

А.А. Управителей

Санкт-Петербургский государственный университет, Санкт-Петербург; European University Institute, Italy

Неоклассические корни поведенческой экономики¹

Аннотация. Несмотря на то что с работами в сфере поведенческой экономики связаны имена ряда нобелевских лауреатов, многие исследователи рассматривают это направление науки как неортодоксальную, авангардную исследовательскую программу. Цель данной статьи – заполнить этот теоретический пробел и связать поведенческую экономику с одним из направлений экономической мысли, а также разделить ее на несколько направлений относительно исходных предпосылок, – либо теорий неоклассики. В первом ряду случаев нами были собраны те предпосылки неоклассической экономики об экономическом поведении, в отношении которых поведенческой экономикой были предложены альтернативные концепции. Этими неоклассическими предпосылками являются рациональность, информированность, эгоистичность поведения. Другой ряд теорий – неоклассические, в диалоге с которыми возникли теории поведенческой экономики. Это теории ожидаемой полезности, игр, экспоненциальной дисконтированной полезности, выявленных предпочтений, рациональных ожиданий, эффективного рынка. Приведенная нами последовательность развития поведенческой экономики связывает ее, некогда неортодоксальное направление науки, с неоклассическими направлениями экономической мысли. Статья может служить справочником по наиболее важным сферам исследования поведенческой экономики. Данное исследование систематизирует научное знание и устанавливает внутри него методологические связи.

Ключевые слова: поведенческая экономика, экономическая методология, ограниченная рациональность, теория перспектив, нейроэкономика.

Классификация JEL: B41, B59, D01.

Для цитирования: **Управителей А.А.** (2023). Неоклассические корни поведенческой экономики // *Журнал Новой экономической ассоциации*. № 1 (58). С. 110–140. DOI: 10.31737/22212264_2023_1_110

Введение

Поведенческая экономика – относительно молодое направление экономической науки. Многие исследователи рассматривают его как авангардную исследовательскую программу в противовес неоклассической экономике, имеющей статус ведущего направления. Однако мы устанавливаем связи ключевых теорий поведенческой экономики с основным направлением экономической мысли, а также выделяем в поведенческой экономике несколько ключевых направлений исследований.

Мы считаем обоснованным показать в теории поведенческой экономики два больших направления, и такое разделение мы проводим относительно характера взаимодействия этой теории с неоклассикой. В первом разделе собраны предпосылки неоклассической экономики, в отношении которых поведенческой экономикой были предложены альтернативные концепции. Этими неоклассическими предпосылками являются субстанциальная рациональность, полная информация, эгоистичность поведения. Эти предпосылки были изменены или ослаблены в рамках поведенческой экономики.

¹ Автор выражает благодарность и глубокую признательность А.В. Белянину за советы и ценные замечания при работе над данной статьей, а также А.Н. Дубянскому и В.М. Остапенко за полезную научную дискуссию.

В разд. 2 представлены неоклассические теории, в диалоге с которыми возникла другая часть теорий поведенческой экономики; под диалогом мы здесь подразумеваем уточнение, дополнение, развитие первоначальной концепции. Это – теории ожидаемой полезности, игр, экспоненциальной дисконтированной полезности, выявленных предпочтений, рациональных ожиданий, эффективного рынка.

Поскольку наше деление на две категории опирается на различие понятий «предпосылки» и «теории», стоит уточнить эти два термина. «Предпосылка» (assumption), согласно Кембриджскому словарю² – нечто, считающееся достоверным без сомнений или доказательств. В то время как «теория», согласно этому же словарю, нечто предложенное в качестве обоснованного объяснения фактов, состояния или события. В этих определениях для нас наиболее важным является, что предпосылки принимаются без доказательств, в то время как теория строится на доказательствах.

1. Предпосылки неоклассики об экономическом поведении человека

1.1. Рациональность поведения

В неоклассической экономике человек рассматривается как субъект, обладающий полнотой информации, имеющий достаточные когнитивные способности, чтобы сделать рациональный выбор, максимизирующий для него полезность. Однако американский исследователь Герберт Саймон указал на важные недостатки этой модели: процесс поиска альтернативных вариантов затрачивает ресурсы, тогда как реальные люди стремятся к удовлетворяющему варианту, сложность задачи принятия решения субъективна, а внимание зачастую ограничено. В качестве альтернативы была предложена концепция ограниченной рациональности (Simon, 1978).

Д. Канеман и А. Тверски значительно уточнили многие идеи Саймона об ограниченности внимания и когнитивных способностей, внедрив в экономику психологическое понятие «фрейминг» (framing), под которым они понимают концепцию принятия решений, на которую влияют как формулировка проблемы, так и нормы, и привычки лица, принимающего решения (Tversky, Kahneman, 1985). Также благодаря Канеману и Тверски в экономике появилось понятие «эвристик» (heuristics), или простых стратегий, с помощью которых происходит принятие решений в сложных ситуациях (Tversky, Kahneman, 1974), также опирающееся на идею Г. Саймона о стремлении упрощать вычисления (Simon, 1955).

Основываясь на этих концепциях ограниченности внимания, И. Гилбоа и Д. Шмейдлер предложили **теорию решений, принятых на основе прошлых случаев** (case-based decision theory). Они утверждают, что прошлый опыт влияет на дальнейшие предпочтения. Принимая решение, человек сначала идентифицирует окружающую среду, или ситуацию. Затем человек смотрит на прошлые действия, предпринятые в различных средах, и сравнивает текущие пары «среда (действие со всеми парами) – среда (действие в прошлом)», используя функцию подобия. Такое сравнение может подсказать оптимальные действия, которые следует предпринять в актуальной на данный момент ситуации (Gilboa, Schmeidler, 1995, 1997). И. Гилбоа и Д. Шмейдлер также используют понятие сходств – в том ключе, что люди в ситуации выбора сравнивают конкретный случай со всеми другими случаями в своей базе данных. Степень сходства с другими

² См. Cambridge Dictionary (<https://dictionary.cambridge.org/dictionary/>).

случаями направляет людей в процессе формирования убеждений относительно их нового выбора.

Также к идеям Саймона об ограниченности внимания и Канемана–Тверски о фрейминге восходит **теория значимости контекста** (context salience) (Bordalo, Gennaioli, Shleifer, 2012). Нередки случаи, когда внимание лица, принимающего решение, привлекают точно определенные варианты выигрышей, имеющие некую индивидуальную значимость. Это приводит к контекстно-зависимому пониманию лотерей, где выбор тех или иных вероятностей значительно искажается в пользу тех или иных определенных вариантов. Лицо, принимающее решения, сравнивает результаты и присваивает им ту или иную значимость, затем результаты ранжируются и более значимые результаты получают больший вес решения.

К. Габэ предложил **модель ограниченной рациональности**, основанной на разреженной максимизации (sparse max) полезности, что позиционируется как поведенческая версия максимизации полезности. При разреженной максимизации люди не придают много значения или вообще не придают значения некоторым деталям имеющейся проблемы, что также можно объяснять ограниченностью внимания. Агент строит собственную картину мира, значительно более простую, чем то, каким мир является на самом деле. Понятие разрежения более содержательно, чем упрощение, так как предполагает в векторе характеристик поведения немалое число нулей, обозначающих отсутствие той или иной активности. Модель также основывается на предпосылке ограниченности внимания, ввиду того что данные разрежены, а экономические агенты полагаются на некие значения по умолчанию, которые являются ожидаемыми значениями тех или иных переменных. В то же время внимание целенаправленно распределено к тем деталям, которые кажутся имеющими значение. Принимая во внимание информацию, человек привязывается к значениям по умолчанию и потом уже делает подстройку и корректировку (Gabaix, 2014).

Также на основе предпосылки об ограниченности внимания сформировалась **теория сходств** (similarities) А. Рубинштейна. Он утверждал, что в лотереях суммы выигрыша, так же как и вероятности, могут выглядеть схожими – например 0,2 и 0,25 – что серьезно влияет на принятие решения. Согласно теории сходств, если вероятности схожи между собой, то выбирать нужно тот вариант, где выигрыш больше. В ситуации, когда схожи выигрыши, выбирать следует тот вариант, где выше вероятность (Rubinstein, 1988).

Феномен ментальной бухгалтерии (mental accounting) является примером фрейминга (framing), т.е. зависимости решений от условий, в которых они принимаются. Ментальная бухгалтерия представляет собой совокупность когнитивных операций, используемая экономическими агентами для организации, оценки и упорядочивания своей деятельности. Термин «бухгалтерия» используется метафорически – в том смысле, что в сознании агента есть некий «план счетов», идея разных денежных потоков, пользуясь которыми агент организует свое экономическое поведение (Thaler, 1999).

Комплекс этих теорий закрепляет в основном направлении экономической науки модель ограниченно рационального экономического субъекта и значительно снижает актуальность неоклассической модели человека как рационального максимизатора полезности.

1.2. Информированность поведения

Неоклассическая экономика исходит из предпосылки, что человек обладает полнотой информации. Эта идея вполне удобна, однако очень сильно упрощает экономическую действительность. Дж. Акерлоф исследовал феномен асимметрии информации на примере рынка подержанных автомобилей и пришел к выводу, что если покупатели не владеют информацией о качестве товара в той же мере, что и продавцы, то плохие товары могут вытеснить хорошие вплоть до полного исчезновения рынка (Akerlof, 1970). Если в этой работе Акерлоф уделит внимание тому, какие последствия асимметрия информации будет иметь для рынка, то поведенческие экономисты обратили свой взор на экономических агентов и создали теории принятия решений в условия неполноты информации — и, соответственно, неопределенности и риска, которые эта неполнота создает.

Так, А. Тверски и Д. Келер предложили **теорию поддержки** (support theory), где утверждают, что агенты систематически нарушают условие экстенциональности (extensionality), или объемности множеств (Tversky, Koehler, 1994). Согласно математической аксиоме экстенциональности (объемности) два множества равны в том и только в том случае, когда они состоят из одних и тех же элементов — то есть, например, два набора цветных карандашей равны только в том случае, когда в каждый набор входит одинаковое число карандашей одних и тех же цветов. Тверски и Келер приводят результаты эксперимента, проведенного авторами (Johnson et al., 1993) на тему выбора вариантов страховки. Люди склонны больше платить за страховку, покрывающую госпитализацию по причине любой болезни или происшествия, чем за страховку, просто покрывающую любой, без уточнения, случай госпитализации. Точное упоминание болезни, которую покрывает страховка, увеличивает в сознании страхуемого возможную вероятность госпитализации, а значит, и повышает привлекательность такой страховки. По мнению Тверски и Келера, в нарушениях экстенциональности состоит базовое свойство суждений о вероятности — они зависят не от самих событий, а от описаний событий. И суждение о вероятности, как и решение в условиях неопределенности, зависит от точности описания.

Ч. Хит и А. Тверски (Heath, Tversky, 1991) вводят **гипотезу компетентности** (competence). Она заключается в том, что готовность сделать выбор в условиях неопределенности зависит не только от предполагаемой вероятности события и точности этой оценки, но также зависит и от точных знаний и понимания соответствующего контекста. При равной оценочной вероятности люди предпочитают делать выбор в тех условиях, когда чувствуют себя компетентными, а чувство компетентности определяется тем, что человек знает.

