
МАТЕМАТИЧЕСКИЙ АНАЛИЗ ЭКОНОМИЧЕСКИХ МОДЕЛЕЙ

СОЦИАЛЬНЫЕ ПРЕДПОЧТЕНИЯ В ЭКОНОМИЧЕСКОМ ПОВЕДЕНИИ: МЕТОДЫ ИЗМЕРЕНИЯ И МОДЕЛИРОВАНИЯ

© 2010 г. И.В. Галочкин

(Москва)

Приводится обзор использующихся в поведенческой экономической школе методов исследования социальных предпочтений, т.е. учета индивидом не только своего выигрыша, но и выигрышей (проигрышей) других. Описаны основные модели социальных предпочтений. Объясняется эволюционная стабильность социальных предпочтений и их роль в снижении транзакционных издержек.

Ключевые слова: экономическое поведение, социальные предпочтения, выигрыш, моделирование, транзакционные издержки.

В экономической теории в основе моделирования поведения индивида лежит принцип методологического индивидуализма, т.е. принятие решения исходя из собственных интересов и предпочтений путем взвешивания предельных выгод и предельных затрат для максимизации выигрыша. Такое поведение называют эгоистическим. Однако в реальности человек учитывает не только свои интересы, но и интересы других людей (Смит, 1997; Becker, 1974; Arrow, 1981; Samuelson, 1993). С 1970-х годов наличие альтруистических и других социальных мотивов в поведении было доказано экспериментально: многие испытуемые систематически соглашались уменьшить свой выигрыш для того, чтобы повысить или, наоборот, понизить выигрыш других участников эксперимента. Это противоречит неоклассической модели поведения: человек рациональный должен быть заинтересован только в максимизации своего выигрыша независимо от выигрыша других. Для описания обнаруженных феноменов поведения был введен термин “*социальные предпочтения*” (social preferences). Социальные предпочтения – это учет индивидом выигрышей и проигрышей других в своей функции полезности.

МЕТОДЫ ИЗУЧЕНИЯ СОЦИАЛЬНЫХ ПРЕДПОЧТЕНИЙ

С 1970-х годов проводятся измерения социальных предпочтений в экспериментальных условиях. В основе эксперимента лежит теория игр – область математики, исследующая стратегические взаимодействия и их возможные исходы. Теория игр не только формализует экономическую ситуацию, но может давать прогнозы поведения участников. В типичном эксперименте испытуемые (обычно студенты младших курсов университетов) помещаются в искусственную игровую ситуацию, где должны выбрать один из вариантов действий. От их выбора и выбора других игроков зависит вознаграждение за участие в эксперименте. Условия игры описываются абстрактным языком. Участники обычно не знают, с кем они играют, и не могут переговариваться с партнером, т.е. не могут действовать согласованно и договориться о кооперации. Выборы испытуемых затем сравниваются с прогнозом поведения, который дает теория игр для абстрактного рационального индивида. Для игр со сложными условиями могут использоваться повторные игры (в каждом следующем раунде партнеры меняются случайным образом), в этом случае интерес представляет и изменение поведения при обучении.

За три десятилетия развития поведенческой школы распространение получили семь экспериментальных игр. Поскольку социальные предпочтения тесно связаны с проблемой кооперации, все эти игры развивались из *дилеммы заключенного*, которая к моменту возникновения поведенческого направления в экономической науке уже была хорошо изучена как базовая модель некооперативного поведения (фон Нейман, Моргенштерн, 1970). Другие шесть игр называются так:

“общественное благо”, “ультиматум”, “диктатор”, “доверие”, “ты мне – я тебе”, “судья”. Ниже опишем условия этих игр и типичные результаты.

Дилемма заключенного (prisoner's dilemma). В этой игре каждый игрок может выбрать одну из стратегий: вести себя кооперативно или некооперативно. Выигрыши игроков определяются матрицей платежей (в общем виде – табл. 1, где $T > H > L > S$; в табл. 2 приведен вариант матрицы). Игра используется для моделирования производства отрицательных экстерналий (загрязнение предприятием окружающей среды и т.п.), рекламной политики, назначения цен в олигополии, обмена без жестких контрактов, соревнования за статус.

Таблица 1

Игрок 1	Игрок 2	
	Кооперировать	Не кооперировать
Кооперировать	H / H	S / T
Не кооперировать	T / S	L / L

Таблица 2

Игрок 1	Игрок 2	
	Кооперировать	Не кооперировать
Кооперировать	3/3	0/5
Не кооперировать	5/0	1/1

Рациональным поведением обоих игроков является стратегия “не кооперировать”, даже если игра состоит из многих повторяющихся раундов. Отсутствие кооперации будет доминантной стратегией, т.е. стратегией, наилучшей при любых действиях противника. Рассмотрим игру с позиции первого игрока. Если второй игрок выберет “кооперировать”, то для первого игрока выигрыш 5 лучше, чем выигрыш 3, поэтому выгоднее выбрать стратегию “не кооперировать”. Если второй игрок выберет стратегию “не кооперировать”, то выигрыш 1 лучше, чем выигрыш 0, поэтому опять лучше выбрать “не кооперировать”. Ситуация для второго игрока аналогична.

Эксперименты показывают, что примерно 50% игроков принимает стратегию “кооперировать” (Dawes, 1980). Причем частота кооперирования возрастает, если разрешаются переговоры между сторонами между раундами игры. Это объясняется тем, что игроки ожидают кооперации и заблаговременно благодарят за нее, поднимая свою репутацию.

