lambda(x^* - x) + [(x^*)^2 - (x)^2] / 2r - y^2 / 2w\}.$$

По аналогии с тем, как это делается в [3, 13], можно доказать, что область компромисса непуста тогда и только тогда, когда

$$(17) \Phi_0^* \geq \Phi_1^* + \Phi_2^*.$$

Приведем пример: пусть $H_1(x, y) = \alpha x + (1 - \alpha)y$, $H_2(x, y) = (1 - \alpha)x + \alpha y$, где $\alpha \in [0; 1]$ – константа, отражающая степень согласованности интересов центров, т.е. «пропорцию», в которой учет экономических и экологических показателей входит в их функции выигрыша (если $\alpha = 0$ или $\alpha = 1$, то один из центров заинтересован только в экономических показателях – объеме производства, а другой только в экологических – уровне безопасности).

Находим из (14) и (16):

$$\Phi_1^* = r\alpha [2\lambda + \alpha] / 2 + w(1 - \alpha)^2 / 2,$$

$$\Phi_2^* = r[\alpha^2 + 1 + 2\lambda - 2\alpha - 2\alpha\lambda] / 2 + w\alpha^2 / 2,$$

$$\Phi_0^* = r[2\lambda + 1] / 2 + w / 2,$$

при этом (17) выполняется как тождество при любых значениях параметров модели (α, r, w). Следовательно, в рассмотренном случае при любой информированности и/или взаимной информированности центров они могут быть уверены, что область компромисса не пуста!

6. Заключение

В настоящей работе рассмотрен ряд простых моделей, иллюстрирующих возможность и целесообразность использования аппарата рефлексивных игр для решения задач описания совместного принятия решений участниками эколого-экономических систем.

Проведенный анализ свидетельствует, что взаимная информированность участников эколого-экономических систем существенным образом влияет на совместно принимаемые ими решения, и, оказывая управленческие воздействия (т.е. изменяя эту информированность), можно изменять равновесные состояния подобных систем.

Литература

1. AUMANN R.J. *Agreeing to Disagree* // The Annals of Statistics. – 1976. – Vol. 4. – No. 6. – P. 1236–1239.
2. AUMANN R.J. *Interactive Epistemology I: Knowledge* // International Journal of Game Theory. – 1999. – No. 28. – P. 263–300.
3. BURKOV V., NOVIKOV D., SHCHEPKIN A. *Control Mechanisms for Ecological-economic Systems*. – Berlin: Springer, 2015. – 174 p.
4. CHKHARTISHVILI A.G. *Concordant Informational Control* // Automation and Remote Control. – 2012. – Vol. 73, No. 8. – P. 1401–1409.
5. GERMEIER YU. *Non-antagonistic Games*. - Dordrecht, Boston: D. Reidel Pub. Co., 1986. – 331 p.
6. LEFEBVRE V. *Lectures on the Reflexive Games Theory*. – N.Y.: Leaf & Oaks Publishers, 2010. – 220 p.
7. *Mechanism Design and Management: Mathematical Methods for Smart Organizations* / Ed. by Prof. D. Novikov. – N.Y.: Nova Science Publishers, 2013. – 204 p.
8. MYERSON R. *Game Theory: Analysis of Conflict*. – London: Harvard Univ. Press, 1991. – 568 p.
9. NOVIKOV D.A., CHKHARTISHVILI A.G. *Information Equilibrium: Punctual Structures of Information Distribution* // Au-

- tomation and Remote Control. – 2003. – Vol. 64, No. 10. – P. 1609–1619.
10. NOVIKOV D.A., CHKHARTISHVILI A.G. *Graph of a reflexive game and belles-lettres* // *Studia Humana*. – 2014. – Vol. 3:3 – P. 11–15.
 11. NOVIKOV D.A., CHKHARTISHVILI A.G. *Mathematical Models of Informational and Strategic Reflexion: a Survey* // *Advances in Systems Science and Applications*. – 2014. – No. 3. – P. 254–277.
 12. NOVIKOV D., CHKHARTISHVILI A. *Reflexion and Control: Mathematical Models*. – Leiden: CRC Press, 2014. – 298 p.
 13. NOVIKOV D.A. *Theory of Control in Organizations*. – N.Y.: Nova Scientific Publishing, 2013. – 341 p.
 14. OUGOLNITSKY G. *Sustainable Management*. – N.Y.: Nova Scientific Publishing, 2012. – 287 p.
 15. WHITELOW K. *Environmental Systems Handbook*. – Oxford: Elsevier, 2004. – 237 p.

MODELS OF REFLEXIVE GAMES IN CONTROL PROBLEMS OF ECOLOGICAL-ECONOMIC SYSTEMS

Dmitry Novikov, Institute of Control Sciences of RAS, Moscow, Doctor of Science, Professor, corresponding member of RAS, deputy director (novikov@ipu.ru).

Alexander Chkhartishvili, Institute of Control Sciences of RAS, Moscow, Doctor of Science, principal research scientist (sandro_ch@mail.ru).

Abstract: We substantiate the feasibility and reasonability of employing the framework of reflexive games for describing decision-making and control problems of ecological-economic systems.

Keywords: ecological-economic system, reflexive game, awareness structure, informational equilibrium.

*Статья представлена к публикации
членом редакционной коллегии М.В. Губко*

Поступила в редакцию 03.02.2015.

Опубликована 31.05.2015.