Также А. Тверски, но уже в соавторстве с К. Фоксом, предлагает **гипотезу сравнительной неосведомленности** (comparative ignorance), согласно которой неприятие двусмысленности возникает из-за сравнения какого-либо исхода с менее неоднозначными событиями, или же когда человек, принимающий решения, сравнивает себя с более осведомленными людьми. Неприятие двусмысленности также возникает, когда человек сравнивает ясную и туманную перспективы, однако это неприятие значительно уменьшается, когда человек сравнивает каждую перспективу в отдельности. Следовательно, выбор зависит не только от степени неопределенности, но и от источника неопределенности (Fox, Tversky, 1995).

1.3. Эгоистическое поведение

Неоклассическая экономика склонна рассматривать человека как эгоиста – известно утверждение К. Эрроу, что человек руководствуется «разнообразными формами жадности» и склонен к максимизации полезности (Arrow, 1980). Эта неоклассическая предпосылка была опровергнута результатами множества экономических экспериментов, о которых речь пойдет ниже. Они демонстрируют, что индивиды стремятся к поведению, не максимизирующему, но удовлетворяющему условиям социальной и внутриличностной приемлемости. И реальному поведению, в противовес неоклассической предпосылке об эгоистичности, свойственны стремление к справедливости, избегание неравенства, взаимобразное поведение, альтруизм, доверие. Модели, описывающие социальные факторы экономического поведения, получили название моделей социальных предпочтений (social preferences).

Экономическое поведение, отклоняющееся от неоклассических предпосылок под воздействием социальных предпочтений, фиксируется в различных экспериментальных играх, таких как «Ультиматум», «Диктатор», «Обмен дарами» и «Общественное благо», существующих в различных модификациях. «Ультиматум» является одной из наиболее популярных игр, она выглядит следующим образом: есть два участника, предлагающий и принимающий (Güth, Schmittberger, Schwarze, 1982). Первому дается денежная сумма, которую нужно одним решением разделить на двоих. Если второй принимает деньги, то они делятся согласно предложению первого. Если второй отклоняет условия раздела, то деньги не получает ни один участник эксперимента. Исследования показывают, что если принимающему предлагается сумма менее 30% общей, то с очень высокой долей вероятности он отклонит это предложение. Примечательно, что результаты экспериментов практически не зависят от предлагаемой суммы и национальности участников (Smith, Suchanek, Williams, 1988). Это же подтверждает и межстрановое исследование (Roth et al., 1991). В. Гут, создатель игры «Ультиматум», по прошествии более 30 лет в соавторстве с М. Кохером опубликовал метаисследование, в котором обобщил различные модификации этой игры, влияющие на результаты: несколько раундов, более двух игроков, зависимость от размера суммы, принадлежность к национальным и религиозным группам, количество тех или иных гормонов в крови, пол, возраст участников и т.д. (Güth, Kocher, 2014).

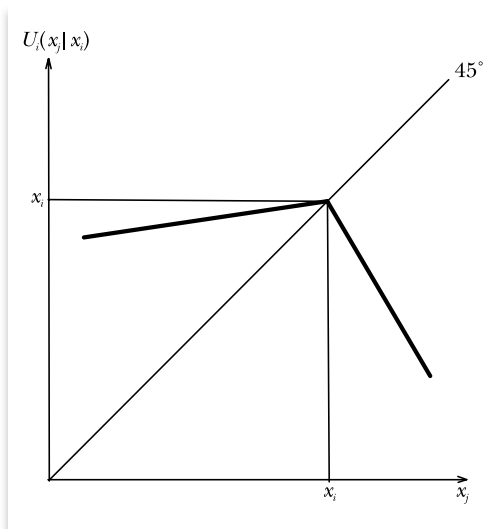
Появившиеся теории и модели имели цель обобщить богатые эмпирические данные. Одной из важнейших групп теорий социальных предпочтений являются теории справедливого поведения. Так, М. Рабин создал теорию честности, опираясь на предпосылку о том, что в экономических отношениях люди стремятся к равновесию справедливости (т.е. они стремятся помочь тем, кто помогает им, и стремятся нанести вред тем, кто нанес вред им, даже в ущерб собственной выгоде). Он приложил эту идею к теории игр и получил следующие результаты: любое равновесие Нэша, в котором результат взаимно-максимальный или взаимно-минимальный, является равновесием честности. Если материальные выплаты невелики, то результат является равновесием справедливости только тогда, когда это взаимно-максимальный или взаимно-минимальный исход. Если

материальные выплаты велики, то результат является равновесием справедливости тогда и только тогда, когда это равновесие по Нэшу (Rabin, 1993).

Эти подходы подтвердились в реальной жизни. Так, многие фирмы даже во время кризиса стремятся не сокращать заработных плат, ввиду того что работники воспринимают это как несправедливое отношение к ним, что сильно влияет на моральный настрой в коллективе (Blinder, Choi, 1990; Agell, Lundborg, 1995). В работе (Kahneman, Knetsch, Thaler, 1986) авторы доказали, что люди очень остро реагируют на резкое повышение цены продукции, считая это несправедливым, что не позволяет фирмам-монополистам пользоваться своей властью на рынке.

Э. Фер и К. Шмидт построили **теорию справедливости, конкуренции и кооперации** (Fehr, Schmidt, 1999). Они определили справедливость как эгоцентрическое избегание неравенства. А избегание неравенства, в свою очередь, означает, что люди избегают тех исходов, которые считают несправедливыми, т.е. они даже готовы понести какие-то материальные издержки, чтобы исход был более справедливым. Эгоцентричность в данном случае означает, что люди не беспокоятся о несправедливости, которая существует среди других людей, но заинтересованы в том, чтобы их материальная выгода была бы честной по отношению к выгоде других людей. Иными словами, люди, помимо собственного благосостояния, ценят равномерность распределения богатства в обществе: человеку не нравится ни когда он сильно беднее других, ни когда он значительно богаче других. Рисунок иллюстрирует полезность (U) полученных выплат игрока i . Полезность становится максимальной тогда, когда выплата x , которую получает игрок j , становится равной выплате, которую получает игрок i . Падение полезности сильнее от неблагоприятного для игрока i неравенства ($x_i < x_j$), чем падение полезности в том случае, когда игрок i получает больше, чем игрок j ($x_i > x_j$) (Fehr, Schmidt, 1999).

Другая группа моделей связана с понятием «взаимообразность» (reciprocity). В качестве недостатка модели Фера–Шмидта указывается, что она не учитывает такого фактора, как намерения людей. Люди склонны повышать благосостояние тех, кто щедр к ним, и снижать благосостояние тех, кто в отношении них проявляет жадность. Э. Фер и С. Гэхтер (Fehr, Gächter, 2000) вводят понятия «положительной» и «отрицательной взаимобразности». Взаимообразное поведение не противоречит эгоистичному – при одних обстоятельствах одни и те же индивиды принимают эгоистичные решения, при других – взаимобразные. Теория взаимобразности А. Фалька и У. Фишбахера учитывает, что люди оценивают не только последствия действия, но и намерения. Одни и те же последствия могут вызвать разные



Рисунок

Предпочтения при избегании неравенства

ответные реакции в разных средах. Также теория объясняет, почему результаты двусторонних взаимодействий обычно справедливы, тогда как на конкурентных рынках может возникнуть крайне несправедливое распределение (Falk, Fischbacher, 2006).

Понятие взаимообразного поведения лежит в стороне от поведения сотрудничающего и враждебного, так как в этих случаях индивиды ожидают какую-либо выгоду от своих действий, а взаимообразное поведение – ответ на дружелюбное или жадное поведение даже в тех случаях, когда не ожидается какой-либо материальной выгоды. Взаимность также не является формой альтруизма, так как альтруизм – это безоговорочная доброта, и он не может возрастать или сокращаться в зависимости от ответа на альтруизм. Примером негативного взаимообразного поведения может быть месть, развязывание войн часто тоже имеет взаимообразный характер. Позитивной взаимообразностью объясняется тот факт, что улыбающимся официанткам платят больше чаевых, чем менее дружелюбным. Сборщики благотворительности знают, что если приложить маленький подарок к просьбе внести деньги в фонд, повышается вероятность того, что человек захочет сделать взнос.

Одна из ключевых моделей взаимообразного поведения принадлежит перу Г. Болтона и А. Окенфельса (Bolton, Ockenfels, 2000). Эта модель получила название ERC по первым буквам феноменов, которые она объясняет – equity, reciprocity, competition, или – справедливость, взаимообразность, конкуренция соответственно. Она позволяет ответить на вопрос, почему предлагающие в игре «Ультиматум» дают больше, чем «Диктатор». Она предсказывает, что в игре «Диктатор» предлагающий оставит себе между половиной и всей суммой. Касательно игры «Ультиматум» она предсказывает, что нулевые предложения будут отклонены, и процент отклонений будет расти, по мере того как предложение предлагающего падает. Другую модель взаимообразности, основанной на намерениях, предложили Г. Чарнесс и М. Рабин – в результате экспериментов они обнаружили, что субъекты отказываются от готовности жертвовать ради достижения справедливого результата в тех случаях, когда другие не желают жертвовать, также они склонны наказывать других за несправедливое поведение (Charness, Rabin, 2002). Исследования показывают, что доля участников, беспокоящихся о справедливости и ведущих себя взаимообразно, колеблется от 40 до 66% (Abbinck, Irlenbusch, Renner, 2000).

Людам также свойственно и избегать обмана. Согласно неоклассическим предпосылкам о поведении людям свойственно обманывать, поскольку это максимизирует прибыль. Однако исследования показывают, что люди, даже точно зная, что их обман не будет раскрыт, стремятся к немаксимизирующему поведению. Широко известен эксперимент У. Фишбахера и Ф. Фельми-Хеуси (Fischbacher, Föllmi-Heusi, 2013): испытуемый должен был бросить кубик и сообщить, сколько выпало. Выплата была кратной выпавшему результату, что создает очевидные стимулы сообщить наиболее выигрышный исход, т.е. солгать. Дизайн эксперимента гарантировал полную анонимность, тем не менее около 40% участников были абсолютно честными, причем немалая часть тех, кто обманул, стремились к обману, немаксимизирующему выгоду, т.е. самым популярным был ответ: на кубике выпало «5» при максимуме «6». В работе (Abeler, Nosenzo, Raymond, 2019)

авторы, опираясь на (Fischbacher, Föllmi-Heusi, 2013), сделали метаисследование (участвовали более чем 44 000 человек из 47 стран) о предпочтениях сообщать правду или ложь. Эксперимент показал, что люди склонны сообщать не те данные, которые максимизируют их материальную выгоду, а те, что соответствуют предпочтению говорить правду. Предпочтения выглядеть честным и быть честным – ключевая мотивация сообщения правды. И политические меры, основанные на добровольном сообщении правды, могут быть очень успешными, особенно если будет сложно лгать, сохраняя при этом хорошую репутацию.

Масштабное исследование (Falk et al., 2018) обобщило данные Глобального опроса предпочтений, охватывающего 80 тыс. человек из 76 стран, в которых проживает 90% населения мира. Опросы были посвящены предпочтениям времени, риска, положительной и отрицательной взаимообразности, альтруизма и доверия. Данные показали, что предпочтения очень разнородны и на внутрискановом, и на межскановом уровне. Так, в частности, жители Африки и Ближнего Востока более склонны к риску, жители Европы и англоговорящих стран более терпеливы. Социально ориентированные предпочтения особенно ярко выражены в странах Азии и относительно слабы в странах Африки к югу от Сахары. Женщины более нетерпеливы, более социально ориентированы и менее склонны к риску, чем мужчины. Несмотря на то что различия в предпочтениях между странами значительны, различия внутри страны еще больше. Это позволяет предположить, что индивидуальные характеристики даже более важны для объяснения различий в предпочтениях, чем национальные границы (Falk et al., 2018).

