УДК 519.87,51.77 DOI: http://doi.org/10.25728/pu.2020.4.4

КОНСТАНТНОЕ ПОВЕДЕНИЕ В ДЕЛОВЫХ ИГРАХ РАСПРЕДЕЛЕНИЯ РЕСУРСА: УСТОЙЧИВОСТЬ К ДИЗАЙНУ ИГР И МОДЕЛЬ[1]

В.О. Корепанов

Аннотация. В проведенных деловых играх распределения ресурсов обнаружено «константное» поведение (КП) игроков, когда игроки не меняют своих действий в течение нескольких шагов игры. Особенностью игровых данных является то, что КП занимает в них большую долю. Предложен измененный дизайн игр, где выплаты производятся за каждый из пяти игровых шагов и интерфейс не позволяет игроку легко вводить неизменную заявку. Результаты игр с измененным дизайном показали, что доля КП уменьшилась, но остается на достаточно высоком уровне. Далее представлены результаты поиска причины данного поведения с помощью статистических гипотез и решения задач классификации решений игроков. Среди статистических гипотез приняты гипотезы о случайной природе КП и остановки КП в случае падения выигрыша. Задача классификации позволяет провести отбор информативных признаков (параметров и истории игры) и правило принятия решения игроками (классификатор) об остановке КП на их основе. Результаты построения классификатора дали идею усложнения модели: у игрока есть не только принцип остановки/продолжения КП, но и принцип начала КП. Также показано, что результаты статистического исследования и классификации имеют общие черты и могут в дальнейшем дополнить друг друга.

Ключевые слова: деловые игры, экспериментальная экономика, модели поведения, задача распределения ресурса.

ВВЕДЕНИЕ

В работе [1] описано проведение деловых игр по распределению ограниченного ресурса на основе механизма Гровса — Лейдярда при трансферабельной полезности (полезность можно «переносить» между игроками). Эксперимент тогда и далее проводился в лаборатории, участники знакомились с условиями игры, с самой игрой и играли без разговоров между собой, каждый за своим компьютером. Для принятия решений от игроков и сбора данных игры применялась система создания и проведения экономических экспериментов zTree. Проводились игры со студентами экономических и технических специальностей.

В работе [1] была обнаружена большая доля константного поведения (КП) у игроков — когда в течение нескольких шагов игры действие игрока не меняется. Причем в большинстве случаев такое поведение не соответствует ни действию, представляющему собой наилучший ответ на предыдущие действия, ни даже движению в сторону такого действия [1].

Понять причины такого поведения важно, так как:

— итоговые распределения ресурсов в играх неэффективны, хотя используемые механизмы распределения ресурсов в основном обладают очень хорошими теоретическим свойствами: единственность и устойчивость равновесия Нэша, сходимость к равновесию Нэша моделей поведения и обучения [1, 2],

— обнаружена проблема манипулируемости механизмов распределения ресурсов [3, 4],

— важно фундаментальное понимание поведения людей в деловых играх.

В общем случае причина того или иного поведения кроется либо в самих игроках, либо в элементах дизайна игры, в которой участвуют игроки. К элементам дизайна игр относятся, например,

[1] Работа выполнена при поддержке гранта РФФИ № 17-07-01550 А.

наличие ознакомительной лекции, число ознакомительных шагов игры, число основных шагов игры, правила вознаграждения за участие в игре, информация, которая дается игрокам в игре и т. п. В § 3 данной работы мы оценим влияние дизайна проведенных игр на наличие КП.

Если искать источник поведения в самих игроках, наиболее близко к исследуемой модели КП находятся модели обучения из экспериментальной экономики (см., например, работы [5—7]), которые можно поделить на два класса.

К первому классу относятся модели на основе представлений (beliefs-based) — игроки формируют представления о поведении оппонентов и на их основе выбирают свое поведение. К этому классу принадлежат модели Курно [8] и Fictitious Play [9, 10] и дальнейшие их обобщения [11—13], построенные для преодоления слабых сторон этих моделей.

Параллельно развивались модели, не основанные на представлениях, а развивавшиеся на идеях обучения с подкреплением (reinforcement learning, RL). Идея обучения с подкреплением в том, что игрок смотрит не на поведение других игроков, а на свой выигрыш от выбранных действий и с большей вероятностью выберет такое действие, которое давало ему больший выигрыш в прошлом. В данных моделях игроку не нужно знать информацию об остальных игроках, при достаточных предположениях знание только своего выигрыша от истории игры позволяет корректировать выбор действия на следующем шаге и находить оптимальную стратегию, по крайней мере, в пределе по шагам. Примеры моделей данного класса в работах [14—17].

Модели обучения допускают, чтобы игроки не меняли свое действие в течение некоторого времени, но идея, из-за которой это происходит, одна — «стремление к действию» или «привлекательность» действия, приносившего в прошлом больший выигрыш. Другими словами, чаще должно выбираться то действие, которое в прошлом приносило выигрыш больше, чем другие действия, поэтому, в том числе, оно может выбираться неизменно в течение нескольких шагов игры. Такой режим в моделях обучения называют «эксплуатацией». При этом выбор действия вероятностен, а вероятности действий определяются через привлекательности действий — на каждом шаге режима, в принципе, может быть выбрано разное действие.

Наиболее близкая модель, которая явно вносит режим КП — Inertia sampling and weighting (I-SAW) [18]. В ней есть три режима: Исследование, Эксплуатация и Инерция. Как и в основной массе моделей обучения, в режиме исследования равновероятно выбирается любое из возможных действий игрока. Константное поведение в модели I-SAW может ассоциироваться с режимом инерции, в котором есть вероятность повторения именно последнего действия, в отличие от режима эксплуатации, в котором последнее действие может повториться только случайно (хотя, может быть, и с большой долей вероятности, если оно приносило в прошлом большой выигрыш). При этом вероятность оказаться в этом режиме уменьшается в степенной зависимости при увеличении размера «сюрприза»: оценки различия между ожидаемым выигрышем и полученным на прошлом шаге.

Важно отметить, что в данной работе не предлагается полная модель поведения, в отличие от упомянутых выше, а предлагаются результаты по поиску только моделей, объясняющих моменты КП игрока: сначала вероятностных, затем, основанных на использовании информации, доступной игрокам, и построении на ее основе классификаторов, которые были бы моделью принятия решения на следующем шаге: «продолжать КП — не продолжать КП». Эти результаты могут быть далее встроены, например, в модель I-SAW на место режима Инерция, и их эффективность проверена с помощью стандартных методов.

Из неполноты модели вытекает и то, что мы пока не применяем методы идентификации правил обучения, принятые в экономической литературе [19, 20] — эти методы применяются для моделей поведения, которые стремятся описать поведение на каждом шаге игры, мы же берем только шаги с неизменными действиями с заданной точностью и на них проверяем модель. Данные методы, наверное, можно применить и для модели КП на части шагов, но будет под вопросом сравнение полученного значения критерия эффективности модели КП на части шагов со значениями критерия для других моделей, посчитанным на данных всех шагов.