Игра “общественное благо” (public goods game). В игре участвует n игроков. Каждому игроку в начале игры выдается u денег. В каждом раунде игры все игроки независимо друг от друга и одновременно принимают решение об инвестировании в общественное благо и вкладывают некоторую сумму. Выигрыш игрока равен сумме всех вложенных в данном раунде средств в общественное благо, умноженное на некоторый коэффициент m , причем $m < 1$ и $mn > 1$. Численный пример такой игры: пять игроков получают по 10 долларов, $m = 0.4$. Если игроки вложат все полученные деньги, то доход каждого составит по $50 \times 0.4 = 20$ долл.

Эта игра является расширением дилеммы заключенного и иллюстрирует ситуацию, когда возможность пользоваться бесплатно благами, произведенными другими, снижает мотивацию к производству этих благ и общий выигрыш. Игра предназначена для моделирования командной работы в трудовых коллективах, охоты в первобытных обществах, производства общественных благ в рыночной экономике.

Прогноз с точки зрения рационального поведения неутешителен: никто из игроков ничего не инвестирует (это их доминантная стратегия), а общественное благо не производится.

Эксперименты с этой игрой показывают, что в среднем игроки вкладывают 50% начальных средств в игре с одним раундом (Ledyard, 1995). При повторных раундах вклады уменьшаются,

и в конце игры большинство игроков уже ничего не вкладывает в общественное благо. Наличие переговоров, однако, значительно увеличивает вклады. Возможность наказывать других игроков за обман (когда игрок может заплатить небольшую сумму за то, чтобы на нарушителя был наложен большой штраф) значительно увеличивает кооперацию (Fehr, Gächter, 2000). При большем коэффициенте m (производительности общественного блага) вклады обычно выше (Sally, 1995).

Очевидно, что реальные люди вкладывают деньги в производство общественных благ, ожидая кооперации от других участников. Однако при наличии в группе одновременно эгоистичных и кооперативных игроков кооперативные постепенно “разочаровываются” и уменьшают свои вклады.

Игра “ультиматум” (ultimatum game). В этой игре один игрок называется предлагающим, другой – решающим. Экспериментатор дает предлагающему игроку сумму S , от которой тот может оставить себе часть, а остальное (x) должен отдать решающему игроку. Решающий игрок может согласиться на предложение или отказаться: если он соглашается, то получает сумму x , а предлагающий игрок – сумму $(S - x)$, а если отказывается, то все деньги возвращаются к экспериментатору и игроки ничего не получают. Игра предложена в работе (Gueth, Schmittberger, Schwarze, 1982).

Игра отражает монополистическое назначение цены на скоропортящийся товар или переговоры об условиях контракта в условиях ограниченного времени, отведенного на переговоры.

Если игроки ведут себя рационально, то решающий игрок должен соглашаться на любое предложение больше нуля. При этом его не должно интересовать, сколько получил предлагающий игрок. Зная это, предлагающий игрок должен дать решающему минимальную возможную сумму (например, 1 цент), чтобы тот согласился. Тем не менее данные экспериментов далеки от такого прогноза (Camerer, 2003): предложения менее $0.2S$ отвергаются решающим игроком в 50% случаев. Интересно, что игроки, оказывающиеся в роли предлагающих, обычно предвидят месть решающих игроков и в среднем дают сумму x , лежащую между $0.3S$ и $0.5S$. Эти результаты в целом сходны для разных социальных групп и развитых стран.

Если же ввести в игру конкуренцию между предлагающими игроками, то x повышается до $0.75S$, а если между решающими, то понижается примерно до $0.2S$ (Roth, Prasnikar, Okuno-Fujiwara, Zamir, 1991). В среднем решающие игроки считают справедливым $x = 0.5$, тогда как предлагающие – $x = 0.25$.

Несложно объяснить подобные данные экспериментов. Решающие игроки наказывают предлагающих за несправедливые предложения. Это называется местью или “негативной благодарностью” (negative reciprocity). Предлагающие игроки предвидят подобную месть, но могут проявлять альтруизм и стремление к равенству.

Игра “диктатор” (dictator game). Эта игра похожа на игру “ультиматум”, но решающий игрок не может отказаться от предложения, т.е. предлагающий игрок навязывает распределение $(S - x, x)$. Игра была разработана для того, чтобы проверить, насколько страх отказа решающего игрока в игре “ультиматум” влияет на выбор предлагающего игрока (Kahneman, Knetsch, Thaler, 1986).

Реальным примером игры является благотворительность, осуществляемая за счет недавно полученного и неожиданного наследства или выигрыша в лотерею.

Рациональный игрок в роли диктатора не должен делиться с другим игроком, т.е. $x = 0$. Однако эксперименты показывают, что в среднем диктатор предлагает $0.2S$, хотя сумма пожертвования существенно различается в разных экспериментах и у разных индивидов. Так, в роли диктатора студенты предлагают около $0.2S$, рабочие в Канзасе – около $0.5S$ (Carpenter, Burks, Verhoogen, 2004), а в некоторых примитивных обществах часто появляются предложения выше $0.5S$ (Henrich et al., 2005). Исход также зависит от того, видит ли испытуемый в роли диктатора своего партнера (Bohnet, Frey, 1999). Эти данные доказывают, что диктатор проявляет чистый альтруизм.