Таким образом, поведенческая экономика противопоставляет предпосылке рациональности экономического субъекта его ограниченную рациональность, информированности поведения – неполноту информации и связанную с ней неопределенность, а эгоистичности – социально ориентированное поведение. Внедрение этих новых концепций значительно повышает реалистичность моделей поведения субъектов в экономической науке.

2. Теории неоклассической экономики

Другую группу теорий поведенческой экономики мы объединяем согласно самому факту внесения теориями дополнения в исходную теорию неоклассики. Толчком к возникновению этих теорий в большинстве случаев становились данные полевых или экспериментальных исследований, указывающие на слабые места неоклассических теорий экономического поведения. Если в первом разделе нашей статьи поведенческие концепции вступали в диалог лишь с общими предпосылками, допущениями неоклассики об экономическом поведении, то в данном разделе мы постарались указать те полновесные теории, в диалоге с которыми возникли теории поведенческой экономики.

2.1. Теория ожидаемой полезности

Одной из ключевых теорий неоклассики, характеризующих экономическое поведение, является теория ожидаемой полезности (ОП). Она принадлежит перу Джона фон Неймана и Оскара Моргенштерна. Впервые она была изложена в 1947 г. во втором издании книги «Теория игр и экономическое поведение». Эта теория включала ряд основополагающих аксиом – полноты, транзитивности,

независимости и непрерывности. Она исходит из того, что в ситуации неопределенности люди оценивают перспективы в соответствии с ожидаемым уровнем полезности и рациональный выбор состоит в выборе действия с наивысшей ожидаемой полезностью.

Однако вскоре после своего выхода эта теория встретила яркие эмпирические подтверждения несостоятельности своих аксиом, таким образом был сформирован ряд поведенческих расширений на обобщенные функционалы полезностей. Так, Морис Алле, один из представителей раннего периода экспериментальной экономики, указал на систематические нарушения аксиомы независимости. Стандартная аксиома независимости выглядит так: если $X > Y$, тогда $pX + (1-p)Z > pY + (1-p)Z$ для всех Z и значениями вероятности p (0, 1).

Иными словами, если банан (X) предпочтительнее яблока (Y), то 70% вероятности получения банана и 30% вероятности получения груши (Z) лучше, чем 70% вероятности получения яблока и 30% вероятности получения груши. Нарушение этой аксиомы делает невозможным представить предпочтения в виде функционала ожидаемой полезности. Первый контрпример – эффект общих последствий (common consequence effect). Представим, что индивиду предлагается выбор – ситуация А: можно получить 1 млн франков с вероятностью 100% – и ситуация В: 5 млн с вероятностью 10%; 1 млн франков – с вероятностью 89%; а также 0 франков – 1%. Другая пара выбора: ситуация С – 1 млн с вероятностью 11% – и 0 с вероятностью 89%, ситуация D – 5 млн с вероятностью 10%; и 0 франков – с вероятностью 90%. Можно увидеть, что в первой паре в каждом варианте есть общее последствие выигрыша 1 млн франков с вероятностью 89%, а ситуации С и D получаются путем вычитания этого общего последствия из А и В соответственно. Согласно ожидаемой полезности индивид, выбравший вариант А в первой паре, выбрал бы D во второй, или же, выбрав В, выбрал и С. Однако эксперименты показали, что индивиды склонны выбирать А в первой паре и С во второй (Allais, 1953).

Такое же нарушение было зафиксировано М. Алле и в схожем эффекте общего отношения (common ratio effect). Есть две пары выбора. Ситуация А: 3000 долл. с вероятностью 100%, – В – 4000 долл. с вероятностью 80% и 0 долл. – с вероятностью 20%. Вторая пара – ситуация С: 1 млн долл. – с вероятностью 5%, – 0 долл. – 95%. Ситуация D: 5 млн долл. – с вероятностью 4%, – 0 долл. – 96%. Ситуации выигрыша находятся в обеих парах в одинаковом отношении: 100 к 80 и 5 к 4 соответственно. Согласно ожидаемой полезности индивиды будут выбирать варианты А и С, однако эксперименты показывали склонность к вариантам А и D. Совокупность этих систематических нарушений аксиомы независимости ожидаемой полезности («аномалий»), получила название **парадоксов Алле**. Эксперименты Алле позднее были реплицированы, К. МакКриммон получил 40%, Д. Моррисон – 30% отклонений от ожидаемой полезности (MacCrimmon, 1965; Morrison, 1967).

Эти феномены, указывающие на слабые места теории ожидаемой полезности, подтолкнули исследователей создать различные версии модифицированной ожидаемой полезности, или иначе, – теорий обобщенной полезности (generalized utility theories). В основном эти неоклассические модификации пошли по пути обобщения ожидаемой полезности с ослаблением аксиомы независимости.

Одной из первых была **теория взвешенной полезности** (weighted utility theory) С. Чу (Chew, 1983). Взвешенная полезность ослабляет аксиому независимости посредством ввода коэффициента q , она выглядит следующим образом: если $X > Y$, тогда для каждого $p(0,1)$ есть и $q(0,1)$, и $pX + (1-p)Z > qY + (1-q)Z$ для всех Z .

Эта аксиома учитывает, что восприятие вероятностей человеком, принимающим решение, может находиться под влиянием различных индивидуальных факторов. Покажем это на простом примере: если банан (X) предпочтительнее яблока (Y), тогда для каждого коэффициента вероятности p , находящегося в промежутке от 0 до 1, существует и коэффициент q , находящийся в таком промежутке. При $p = 0,7$ и $q = 0,6$ получается следующее: вероятность 70% получения банана и 30% получения груши (Z) лучше, чем 60%-ная вероятность получения яблока и 40%-ная вероятность получения груши. Таким образом, аксиома независимости является частным случаем этой аксиомы. А это означает, что ожидаемая полезность является лишь частным случаем теории взвешенной полезности и эта последняя успешно разрешает парадоксы Алле.

Исследования Д. Белла (Bell, 1985), а также Г. Лумса и Р. Сагдена (Loomes, Sugden, 1986) формируют **теорию сожаления** (regret theory), т.е. несоответствия ожиданиям при принятии решений в условиях неопределенности и влияния разочарования на дальнейшее принятие решений. Так, в их модели экономический агент, рассматривая неопределенную перспективу, формирует некие предварительные ожидания. После того как неопределенность разрешается, агент получает следствие от перспективы: если следствие меньше предварительных ожиданий, то вместе с полезностью, полученной от ожидаемого блага, человек получает разочарование. Если следствие превосходит ожидание, то агент испытывает радость, что влияет на его дальнейшие экономические решения. Полезность зависит от того, насколько успешным может быть исход альтернативного действия – индивидам свойственно избегать сожаления.

Теория разочарования (disappointment theory), как и теория сожаления, обобщает идею ожидаемой полезности, делая аксиомы, описанные Нейманом–Моргенштерном лишь частным случаем своей теории. В этой теории, автором которой является Ф. Гул, вводится точка отсчета ожидаемого результата и к реальному исходу формируется отношение согласно имеющимся предварительным ожиданиям. Исход меньше ожидаемого кодируется как проигрыш, больше – как выигрыш. Это неоклассическое обобщение ожидаемой полезности содержательно связано с эффектом точки отсчета из теории перспектив Канемана–Тверски (Gul, 1991).

Теория неявной ожидаемой полезности (implicit expected utility) Э. Декеля (Dechel, 1986), или скрытая взвешенная полезность (implicit weighted utility) С.Х. Чу (Chew, 1989), вводит ослабленную форму независимости, которая была названа «промежуточность» (betweenness). Если $X > Y$, тогда $X > pX + (1-p)Y > Y$ при вероятности p , находящейся в промежутке от 0 до 1. Предпочтение смешения вероятностей двух лотерей находится между предпочтениями любой из лотерей. Промежуточность подразумевает нейтральность относительно одинаково хороших результатов. В теории скрытой полезности кривые безразличия выглядят как прямые линии, но они не являются параллельными, как в теории ожидаемой полезности.

Дж. Бекер и Р. Сарин предложили **теорию лотерейно-зависимой полезности** (lottery-dependent utility theory), в которой полезности зависят от лотерей. Зависимость полезности от оцениваемой лотереи достигается за счет ограничения меры полезности параметрическим семейством функций. Идея состоит в том, чтобы использовать каждую лотерею для определения конкретного значения параметра, характеризуя таким образом функцию полезности для каждой конкретной лотереи. Ожидаемое значение этой функции полезности, зависящей от лотереи, представляет собой общую меру предпочтения. Модель сохраняет свойства транзитивности, стохастической доминантности и непрерывности и дает возможность учитывать неприятие риска (Becker, Sarin, 1987).

Двойственная теория (dual theory) М. Яри разрешает некоторые парадоксы ожидаемой полезности (парадокс Алле), в то же время как и ожидаемая полезность разрешает некоторые парадоксы двойственной теории. Перспективы риска в этой теории оцениваются по числовой шкале, что сходно с теорией полезности, однако роли выплат и вероятностей меняются местами. Эта теория смещает фокус функции полезности по выборам в условиях неопределенности с линейности по вероятностям и нелинейности по выплатам – как в ожидаемой полезности – на нелинейность по вероятностям и линейность по выплатам. Также двойственная теория позволяет отделить отношение агента к риску от отношения к богатству (Yaari, 1987).

М. Мачина заменяет аксиому независимости предпосылкой гладкости (smoothness) предпочтений по сравнению с альтернативным распределением вероятностей. Это означает, что, хотя аксиома независимости и, следовательно, ожидаемой полезности, могут быть эмпирически невалидными, выводы и прогнозы теоретических исследований, в которых используется аксиома ожидаемой полезности, будут действительными при условии, что предпочтения являются гладкими – в том смысле, что функционал предпочтений может быть дифференцируемым (Machina, 1982).

Теорию ожидаемой полезности, зависящей от ранга (rank-dependent expected utility, RDEU) Дж. Куиггина можно рассматривать как расширение подхода к взвешиванию вероятностей на случай множественных исходов. Также ее можно интерпретировать как альтернативное представление неприятия риска и предпочтения риска, которое зависит только от наличия риска, а не от природы пространства результатов. Автор указывает, что, с одной стороны, люди стремятся к риску, приобретая лотерейные билеты, с другой – избегают риска, покупая страховки. Основная идея теории заключается в том, чтобы сделать взвешивание вероятностей зависимым от ранжирования результатов. Ключевое замечание состоит в том, что аргументы для чрезмерно высокого взвешивания событий с низкой вероятностью в основном применимы к экстремальным событиям с низкой вероятностью. В этой модели применяется взвешенная функция, в чем она наследует теории взвешенной полезности С. Чу, однако не к вероятностям отдельных событий, а к совокупной функции распределения (Quiggin, 1982, 1993).

Поворотным событием в развитии теорий, вносящих дополнения в ожидаемую полезность, стало появление **теории перспектив** (prospect theory) Канемана–Тверски. Она включила не только нелинейное взвешивание из

RDEU, но также добавила и неприятие потерь – разное отношение к потерям и выигрышам, которые оцениваются относительно субъективной точки отсчета (Kahneman, Tversky, 1979).