В данной работе будет показано, что у КП, обнаруженных на имеющихся данных, моделей для объяснения КП может быть две — одна вероятностная, другая на основе классификатора. Во втором случае КП не может быть сведено полностью к известным моделям обучения.

Работа организована по такой схеме:

— описание деловых игр;

— модель константного поведения;

— проверка влияния ряда факторов дизайна игры на наличие константного поведения. Проводится деловая игра с измененным дизайном — экспериментальным (для сравнения с базовым дизайном):

— модель КП с режимами, проверка ряда гипотез о КП;

— задачи классификации для модели КП с режимами, попытка выделить значимые признаки для объяснения КП.

В работе использовались данные деловых игр (действия игроков на шагах игры) для базового дизайна игры, деловых игр для экспериментального дизайна игры, анкетирования игроков после игры.

1. ОБЩЕЕ ОПИСАНИЕ ДЕЛОВЫХ ИГР

Проводились игры по распределению делимого ограниченного ресурса $R = 115$ между тремя игроками. Каждый игрок i делает заявку $s_i(\tau)$ на каждом шаге игры $\tau \in \{1, ..., T\}$, после подачи игроками своих заявок $s(\tau) = (s_1(\tau), s_2(\tau), s_3(\tau))$ происходит расчет: ресурса, назначаемого игрокам $x_i(s(\tau))$ (шаг τ далее опускаем, если понятно о чем речь), полезностей игроков $u_i(x_i(s)) \in \mathbb{R}$, штрафов $p_i(s) \in \mathbb{R}$ и выигрышей игроков $g_i = u_i - p_i$. Далее игра переходит на следующий шаг или заканчивается. В качестве результата игры для игрока брался его выигрыш на последнем шаге.

В качестве функций полезности игроков принимается функция от полученного ресурса $u_i(x_i) = \sqrt{r_i + x_i}$, где r_i — тип игрока i или базовое количество ресурса, имеющееся у игрока. У игроков были типы 1, 9 и 25, но типы игрокам назначались случайно в каждой игре.

Рассматриваемый в данной работе механизм распределения ресурса Yang-Hajek (YH) [2], представляется таким образом:

$$s_i \in [0, 1000],$$

$$x_i = \begin{cases} \dfrac{s_i}{S} R, & S > 0, \\ 0, & S = 0, \end{cases}$$

где $S = s_1 + s_2 + s_3$,

$$p_i = \beta s_i S_{-i},$$

где $S_{-i} = S - s_i$, $\beta = 5 \cdot 10^{-4}$.

В базовом дизайне число шагов в игре ограничивалось 60 шагами, но игра могла закончиться досрочно, если никто из игроков не поменял свои заявки. Всего было проведено 13 игр с механизмом YH с общим числом шагов 330, а если считать число шагов для каждого игрока в отдельности, то получится $330 \cdot 3 = 990$. На рис. 1 представлен график продолжительности игр (число шагов в каждой игре). Общее число участников игр — 39.

Общий процесс деловой игры состоял в следующем: рассказ группе студентов об игре и процессе экспериментов, ознакомительная игра, показ результатов игры, основная игра, обсуждение итогов и выигрышей, выплата вознаграждений. Перед ознакомительной и основной игрой студенты садятся каждый за свой стол с компьютером, затем, с помощью функционала zTree, они случайным образом группировались по три человека в группе и случайным образом в каждой группе игрокам назначались их типы из множества {1, 9, 25} без повтора. В ознакомительной игре студентам разрешалось устное общение для обмена опытом, совместного исследования игры, в основной игре — было запрещено. Вознаграждение выплачивалось в размере 20—30 руб. за единицу выигрыша.

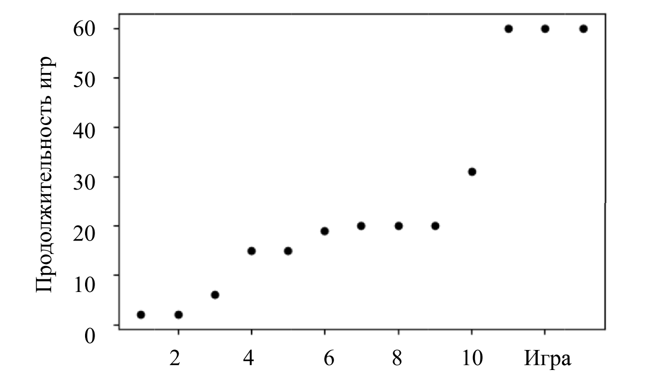

Рис. 1. Продолжительность игр

В процессе ознакомительной или основной игры игроки знали: общий вид функций полезности друг друга, функции распределения ресурса и функции штрафов, свои типы, распределения ресурса и собственные выигрыши на предыдущих шагах игры в числовой форме и в форме графиков, условия остановки игры. Игроки не знали: заявки, поданные другими, типы других игроков (кроме информации, что они из множества {1, 9, 25} и не равны типу самого игрока), выигрыши, штрафы и др.

Игры создавались и проводились с помощью программы zTree [21], обработка данных осуществлялась в основном с помощью MS Excel и библиотек numpy, pandas, scikit-learn на языке программирования python3 в среде Jupyter.

2. КОНСТАНТНОЕ ПОВЕДЕНИЕ

При анализе поведения людей в условиях дискретного пространства действий зачастую в модель добавляют некую случайность, которая позволяет учесть случайность в экспериментальных данных, устойчивость стратегий. Мы учтем случайность поведения в константном поведении с помощью параметра погрешности.

Интуитивно КП можно определить как шаги, при которых игрок не изменил свою заявку с прошлого шага с заданной абсолютной погрешностью $\varepsilon \geq 0$. Это дает понять суть КП, но при погрешности $\varepsilon = 1$ последовательность заявок (1, 2, 3, 4, ...) будет представлять КП. Скорее, нужно говорить о том, что КП — это набор заявок, не меняющийся *в целом* больше, чем заданная погрешность. Такая модель похожа на модели «дрожания» (tremble) при выборе игроком заявки [20].

Обозначим $s_i(a, b) = \{s_i(\tau) | a \le \tau \le b\}$, $[a, b] = \{a, a+1, ..., b\}$, $|A| = \max(A) - \min(A)$, для конечных множеств $A \subseteq \mathbb{R}^1$. Дадим более точное определение.

Определение 1. Игрок i проявляет константное поведение $CB(i, t_s, t_e, \varepsilon)$ с погрешностью $\varepsilon \ge 0$ с шага $t_s + 1$ до шага $t_e > t_s$, если:

$$|s_i(t_s, t_e)| \le \varepsilon, \quad (1)$$

$$\forall a < t_s, \quad b > t_e: |s_i(a, t_e)| > \varepsilon \vee |s_i(t_s, b)| > \varepsilon, \quad (2)$$

не существует другого КП $CB(i, t_a, t_b, \varepsilon)$, удовлетворяющего условиям (1) и (2) такого, что

$$t_a > t_s \text{ и } [t_a, t_b] \cap [t_s, t_b] \ne \varnothing. \ \blacklozenge \quad (3)$$

Смысл условий (1) и (2) прозрачен, условие (3) устраняет неоднозначность выбора КП условиями (1) и (2) в сторону первого шага. Для пояснения возьмем последовательность заявок (10, 11, 12, 13, 14). Для погрешности $\varepsilon = 2$ в этой последовательности можно выделить два набора заявок, удовлетворяющих условиям (1) и (2): {(10, 11, 12), (13, 14)} и {(11, 12, 13)}. Условие (3) ограничивает выбор только заявками КП {(10, 11, 12), (13, 14)}.

