

Ключевые слова:

теория хаоса,
гипотеза эффективного рынка,
прогнозы ценовой динамики,
случайное блуждание цен

Л. Н. Каваленя, аспирант кафедры «Фондовые рынки»
Государственного университета Минфина России (ГУМФ),
главный актуарий ПСГ «Основа»
(e-mail: leonardokavaleni@ya.ru)

Изменение цен: случайное блуждание или хаотический процесс?

С момента опубликования в 1952 г. Г. Марковицем работы «Выбор портфеля» процесс изменения цен финансовых активов принято рассматривать как стохастическое случайное блуждание. Эта парадигма оказалась исключительно плодотворной (достаточно вспомнить CAPM и формулу Блэка — Шоулза). Тем не менее она не является безальтернативной. Принципиально иной подход к анализу ценообразования предлагает теория хаоса, согласно которой эволюция цен — нелинейный динамический процесс. Насколько обоснованна данная альтернатива? Мы предприняли попытку «тестирования» этой гипотезы для российского фондового рынка в период кризиса 2008 г.

Прежде чем перейти к изложению полученных результатов, остановимся на экономическом смысле двух указанных парадигм. Для иллюстрации различия между ними вначале рассмотрим два утверждения:

1. «Ряд 11111110 менее случайный, чем 10011010».
2. «...Падение (котировок) выглядит на удивление логичным — после семи сессий подряд роста инвесторы взяли паузу. Рынку нужна консолидация...»¹.

Первое утверждение кажется вполне разумным. Однако эта разумность обманчива. Все дело в распространенном заблуждении: свойство всей совокупности (частота появления 1–50 %) должно проявляться даже в малой выборке из этой совокупности. Ряды, в которых это свойство нарушается (например, в них содержатся длинные серии), воспринимаются человеком как **неслучайные**. Д. Канеман, П. Словик и А. Тверски описали массу увлекательных примеров, когда даже подготовленные испытуемые поддаются подобным статистическим «миражам»².

¹ Полную версию «прогноза от УК» см.: Дальнейшие «успехи» российского рынка будут во многом зависеть от динамики цен на нефть и притока «западных» денег / Финанс.ру (<http://www.finam.ru/analysis/marketnews4798E/default.asp>).

² Канеман, Д., Тверски, А. Рациональный выбор, ценности и фреймы // Психологический журнал. — 2003. — Т. 24. — № 4. — С. 31–42.

Теперь рассмотрим второе утверждение. Это типичный «комментарий технического аналитика». И такими комментариями буквально переполнены СМИ. Но ведь если исходить из **стохастической природы динамики цен**, то приведенные выше утверждения имеют примерно равную, т. е. отрицательную, научную ценность. Приходится признать, что за каждым из них не стоит ничего кроме свойственной любому человеку склонности находить псевдозакономерности даже в абсолютно случайных данных.

Если же рассматривать эволюцию цен не как стохастическое случайное блуждание, а как нелинейный динамический, или хаотический, процесс, то второе утверждение может иметь статистический смысл. Поясним принципиальное содержание этих двух подходов.

Отличие хаотических процессов от стохастических заключается в количестве факторов, объясняющих эволюцию системы. В стохастической системе их очень много. Связи между факторами и объясняемой переменной **вероятностные**: значению объясняющей переменной соответствует распределение значений зависимой. В хаотической системе все иначе: факторов мало, из них доминирующих может быть вообще два или три. Кроме того, зависимости между переменными функциональны: точно зная текущее положение (значения всех переменных) можно однозначно предсказать (просто рассчитать) положение на следующем шаге.

Однако из-за наличия **обратной связи** даже микроскопической ошибки в точности определения текущего положения достаточно, чтобы реальная будущая траектория **экспоненциально быстро** разбежалась с траекторией построенного прогноза. Хаотические системы даже с малым числом переменных демонстрируют это свойство, что делает их эволюцию очень похожей на действительно непрогнозируемое случайное блуждание.

Выше мы приводили абсурдность «комментария технического аналитика», если искать ему аналогию в рамках **стохастического подхода**. Если же искать аналогию в классических динамических моделях, то второе утверждение может оказаться вполне содержательным. Рассмотрим простейшую модель ценовой эволюции:

$$\frac{dx}{dt} = \lambda \cdot (V_t^+ - V_t^-), \tag{1}$$

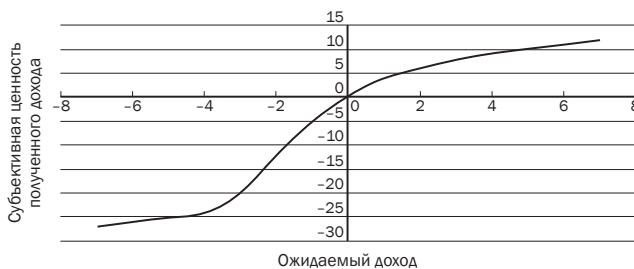
где x_t — цена в момент t ; V_t^+ — заявки на покупку; V_t^- — заявки на продажу; λ — параметр ликвидности рынка.