Игра “доверие” (trust game). В этой игре один игрок называется инвестором, другой – агентом. Экспериментатор дает инвестору сумму S , инвестор может вложить какую-то часть y этой суммы, т.е. отдать ее агенту, после чего экспериментатор утраивает сумму в руках агента. Теперь

агент может из имеющихся у него денег отдать инвестору сумму x (которая может быть от 0 до $3y$). Таким образом, инвестор в конце игры получит $S - y + x$, а агент – $3y - x$. Игра предложена в работе (Berg, Dickhaut, McCabe, 1995) и соответствует в реальной экономике обмену или совместному ведению дел без письменных контрактов.

С точки зрения рациональности прогноз негативный: агент ничего не будет возвращать инвестору, поскольку он выступает в роли диктатора. Предвидя это, инвестор ничего не посылает агенту.

Эксперименты показывают, что в среднем инвесторы доверяют агентами и посылают $y = 0.5S$, а агент возвращает x чуть меньше $0.5S$, т.е. в среднем инвестор проигрывает, оказывая доверие. Чем выше y , тем выше x , т.е. агент чувствует большую благодарность, если ему больше доверяют.

Игра “доверие” показывает, что игроки в роли инвестора склонны оказывать доверие, а агенты демонстрируют благодарность за доверие.

Игра “ты мне – я тебе” (gift-exchange game). Игра “доверие” может оформляться в виде “трудовых отношений” между игроками и тогда превращается в игру “ты мне – я тебе”. Инвестор выступает в роли работодателя, а агент – в роли работника. Работодатель объявляет заработную плату w и желаемый уровень интенсивности труда e^* . Работник может согласиться или отказаться. Если он отказывается, то никто ничего не получает. Если же соглашается, то работник может выбрать уровень интенсивности своего труда от 1 до 10 (независимо от того, что требует работодатель). После этого работодатель получает выигрыш, равный $10e - w$, а работник – выигрыш $w - c(e)$, где $c(e)$ – функция затрат на труд; $c(e)$ – строго возрастающая функция, т.е. чем большую интенсивность труда выберет работник, тем больше его издержки. Игра предложена в работе (Fehr, Kirchsteiger, Riedl, 1993). Игра моделирует трудовые договоры или торговые отношения, в которых невозможно подписание контракта или невозможен контроль за качеством труда или производимой продукции.

Поскольку работник получит заработную плату независимо от своих трудозатрат, его рациональным поведением будет выбирать минимальный уровень усилий ($e = 1$). Рациональный работодатель, предвидя это, должен назначить минимальную заработную плату, равную расходам работника при минимальной интенсивности труда, плюс минимальную денежную единицу (например, 1 цент).

Эксперименты показали, что интенсивность труда работника возрастает с ростом заработной платы w (Fehr, Gächter, Kirchsteiger, 1997). Это дает работодателям стимул ставить заработную плату выше нуля. В среднем работодатели требуют трудозатрат $e^* = 7$, а предлагаемая заработная плата w устанавливается так, что при этих трудозатратах работника заработная плата составит 44% прибыли. В среднем работники устанавливают реальные трудозатраты на уровне 4.4 (30% работников ставит минимальный уровень трудозатрат независимо от уровня заработной платы). В отличие от ситуации конкуренции между решающими игроками в игре “ультиматум” (где предлагаемая сумма снижалась) конкуренция между работниками не ведет к снижению заработной платы.

Эта несколько сложная игра показывает, что игроки в роли работников часто, но не всегда благодарят за высокую заработную плату. Работодатели стремятся вызвать благодарность работников, назначая высокую заработную плату.

Игра “судья” (third-party punishment game). Игра проверяет, насколько игроки ценят норму справедливости. Один игрок называется диктатором, второй – получателем, третий – судьей. Диктатор и получатель играют в игру “диктатор”. Судья наблюдает за игрой и видит, какая доля от суммы S передана получателю. Если сочтет нужным, судья может наказать диктатора, т.е. заплатить сумму x экспериментатору для того, чтобы тот уменьшил выигрыш диктатора на $10x$. Игра предложена в работе (Fehr, Fischbacher, 2004).

Аналогом данной игры в реальной экономике является общественное порицание вреда, нанесенного другим, правосудие в малых обществах. Если судья рационален, то он никогда не будет наказывать диктатора, ведь это снижает его выигрыш. Поэтому диктатор не боится наказания судьи и ничего не дает получателю.

Данные экспериментов показывают, что судья сильнее наказывает диктатора (и тратит на это больше денег), чем меньше диктатор дает получателю. Игра подтверждает, что игроки в роли судьи наказывают диктатора за нарушение общественной нормы справедливо делить ресурсы.

МОДЕЛИРОВАНИЕ СОЦИАЛЬНЫХ ПРЕДПОЧТЕНИЙ

Можно определить социальные предпочтения формально. Пусть в игре участвуют N индивидов и имеется некоторая функция распределения выигрышей $x = (x_1, \dots, x_N)$. Индивид i эгоистичен, если его полезность u зависит только от x_i , и имеет социальные предпочтения, если u является функцией от всего распределения выигрышей. Вид этой функции определяет тип социальных предпочтений.