Определение 1 аналогично результатам «алгоритма» последовательного поиска КП от первого шага каждой игры к последнему: ищем, начиная с первого шага, последовательности заявок игроков, максимум и минимум которых отличаются не больше, чем на ε.

Последовательность $s_i(t_s, t_e)$ будем называть *заявками КП*. Множество всех КП погрешности ε обозначим $C(\varepsilon)$, длиной КП назовем $L = t_e - t_s$. Заметим, что длина КП, например, $CB(1, 2, 5, 1)$ будет равна 3 (шаги 3, 4 и 5), и шаг 2 не входит в КП, но важен как отправная точка, от которой началось КП.

В проведенных играх число КП и суммарное число шагов в них в зависимости от точности представлено в табл. 1.

Уже при погрешности 0 число шагов в КП достигает почти треть всех 978 шагов, т. е. КП важно для описания поведения игроков в базовом дизайне игры.

3. ВЛИЯНИЕ ДИЗАЙНА ИГРЫ НА НАЛИЧИЕ КОНСТАНТНОГО ПОВЕДЕНИЯ

В работе [1] были проанализированы данные для двух механизмов распределения ресурсов из одного класса, позднее были проведены игры с механизмами других классов [4], и КП также занимало большую долю среди действий игроков.

Можно проверить гипотезу, что сам базовый дизайн игры стимулировал подобное поведение — игроки получали итоговый выигрыш только по результату последнего шага игры, и последний шаг игры мог определяться динамически — шаг, в котором никто из игроков не поменял своего действия — условие, позволявшее игрокам прийти к удовлетворяющему всех распределению ресурса. Поэтому встал вопрос о влиянии дизайна игры на полученные результаты. Помимо введения фиксированного числа шагов, решили назначать итоговый выигрыш по суммарному выигрышу за каждый шаг игры, чтобы избежать аномального поведения игроков на последнем шаге. Возможно, также повлиял интерфейс игры, так как по умолчанию в поле ввода заявки игроку ставилась заявка с прошлого шага, таким образом, у игрока мог быть стимул для «лени».

Были проведены игры с измененным дизайном, игроки были другие, но все также студенты экономических и технических специальностей. Назовем дизайн «экспериментальным», так как он ближе к дизайну игр экспериментальной экономики [22, 23]:

— пять обучающих шагов;

— пять основных шагов;

— вознаграждение игроку за игру было пропорционально суммарному выигрышу за *основные шаги* игры;

— в поле ввода заявок устанавливался 0 по умолчанию на каждом шаге.

Результаты по числу константного поведения и числу шагов в КП для игр с механизмом Yang-Hajek (YH) [2] и Grovse-Ledyard (GL) [1] представлены в табл. 2 и 3 и на рис. 2.

Хотя общее число шагов в случае экспериментального дизайна на два порядка меньше, качественно можно сказать, что:

— доля КП с погрешностью 0 становится меньше в 1,5—4,5 раз (насколько меньше, зависит так-

Таблица 1

Число контактного поведения и суммарное число шагов в КП

ε	0	1	2	3	4	5	6	7	8	9	10
Число КП	131	180	192	198	203	229	221	214	212	202	202
Число шагов в КП	295	415	465	498	526	607	613	620	629	635	688

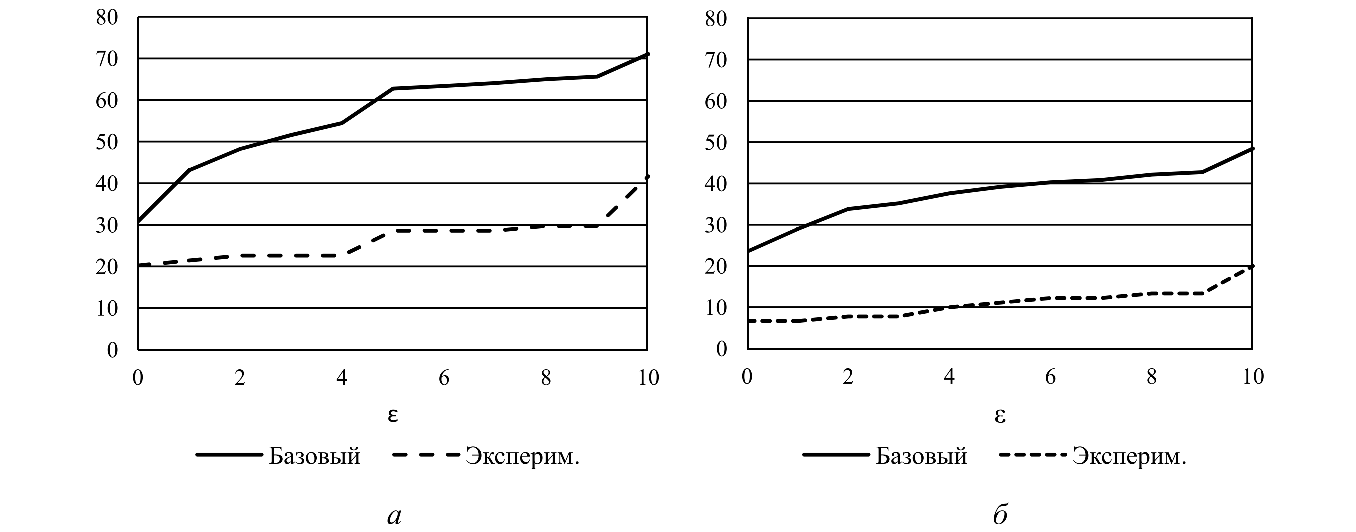

Рис. 2. Процент шагов во всех КП для базового и экспериментального дизайна в механизмах YH (*а*) и GL (*б*) в зависимости от погрешности КП

же от механизма), но остается существенной в поведении игроков для механизма YH;

— доля КП быстро растет при увеличении погрешности от 0 до 3 в случае базового дизайна и почти не меняется в случае экспериментального дизайна.

Первый вывод говорит о том, что обнаруженное КП сильно зависело от дизайна экспериментов, но не полностью им объясняется, т. е. оно важно для описания поведения игроков. Второй вывод наталкивает на мысль о влиянии базового дизайна — можно предложить гипотезу, что рост доли КП с погрешностью больше 0 связан с моделью поведения, в которой игроки меняли немного свои действия, чтобы посмотреть на его эффективность, не останавливая игру, так как если все игроки не поменяют своих действий в базовом дизайне, то игра остановится.