Смысл уравнения прозрачен: любое изменение цены — результат разбалансирования спроса и предложения. Избыток спроса — цена растет; недостаток — падает. Цена перестанет меняться, только когда наступит равновесие: $V_t^+ = V_t^-$.

Усложним модель, добавив в нее обратные связи. На наш взгляд, для их возникновения достаточно принять всего одно «мягкое» допущение. А именно — о нелинейности функции полезности, по крайней мере, у части инвесторов (рис. 1).

Рисунок 1

**Эмпирическая функция полезности
ожидаемой доходности по А. Тверски и Д. Канеману**



Источник: Канеман Д., Тверски А. Рациональный выбор, ценности и фреймы // Психологический журнал. — 2003. — Т. 24. — № 4. — С. 31–42.

Модели «рационального инвестора» это допущение противоречит (функция — не-прямая линия). Зато эмпирическим наблюдениям соответствует очень хорошо. Как показали опыты нобелевского лауреата Д. Канемана, отношение реальных инвесторов (и людей вообще) к ожидаемой доходности и риску **зависит от точки отсчета**, т. е. прибыльности их текущей открытой позиции.

В таком случае на решения трейдеров кроме настоящего начинает влиять и прошлое — то, как цена актива пришла к нынешнему значению. И процесс принятия решений о заключении сделок становится немарковским:

$$\frac{dV_t^+}{dt} = f_1(V_{t-1}^-); \quad \frac{dV_t^-}{dt} = f_2(V_{t-1}^+). \quad (2)$$

Объединяя (1) и (2), получаем классическую нелинейную модель с обратными связями, в которой «все зависит от всего»: текущее изменение цены влияет на будущее поведение инвесторов, которые своими действиями приводят к очередному изменению цен, и т. д.

Представляется любопытным интуитивное сходство данной системы со знаменитой метеорологической моделью Э. Лоренца:

$$\begin{cases} \dot{x} = -sx + sy \\ \dot{y} = -xz + rx - y, \\ \dot{z} = xy - bz \end{cases} \quad (3)$$

где x — скорость течения; y — давление; z — температура.

Как видно, уравнение для скорости течения повторяет модель цены в (1). Всего трех простых уравнений в (3) оказывается достаточным, чтобы долгосрочный прогноз системы был невозможен. Ошибка в определении даже десятой цифры после запятой у любого параметра сразу распространяется на все три координаты и очень быстро нарастает, что уничтожает ценность любого долгосрочного прогноза. Из-за этой непрогнозируемости хаотическую систему легко спутать со стохастическим блужданием.

Описывать хаотическую систему можно и с помощью аппарата теории вероятностей — устойчивость частот событий присутствует и здесь. Более того, на долгосрочных периодах вероятностные прогнозы вообще единственно возможны. Однако на краткосрочных интервалах точность динамических моделей может быть гораздо выше. Поэтому знание о том, что изучаемая система генерируется хаосом (несколько доминирующих факторов), а не случайностью (очень много сравнительно равновесных факторов), дает теоретическую возможность существенно улучшить краткосрочный прогноз.

Возможно, финансовые ряды порождаются именно хаотическими системами. И существует небольшое число факторов, которые объясняют движения мировых индексов, валют и т. д. В данной статье осуществлена попытка верификации этой крайне заманчивой гипотезы.

Практическая необходимость в различении случайных и хаотических временных рядов существует во многих сферах — от медицины до астрономии. Поэтому уже разработан ряд математических методов, позволяющих анализировать степень хаотичности. Мы использовали тесты, построенные на основе следующих подходов:

- Определение стохастической сложности объекта (по А. Колмогорову).
- Вычисление корреляционного интеграла.

³ Чтобы ошибка «добралась» из «15 знака после запятой» до целых значений, потребуется всего лишь несколько десятков итераций. Получается, что повышение точности измерений в миллионы раз практически не улучшает прогнозируемости процесса.

Статистический смысл определения стохастической сложности объекта заключается в тестировании ряда на возможность его представления в более компактной форме. Абсолютно случайный ряд невозможно сжать. Для его передачи без потерь нельзя использовать более короткую программу, чем он сам, что вполне возможно для рядов, порожденных хаотической системой.