Кумулятивная теория перспектив (Tversky, Kahneman, 1992), известная как теория перспектив второго поколения, вносит дополнения в общую теорию перспектив. Эта теория включает два независимых обобщения теории ожидаемой полезности – зависимость от ранга и зависимость от признака. Зависимость от ранга влияет таким образом, что события, которые ассоциируются с наименее желательным исходом, оцениваются как более тяжелые. Зависимость от признака проявляется так, что события, ассоциирующиеся с прибылью, и события, ассоциирующиеся с потерями, могут иметь неодинаковые оценки сложности (Tversky, Kahneman, 1992).

Теория перспектив третьего поколения, разработанная У. Шмидтом, К. Стармером и Р. Сагденом (Schmidt, Starmer, Sugden, 2008), сохраняет прогностическую силу предыдущих версий теории перспектив, но расширяет эту теорию, позволяя точкам отсчета быть неопределенными, в то время как веса решений указываются в зависимости от ранга. В их модели рассматриваются ситуации, когда люди должны выбирать между альтернативными действиями, которые обычно вызывают различное распределение вероятностей по результатам. Например, лицо, принимающее решение, покупает дом и сталкивается с условием, что его дом с некоторой вероятностью может пострадать в результате пожара или других стихийных бедствий. Покупка страховки этим человеком вызывает другое распределение вероятностей по релевантным результатам. Этой теории предшествовала теория Р. Сагдена о субъективной ожидаемой полезности (reference-dependent subjective expected utility), зависимой от точки отсчета, в которой предпочтения между действиями зависят как от конечных результатов, так и от точек отсчета, которые могут быть неопределенными (Sugden, 2003).

Д. Канеман и А. Тверски привнесли из психологии в экономику важное понятие – «точка отсчета» (reference point). Они утверждали, что носителями полезности являются не конечные уровни богатства, а отклонения фактических уровней богатства от точки отсчета (Kahneman, Tversky, 1979). Б. Козжеги и М. Рабин (Kószegi, Rabin, 2006) развили теорию точки отсчета Канемана и Тверски. Так, в модели Козжеги–Рабина точки отсчета соответствуют ожиданиям агентов относительно исходов уже принятого решения. Полезность складывается из полезности потребления и полезности от соотношения выгод и потерь. При этом полезность от соотношения выгод и потерь зависит от точек отсчета – рациональных ожиданий относительно потребления. М. Рабин и Б. Козжеги утверждают, что люди скорее беспокоятся по поводу потерь относительно некоторой точки отсчета, чем радуются своим выигрышам. Агент ведет себя в соответствии с предпочтительным для него личным равновесием – он использует такой свой излюбленный план потребления, что если у него есть точка отсчета, в которой он ожидает от себя следования этому плану, то он так и будет делать. В статических ситуациях и в отсутствие неопределенности эта модель дает те же прогнозы, что и классическая теория потребления, но, помимо этого, модель Козжеги–Рабина дает точные альтернативные прогнозы в ситуациях, когда существует динамический выбор, присутствует внутренняя неопреде-

ленность или нечто неожиданное. Также эта модель расширяет и само понятие точки отсчета за пределы неопределенности (Kőszegi, Rabin, 2006).

2.2. Теория игр

Дж. фон Нейман и О. Моргенштерн обогатили экономическую науку не только теорией ожидаемой полезности, но и теорией игр, охватывающей стратегическое поведение. Поскольку эмпирическим путем была доказана несостоятельность некоторых аксиом ожидаемой полезности, логично, что это открыло пространство и для эмпирических проверок теории игр и создало поведенческую теорию игр (behavioral game theory).

Существует очень много экспериментальных тестов, результаты которых опровергают предсказания классической теории игр. Рассмотрим некоторые яркие случаи.

Так, в игре «Сороконожка» (centipede game), предложенной Р. Розенталем (Rosenthal, 1981), есть два игрока. Начальная сумма равняется 1 доллару, если первый игрок соглашается, игра заканчивается, если нет – сумма увеличивается в 10 раз, однако выбор делает уже второй игрок, который также имеет выбор: либо забрать сумму, либо отказаться – тогда сумма снова увеличится в 10 раз, но выбор перейдет к первому игроку. Таких раундов всего 100, финальная сумма доходит до 100 млн долл., если она отвергается, то каждый игрок получает 0. Если игроки не имеют возможности сговориться, то согласно обратной индукции выгоднее всего первому игроку взять 1 доллар и завершить игру. Однако исследователи Р. МакКелви и Р. Пэлфри провели три экспериментальные игры типа «Сороконожка» со своими студентами: 4 и 6 раундов со стандартными ставками или 4 раунда с повышенными. Они обнаружили, что завершение игры первым игроком во время первого раунда выбирают всего лишь 7%, 1% и 15% участников соответственно (McKelvey, Palfrey, 1992). Но когда в эту же игру сыграли профессиональные игроки в шахматы, этот вариант выбрали 69%, а среди гроссмейстеров – 100% (Palacios-Huerta, Volij, 2008). МакКелви и Пэлфри пришли к выводу, что в данном несоответствии поведения участников эксперимента с теорией игр большое значение имеет неполнота информации, поскольку есть вероятность того, что другой игрок окажется альтруистом, а не человеком, обеспокоенным лишь собственной материальной выгодой.

Эти и многие другие эмпирические данные, идущие вразрез с предсказаниями классической теории игр, открыли пространство для создания новых, более точных теорий и моделей, учитывающих отклонения наблюдаемого поведения от отклонения Нэша. Совокупность теорий в этой сфере получила название поведенческой теории игр.

Одна из поведенческих альтернатив совершенного равновесия Байеса–Нэша – равновесие квантового отклика (quantal response equilibrium, QRE). В этой модели игроки, как и в стохастических моделях, делают лучший выбор исходя из представлений о стратегиях других игроков, причем их выбор находится под влиянием «шума» – предпочтений вкуса. Однако игроки могут предвидеть шумные ответы других игроков. Шумные ответы не являются результатом случайного распределения, поскольку действия игроков, дающие больший выигрыш, имеют большую вероятность. То есть по сравнению с равновесием Нэша

в QRE ослабляется предпосылка того, что игроки каждый раз делают свой лучший выбор (McKelvey, Palfrey, 1995).

Однако в первых раундах игр игроки не могут иметь убеждений о том, какие выборы сделают другие игроки, и это условие предусматривает модель k -уровней (level- k). В них игроки обладают иерархией убеждений (hierarchy of beliefs). Игрок уровня L0 не обладает стратегией и во многих приложениях выбирает равномерно случайным образом из набора доступных действий. Игрок уровня L1 выбирает наиболее оптимальный ответ на действия игрока L0 и т.д. для всех игроков, уровня выше L0. Игрока уровня L0 может даже не существовать, но его присутствие в модели демонстрирует иерархичность игроков. Убеждения игроков не находятся в равновесии, поэтому модели k -уровня применимы для новых, нетипичных ситуаций, в которых оказываются игроки; в играх, которые происходят нерегулярно (Stahl, Wilson, 1994). Модель k -уровней помогает сделать предсказания в игре типа « p -конкурс красоты» (p -beauty contest). Эта игра основывается на концепции конкурса красоты, которую ввел Дж.М. Кейнс. В концепции классического вида победителем конкурса становится тот, кто смог угадать самые популярные лица участников конкурса. Возникает иерархия убеждений: L0 выбирает фотографии с самыми красивыми лицами; L1 предполагает, как L0 будут делать выбор и на их основе делает свой выбор; L2 делает выбор на основе убеждений о выборе L1 и L0 – и так далее: каждый следующий уровень учитывает выбор всех предыдущих уровней. В модели p -конкурса красоты игроки одновременно делают выбор на промежутке от 1 до 100, а победителем является тот, чей выбор – средний из всех выборов, умноженный на предварительно заданный параметр p (в модели Кейнса $p = 1$). Модель k -уровней помогает сделать точные предсказания в этой игре. Допустим, игроки L0 делают выбор от 0 до 100, а $p = 2/3$. L1 уверены, что все игроки – это L0 и их средний выбор равен 50. Так что они уверены, что выигрышный вариант – $50 \times 2/3 = 33,33$. Игроки L2 уверены, что все игроки – L1, следовательно, они считают, что выигрыш – $50 \times (2/3)^2 = 22,22$. Если обобщить выше сказанное, то убеждение игрока уровня L k о выигрышном варианте равняется $50 \times (2/3)^k$ (Nagel, 1995; Но, Camerer, Weigelt, 1998). Модель когнитивной иерархии (cognitive hierarchy model) близка модели k -уровней. Концептуальная разница состоит лишь в том, что рациональность игроков возрастает с каждым уровнем и каждый игрок делает выбор, исходя из убеждения, что уровни других игроков – ниже его (Camerer, Но, Chong, 2004).

Игроки могут иметь представления об убеждениях других о своих действиях, и им может не нравиться отклоняться от этих убеждений – они могут испытывать чувство вины (guilt). Совокупность моделей, предусматривающих влияние этого чувства и сопутствующего ему, на поведение в играх получило название психологической теории игр (psychological game theory), это понятие ввели исследователи Дж. Геанакоплос, Д. Пирс и Э. Стаккетти (Geanakoplos, Pearce, Stacchetti, 1989). Простая модель игры с двумя игроками, например, когда водитель такси ожидает чаевые, а пассажир уверен в этом ожидании таксиста, была рассмотрена П. Баттигалли и М. Дюфвенбергом (Battigalli, Dufwenberg, 2007). В игре «Общественное благо» с тремя участниками игрок может испытывать чувство вины, если думает, что его вложение ниже, чем среднее ожидание других о его вложении. В таком случае существует параметр избегания чувства вины (guilt aversion),

и, если он равен нулю, тогда чувство вины отсутствует и игрок выбирает стратегию безбилетника (*free ride*), как в классической теории игр. Существование избегания чувства вины и его влияние на поведение было зафиксировано в экспериментальных играх по схемам «Потерянный кошелек» (Dufwenberg, Gneezy, 2000), «Доверие» (Charness, Dufwenberg, 2006) и «Диктатор» (Khalmetski, Ockenfels, Werner, 2015).

Классическая теория игр не предусматривает способности человека к обучению, хотя обучение, т.е. накопление опыта и изменение поведения, — свойственно не только человеку, но и животным.

Одна из базовых моделей обучения — обучение с подкреплением (*reinforcement learning*). Идея этой модели состоит в следующем: если стратегия приводит к более высокому результату, то более вероятно, что именно она будет реализована в будущем (Bush, Mosteller, 1955). И. Эрв и А. Рот (Ergev, Roth, 1998) дополнили эту модель двумя параметрами. Первый — экспериментирование (обучаясь, игроки экспериментируют с вариантами, подобными тем, которые были успешными в прошлом), второй — недавние события, которые вместе с опытом влияют на поведение в большей степени, чем события из более далекого прошлого. Эта модель получила название трехпараметрической модели обучения с подкреплением. Авторы применили ее к матричным играм с уникальным равновесием смешанной стратегии, в которых повторение не создает возможностей для сотрудничества игроков, и обнаружили, что эта модель значительно лучше описывает и предсказывает поведение реальных игроков, чем предсказания статического равновесия.