В итоге получается, что доля КП оказалась завышена в базовом дизайне; после изменения дизайна игры видим, что доля КП снизилась, оставшись при этом на заметном уровне для механизма YH (минимум 20 %). Интересна зависимость КП от применяемого механизма — в механизме GL

Таблица 2

Сравнение числа шагов в КП в играх с механизмом Yang-Hajek в случаях базового (Б) и экспериментального (Э) дизайна

Дизайн	Всего шагов	КП										
		ε = 0	ε = 1	ε = 2	ε = 3	ε = 4	ε = 5	ε = 6	ε = 7	ε = 8	ε = 9	ε = 10
Б	978	295	415	465	498	526	607	613	620	629	635	688
	%	30	42	48	51	54	62	63	63	64	65	70
Э	90	17	18	19	19	19	24	24	24	25	25	35
	%	19	20	21	21	21	27	27	27	28	28	39

Таблица 3

Сравнение числа шагов в КП в играх с механизмом Grovse-Ledyard в случаях базового (Б) и экспериментального (Э) дизайна

Дизайн	Всего шагов	КП										
		ε = 0	ε = 1	ε = 2	ε = 3	ε = 4	ε = 5	ε = 6	ε = 7	ε = 8	ε = 9	ε = 10
Б	2346	545	672	785	817	873	909	935	948	978	992	1125
	%	23	29	34	35	38	39	40	41	42	43	48
Э	96	6	6	7	7	9	10	11	11	12	12	18
	%	7	7	8	8	10	11	12	12	13	13	20

доля КП в экспериментальном дизайне резко падает, возможно, это связано с качественным отличием пространств заявок: $\mathbb{R}^3$ в GL и $\mathbb{R}$ в YH.

Доли КП с погрешностью 0 меньше всего отличаются в базовом и экспериментальном дизайнах, к тому же при таком поведении ясно, что на игрока не влияет условие остановки игры базового дизайна, если бы игрок боялся остановки игры, то воспользовался бы КП с погрешностью больше 0. Поэтому далее будет проводиться анализ КП с погрешностью 0, если иное не оговорено особо.

4. МОДЕЛЬ КОНСТАНТНОГО ПОВЕДЕНИЯ С РЕЖИМАМИ. ГИПОТЕЗЫ

В данном параграфе предлагается модель КП и базовые гипотезы для ее описания. В § 3 речь шла только об анализе тех шагов, на которых игроки не меняли своих действий с заданной погрешностью. Здесь предлагается модель, в которой КП только часть более сложного поведения, и нужно понять, когда игрок переходит к ней и когда уходит от нее.

Базовая идея сложного поведения игрока состоит в том, что оно может быть разбито на два режима: наблюдение и выбор. В режиме наблюдения игрок не меняет (с заданной погрешностью) свое действие в течение нескольких шагов (собственно, проявляет КП), а в режиме выбора он выбирает следующее действие. В этом случае игрок должен иметь, как минимум, два принципа принятия решения: *принцип окончания КП* и *принцип выбора действия;* т. е. предполагаем, что игрок действует таким образом:

— на первом шаге выбирает свое первое действие с помощью принципа выбора действия;

— на втором шаге применяет принцип окончания КП; если он сработал, то опять выбирает действие с помощью принципа выбора действия, если он не сработал — меняет свое действие случайно в рамках заданной погрешности КП;

— остальные шаги аналогичны второму шагу.

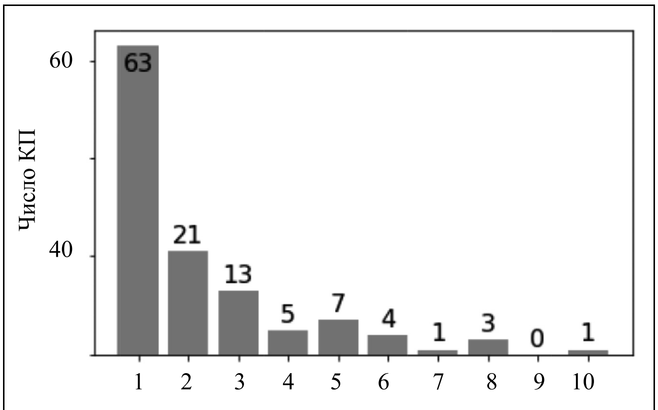

Рис. 3. Гистограмма длин КП

Принцип выбора действия в данной работе не рассматривается. Для принципа окончания КП было предложено несколько гипотез: Г1 — фиксированное число шагов; Г2 — случайное число шагов; Г3 — уменьшение выигрыша; Г4 — уменьшение выигрыша ниже заданного порога.

Заметим, что гипотезы Г2 и Г4 в каком-то смысле уточняют гипотезы Г1 и Г3.

В работе [24] показано, что гипотеза Г1 не верна, гипотеза Г2 кажется правдоподобной (но не была доказана статистически), а также гипотеза Г3 для КП с погрешностью 0 кажется правдоподобна. На рис. 3 представлена гистограмма длин КП.

Статистически гипотеза Г2 действительно верна. Мы подобрали визуально, что логарифмическое распределение Log(p) хорошо описывает распределение длин КП, метод максимального правдоподобия позволил найти оптимальный параметр распределения $p \approx 0{,}7646$. Далее для проверки гипотезы, что эмпирическое распределение совпадает с логарифмическим (по критерию согласия Пирсона, см., например, книгу [24]), длины КП были сгруппированы в 5 групп ($\{L = 2\}$, $\{3\}$, $\{4\}$, $\{5\}$, $\{6$ и более$\}$), чтобы в каждой группе ожидаемое и наблюдаемое число реализаций было не меньше 5. Тест Σ^2 с уровнем значимости 0,05 и степенью свободы 3 (5 — 1 — 1 (число параметров теоретического распределения)) показал значение статистики $\approx$1,844, что меньше расчетного значения 11,07, а p-критерий $\approx$0,61. Таким образом, гипотеза, что распределение длин КП имеет логарифмическое распределение, не может быть отвергнута.

Обозначим через $g_i(\tau)$ — выигрыш игрока i на шаге τ. Для проверки гипотезы Г3 примем, что различие выигрышей игрока в начале и в конце КП отражается на различии усредненных выигрышей первой и второй половины КП. Поэтому разделим каждое константное поведение $i(i, t_s, t_e, \varepsilon)$ и его первый шаг t_s на первую половину M_1 и вторую M_2 и возьмем от них средние:

$$h = \left\lfloor \frac{t_e - t_s + 1}{2} \right\rfloor, \quad H = \left\lceil \frac{t_e - t_s + 1}{2} \right\rceil + 1,$$

$$M_1 = \{\overline{g_i(t_s, h)}\,|\forall(i, t_s, t_e, \varepsilon) \in C(\varepsilon)\},$$

$$M_2 = \{\overline{g_i(H, t_e)}\,|\forall(i, t_s, t_e, \varepsilon) \in C(\varepsilon)\},$$

$$M = \{\overline{g_i(t_s, h)} - \overline{g_i(H, t_e)}\,|\forall(i, t_s, t_e, \varepsilon) \in C(\varepsilon)\},$$

где $g_i(a, b) = \{g_i(\tau) | a \le \tau \le b\}$.