Альтруистические предпочтения моделируются функцией $u(x_1, \dots, x_N)$ с положительными частными производными, т.е. полезность индивида строго возрастает при увеличении его выигрыша и выигрышей других индивидов. Это определение несколько отличается от понимания альтруизма в этике как стремления индивида повышать выигрыш *только* других индивидов. Г. Беккер использует альтруистические предпочтения при анализе пожертвований и добровольного участия в производстве общественных благ (Becker, 1974).

Игра “диктатор” выявляет альтруистические предпочтения. В одном из экспериментов испытуемый-диктатор распределял фишки, который затем конвертировались в деньги по разному курсу (Andreoni, Miller, 2000). Около 30% испытуемых распределяло фишки так, чтобы конечный денежный выигрыш оказался одинаковым у обоих игроков, около 20% максимизировало конечную взвешенную сумму выигрышей, а 50% оставляло все или почти все фишки себе.

Г. Чарнесс и М. Рабин (Charness, Rabin, 2000) описывают особую форму альтруизма – квазимаксиминные предпочтения (quasi-maximin preferences). Они вводят функцию общественного благосостояния как линейную комбинацию роулсианской и утилитаристской функций благосостояния:

$$W(x_1, \dots, x_N) = \delta \min \{x_1, \dots, x_N\} + (1 - \delta) \sum_{i=1}^N x_i,$$

где $\delta \in (0, 1)$. Полезность индивида зависит как от его собственного выигрыша, так и от благосостояния общества:

$$U_i(x_1, \dots, x_N) = (1 - \gamma)x_i + \gamma \left[\delta \min \{x_1, \dots, x_N\} + (1 - \delta) \sum_{i=1}^N x_i \right].$$

Для двух игроков эта запись упрощается:

$$U_i(x_1, x_2) = \begin{cases} x_i + \gamma(1 - \delta)x_j, & \text{если } x_i < x_j; \\ (1 - \gamma\delta)x_i + \gamma x_j, & \text{если } x_i \geq x_j, \end{cases}$$

откуда видно, что индивид больше заботится о партнере, если выигрыш партнера меньше его собственного.

Альтруистические предпочтения объясняют поведение испытуемых в играх “диктатор”, “ультиматум”, “доверие” и “ты мне – я тебе”, но не они влияют на желание людей наказывать в игре “судья”.

Термин “альтруизм” используется в литературе и в других значениях.

Чистый альтруизм (pure altruism) – желание испытуемого понизить свой выигрыш ради повышения выигрыша другого независимо от предыдущих действий реципиента альтруистического действия.

Реципрокный альтруизм (reciprocal altruism) – вознаграждение партнера за кооперативные действия, а также кооперативное поведение с расчетом на увеличение собственного выигрыша в будущем за счет повышения своей репутации.

Альтруистическим наказанием (altruistic punishment) обозначается случай, когда индивид уменьшает свой выигрыш ради того, чтобы наказать нарушителя некоторой социальной нормы, например справедливости. Альтруистическое наказание выявляется в играх “ультиматум” и “судья”.

Альтруистическим вознаграждением (altruistic rewarding) называют желание индивида снизить свой выигрыш, чтобы наградить другого индивида за выполнение социальной нормы. Например, в игре “ты мне – я тебе” агент вознаграждает инвестора за доверие.

Завистью называют учет индивидом не только своего собственного выигрыша, но и своего положения по отношению к другим индивидам. Идея “относительного дохода” и важности относительного статуса появляется уже у Т. Веблена (Веблен, 1984). Г. Болтон (Bolton, 1991) предложил моделировать поведение завистливого индивида функцией

$$U_i(x_1, x_2) = u_i(x_i, x_i/x_j),$$

где

$$\frac{\partial U_i}{\partial x_i} > 0, \frac{\partial U_i}{\partial x_i/x_j} > 0 \text{ при } x_i < x_j,$$

$$\frac{\partial U_i}{\partial x_i/x_j} = 0 \text{ при } x_i \geq x_j,$$

т.е. индивид недоволен, если его выигрыш меньше выигрыша другого, но ему безразличен выигрыш другого, если его собственный выигрыш больше. Получается, что $\partial U_i/\partial x_j \leq 0$, т.е. ситуация, обратная альтруизму.

Неприятием неравенства (inequity aversion) называют предпочтение индивидом уравнительного распределения выигрышей. Е. Фер и К. Шмидт (Fehr, Schmidt, 1999) предлагают следующую модель: индивид альтруистичен, пока выигрыш другого индивида ниже некоторого заданного уровня, и завистлив, когда выигрыш выше этого уровня. Обычно за такой уровень берется выигрыш, который получили бы игроки, если их суммарный выигрыш распределялся бы поровну. Полезность индивида может задаваться функцией

$$U_i(x_1, \dots, x_N) = x_i - \alpha_i \frac{\sum_{j \neq i} \max\{x_j - x_i, 0\}}{N-1} - \beta_i \frac{\sum_{j \neq i} \max\{x_i - x_j, 0\}}{N-1},$$

где $\beta_i \leq \alpha_i$ и $\beta_i \leq 1$; $\partial U_i/\partial x_j \geq 0$ только при $x_i \geq x_j$, т.е. индивид приветствует повышение выигрыша другого, пока его собственный выигрыш выше, и недоволен неравенством, когда его собственный выигрыш ниже выигрыша других.