Опираясь на эту логику, А. Лемпель и Я. Зив⁴ в 1976 г. ввели количественную меру стохастичности процесса — C_{LZ} , основанную на возможности сжатия данных. Алгоритм ее расчета выглядит следующим образом. Временной ряд преобразуется в последовательность нулей и единиц (больше/меньше среднего/медианы). Полученная последовательность разбивается на «слова» минимальной длины таким образом, чтобы ряд не содержал одинаковых «слов». Например, 010101000011 будет преобразовано в 0*1*01*010*00*011 («*» означает пробел). Суть теста заключается в том, что при стремлении избежать повторения «слов» приходится увеличивать их длину. Если ряд описывает детерминированный процесс, то будет проявляться тенденция повторения в нем одних и тех же «слов». Это приведет к необходимости разбивать последовательность на пакеты большей длины, чем если бы ряд был абсолютно случайным. В результате количество использованных «слов» при разбиении детерминированного (и хаотического) ряда всегда меньше, чем при разбиении случайного той же длины. Соответственно, различие в числе «слов», необходимых для разбиения данного ряда и случайного ряда той же длины, может служить оценкой его нестохастичности. Обозначим через N длину анализируемой последовательности: N_w — количество слов, использованных для разбиения ряда. Тогда

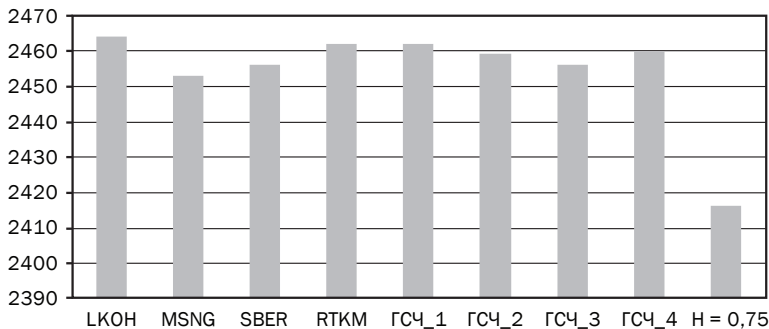
$$C_{LZ} = \lim_{N \rightarrow \infty} \frac{L(N)}{N}, \tag{4}$$

где $L(N) \sim N_w \cdot [\log_{N_w}(N) + 1]$.

У случайных процессов новые «слова» появляются чаще, чем у детерминированных, и C_{LZ} растет быстрее. Рассмотрим статистику N_w для некоторых акций ММВБ, случайных последовательностей и фрактального броуновского шума ($H = 0,75$) той же длины (рис. 2).

Рисунок 2

Статистика N_w для рядов часовой доходности акций в период 05.01.2000–24.12.2010 (23 524 наблюдения), четырех случайных рядов (ГСЧ) и броуновского шума ($H = 0,75$)



Источник: рассчитано автором.

⁴ Lempel A., Ziv J. On the complexity of finite sequences / IEEE Trans. Infor. Theory, IT-22. — 1976. — P. 75-81. Эта теоретическая работа легла в основу разработки всех современных архиваторов данных: Zip, RAR и др.

Приведенный алгоритм ориентирован на анализ больших рядов (сотни тысяч значений). Но и по 23 524 наблюдениям он обнаружил наличие зависимости в сгенерированном ряде фрактального блуждания с параметром Харста $H = 0,75$ (что и ожидалось). В то же время существенных отклонений от чистого стохастического процесса данный метод не зафиксировал ни для одной акции. Возможно, потому, что ряды действительно оказались несжимаемыми. Но причиной могут быть и недостатки самого метода.

Помимо указанного выше требования к объему данных, тест обладает тем недостатком, что анализирует только знак приращений. Более тонкие закономерности он не улавливает, что заставляет использовать другие тесты стохастичности. Например, IFS clumpiness test. Этот эмпирический тест также позволяет отличать стохастический белый шум от коррелированных шумов и хаотических процессов. Однако вместо двух состояний переменной (больше/меньше среднего) он оперирует уже четырьмя (квартили), что существенно повышает его точность.

Алгоритм проверки заключается в следующем:

1) Область значений переменной разбивается на четыре квартиля. Каждое наблюдение получает номер квартиля, к которому оно принадлежит. Получается новый временной ряд.