Говоря по-простому, если во множестве M больше положительных, чем отрицательных элементов, то в первой половине большинства КП сред-

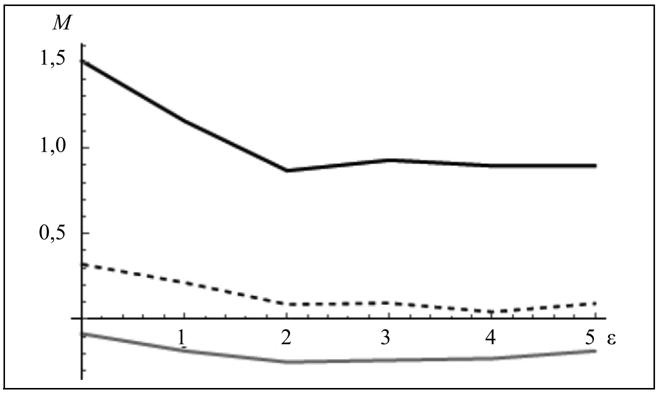

Рис. 4. Квантили и медиана для множества *M*: 0,25-квантиль (*серая линия*); медиана (*пунктирная линия*); 0,75-квантиль (*черная линия*)

ний выигрыш игроков был больше, чем во второй, поэтому, мы не отвергаем гипотезу Г3.

На рис. 4 представлены линии 0,25- и 0,75-квантилей и медиана для множества M. Видно, что для $\varepsilon > 0$ медиана близка к нулю, хотя нижний квартиль примерно в 4 раза меньше верхнего по абсолютной величине. В случае $\varepsilon = 0$ даже нижний квартиль почти равен нулю, что говорит о заметном падении выигрыша в этом случае.

Статистический U-критерий Манна — Уитни для всех ε подтверждает нулевую гипотезу, что M_2 меньше, чем M_1 при уровне значимости 0,05. Таким образом, даже в целом, если рассматривать M_1 и M_2 как независимые выборки, нельзя отвергнуть гипотезу, что $M_2 < M_1$. Но на самом деле от каждого КП в M_1 и M_2 идут средние первой и второй половины, соответственно, поэтому это не независимые выборки. Поэтому мы применили критерий Уилкоксона для связных выборок (см., например, книгу [25]), который при таком же уровне значимости 0,05 подтверждает нулевую гипотезу, что среднее значение выигрыша в КП меньше во второй его половине (значения в M_2 меньше соответствующих им значений из M_1).

С гипотезой Г4 в работе [24] была допущена ошибка, после устранения которой она кажется правдоподобной. Будем считать, что КП имеет нижний порог δ_τ по выигрышу, с периодом $\tau > 0$, если только в конце КП наблюдаются выигрыши меньше δ_τ в течение τ шагов; формально:

$$\forall t > t_\delta = t_e - \tau\colon g_i(t) < \delta_\tau, \tag{4}$$

$$\forall [k_s, k_e] \subset [t_s, t_e - \tau]\colon$$
$$(\forall t \in [k_s, k_e]\colon g_i(t) < \delta_\tau) \Rightarrow (k_e - k_s - 1 < \tau). \tag{5}$$

Практически, для заданного извне периода τ, кандидатом на нижний порог по выигрышу КП будем брать значение $\delta_\tau = \max(g(t_e - (\tau - 1)), ..., g(t_e))$ и проверять его на выполнение условий (4)—(5). Важно заметить, что нижний порог в гипотезе Г4 не одинаков для всех КП — в каждом КП он может быть разным.

На рис. 5, *а* показан график доли КП в множестве $C(0)$, имеющих нижний порог по выигрышу в соответствии с выражениями (4)—(5) в зависимости от заданного периода. При расчете доли рассматриваются только КП с длиной больше, чем заданный период (рис. 5, *б*).

Гипотеза Г4 выглядит правдоподобно — число КП, имеющих нижний порог по выигрышу, колеблется от 60 до 90 % для значений периода от 1 до 5. Сами значения порогов представлены на рис. 6 в форме гистограммы.

Таким образом, не подтвердилась гипотеза Г1, статистически верна гипотеза Г2, расширяющая гипотезу Г1. Также статистически верна гипотеза

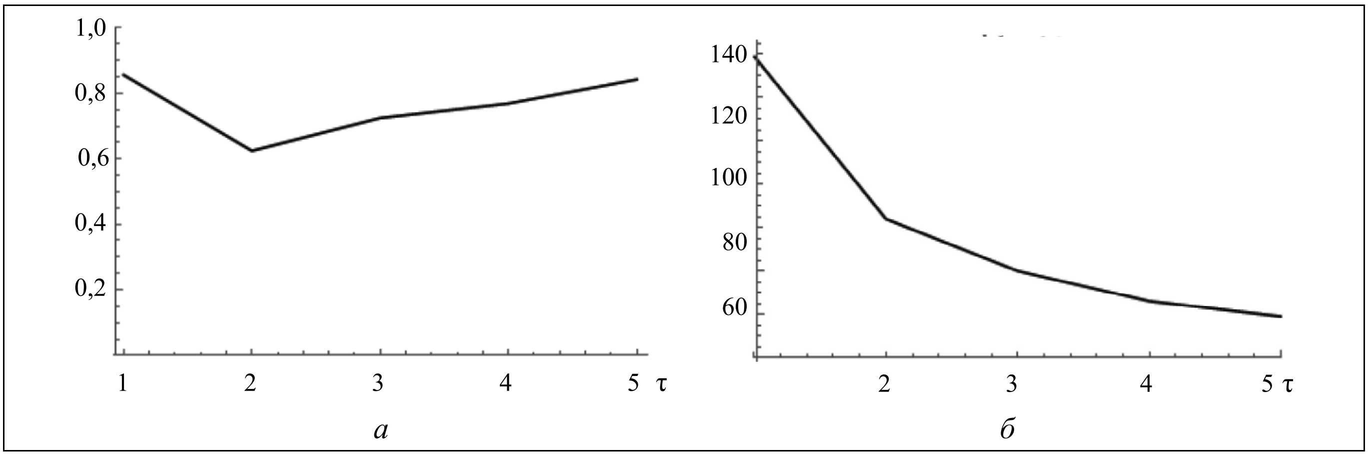

Рис. 5. Доля КП, имеющих порог по выигрышу в множестве *C*(0) (*а*) и число КП, длина которых больше заданного периода (*б*)

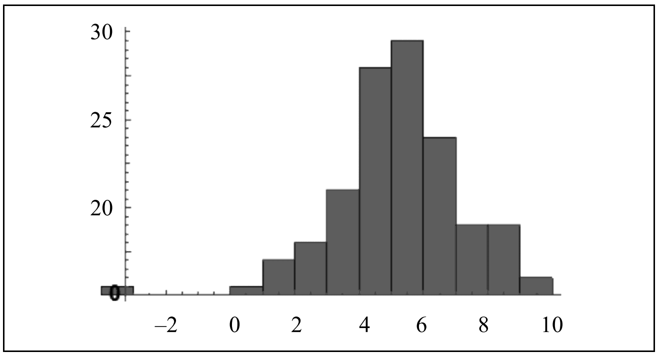

Рис. 6. Гистограмма порогов по выигрышу для периода 1

Г3 и выглядит правдоподобно гипотеза Г4. При этом гипотезы Г2 и Г4 позволяют (в отличие от гипотезы Г3) построить модель принципа окончания КП: в гипотезе Г2 игрок выбирает длину КП как значение реализации случайной величины, в гипотезе Г4 игрок выбирает сначала порог по выигрышу и период, затем, если в течение выбранного периода получаемый выигрыш меньше порога, то игрок заканчивает режим наблюдения. Для второго случая открытым остается вопрос о том, как игрок выбирает порог и период для КП.