Данная модель объясняет, почему испытуемый-диктатор в игре “диктатор” все-таки отдает часть суммы партнеру: он пытается выравнять выигрыши. Также она косвенно объясняет поведение игроков в играх “доверие” и “ты мне – я тебе”: игроки стремятся получать одинаковый выигрыш в игре. Этим же объясняется отказ решающего в игре “ультиматум”, если предлагаемая ему сумма меньше 0.2 от распределяемой суммы. Предпочтение равенства объясняет и наказание эгоистов испытуемым-судьей в игре “судья” и в остальных играх, если одна из сторон имеет право платного наказания другой стороны. Е. Фер и К. Шмидт указывают на вариацию социальных предпочтений между индивидами и моделируют ее коэффициентами α и β в функции полезности.

Недоброжелательность (spite) – поведение, направленное на снижение выигрыша других независимо от выигрыша самого индивида. Альтруизм и недоброжелательность комбинируются в модели благодарности (reciprocity) Д. Левайна (Levine, 1998). *Благодарность* – поведение, задаваемое желанием вознаградить дружественные действия партнера и наказать враждебные. Индивид имеет следующую функцию полезности:

$$U_i = x_i + \sum_{j \neq i} (a_i + \lambda a_j)/(1 + \lambda),$$

где $0 \leq \lambda \leq 1$ и $-1 < \alpha_i < 1$ для всех $i \in \{1, \dots, N\}$. Индивид склонен к альтруизму, если a_i ближе к 1, и склонен к недоброжелательности, если a_i ближе к -1 . Коэффициент λ определяет, насколько выбор индивида между альтруизмом и недоброжелательностью зависит от альтруизма или недоброжелательности другого индивида a_j . Так, если $\lambda = 0$, то индивид i всегда альтруистичен при $a_i > 0$ и всегда недоброжелателен при $a_i < 0$. Если $\lambda > 0$, то чем выше склонность к альтруизму другого индивида a_j , тем альтруистичнее поведение данного индивида. Модель хорошо предсказывает отказ решающего игрока от малых сумм в игре “ультиматум”.

Ситуации *благодарности с учетом мотивов* (intention-based reciprocity) наблюдаются, когда игроки выражают благодарность партнеру не за абсолютные выигрыши, а за разницу в выигрышах

шах, полученную в результате действий партнера. Пусть в игре “ультиматум” в одном эксперименте выбор предлагающего игрока сведен только к двум альтернативам: предложить разделение 50/50 или 80/20, о чем известно решающему игроку. В другом эксперименте предлагающий игрок может выбирать только между распределениями 80/20 и 20/80. Если решающий игрок оценивал бы только абсолютное значение предлагаемой ему суммы, то в обоих экспериментах он одинаково часто должен был бы отказываться от предложения 80/20. Однако эксперименты показали, что во втором случае решающий игрок отказывается реже, поскольку видит, что у предлагающего игрока не было возможности предложить справедливое распределение (Falk, Fehr, Fischbacher, 2000).

Описанные до этого момента модели не могут объяснить это явление. В 1993 г. М. Рабин предложил “психологический” подход к теории игр, добавив в функцию полезности психологическую оценку ситуации игроком (Rabin, 1993).

Пусть в игре с двумя игроками A_1 и A_2 – наборы смешанных стратегий игрока 1 и 2, соответственно, $a_i \in A_i$ – выбираемая игроком i стратегия, а π_i – платеж игрока i . Принимая решение, игрок i ожидает, что игрок j выберет некоторую стратегию, которую обозначим b_j . Игрок i знает, что игрок j имеет некоторые ожидания – представления игрока i об ожидаемой от него игроком j стратегии обозначим c_j .

М. Рабин вводит понятие “функция доброты” $f_i(a_i, b_j)$, которая показывает, насколько “добр” игрок i к игроку j . Если игрок i ожидает от игрока j выбора стратегии b_j , то, выбирая свою стратегию, игрок i определяет не только свой выигрыш π_i , но и выигрыш партнера π_j . Обозначим через $\pi_j^h(b_j)$ максимально возможный выигрыш игрока j после того, как он выбрал стратегию b_j , а через $\pi_j^l(b_j)$ – минимально возможный выигрыш. М. Рабин обозначает “справедливым” выигрышем $\pi_j^f(b_j)$ арифметическое среднее между максимальным выигрышем и минимальным Парето-эффективным выигрышем. Этот справедливый выигрыш не зависит от действий игрока i , он задается стратегией b_j . “Доброта” игрока i к игроку j является функцией относительного отклонения выигрыша игрока j от “справедливого” выигрыша после выбора игроком i своей стратегии. Она вычисляется как отношение разности между реальным выигрышем игрока j и “справедливым” выигрышем к разбросу возможных выигрышей:

$$f_i(a_i, b_j) \equiv \frac{\pi_j(b_j, a_i) - \pi_j^f(b_j)}{\pi_j^h(b_j) - \pi_j^l(b_j)},$$

где $j = 3 - i$, причем $f_i(a_i, b_j) = 0$ при $\pi_j^h(b_j) - \pi_j^l(b_j) = 0$. Значения “функции доброты” ограничены отрезком $[-1; +0.5]$. Игрок i оказывается “добрым”, если обеспечивает игроку j выигрыш, больше “справедливого” выигрыша.