2.3. Теория экспоненциальной дисконтированной полезности

Теория ожидаемой полезности — одна из самых крупных теорий неоклассики, на диалоге с которой сформировалась значительная часть поведенческой экономики. Другой такой теорией является теория экспоненциальной дисконтированной полезности (*exponential discounting*). Широко известен феномен временного дисконтирования (*time discounting*): полезность одного и того же блага уменьшается в том случае, если индивид получает его не сейчас, а по прошествии какого-то интервала времени. Существование межвременного дисконтирования лабораторно проявляется следующим образом: участникам эксперимента дается гипотетический выбор — получить 100 долл. сейчас или 200 долл. через два года. Более предпочтительным является первый вариант. Вторым эксперимент предлагает сделать выбор: получить 100 долл. через 6 лет или 200 долл. через 8 лет. Здесь более предпочтительным оказывается уже второй вариант — несмотря на то что и в первом, и во втором эксперименте разница до получения большего вознаграждения равна двум годам. Таким образом, происходит изменение полезности: 100 долл. сейчас оказываются полезнее 200 долл. через 2 года, но в отдаленном промежутке полезность суммы, которую можно получить в ближайший момент, исчезает.

Оригинальная неоклассическая теория была предложена П. Самуэльсоном (Samuelson, 1937). Эта модель предусматривает временную согласованность предпочтений (*time consistent preferences*). Ключевая предпосылка его модели состоит в том, что все несопоставимые мотивы, определяющие межвременной выбор, могут быть сведены к одному мотиву — ставке дисконтирования.

Однако поведенческие экономисты указали на ряд аномалий, существование которых не может быть объяснено теорией межвременного дисконтирования Самуэльсона.

К поведенческим эффектам межвременного дисконтирования относятся:

- эффект магнитуды (*magnitude effect*) – малые суммы дисконтируются сильнее больших;

- эффект знака (*sign effect*) – потери дисконтируются меньше, чем приобретения (Thaler, 1981);

- субаддитивное дисконтирование (*subadditive discounting*) – разделение временного интервала на подынтервалы уменьшает степень дисконтирования;

- очевидная нетранзитивность предпочтений (*apparent intransitivity of preferences*) – временной выбор может быть циклическим;

- эффект общего различия (*common difference effect*) и эффект смещения к настоящему (*present bias*) – перемещение по дате к настоящему моменту изменяет ставку дисконтирования;

- предпочтение повышения (*preference for increasing*) – работники настроены на то, чтобы их заработная плата росла постепенно, даже если в совокупности это дает меньший выигрыш (Loewenshtein, Sicherman, 1991).

Также ставка дисконтирования может зависеть от типа индивидуального потребления, возраста, уровня интеллекта, наличия алкогольной, никотиновой, наркотической и игровой зависимостей.

В качестве модели, существенно дополняющей неоклассическую теорию экспоненциальной дисконтированной полезности, поведенческими экономистами была предложена модель гиперболического дисконтирования (*hyperbolic discounting*), которая учитывает наибольшую часть указанных аномалий (Laibson, 1997). Кратко идею гиперболического дисконтирования можно выразить следующим образом: разница в один день значит тем меньше, чем дальше в будущем возникнут события (Loewenshtein, Thaler, 1989). Модель, объединяющая гиперболическое дисконтирование и субаддитивное дисконтирование, названа теорией отсчетного времени (*reference time theory*); она включает не только перечисленные выше аномалии межвременного дисконтирования, но также и модели неопределенности и атрибутов. Предпочтения в этой модели транзитивны, поскольку объясняются изменением точки отсчета времени (al-Nowaihi, Dhami, 2008).

Впрочем, нужно учитывать, что деньги не являются универсальным стимулом, эквивалентом полезности. Такие внешние стимулы, как денежное вознаграждение, могут конфликтовать со внутренними стимулами, т.е. желанием сделать то или иное дело само по себе. Так, Р. Бенабу и Ж. Тироль предложили теорию саморегулирования, в которой показали, что денежное вознаграждение за какую-либо социально ориентированную активность может быть контрпродуктивным. К тому же внутренней мотивации может хватить очень надолго, а при внешней мотивации возникает сильный, но краткосрочный эффект. Денежная награда может иметь скрытые издержки (например, индивид переключит внимание с процесса выполнения задачи на ее результат, т.е. на будущее вознаграждение (Bénabou, Tirole, 2003)).

2.4. Теория выявленных предпочтений

В основе неоклассической теории выявленных предпочтений (revealed preference theory) П. Самуэльсона лежит идея о том, что предпочтения потребителей могут быть выявлены по их покупательским привычкам (Samuelson, 1938). Это значит, что действия человека напрямую зависят от его убеждений.

Однако нейроэкономические исследования опровергают эту неоклассическую теорию. Так, людям свойственно сначала аффективно принимать решение, а потом уже подгонять его под какую-либо логику. На принятие решений значимо воздействуют такие нейромодуляторы, как допамин, серотонин и норадреналин. Они влияют на предпочтения риска, времени, а также на социальные предпочтения (Crockett, Fehr, 2014). Гормоны также воздействуют на экономическое поведение. Эксперименты показывают, что уровень окситоцина отражается на межличностном доверии при инвестировании, причем окситоцин увеличивает доверие именно при социальных взаимодействиях, а не просто усиливает стремление преодолеть риск (Kosfeld et al., 2005). Мужчины, имеющие более высокий уровень тестостерона, чаще совершают отказы в игре «Ультиматум» (Burnham, 2007). Мужчины-трейдеры с более высоким уровнем тестостерона получают более высокую прибыль. Успешно проведенная сделка подталкивает к более легкому принятию риска, к открытию новой сделки. Так называемый «эффект победителя» является следствием повышения уровня тестостерона после достижения успеха (Coates, Herbert, 2008).

Дж. Левенштейн ввел понятие внутренних (visceral) факторов экономического поведения, среди которых – такие физиологические состояния, как голод, жажда, сексуальное влечение, боль, тяга к наркотику, от которого есть зависимость. Эти состояния сильно влияют на эмоциональный фон человека, а значит, и на принятие им экономических решений (Loewenstein, 1996). Также на принятие экономических решений влияют такие факторы, как гнев, ненависть, сожаление, разочарование, страх, наслаждение, влюбленность и многие другие (Elster, 1998).

Ф. Гул и В. Пезендорфер в качестве поведенческой альтернативы неоклассической теории выявленных предпочтений предложили **теорию соблазнов** (temptation) (Gul, Pesendorfer, 2001). В их понимании, соблазн – динамическая непоследовательность, т.е. изменение предпочтений под действием, в частности, висцеральных факторов. Например, предпочтением человека может быть диетическая пища, но в момент голода он может поддаться соблазну и съесть гамбургер. Агенты используют самоконтроль для сопротивления соблазнам. Отличие от моделей предпочтений состоит в том, что агент, склонный к соблазнам, тратит ресурсы самоконтроля на то, чтобы убрать заманчивые альтернативы из своих вариантов выбора.

Полевые эксперименты показывают, что люди склонны подгонять логику под уже сформировавшееся решение. Например, в США было проведено исследование среди менеджеров по персоналу: были разосланы два идентичных резюме, одно с именем, характерным для белого американца, другое – для афроамериканца. И резюме с именем, характерным для белого, получало в среднем на 50% больше звонков, хотя эти менеджеры были уверены, что их выбор не обусловлен ксенофобией, которая в США порицается (Bertrand, Mullainathan, 2004). Другой пример,

врачи уверены, что подарки, которые они получают от фармацевтических компаний, не влияют на принятие решений о том, какое лекарство назначить пациенту, однако исследование (Dana, Lowenstein, 2003) утверждает, что это не так.

Согласно неоклассической экономике наилучшее решение о собственном благополучии может принять лишь сам человек. И, как гласит теория выявленных предпочтений, решения напрямую связаны с убеждениями.

Однако поскольку поведенческая экономика дала нам тот факт, что убеждения и решения могут находиться под воздействием различных когнитивных искажений, следовательно, и решения о собственном благополучии человек может принимать ограниченно рациональные.

На основе данных об этих искажениях были сформированы концепции поведенческой экономической политики, направленной на сокращение масштаба влияния когнитивных искажений на всеобщее благосостояние.

Наиболее известной концепцией поведенческой экономической политики является **либертарианский патернализм** (libertarian paternalism) Р. Талера и К. Санстейна. Под патернализмом в данном случае понимается вмешательство государства в экономические и социальные процессы. Либертарианский патернализм с его ключевым инструментом подталкивания противопоставляется классической патерналистской политике государства, инструментами которой являются запреты и штрафы. Одно из важнейших понятий либертарианского патернализма – архитектура выбора, т.е. формирование для экономического агента такой ситуации, когда наиболее полезный для него вариант идет в первую очередь и выбор происходит по пути наименьшего сопротивления (Thaler, Sunstein, 2003). Другие поведенческие концепции патернализма – асимметрический патернализм (Camerer et al., 2003), добрый патернализм (Benjamin, Laibson, 2003), легкий патернализм (Loewenstein, Haisley, 2006).

Концепции подобного подталкивания экономических агентов к тем или иным экономическим решениям подняли массу вопросов об этичности таких методов, поскольку государственные деятели, принимающие решения, способны допускать ошибки и поступать корыстно (Frye, Shleifer, 1997), а у людей падает мотивация принимать взвешенные решения (Klick, Mitchell, 2006). Тем не менее, аналитические центры, занимающиеся разработкой и реализацией мер поведенческой экономической политики, существуют в США, Великобритании, Австралии, Дании и многих других странах мира.

2.5. Теория рациональных ожиданий

Макроэкономическая теория рациональных ожиданий (ТРО) – еще одна теория неоклассической экономики, на диалоге с которой возник ряд теорий поведенческой экономики. Изначально она была разработана Дж. Мутом (Muth, 1961) и утверждает, что люди всегда учатся на ошибках прошлого, беспристрастны, используют всю доступную информацию и понимают, как работает экономическая система.

Однако вскоре Дж. Катона указал, что ожидания формируются под воздействием прошлого поведения и полученного опыта (Katona, 1980). В силу ограниченности внимания для полного анализа той или иной ситуации человек не может сделать суждения о наиболее рациональном в данный момент действии – человек использует простые правила принятия решений (Hommes,

2013). Действия и суждения находятся под влиянием предварительно сформированных ожиданий. Нужно учитывать и фактор способности экономических агентов к обучению, индивидуальная предрасположенность к которому также различна. Поскольку каждый индивид обладает разным полученным опытом, у него формируются разнородные ожидания и стратегии экономического поведения.

Чтобы сделать теорию рациональных ожиданий более реалистичной, неоклассик Т. Сарджент внедрил в нее **концепцию ограниченной рациональности**. Он указал на яркое противоречие – экономические агенты в моделях теории рациональных ожиданий выглядят более рационалистичными, чем экономисты, которые эти модели рассчитывают (Sargent, 1993). Он добавил в модель возможность агентов обучаться – по его мнению, эта предпосылка приближает модель к равновесию рациональных ожиданий.

В качестве поведенческой концепции, альтернативной неоклассической теории рациональных ожиданий, поведенческие экономисты предложили **теорию разнородных ожиданий** (heterogenous expectations), согласно которой экономические агенты могут иметь неоднородные ожидания о будущем, обладают разным уровнем ограниченной рациональности, способности к обучению и вниманию к актуальной ситуации.