В § 5 также проводится попытка построить модель принципа окончания КП, но только с помощью другого подхода — задачи классификации.

5. ЗАДАЧИ КЛАССИФИКАЦИИ ДЛЯ МОДЕЛИ КОНСТАНТНОГО ПОВЕДЕНИЯ С РЕЖИМАМИ

В связи с разделением поведения игрока на «режимы», для выхода из режима наблюдения игрок может на основании доступной ему информации в игре (полученный ресурс, выигрыш, штраф, число шагов и т. п.) принять решение об окончании КП или его продолжении. Причем, из имеющихся данных нам известен выбор игроков (завершить/продолжить КП), но неизвестны точно признаки (информация), на основе которых они принимают решение и сама модель выбора.

Задача построения правила для отнесения набора признаков к одному из решений (завершить/продолжить КП) на основе имеющихся данных (набор признаков, решение) относится к задаче классификации, рассматриваемой среди задач машинного обучения [25], в области машинного обучения также есть задача отбора признаков по их информативности. Мы воспользуемся методом случайных лесов как для отбора признаков, так и для построения классификатора. Этот метод, вероятно, не лучший, но с его помощью попробуем провести начальный анализ.

Простыми словами, для рассматриваемой задачи классификатор — это функция, которая преобразует набор признаков на входе (например, действие, полученный ресурс и выигрыш на прошлом шаге) в выбор игрока на текущем шаге. Эффективность построенного классификатора оценивается как доля посчитанных классификатором решений игроков, которые совпали с реальными решениями игроков в игре. Информативность признака оценивается как вклад в эффективность, который дает признак, если его учитывать при построении классификатора по сравнению с классификатором, не использующим данный признак.

Рассмотрим набор признаков w_1 на основе доступных игроку данных: выигрыши $g_i(t)$, штрафы $p_i(t)$, получаемый ресурс $x_i(t)$ и их изменение на текущем шаге и на предыдущих (вплоть до трех шагов назад): $\Delta g_i(t) = g_i(t) - g_i(t-1)$, $\Delta p_i(t) = p_i(t) - p_i(t-1)$, $\Delta x_i(t) = x_i(t) - x_i(t-1)$, $\Delta g_i(t-1)$, $\Delta p_i(t-1)$, $\Delta x_i(t-1)$, $\Delta g_i(t-2)$, $\Delta p_i(t-2)$, $\Delta x_i(t-2)$. Ранее мы, зная начало каждого КП, пытались построить классификатор, который бы по этим признакам отвечал на вопрос — был ли прошлый шаг началом очередного КП или, другими словами, будет ли действие игрока на следующем шаге равно предыдущему действию или нет? Точнее, пусть множество P_1 содержит только шаги игр, начиная с 4-го (для расчета, например, $\Delta g_i(t-2)$) и заканчивая предпоследним шагом в игре (чтобы всегда знать действия игрока на следующем шаге), которые либо не принадлежат ни одному КП, либо являются первым шагом какого-то КП. Для построения классификатора брались признаки w_1 для каждого шага $t \in P_1$ и выбор игрока на шаге $(t+1)$ — продолжать или нет КП.

В работе [24] представлены результаты построения такого классификатора и оценок информативности этих признаков для классификации. Оказалось, что эффективность классификаторов низка, как и информативность признаков. Причем была попытка улучшить классификатор с помощью различных модификаций признаков (например, вместо прироста брать точное значение на прошлом шаге, стандартизация признаков, расчет полного ресурса (тип + ресурс) игроков и т. п.).

Попробуем провести подобную же классификацию на других исходных данных: множество P_2 содержит шаги всех КП и первый шаг после каждого КП, если он есть (игра не закончилась на КП). Набор признаков w_1 оставляем тем же. Эти данные содержат в себе как шаги, после которых игроки решили продолжить КП, так и шаги, после которых прервали КП. Таким образом, можно пытаться понять, на какие признаки игроки опирались при принятии этих решений об окончании КП.

Отбор признаков на основе алгоритма «случайного леса» [26] выявил три признака, выделяющиеся информативностью: прирост штрафа, прирост ресурса и прирост выигрыша на предыдущем шаге (рис. 7).

Построенный по алгоритму «случайного леса» классификатор на этих трех признаках дает точность классификации приблизительно 75 % на тестовой выборке. При этом для двух наиболее информативных признаков данные неплохо разделены на плоскости (рис. 8): черные точки — признаки, после которых КП продолжается, серые — КП прерывается. Шаги, продолжающие КП, находятся в областях, где $\Delta x(t) \cdot \Delta p(t) < 0$, а больше половины шагов, прерывающих КП, находятся в противоположных областях.

При попытке построить классификатор на шагах множества P_1 мы скорее пытались найти принцип «начала наблюдения» в КП, а не принцип окончания наблюдения. Данные результаты говорят, что, прежде всего, у игроков есть не только принцип окончания наблюдения, но и, видимо, принцип старта наблюдения, поскольку один и тот же подход к классификации дал разные результаты для первых (см. работу [24]) и последних шагов КП, а также, что принцип окончания наблюдения может основываться на небольшом числе признаков «прирост штрафа», «прирост ресурса» и «прирост выигрыша на прошлом шаге».

Дальнейшее уточнение этих результатов может идти в сторону отбора лучших признаков, так как случайные деревья снижают информированность коррелированных признаков, или в сторону применения методов выделения признаков (сокращения размерности данных).