У игрока i формируется оценка доброты игрока j . Она вычисляется так же, как определяется “доброта” игрока i , но на основе предполагаемого выбора b_j , который игрок j делает в ответ на свое представление c_i о выборе игрока i :

$$f'_j(b_j, c_i) \equiv \frac{\pi_i(c_i, b_j) - \pi_i^f(c_i)}{\pi_i^h(c_i) - \pi_i^l(c_i)},$$

где $j = 3 - i$ и $f'_j(b_j, c_i) = 0$ при $\pi_i^h(c_i) - \pi_i^l(c_i) = 0$.

Наконец, из указанных выше оценок “доброты” и реального выигрыша складывается функция ожидаемой полезности индивида:

$$U_i(a_i, b_j, c_i) = x_i(a_i, b_j) + f'_j(b_j, c_i)[1 + f_i(a_i, b_j)].$$

Если игрок i считает игрока j “злым” ($f'_j < 0$), то будет и сам “злым”, чтобы минимизировать выражение $1 + f_i(a_i, b_j)$. Если же игрок j воспринимается игроком i как “добрый”, то в интересах игрока i вести себя как можно “добрее”, т.е. увеличивать выигрыш игрока j . Поскольку “функции доброты” не зависят от абсолютного размера выигрышей, при увеличении выигрышей уменьшается роль психологического эффекта от “доброты” индивида и воспринимаемой “доброты” партнера. На основе “функций доброты” находится “справедливое равновесие” (fairness equilibrium), которое задается как пересечение наилучших ответов.

Достоинства модели Рабина в том, что она предсказывает отказы решающего игрока в игре “ультиматум”. Однако у нее есть и серьезный недостаток: зачастую она предсказывает много равновесных ситуаций, многие из которых не наблюдаются в реальности.

В 1990-х годах модель Рабина получила развитие. М. Дуфвенберг и Г. Кирхштайгер (Dufwenberg, Kirchsteiger, 1998) расширили модель для повторяющейся игры N лиц. Индивиды запоминают, какие мотивы руководили действиями других игроков в прошлых раундах. Так, условная кооперация (conditional cooperation) в динамической дилемме заключенного оказывается случаем “реципрокного равновесия”.

Э. Фальк и Ю. Фишбахер (Falk, Fischbacher, 1999) предложили усложненный вариант модели Рабина для N лиц с неполной информацией. Индивид воспринимается “добрым”, если его действия повышают выигрыш другого игрока, если тот игрок уже имеет больший выигрыш, чем сам индивид. У М. Рабина, напротив, “доброта” индивида не зависела от его собственного выигрыша.

Еще один вариант развития модели Рабина – модель Чанесса и Рабина (Charness, Rabin, 2000), в которой каждый игрок имеет репутацию и от нее зависит вес полезности игрока в функции полезности других игроков.

ЭВОЛЮЦИОННОЕ ОБЪЯСНЕНИЕ СОЦИАЛЬНЫХ ПРЕДПОЧТЕНИЙ

Вплоть до XX в. считалось, что альтруизм – феномен культуры, который воспитывается в человеке. Однако сейчас альтруистическое поведение найдено у многих животных, например дельфинов (Connor, Norris, 1982), муравьев (Bourke, Franks, 1995), птиц (Fackelmann, 1989), шимпанзе (de Waal, 1997) и др. Противники теории Ч. Дарвина вначале представили эти наблюдения как опровержение гипотезы, что индивиды стремятся передать только свои наследственные признаки. Тем не менее эволюционисты предложили по крайней мере три объяснения, которые не противоречат дарвинизму. Эти объяснения показывают, что альтруизм, хотя на первый взгляд может снижать жизнеспособность (fitness) индивида, на самом деле его повышает и поэтому оказывается эволюционно устойчивым.

Альтруизм между родственниками. С точки зрения британских зоологов У. Гамильтона и Р. Доукинза, альтруизм существует *между родственниками* (kin-selective altruism). Одна из центральных идей Р. Доукинза в том, что “поведение животного стремится максимизировать выживание генов, определяющих это поведение, независимо от того, находятся ли эти гены в теле этого конкретного животного” (Dawkins, 1999). Таким образом, подчеркивается разделение понятий “индивид” и “ген”, а термин “выживаемость” (fitness) означает не шансы выживания данного индивида, а шансы сохранения гена в следующем поколении. Формализация этой идеи приходится на 1950-е и 1960-е годы (Haldane, 1955; Hamilton, 1963) и выражается *правилом Гамильтона*: альтруистическое действие совершается, если $rB > C$, где r – степень совпадения генотипа реципиента альтруистического действия с генотипом индивида; B – увеличение шансов выживания реципиента; C – снижение шансов выживания индивида от альтруистического действия. Для детей, братьев, сестер и родителей индивида r равно 0.5; для внуков, племянников и бабушек, дедушек – $r = 0.25$, а для двоюродных братьев и сестер – $r = 0.125$ и т.д.

Гамильтон показывает, что альтруизм между родственниками чаще проявляется у видов, где легко отличить родственников (в основном высшие животные), а также у редко перемещающихся животных (большинство соседей индивида, к которым он может проявить альтруизм, скорее всего являются его родственниками).

Групповой отбор. Альтруизм между родственно несвязанными особями объясняется *групповым отбором* (group selection). Как и в дилемме заключенного, альтруизм может давать значительные преимущества в выживании группы индивидов. Эти преимущества перекрывают проигрыши, которые могут при этом испытывать отдельные индивиды (Wynne-Edwards, 1962, 1986). В 1960-х годах было экспериментально показано, что влияние группового отбора сравнимо с влиянием индивидуального (Wade, 1964). Были предложены модели многоуровневого отбора, в частности, основатель социобиологии Э. Уилсон предложил модификацию правила Гамильтона: альтруистическое действие совершается, если $rB_k + B_e > C$, где r – степень генетической общности; B_k – выигрыш родственника; B_e – выигрыш группы; C – издержки индивида, совершающего альтруистическое действие.