Так, Дж. Фрут и К. Френкель в своем исследовании (Frankel, Froot, 1990) поведения на бирже определили два типа игроков с совершенно разными ожиданиями – фундаменталисты и чартисты. Фундаменталисты ориентируются на такие фундаментальные показатели, как макроэкономический рост, ставка безработицы и т.д. Они делают прогнозы, используя правило негативного отзыва, сравнивают прошлый обменный курс с параметрами и прогнозируют, что будущие показатели обменного курса также будут зависеть от них же. Такое поведение ведет к стабилизации рынка. Чартисты – сторонники технического анализа строят свои ожидания и прогнозы, основываясь только на текущем тренде. Чартисты делают прогнозы согласно правилу позитивного отзыва: на основе прошлых изменений курса они делают вывод о будущих изменениях курса. Такое поведение ведет к дестабилизации рынка (Frankel, Froot, 1990). Данные опросов показывают, что трейдеры, работающие с долгосрочной перспективой, ориентируются больше на макроэкономические показатели, а трейдеры, работающие с краткосрочной перспективой, стараются воспользоваться актуальным трендом (Frankel, Froot, 1987).

Идея о том, что на рынке могут присутствовать агенты с разнородными ожиданиями и поведением, открыла пространство для широкого ряда моделей поведенческой макроэкономики.

П. Де Грауве и М. Гримальди предлагают **модель обменного курса**, которая включает разнородные ожидания фундаменталистов и чартистов о ставке обменного курса, что привносит нелинейность в динамику курса. Эта модель позволяет решить такой важный вопрос макроэкономики, как загадку отсутствия взаимосвязи (disconnect puzzle) – колебания валютного курса в краткосрочной перспективе почти не зависят от таких макроэкономических величин, как ставка процента, темп инфляции и других (Grauwe, Grimaldi, 2005).

У. Бранч и Б. Макгоф инкорпорировали разнородные ожидания в **Новую Кейнсианскую модель** – так, совокупное предложение может зависеть от убеж-

дений о будущей инфляции и о текущем объеме производства (Branch, McGough, 2009). Подобную модель вывел и Д. Массаро, разница состоит в том, что в первом случае авторам пришлось сделать некоторые ограничительные допущения в отношении убеждений, чтобы получить агрегирование, во втором – не налагаются ограничения на набор возможных правил прогнозирования (Massaro, 2013). У. Брок и К. Хомс предложили свою модель динамики в простой модели ценообразования активов с дисконтированной стоимостью, где участвуют агенты с разнородными убеждениями (Brock, Hommes, 1998). В. Сюн и Х. Ян предложили модель динамического равновесия рынков облигаций, где также присутствуют разнородные ожидания агентов (Xiong, Yan, 2010). Динамика обучения создает теорию эволюции ожиданий и выбора между альтернативными равновесиями, что имеет последствия для бизнес-циклов, волатильности цен на активы и политики (Evans, Honkapohja, 2001).

Еще классик экономического мысли Дж. М. Кейнс указывал на то, что экономическое поведение человека может находиться под влиянием иррациональных ожиданий, оптимизма или пессимизма по отношению к экономической активности. Он ввел понятие *animal spirits*, или жизнерадостности, что в экономическом смысле означает – спонтанное стремление действовать, а не бездействовать, а также волны оптимизма и пессимизма, которые охватывают инвесторов.

П. де Грауве применяет концепцию жизнерадостности к кейнсианской модели динамического стохастического общего равновесия (DSGE), в которой идея волн оптимизма и пессимизма заменяет рациональные ожидания агентов. Его поведенческая DSGE-модель обладает рядом преимуществ.

Во-первых, из-за свойства нелинейности, поведенческая модель устанавливает степень неопределенности в отношении переноса шоков монетарной политики, которые отличны от неопределенности, представленной в DSGE-моделях. В них неопределенность от эффектов шоков монетарной политики возникает только по причине недостаточной точности оценок структурных параметров модели. В поведенческой модели вводится дополнительная неопределенность – одна и та же монетарная политика может по-разному влиять на экономических агентов, исходя из жизнерадостности, т.е. уровня оптимизма или пессимизма относительно будущего.

Во-вторых, режим инфляционного таргетирования приобретает в поведенческой модели очень большое значение для стабилизации экономики. В ситуации, когда инфляционное таргетирование заслуживает доверия, непостоянство инфляции и выпуска производства невелико. Причина состоит в том, что способность внушать доверие также помогает уменьшить взаимосвязь между убеждениями и последующими волнами оптимизма и пессимизма.

В-третьих, строгое инфляционное таргетирование не является оптимальной политикой. Стабилизация выпуска производства, являющаяся заслуживающей доверия инфляционной целью, также помогает сокращать влияния предвзятых убеждений, что таким образом сокращает масштаб волн оптимизма и пессимизма, которые дестабилизируют выпуск производства и инфляцию.

В-четвертых, поведенческая модель использует совершенно другую модель бизнес-цикла в сравнении с моделью бизнес-цикла, используемой в DSGE-

модели, в которой выпуск производства возрастает только потому, что рациональные агенты не могут мгновенно приспособить свои оптимальные планы после экзогенных возмущений. Цены и негибкость заработной платы препятствуют такой мгновенной подстройке. В результате экзогенные шоки создают инерцию и движения бизнес-цикла. Агенты в поведенческой модели не только не могут мгновенно подстраивать свои цены, но и испытывают информационные проблемы. Они не знают ни природы шока, ни то, как он повлияет на среду. Методом проб и ошибок они получают значимую информацию. Эта когнитивная проблема создает инерцию выпуска производства и цен в дополнение к инерции цен, происходящей от того что контракты не могут быть стремительно откорректированы. Таким образом, получена значительно более глубокая теория бизнес-циклов (Grauwe, 2011).

П. де Грауве совместно с К. Маккиарелли применяет концепцию «жизнелюбия», «жизнерадостности», спонтанного стремления действовать (*animal spirits*) к банковской сфере. В их поведенческой DSGE-модели возникают эндогенные и самовоспроизводящиеся волны оптимизма и пессимизма. Существование банков усиливает эти движения, создавая больше возможностей для бумов и провалов (Grauwe, Macchiarelli, 2015).

Р. Фармер, используя концепцию «жизнелюбия», «жизнерадостности», спонтанного стремления действовать взамен неоклассической модели равновесия на товарном и денежном рынках IS-LM, предлагает альтернативу – модель IS-LM-NAC, где добавляется ещё и третья кривая NAC (*no order condition*), отражающая отсутствие арбитража между физическим капиталом и финансовыми активами. Ее ключевые отличия состоят в том, что, во-первых, равновесия стационарного состояния демонстрируют динамическую неопределенность. Для каждого равновесия устойчивого состояния есть множество динамических путей, каждый из которых ведет к одному и тому же устойчивому состоянию. Модель применяется, для того чтобы объяснить, как изменения в предложении денег могут быть ассоциированы с изменениями в реальной экономической активности без введения искусственных барьеров для изменения цены. Также эта модель демонстрирует неопределенность устойчивого состояния. Фармер использует модель поиска труда, в которой существование внешних эффектов генерируется множественными равновесиями стационарного состояния. Классические модели поиска утверждают, что потенциальные работники и фирмы торгуются по заработной плате. Вместо этого Фармер предлагает модель, где фирмы и потенциальные работники берут заработные платы и цены как данность, а безработица определяется совокупным спросом. Она используется, для того чтобы объяснить высокий уровень безработицы, который в модели отражает постоянные потенциальные отклонения рыночного равновесия от социального оптимума (Farmer, Platonov, 2019).

По мнению Фармера, «жизнелюбие», «жизнерадостность», спонтанное стремление действовать имеет очень большое влияние на равновесие, так как убеждения имеют фундаментальный статус, наряду с предпочитаемыми технологиями. В его модели уверенность – независимая движущая сила, которая определяет долгосрочное стационарное состояние уровня безработицы. Функцией убеждений (*beliefs function*) в своей модели он заменяет кривую Филипса (Farmer, 2010, 2018).

Концепция «жизнелюбия», «жизнерадостности», спонтанного стремления действовать встретила и критику. Так, Р. Барски и Э. Симс создали DSGE-модель, учитывающую уверенность потребителя, под которой они понимают не только жизнерадостность, но и новости, информацию об актуальном и будущем состоянии экономики. Авторы внедряют в модель два вида шоков. Первый – новостные шоки о том, что производство будет расти, сменяется постепенными изменениями макроэкономических переменных. Второй тип шоков – спонтанные волны оптимизма и пессимизма, создающие изменения в потреблении и доходе. Авторы обнаружили, что шоки жизнерадостности слабо воздействуют на переменные и чтобы спонтанные волны оптимизма и пессимизма имели важные экономические последствия, общие силы равновесия, действующие против них, должны быть слабыми. Например, при жестких ценах и центральном банке, не корректирующем процентные ставки с учетом колебаний объема производства. Результаты показали, что фундаментальные новости являются основной движущей силой наблюдаемой взаимосвязи между уверенностью и последующей экономической активностью (Barsky, Sims, 2012).

2.6. Гипотеза эффективных рынков

Поведенческие финансы – одна из наиболее масштабных исследовательских программ поведенческой экономики. Она представляет собой обширный ряд эффектов, вводящих существенные поведенческие дополнения в неоклассическую теорию эффективного рынка Ю. Фама. В ее центре лежит идея о том, что вся существенная информация немедленно и в полной мере отражается на рыночной курсовой стоимости ценных бумаг. Инвесторы рациональны и, таким образом, оценивают ценные бумаги рационально. Если инвесторы иррациональны, то сделки случайны и сводят на нет друг друга, не влияя на цены. Если инвесторы иррациональны схожим образом, то их встречают на рынке трейдеры-арбитражеры, которые извлекают прибыль за счет практически одномоментной купли и продажи одного и того же актива на различных рынках при обнаружении разных цен на этот актив, тем самым устраняя влияние иррациональных инвесторов на цены (Fama, 1970).

Однако исследователи обнаружили массу свидетельств, указывающих на несовершенство теории неэффективного рынка. Так, В. де Бондт и нобелевский лауреат Р. Талер указали на важный **феномен избыточной реакции** (overreaction) – люди чрезмерно реагируют на неожиданные и драматические новости, что делает их менее чувствительными к потрясениям прошлого (De Bondt, Thaler, 1985). Они проанализировали данные Нью-Йоркской фондовой биржи и выделили 35 наиболее успешных акций в «удачный портфель» и 25 наименее успешных в «неудачный портфель». Авторы исследовали успешность каждого портфеля на протяжении трех лет и с удивлением обнаружили, что акции «неудачного портфеля» превысили рыночный индекс, а акции «удачного», наоборот, не достигли его. По прошествии трех лет разница между успешностью этих портфелей была около 25%. Компании из «удачного портфеля» становятся неудачными, и наоборот. По мнению де Бондта и Талера, причина этого явления состоит в том, что в обоих случаях инвесторы реагировали избыточно. В случае неудачных акций инвесторы реагировали на плохие новости и непропорционально сильно понижали цены, но через некоторое время обнаруживали,

что были слишком пессимистичными, и неудачные акции показывали хороший рост. Такой же процесс происходил и с успешными акциями, только в обратную сторону. При этом срабатывает искажение доступной информации (availability bias) – люди активнее реагируют на новости, которые получили недавно, чем на весь массив прошлых новостей, и принимают решения по большей части на основе последних новостей.