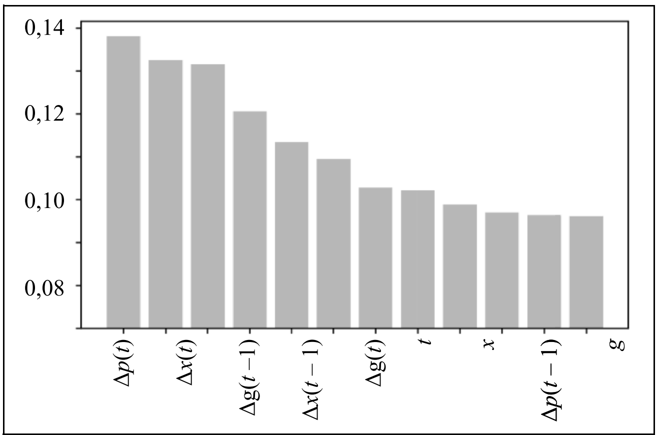

Рис. 7. Информативность признаков

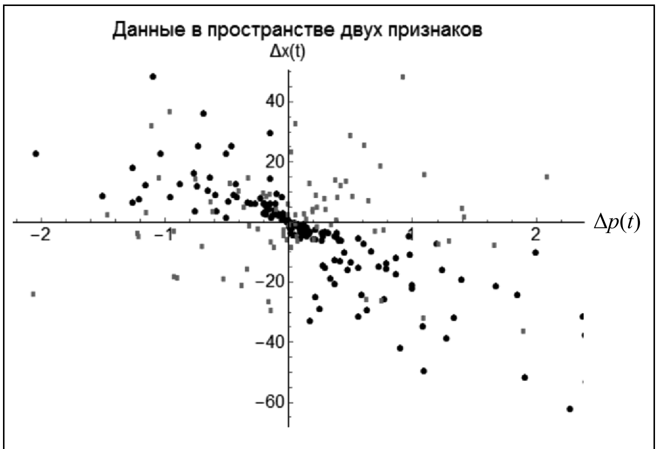

Рис. 8. Данные шагов в пространстве двух наиболее информативных признаков

6. ОБСУЖДЕНИЕ

В § 2 и 3 показано, что КП занимает важную долю в поведении игроков, но в то же время сильно зависит от дизайна игры и, что оказалось побочным результатом, менее сильно от применяемого механизма.

Далее в работе введена сложная модель поведения с двумя режимами: наблюдение (КП) и выбор, исследования пока проведены только для режима «наблюдение», предполагается, что режим «выбор» может быть взят из известных моделей поведения. Автор благодарен рецензенту за вопрос об исследовании содержательности модели КП с помощью анкетных данных. В анкетах было текстовое поле о принципах принятия решений, но игроки могли игнорировать его, вводя, например, «1». В этом поле можно найти ответы типа «Выбрав определенную заявку, которая дает удовлетворительный результат, пытался сохранить ее, отклоняясь незначительно в стороны», но даже этот игрок давал такие заявки: «40, 40, 45, 45, 70, 45, 50, 50, 55, 60, 50, 45, 30, 25, 35» и т. д.; т. е. у игрока может и была «определенная» заявка, но он, видимо, руководствовался чем-то еще, меняя в таких пределах свою заявку. В его действиях только несколько КП с погрешностью 0 (а в работе мы исследовали в основном ее), остальные, как минимум, с погрешностью 5. В будущем можно поставить в анкетах обязательный вопрос, связанный (не)явно с КП.

В § 4 проведен анализ гипотез, которые помогают понять характер КП. Статистически верны оказались гипотезы случайной длины КП и падения выигрыша в КП. Более сложная гипотеза о нижнем пороге выигрыша, после которого игрок прекращает КП, пока не доказана, но выглядит правдоподобной.

В § 5 представлены результаты другого, может быть, более прямого подхода к поиску модели поведения. Сделаны попытки отбора признаков и построения классификаторов, эффективно предсказывающих выбор игроков: продолжать или нет КП. Результаты этого исследования дали идею усложнения модели поведения — у игрока есть не только принцип окончания наблюдения, но и принцип начала наблюдения, причем последний, возможно, основан на случайной модели, так как задача отбора признаков и построения классификатора дали неэффективное решение, в отличие от аналогичных попыток для поиска принципа окончания наблюдения. Интересно, что в близкой модели I-SAW, можно сказать, применяется похожий подход. При классификации есть раздельные задачи отбора и выделения признаков, при выделении создается новый признак на основе данных признаков, а при отборе просто берутся лучшие признаки из данных. Авторы модели I-SAW выделили из данных признак «сюрприз» и на основании этого признака предложили вероятностный классификатор, который позволяет модели переходить в режим «инерция». Возможно, эта работа выполнена с помощью подходов машинного обучения, а может быть, экспертным путем, в настоящей же статье похожий режим строится явно с помощью подходов машинного обучения (метод случайных лесов).

Интересно, как гипотеза о нижнем пороге по выигрышу (см. § 4), согласуется с наиболее информативными признаками, обнаруженными в § 5. Там прирост по выигрышу на прошлом шаге попал в тройку самых информативных признаков. В дальнейшем гипотеза наличия нижнего порога по выигрышу также может учитываться при отборе признаков — при классификации можно добавить новый признак «нижний порог», только пока неясно технически, как строить уровень нижнего порога на основании предыстории игры для конкретного шага.

Другое направление этих исследований — это работа не на полной совокупности данных, а на неких «удобных» подмножествах. Поскольку после некоторых игр проводилось также анкетирование игроков в целях выявления мотивов поведения, оказалось, что анкетные данные указывают на разнообразие мотивов поведения. Это наталкивает на мысль анализировать поведение не всех игроков в целом, а только групп игроков с похожими мотивами (анкетными данными) или с похожим поведением (кластеризация данных игр). Таким образом, можно выполнить ту же работу, что и в § 4 и 5, но при этом брать подмножества исходных данных только для игроков с близкими мотивами. Можно поставить также более сложную задачу — выделение части данных, которые эффективно классифицируются.

И конечно, вторая по важности задача состоит в построении модели принципа выбора действия игроками в КП. Только решив ее, можно построить полную модель поведения игроков для полученных данных игр распределения ресурса. Для решения этой задачи можно воспользоваться регрессионным анализом или известными моделями поведения из экспериментальной экономики.

ЗАКЛЮЧЕНИЕ

Представлены результаты двухстороннего исследования КП игроков. Предложен ряд статистических гипотез, сделаны попытки классификации начала и окончания КП. Описанные подходы могут пролить свет на это явление, с точки зрения моделей случайных величин и моделей принятия решений, основанных на информации об истории игрока. За рамками работы остаются пока причины такого поведения: усталость, непонимание, ожидание ответных реакций оппонентов, забывчивость (recency [6]) и т. п.

В литературе встречаются похожие модели, в дальнейшем логичным шагом было бы встраивание построенного классификатора на место режима «инерция» в модель I-SAW [18] и сравнение эффективности исходной и модифицированной модели I-SAW.