Реципрокный альтруизм. Последний вариант альтруизма – условный альтруизм, т.е. кооперативное поведение индивида при условии, что ему тем же отвечают в будущем. Выгоды реципиента альтруистического действия должны превосходить издержки того, кто это действие осуществляет. Такой альтруизм называется *реципрокным*, или *ответным* (reciprocal altruism) (Brosnan, de Waal, 2002). Реципрокный альтруизм является не чем иным, как стратегией “повторяй действия контрагента на прошлом ходу” (“Tit-for-Tat”), эволюционную устойчивость которой обнаружил Р. Аксельрод при сравнении эффективности различных стратегий в повторяющейся дилемме заключенного (Axelrod, 1984). Реципрокный альтруизм наблюдается, например, у шимпанзе (de Waal, 1997) и летучих мышей (Wikinson, 1990), а в человеческом обществе является, на наш взгляд, основой социальных отношений.

ЗАКЛЮЧЕНИЕ

По мнению швейцарского экономиста Э. Фера, реципрокный альтруизм (в особенности готовность людей жертвовать своим выигрышем не только для выражения благодарности за кооперацию, но и для наказания за оппортунизм) – один из факторов успеха человеческого общества. Оппортунизмом называется “представление неполной, искаженной информации, особенно когда речь идет о преднамеренном обмане, введении в заблуждение, искажении и сокрытии истины или других типах запутывания партнера” (Уильямсон, 1993, с. 43). С точки зрения модели рационального поведения мстить за оппортунизм не имеет смысла, поскольку месть уже не может возместить потери пострадавшей стороны, а только увеличивает их. Однако сильная форма реципрокности выступает как некоторая гарантия соблюдения прав собственности и выполнения обязательств по сделке, дает возможность организовывать и поддерживать сложную кооперацию между большими группами генетически несвязанных индивидов. Общество будет нести меньшие транзакционные издержки на защиту от оппортунизма, чем общество из рациональных индивидов, если индивиды обладают большей реципрокностью (склонны к благодарности и мести) и даже согласны жертвовать своим выигрышем ради справедливости к третьим лицам (в игре “судья”). Значит, благодаря групповому отбору такое общество переживет и вытеснит общества, где доверие и благодарность менее распространены (Fehr, Fishbacher, 2003).

Эксперименты показывают, что на экономическое поведение людей влияют социальные нормы, и зачастую это ведет к выигрышу для всех, хотя в некоторых случаях (например, в ситуации производства общественных благ, моделируемой игрой “общественное благо”) эти социальные нормы могут оказаться нежизнеспособными.

Иногда люди склонны проявлять даже чистый альтруизм, особенно если они еще не привыкли к обладанию благом. Игры “диктатор” и “ультиматум” показывают, что пока индивид не считает благо “своим”, он склонен делиться с другими, хотя представления о “справедливом” разделе различаются у испытуемых таким образом, что более справедливым признается распределение в свою пользу.

Все эти социальные предпочтения оказываются устойчивыми в экономическом поведении, поскольку они выполняют функцию снижения транзакционных издержек в экономической системе. Большой интерес представляет изучение не только их влияния на поведение, но также и их роли в повышении эффективности экономики.

СПИСОК ЛИТЕРАТУРЫ

- Веблен Т. (1984): Теория праздного класса: экономическое исследование институций. М.: Прогресс.
- Нейман фон Дж., Моргенштерн О. (1970): Теория игр и экономическое поведение. М.: Наука.
- Смит А. (1997): Теория нравственных чувств. М.: Республика.
- Уильямсон О. (1993): Поведенческие предпосылки современного экономического анализа // *THESIS*. Вып. 3.
- Andreoni J., Miller J. (2000): Giving According to GARP: An Experimental Test of the Rationality of Altruism. Working Paper. Mimeo, University of Wisconsin and Carnegie Mellon University.