Ш. Бенарци и Р. Талер обнаружили **феномен близорукого избегания потерь** (myopic loss aversion). Они объясняют загадку премии за риск (equity premium puzzle): риски по акциям более высокие, чем по облигациям. Инвесторы, которые стремятся избежать рисков, покупают облигации, несмотря на то что их доходность в среднем значительно ниже. В краткосрочной перспективе риски по акциям выше, чем по облигациям, в долгосрочной – доход по облигациям ниже. Однако инвесторы не желают заглядывать вдаль и делать долгосрочную оценку акций. Ш. Бенарци и Р. Талер объясняют эту закономерность избеганием потерь и ментальной бухгалтерией, т.е. коротким периодом оценки доходности (Benartzi, Thaler, 1995).

Также большое значение имеют настроения инвесторов (investor sentiment), описывающие поведение трейдеров, полагающихся не на фундаментальный анализ, а на некие правила принятия решений (эвристику). Настроения инвесторов могут значимо влиять на отклонения цен от фундаментальных значений. Настроения инвесторов формируются благодаря тому, что люди, делая прогнозы, уделяют слишком большое внимание доказательствам и слишком малое – вероятности того или иного явления. Так, например, серия новостей о хороших доходах компании – убедительная и достоверная информация, но обладает низкой вероятностью относительно продолжения тренда (Barberis, Shleifer, Vishny, 1998).

Помимо избыточной реакции на рынках существует недостаточная реакция (underreaction), представление о которой ввели Х. Хонг и Дж. Стайн (Hong, Stein, 1999). Под недостаточной реакцией понимается то, как позитивные новости (такие как, например, годовой отчет о доходах) влияют на цены. Позитивная информация медленно влияет на цены, эффект может растянуться на 1–12 месяцев, и пока влияние этих позитивных новостей не примет максимального значения, реакция рынка не будет достаточной.

Х. Шефрин и М. Стетман (Shefrin, Statman, 1985) открыли **эффект диспозиции** (disposition effect). Трейдеры склонны быстро продавать растущие акции и неохотно расстаются с акциями, которые теряют в цене. Это объясняется тем, что агенты избегают рисков в плане прибыли, но рискуют в сфере, где могут потерять. То есть при росте убытка от акции они недостаточно чувствуют нарастающие потери и продолжают рисковать. Инвесторы избегают продажи убыточных акций, потому что такое решение указывает, что их первоначальное суждение о прибыльности этой акции было ошибочным, а людям сложно признавать свои ошибки, им свойственны гордость за себя и стремление избегать сожалений. Т. Одеан (Odean, 1998) подтвердил наличие этого эффекта, проанализировав операции более 10 000 трейдинговых аккаунтов с 1987 по 1993 г.

Г. Хьюберман и Т. Регев описывают **эффект толпы** (herding effect) на фондовом рынке на примере акций компании «Entremed» (Huberman, Regev, 2001). В ноябре 1997 г. в «Нью-Йорк Таймс» вышла небольшая заметка о том, что компа-

ния разработала успешное лекарство от рака, что подняло стоимость акций компании приблизительно на 20%. В мае следующего года вышла информация того же самого содержания, но она была опубликована на обложке массовой газеты «Таймс», что подняло стоимость акций компании более чем в три раза, несмотря на то что фундаментально ничего не изменилось – информация, опубликованная в «Таймс», не являлась чем-то новым. В ноябре 1998 г. на обложке «Уолл Стрит Джорнал» вышла информация, согласно которой независимая лаборатория не смогла повторить результаты, полученные «Entremed», т.е. компания не обладает лекарством от рака, как первоначально заявлялось. Эта информация обрушила стоимость акций компании, но все равно цена на них осталась в два раза выше, чем была до публикации об успехах фирмы, вышедшей в мае. На основе этой динамики можно утверждать, что после публикации в ноябре 1997 г. акции были сильно недооценены, а после падения в ноябре 1998 г. акции все равно остались сильно переоцененными. Однако интереснее другой факт. Резкий рост (на 330%) акций «Entremed» вызвал энергичный рост стоимости акций семи других биотехнологических компаний, несмотря на то что к открытию «Entremed» они не имеют никакого отношения. Такой взлет цен можно описать как эффект толпы – случайная публикация, не имеющая в своей основе новой информации, способна создать необычайный рост стоимости акций компании, а он может спровоцировать рост акций компаний, которые не имеют отношения к опубликованным новостям.

На рынках существует **эффект чрезмерной самоуверенности** (overconfidence). Д. Монтье (Montier, 2007) провел опрос трехсот профессиональных менеджеров инвестиционных фондов на тему того, считают ли они, что их профессиональные способности выше или ниже среднего уровня. 74% менеджеров ответили, что их способности выше среднего, а среди оставшихся 26% большая часть ответила, что у них средние способности. Т. Одеан исследовал поведение чрезмерно самоуверенных трейдеров и пришел к выводу, что объем торгов на рынке возрастает из-за их действий и в какой-то мере является чрезмерным: люди торгуют даже тогда, когда им выгоднее было бы просто сохранить акции (Odean, 1999). Чрезмерная самоуверенность может вызвать недооценку рынками информации от рациональных трейдеров, приводя к серийным взаимосвязанным доходам. В некоторых случаях рынки реагируют чрезмерно, например на IPO-компания (Initial Public Offering), и недостаточно, например на извещение о прибыли. Самоуверенность сокращает ожидаемую полезность трейдера – такие игроки держат недиверсифицированные портфели акций, что повышает риски. Чрезмерной самоуверенностью, которая побуждает торговать даже тогда, когда выгоднее было бы сохранить акции, можно и объяснить чрезмерную волатильность рынков (Odean, 1998).

Данные эффекты поведенческих финансов указывают на слабые места неоклассической теории эффективного рынка и открывают пространство для создания новых теорий науки о финансах.

Заключение

Таким образом, наш обзор показывает, как теории поведенческой экономики возникли в противопоставление нереалистическим предпосылкам неоклассической экономики о субстанциональной рациональности, эгоистичности

и информированности поведения человека. Теории поведенческой экономики значительно дополнили или предложили альтернативу теориям ожидаемой полезности, игр, экспоненциальной дисконтированной полезности, выявленных предпочтений, рациональных ожиданий и эффективного рынка.

Наш обзор демонстрирует основные направления развития поведенческой экономики, систематизируя научное знание. Приведенная нами последовательность развития поведенческой экономики связывает ее, ранее неортодоксальное, авангардное направление науки, с основным направлением экономической мысли.

Несмотря на то что с исследованиями в области поведенческой экономики связаны имена таких нобелевских лауреатов, как Д. Канеман, Р. Талер, Г. Саймон, Дж. Акерлоф, на данный момент поведенческая экономика остается перспективным направлением экономической науки, открывая исследователю новые возможности и новые горизонты.

ЛИТЕРАТУРА / REFERENCES

- Abbinck K., Irlenbusch B., Renner E.** (2000). The moonlighting game: An experimental study on reciprocity and retribution. *Journal of Economic Behavior & Organization*, 42, 2, 265–277.
- Abeler J., Nosenzo D., Raymond C.** (2019). Preferences for truth-telling. *Econometrica*, 87, 4, 1115–1153.
- Agell J., Lundborg P.** (1995). Theories of pay and unemployment: Survey evidence from Swedish manufacturing firms. *The Scandinavian Journal of Economics*, 97, 2, 295–307.
- Akerlof G.** (1970). The market for “Lemons”: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84, 3, 488–500.
- Allais M.** (1953). Le comportement de l’homme rationnel devant le risque: Critique des postulats et axiomes de l’école Américaine. *Econometrica*, 21, 4, 503–546.
- al-Nowaihi A., Dhami S.** (2008). A general theory of time discounting: The reference-time theory of intertemporal choice. *University of Leicester, Department of Economics. Working Paper*. 08/22.
- Arrow K.** (1980). Discrimination in the labour market. *Readings in Labour Economics*. Oxford: Oxford University Press.
- Barberis N., Shleifer A., Vishny R.** (1998). A model of investor sentiment. *Journal of Financial Economics*, 49, 3, 307–343.
- Barsky R.B., Sims E.R.** (2012). Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review*, 102, 4, 1343–1377.
- Battigalli P., Dufwenberg M.** (2007). Guilt in games. *American Economic Review*, 97, 2, 170–176.
- Becker J.L., Sarin R.K.** (1987). Lottery dependent utility. *Management Science*, 33, 11, 1367–1382.
- Bell D.** (1985). Disappointment in decision making under uncertainty. *Operations Research*, 33, 1, 1–27.
- Bénabou R., Tirole J.** (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70, 3, 489–520.
- Benartzi S., Thaler R.H.** (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, 110, 1, 73–92.
- Benjamin D., Laibson D.** (2003). Good policies for bad governments: Behavioral political economy. *Conference Series (Proceedings)*, 48.

- Bertrand M., Mullainathan S.** (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review*, 94, 4, 991–1013.
- Blinder A., Choi D.** (1990). A shred of evidence on theories of wage stickiness. *The Quarterly Journal of Economics*, 105, 4, 1003–1015.
- Bolton G.E., Ockenfels A.** (2000). ERC: A theory of equity, reciprocity, and competition. *The American Economic Review*, 90, 1, 166–193.
- Bordalo P., Gennaioli N., Shleifer A.** (2012). Saliency theory of choice under risk. *The Quarterly Journal of Economics*, 127, 3, 1243–1285.
- Branch W., McGough B.** (2009). A new Keynesian model with heterogeneous expectations. *Journal of Economic Dynamics & Control*, 33, 1036–1051.
- Brock W., Hommes C.** (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22, 8–9, 1235–1274.
- Burnham T.** (2007). High-testosterone men reject low ultimatum game offers. *Proceedings: Biological Sciences*, 274, 1623, 2327–2330.
- Bush R.R., Mosteller F.** (1955). *Stochastic models for learning*. New York: John Wiley & Sons, Inc.
- Camerer C., Issacharoff S., Loewenstein G., O'Donoghue T.** (2003). Regulation for conservatives: Behavioral economics and the case for 'asymmetric paternalism'. *University of Pennsylvania Law Review*, 151, 1211–1254.
- Camerer C.F., Ho T.-H., Chong J.-K.** (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 3, 119, 861–898.
- Charness G., Dufwenberg M.** (2006). Promises and partnership. *Econometrica*, 74, 1579–1601.
- Charness G., Rabin M.** (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117, 3, 817–869.
- Chew S.H.** (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica*, 51, 4, 1065–1092.
- Chew S.H.** (1989). Axiomatic utility theories with the betweenness property. *Annals of Operations Research*, 19, 273–298.
- Coates J.M., Herbert J.** (2008). Endogenous steroids and financial risk taking on a London trading floor. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 16, 6167–6172.
- Crockett M.J., Fehr E.** (2014). Pharmacology of economic and social decision making. In: *Neuroeconomics*. Vol. 2. W. Glimcher, E. Fehr (eds.). London: Elsevier Inc., 257–279.
- Dana J., Loewenstein G.** (2003). A psychological perspective on the influence of gifts to physicians from industry. *Journal of the American Medical Association*, 290, 2, 252–255.
- De Bondt W., Thaler R.** (1985). Does the stock market overreact? *The Journal of Finance*, 40, 3, 793–805.
- Dekel E.** (1986). An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic Theory*, 40, 2, 304–318.
- Dufwenberg M., Gneezy U.** (2000). Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior*, 30, 2, 163–182.
- Elster J.** (1998). Emotions and economic theory. *Journal of Economic Literature*, 36, 1, 47–74.