Благодарности. *Автор благодарит рецензентов за содержательные и ценные замечания. С их помощью в статью вошли уточнения описания экспериментов и сравнения с существующими моделями.*

ЛИТЕРАТУРА

1. *Korgin, N.A., Korepanov, V.O.* An Efficient Solution of the Resource Allotment Problem with the Groves — Ledyard Mechanism under Transferable Utility // Automation and Remote Control. — 2016. — Vol. 77, no. 5. — P. 914—942.
2. *Yang, S., Hajek, B.* Revenue and stability of a mechanism for efficient allocation of a divisible good. Working paper, University of Illinois at Urbana-Champaign (2006).
3. *Korgin, N.A., Korepanov, V.O.* Experimental Gaming Analysis of ADMM Dynamic Distributed Optimization Algorithm // IFAC-PapersOnLine. — 2016. — Vol. 49, iss. 12. — P. 574—579.
4. *Korgin, N.A., Korepanov, V.O.* Experimental and Theoretical Comparison of Several Resource Allocation Mechanisms // IFAC-PapersOnLine. — 2017. — Vol. 50, iss. 1. 2017. — P. 15592—15597. Online publication date: 1, Jul. 2017.
5. *Nachbar, J.* Learning in Games / Encyclopedia of Complexity and Systems Science. — N.-Y.: Springer, 2009. — P. 5177—5188.
6. *Fudenberg, D., Levine, D.K.* Whither Game Theory? Towards a Theory of Learning in Games // Journal of Economic Perspectives. — 2016. — Vol. 30, no. 4. — P. 151—170.
7. *Корепанов В.О.* Модели поведения игроков в экспериментальной теории игр / Матер. 14-й Всерос. школы-конф. молодых ученых «Управление большими системами» (УБС—2017, Москва). — Пермь: Пермский нац. исслед. политехн. ун-т, 2017. — С. 14—22. [*Korepanov, V.O.* Models of players behavior in experimental game theory / Contributions of 14th Russian School-Conference of young scientists «Control of big systems». — Perm: Perm National Research Polytechnic University, 2017. — P. 14—22. (In Russian)]

8. *Cournot, A.A.* Recherches sur les Principes Mathematiques de la Theorie de Richesses, 1838. Translated by N.T. Bacon: Researches into the Mathematical Principles of the Theory of Wealth. — London: Hafner, 1960.
9. *Brown, G.W.* Iterative Solution of Games by Fictitious Play // Activity Analysis of Production and Allocation. — 1951. — Vol. 13, no. 1. — P. 374—376.
10. *Robinson, J.* An Iterative Method of Solving a Game // Annals of Mathematics. — 1951. — P. 296—301.
11. *Cheung, Y.W., Friedman, D.* Individual Learning in Normal Form Games: Some Laboratory Results // Games and Economic Behavior. — 1997. — Vol. 19, no. 1. — P. 46—76.
12. *Fudenberg, D., Levine, D.K.* The Theory of Learning in Games. — MIT Press, 1998. — Vol. 2.
13. *Healy, P.J.* Learning Dynamics for Mechanism Design: An Experimental Comparison of Public Goods Mechanisms // Journal of Economic Theory. — 2006. — Vol. 129, no. 1. — P. 114—149.
14. *Chen, Y., Tang, F.F.* Learning and Incentive-Compatible Mechanisms for Public Goods Provision: An Experimental Study // Journal of Political Economy. — 1998. — Vol. 106, no. 3. — P. 633—662.
15. *Roth, A.E., Erev, I.* Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term // Games and Economic Behavior. — 1995. — Vol. 8, no. 1. — P. 164—212.
16. *Camerer, C., Ho, T.H.* Experience Weighted Attraction Learning in Normal Form Games // Econometrica. — 1999. — Vol. 67, no. 4. — P. 827—874.
17. *Arifovic J., Ledyard, J.* A Behavioral Model for Mechanism Design: Individual Evolutionary Learning // Journal of Economic Behavior & Organization. — 2011. — Vol. 78, no. 3. — P. 374—395.
18. *Nevo, I., Erev, I.* On Surprise, Change, and the Effect of Recent Outcomes // Frontiers in Psychology. — 2012. — Vol. 3. — P. 24.
19. *Camerer, C.F., Ho, T.H., Chong, J.K.* Sophisticated Experience-Weighted Attraction Learning and Strategic Teaching in Repeated Games // Journal of Economic Theory. — 2002. — Vol. 104, no. 1. — P. 137—188.
20. *Moffatt, P.G.* Experimetrics: Econometrics for Experimental Economics. — Macmillan International Higher Education, 2015.
21. *Fischbacher, U.* Z-Tree — Zurich Toolbox for Ready-Made Economic Experiments // Experimental Economics. — 2007. — Vol. 10, no. 2. — P. 171—178.
22. *Plott, C.R., Smith, V.L.* (Eds.). Handbook of experimental economics results. Vol. 1. — Elsevier, 2008.
23. *Cassar, A., Friedman, D.* Economics Lab: An Intensive Course in Experimental Economics. — Routledge, 2004.
24. *Корепанов В.О., Коргин Н.А.* Причины константного поведения игроков в экспериментальных играх распределения ресурса // Тезисы XIX Апрельской междунар. науч. конф. по проблемам развития экономики и общества. Москва: ВШЭ (в печати). [*Korepanov, V.O., Korgin, N.A.* The reasons for the constant behavior of players in experimental resource allocation games // Collected Abstracts of XIX April International Academic Conference On Economic and Social Development. Moscow: National Research University Higher School of Economics, 2018. (In Russian)]
25. *Тюрин Ю.Н., Макаров А.А.* Статистический анализ данных на компьютере. — М.: Инфра-М, 1998. — 528 с. [*Tyurin, Y.N., Makarov, A.A.* Statistical analysis of data on a computer. — Moscow: Infra-M, 1998. — 528 p. (In Russian)]
26. *Rashka, S.* Python Machine Learning. — Birmingham, UK: Packt Publishing, 2015. — 454 p.

Статья представлена к публикации членом редколлегии В.Н. Бурковым.

Поступила в редакцию 18.01.2019, после доработки 5.03.2020.
Принята к публикации 18.04.2020.

Корепанов Всеволод Олегович — канд. техн. наук,
Институт проблем управления им. В.А. Трапезникова РАН,
г. Москва, ✉ vkorepanov@ipu.ru.

CONSTANT BEHAVIOR IN RESOURCE ALLOCATION GAMES: RESISTANCE TO THE GAME DESIGN, AND THE MODEL

V.O. Korepanov

V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia
✉ vkorepanov@ipu.ru

Abstract. In the conducted resource allocation games, «constant» behavior (CB) of the players was found when the players do not change their bids during several steps of the game. A feature of the game data is that the CB occupies a large share in them. A modified design of the games is proposed, where payments are made for each of the five game steps and the interface does not allow the player to easily enter an unaltered bid. The results of games with modified design showed that the share of CB has decreased, but remains at a fairly high level. The following are the results of a search for the causes of this behavior using statistical hypotheses and solving problems of classifying decisions of players. Among the statistical hypotheses, those are accepted about the random nature of CB and CB stopping in case of payoffs decrease. The classification task allows selecting informative features (parameters and game history) and the rule of decision-making by players (classifier) to stop the CB on the features basis. The results of constructing the classifier gave the idea of complicating the model: the player has not only the principle of stopping/continuing CB, but also the principle of starting CB. It is also shown that the results of statistical research and classification have common features and can further complement each other.

Keywords: business games, experimental economics, behavior models, resource allocation task.

Funding. The work was performed with financial support of Russian Foundation of Basic Research, grant no. 17-07-01550 A.

Acknowledgments. *The author thanks the reviewers for their informative and valuable comments. With their help, refinements to the description of experiments and comparisons with existing models were included in the article.*