- Arrow K.J.** (1981): Optimal and Voluntary Income Redistribution. In: *"Economic Welfare and the Economics of Soviet Socialism: Essays in Honor of Abram Bergson"*. Cambridge: Cambridge University Press.
- Axelrod R.** (1984): *The Evolution of Cooperation*. N.Y.: Basic Books.
- Becker G.S.** (1974): A Theory of Social Interactions // *Journal of Polit. Econ.* Vol. 82.
- Berg J., Dickhaut J., McCabe K.** (1995): Trust, Reciprocity and Social History // *Games and Economics Behavior*. Vol. X.
- Bohnet I., Frey B.S.** (1999): Social Distance and Other-Regarding Behavior in Dictator Games: Comment // *American Econ. Rev.* Vol. 89.
- Bolton G.E.** (1991): A Comparative Model of Bargaining: Theory and Evidence // *American Econ. Rev.* Vol. 81.
- Bourke A., Franks N.** (1995): *Social Evolution in Ants*. Princeton: Princeton University Press.
- Brosnan S., de Waal F.** (2002): A Proximate Perspective on Reciprocal Altruism // *Human Nature*. Vol. 13.
- Camerer C.** (2003): *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton: Princeton University Press.
- Camerer C., Fehr E.** (2003): Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists. Forthcoming in: J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath (eds.). *"Foundations of Human Sociality"*. Oxford: Oxford University Press.
- Carpenter J., Burks S., Verhoogen E.** (2004): Comparing Students to Workers: the Effect of Social Framing on Behavior in Distribution Games. October. IZA Discussion Paper № 1341.
- Charness G., Rabin M.** (2000): Social Preferences: Some Simple Tests and a New Model. Working Paper. Mimeo. Berkeley: University of California at Berkeley.
- Connor R.C., Norris K.S.** (1982): Are Dolphins Reciprocal Altruists? // *The American Naturalist*. Vol. 119. № 3.
- Dawes R.M.** (1980): Social Dilemmas // *Annual Rev. of Psychology*. Vol. 31.
- Dawkins R.** (1999): *The Extended Phenotype*. Oxford: Oxford University Press (reprint of 1982).
- Dufwenberg M., Kirchsteiger G.** (1998): A Theory of Sequential Reciprocity. Discussion Paper. Tilburg: Center, Tilburg University.
- Fackelmann K.A.** (1989): Avian Altruism: African Birds Sacrifice Self-Interest to Help Their Kin // *Science News*. Vol. 135.
- Falk A., Fehr E., Fischbacher U.** (2000): Testing Theories of Fairness – Intentions Matter. Institute for Empirical Research in Economics, University of Zurich, Working Paper № 63.
- Falk A., Fischbacher U.** (1999): A Theory of Reciprocity. Institute of Empirical Research in Economics, University of Zurich, Working Paper № 6.
- Fehr E., Fischbacher U.** (2004): Third Party Punishment Game and Social Norms // *Evolution and Human Behavior*. Vol. 25. Issue 2.
- Fehr E., Gächter S., Kirchsteiger G.** (1997): Reciprocity as a Contract Enforcement Device: Experimental Evidence // *Econometrica, Econometric Society*. Vol. 65 (4).
- Fehr E., Gächter S.** (2000): Cooperation and Punishment in Public Goods Experiments // *American Econ. Rev.* Vol. 90.
- Fehr E., Kirchsteiger G., Riedl A.** (1993): Does Fairness Prevent Market Clearing? An Experimental Investigation // *Quarterly Journal of Econ.* Vol. CVII.
- Fehr E., Schmidt K.M.** (1999): A Theory of Fairness, Competition and Co-operation // *Quarterly Journal of Econ.* Vol. 114.
- Fehr E., Fishbacher U.** (2003): The Nature of Human Altruism // *Nature*. Vol. 425. № 23.
- Gueth W., Schmittberger R., Schwarze B.** (1982): An Experimental Analysis of Ultimatum Bargaining // *Journal of Econ. Behavior and Organization*. Vol. III.
- Haldane J.B.S.** (1955): Population Genetics // *New Biology*. Vol. 18.
- Hamilton W.D.** (1963): The Evolution of Altruistic Behavior // *American Naturalist*. Vol. 97.
- Henrich J. et al.** (2005): "Economic Man" in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-Scale Societies // *Behavioral and Brain Sciences*. № 28.

- Kahneman D., Knetsch J.L., Thaler R.** (1986): Fairness as a Constraint on Profit Seeking: Entitlements in the Market // *American Econ. Rev.* Vol. LXXVI.
- Ledyard J.** (1995): Public Goods: a Survey of Experimental Research. Chapter 2. In: Roth A., Kagel J. “*Handbook of experimental economics*”. Princeton: Princeton University Press.
- Levine D.** (1998): Modeling Altruism and Spitefulness in Experiments // *Rev. of Econ. Dynamics*. Vol. 1.
- Rabin M.** (1993): Incorporating Fairness into Game Theory and Economics // *American Econ. Rev.* Vol. 83 (5).
- Roth A.E., Prasnikar V., Okuno-Fujiwara M., Zamir S.** (1991): Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study // *American Econ. Rev.* Vol. 81.
- Sally D.** (1995): Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958–1992 // *Rationality and Society*. Vol. 7.
- Samuelson P.A.** (1993): Altruism as a Problem Involving Group Versus Individual Selection in Economics and Biology // *American Econ. Rev.* Vol. 83.
- Waal de F.B.M.** (1997): Food Transfers Through Mesh in Brown Capuchins // *Journal of Comparative Psychology*. Vol. 111. № 4.
- Wade M.J.** (1964): An Experimental Study of Group Selection // *Nature*. Vol. 201.
- Wikinson G.** (1990): Food Sharing in Vampire Bats // *Scientific American*. February.
- Wynne-Edwards V.C.** (1962): Animal Dispersion in Relation to Social Behavior. Edinburgh: Oliver & Boyd.
- Wynne-Edwards V.C.** (1986): Evolution Through Group Selection. Oxford: Blackwell.

Поступила в редакцию
02.02.2009 г.

Social Preferences in Economic Behavior: Methods of Measurement and Modeling

I.V. Galochkin

The author reviews the existing methods of measuring social preferences, which imply that the individual cares not only for his own gains, but also for the gains or losses of others. The common ways of modeling social preferences are described. The article also brings evidence of evolutionary stability of social preferences and their role in decreasing transaction costs.

Keywords: economic behavior, social preferences, gain, modeling, transaction costs.