- Erev I., Roth A.E.** (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88, 4, 848–881.
- Evans G.W., Honkapohja S.** (2001). *Learning and expectations in macroeconomics*. Princeton: Princeton University Press.
- Falk A., Becker A., Dohmen T., Enke B., Huffman D., Sunde U.** (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, 133, 4, 1645–1692.
- Falk A., Fischbacher U.** (2006). A theory of reciprocity. *Games and Economic Behavior*, 54, 2, 293–315.
- Fama E.F.** (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25, 2, 383–417.
- Farmer R.** (2010). Animal spirits, persistent unemployment and the belief function. *NBER Working Papers*, 16522.
- Farmer R.** (2018). Keynesian economics without the Phillips curve. *Journal of Economic Dynamics and Control*, 89, C, 137–150.
- Farmer R., Platonov K.** (2019). Animal spirits in a monetary model. *European Economic Review*, 115, C, 60–77.
- Fehr E., Gächter S.** (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, 14, 3, 159–181.
- Fehr E., Schmidt K.** (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114, 3, 817–868.
- Fischbacher U., Föllmi-Heusi F.** (2013). Lies in disguise: An experimental study on cheating. *Journal of the European Economic Association*, 11, 3, 525–547.
- Fox C.R., Tversky A.** (1995). Ambiguity aversion and comparative ignorance. *The Quarterly Journal of Economics*, 110, 3, 585–603.
- Frankel J., Froot K.** (1987). Short-term and long-term expectations of the yen/dollar exchange rate: Evidence from survey data. *Journal of the Japanese and International Economies*, 1, 3, 249–274.
- Frankel J., Froot K.** (1990). Chartists, fundamentalists and trading in the foreign exchange market. *American Economic Review*, 80, 2, 181–185.
- Frye T., Shleifer A.** (1997). The invisible hand and the grabbing hand. *American Economic Review Papers and Proceedings*, 87, 2, 358.
- Gabaix X.A.** (2014). Sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 129, 4, 1661–1710.
- Geanakoplos J., Pearce D., Stacchetti E.** (1989). Psychological games and sequential rationality. *Games and Economic Behavior*, 1, 1, 60–79.
- Gilboa I., Schmeidler D.** (1995). Case-based decision theory. *The Quarterly Journal of Economics*, 110, 3, 605–639.
- Gilboa I., Shmeidler D.** (1997). Cumulative utility and consumer theory. *International Economic Review*, 38, 4, 737–761.
- Grauwe P. de** (2011). Animal spirits and monetary policy. *Economic Theory*, 47, 423–457.
- Grauwe P. de, Grimaldi M.** (2005). Heterogeneity of agents and the exchange rate: A nonlinear approach. In: *Exchange rate economics: Where do we stand?* Cambridge: The MIT Press, 125–167.
- Grauwe P. de, Macchiarelli C.** (2015). Animal spirits and credit cycles. *Journal of Economic Dynamics and Control*, 59, C, 95–117.

- Gul F.** (1991). A theory of disappointment aversion. *Econometrica*, 59, 3, 667–686.
- Gul F., Pesendorfer W.** (2001). Temptation and self-control. *Econometrica*, 69, 1403–1435.
- Güth W., Kocher M.G.** (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. *Journal of Economic Behavior & Organization*, 108, 396–409.
- Güth W., Schmittberger R., Schwarze B.** (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3, 4, 367–388.
- Heath C., Tversky A.** (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4, 5–28.
- Hommes C.** (2013). *Behavioral rationality and heterogeneous expectations in complex economic systems*. Cambridge: Cambridge University Press.
- Hong H., Stein J.** (1999). Unified theory of underreaction, momentum trading and overreaction in asset markets. *Journal of Finance*, 54, 6, 2143–2184.
- Huberman G., Regev T.** (2001). Contagious speculation and a cure for cancer: A Nonevent that made stockprices soar. *Journal of Finance*, 56, 1, 387–396.
- Johnson E.J., Hershey J., Meszaros J., Kuhnreuther H.** (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7, 35–51.
- Kahneman D., Knetsch J., Thaler R.** (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review*, 76, 4, 28–41.
- Kahneman D., Tversky A.** (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 2, 263–291.
- Katona G.** (1980). How expectations are really formed. *Challenge*, 23, 5, 32–35.
- Khalmetski K., Ockenfels A., Werner P.** (2015). Surprising gifts: Theory and laboratory evidence. *Journal of Economic Theory*, 159, 163–208.
- Klick J., Mitchell G.** (2006). Government regulation of irrationality: Moral and cognitive hazards. *Minnesota Law Review*, 90, 1620–1663.
- Kosfeld M., Heinrichs M., Zak P., Fischbacher U., Fehr E.** (2005). Oxytocin increases trust in humans. *Nature*, 435, 7042, 673–676.
- Kószegi B., Rabin M.** (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121, 4, 1133–1165.
- Laibson D.** (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112, 2, 443–478.
- Loewenshtein G., Sicherman N.** (1991). Do workers prefer increasing wage profiles? *Journal of Labor Economics*, 9, 1, 67–84.
- Loewenshtein G., Thaler R.** (1989). Anomalies: Intertemporal choice. *Journal of Economic Perspectives*, 3, 4, 181–193.
- Loewenstein G.** (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 65, 3, 272–292.
- Loewenstein G., Haisley E.** (2006). The economist as therapist: Methodological ramifications of ‘light’ paternalism. In: *Perspectives on the future of economics: Positive and normative foundations. The Handbook of Economic Methodologies*, 1. Oxford: Oxford University Press.
- Loomes G., Sugden R.** (1986). Disappointment and dynamic consistency in choice under uncertainty. *Review of Economic Studies*, 53, 2, 271–282.
- MacCrimmon K.** (1965). An experimental study of the decision making behavior of business executives. *Unpublished dissertation*. University of California at Los Angeles.

- Machina M.J.** (1982). “Expected utility” analysis without the independence axiom. *Econometrica*, 50, 2, 277–323.
- Massaro D.** (2013). Heterogeneous expectations in monetary DSGE models. *Journal of Economic Dynamics & Control*, 37, 680–692.
- McKelvey R.D., Palfrey T.R.** (1992). An experimental study of the centipede game. *Econometrica*, 60, 4, 803–836.
- McKelvey R.D., Palfrey T.R.** (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 1, 10, 6–38.
- Montier J.** (2007). *Behavioural investing: A practitioner’s guide to applying behavioural finance*. N.Y.: Wiley.
- Morrison D.G.** (1967). On the consistency of preferences in Allais’ paradox. *Behavioral Science*, 12, 5, 373–383.
- Muth J.** (1961). Rational expectations and the theory of price movements. *Econometrica*, 29, 3, 315–335.
- Nagel R.** (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, 85, 5, 1313–1326.
- Odean T.** (1998). Are investors reluctant to realize their losses? *Journal of Finance*, 53, 5, 1775–1798.
- Odean T.** (1999). Do investors trade too much? *American Economic Review*, 89, 5, 1279–1298.
- Palacios-Huerta I., Volij O.** (2008). Experientia docet: Professionals play minimax in laboratory experiments. *Econometrica*, 76, 1, 71–115.
- Quiggin J.** (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3, 4, 323–343.
- Quiggin J.** (1993). *Generalized expected utility theory: The rank-dependent model*. Boston, Dordrecht, London: Kluwer Academic Publishers.
- Rabin M.** (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83, 5, 1281–1302.
- Rosenthal R.** (1981). Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory*, 25, 1, 92–100.
- Roth A., Prasnikar V., Okuno-Fujiwara M., Zamir S.** (1991). Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study. *American Economic Review*, 81, 5, 1068.
- Rubinstein A.** (1988). Similarity and decision-making under risk (is there a utility theory resolution to the Allais paradox?). *Journal of Economic Theory*, 46, 1, 1988, 145–153.
- Samuelson P.** (1937). A note on measurement of utility. *The Review of Economic Studies*, 4, 2, 151–156.
- Samuelson P.A.** (1938). A note on the pure theory of consumer’s behaviour. *Economica*, 5, 17, 61–71.
- Sargent T.J.** (1993). *Bounded rationality in macroeconomics*. Oxford: Oxford University Press.
- Schmidt U., Starmer C., Sugden R.** (2008). Third-generation prospect theory. *Journal of Risk and Uncertainty*, 36, 3, 203–223.
- Shefrin H., Statman M.** (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance*, 40, 3, 777–790.
- Simon H.** (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69, 1, 99–118.

- Simon H.** (1978). Rationality as process and as product of thought. *The American Economic Review*, 68, 2, 1–16.
- Smith V., Suchanek G., Williams A.** (1988). Bubbles, crashes and endogenous expectations in experimental spot asset markets. *Econometrica*, 56, 5, 1119–1151.
- Stahl D., Wilson P.** (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior and Organization*, 25, 309–327.
- Sugden R.** (2003). Reference-dependent subjective expected utility. *Journal of Economic Theory*, 111, 172–191.
- Thaler R.** (1981). Some empirical evidence of dynamic inconsistency. *Economics Letters*, 8, 201–207.
- Thaler R., Sunstein C.** (2003). Libertarian paternalism. *American Economic Review*, 93, 2, 175–179.
- Thaler R.H.** (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, 12, 183–206.
- Tversky A., Kahneman D.** (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 4157, 1124–1131.
- Tversky A., Kahneman D.** (1985). The framing of decisions and the psychology of choice. *Behavioral decision making*. Boston: Springer, 25–41.
- Tversky A., Kahneman D.** (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 4, 297–323.
- Tversky A., Koehler D.J.** (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, 101, 4, 547–567.
- Xiong W., Yan H.** (2010). Heterogeneous expectations and bond markets. *The Review of Financial Studies*, 23, 4, 1433–1466.
- Yaari M.E.** (1987). The dual theory of choice under risk. *Econometrica*, 55, 1, 95–115.

Поступила в редакцию 22.09.2020

Received 22.09.2020

A.A. Upravitelev

St. Petersburg State University, St. Petersburg, Russia; European University Institute, Italy

Neoclassical roots of behavioral economics³

Abstract. Despite the fact that the names of a number of Nobel laureates are associated with behavioral economics, researchers regard this branch of science as an unorthodox, avant-garde research program. This article aims to fill this theoretical gap and link behavioral economics to mainstream of economic thought, as well as to divide behavioral economics into several directions – relative to the initial approaches or theories of neoclassics. In the first row of cases, we collected those approaches of neoclassical economics about economic behavior, in relation to which alternative concepts were proposed by behavioral economics. These neoclassical approaches are rationality, awareness, egoism of behavior. Another series is neoclassical theories, in dialogue with which theories of behavioral economics arose. These are theories of expected utility, game theory, exponential discounted utility, revealed preferences, rational expectations, efficient market. The sequence of development of behavioral economics we presented connects it, a once heterodox branch of science with the mainstream economic thought. The article can serve as a guide to the most important areas of research in behavioral economics. This study systematizes scientific knowledge and establishes methodological links within it.

Keywords: *behavioral economics, economic methodology, research program, bounded rationality, prospect theory, neuroeconomics.*

JEL Classification: B41, B59, D01.

For reference: **Upravitelev A.A.** (2023). Neoclassical roots of behavioral economics. *Journal of the New Economic Association*, 1 (58), 110–140. DOI: 10.31737/22212264_2023_1_110

³ I am grateful to Alexis Belianin, Alexander Dubyanskiy, and Vsevolod Ostapenko for helpful comments and suggestions.