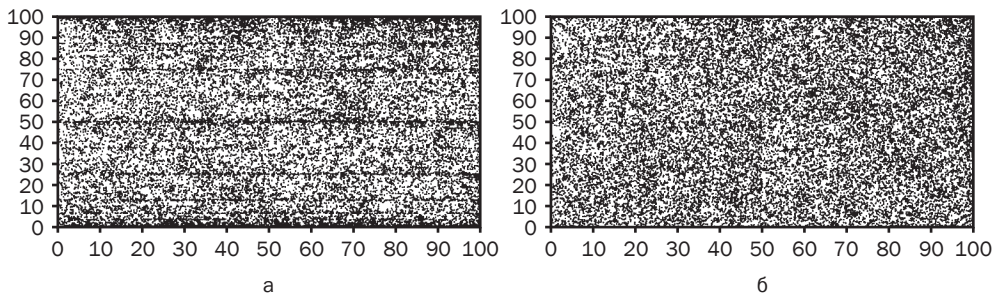
2) На плоскости чертится квадрат, и каждому углу ставится в соответствие квартал.

3) В центре квадрата располагается исходная точка. От нее откладывается половина расстояния до угла, соответствующего квартилю, которому принадлежит первое значение ряда. От полученной точки снова откладывается половина расстояния в направлении угла квартиля второго наблюдения и т. д. В результате исходный ряд отображается на некоторое множество точек внутри квадрата.

Идея теста в том, что характер распределения точек зависит от природы изучаемого процесса. Белый шум должен приводить к равномерному заполнению квадрата. Детерминированные и хаотические процессы отображаются в квадрат неравномерно. Несмотря на свою кажущуюся примитивность тест обладает хорошей диагностической силой (рис. 3).

Рисунок 3

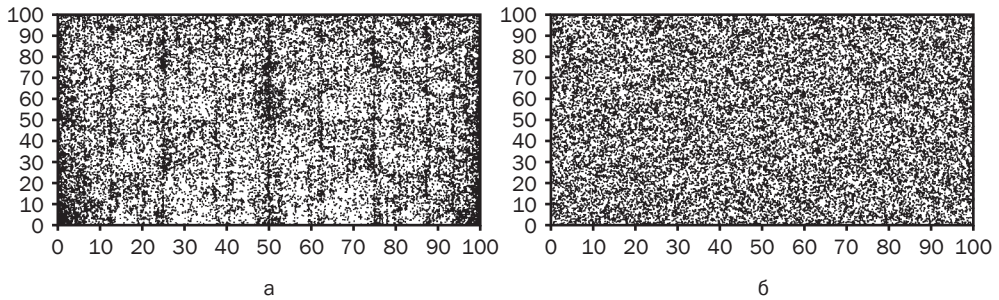
**IFS clumpiness test для ряда часовых приращений
SBER 05.01.2000–24.12.2010 (23 524 наблюдения) (а)
и после его перемешивания (б)**



Примечание: I квартал — левый нижний угол. Далее — по часовой стрелке.
Источник: рассчитано автором.

На левом рисунке хорошо видна группировка значений вдоль горизонтальных уровней $\approx 12,5; 25,0; 50,0; 75,0$. После перемешивания ряда и повторного расчета статистики данный эффект пропадает (справа). Аналогичная неустойчивость к перемешиванию наблюдается у всех процессов, обладающих эффектом статистической памяти. Например, у фрактального случайного блуждания (рис. 4).

IFS clumpiness test для ряда обобщенного Винеровского процесса с $H = 0,75$ (20 000 значений) (а) и после его перемешивания (б)



Источник: рассчитано автором.

Если бы параметр Харста был не 0,75, а 0,50 и корреляция между приращениями отсутствовала (обычный Винеровский процесс), то результат и до, и после перемешивания был бы идентичен изображенному на рис. 3б и 4б. Однако, как показывают расчеты, и для других акций⁵ равномерное заполнение квадрата наступает только после перемешивания последовательности.

Приведенные результаты не противоречат теории хаоса. Они доказывают, что ценовая динамика не является в чистом виде случайным блужданием с независимыми приращениями. Однако тест не позволяет определить причину эффекта памяти. Кроме хаотических процессов слепок (см. рис. 3б) могут оставлять и обычные коррелированные случайные блуждания. Ответить на вопрос, какая из альтернатив верна, тест не позволяет. Не позволяет он также оценивать нестохастичность и прогнозируемость количественно.

IFS clumpiness test отображает временной ряд в двумерное пространство (плоскость). При этом неоднородность заполнения предоставляемого объема (родовая черта хаотических, но не стохастических процессов) приходится оценивать на глаз.

Чтобы устранить эти недостатки, метод можно обобщить:

- отображать исходный процесс не на плоскость, а в пространство произвольной размерности n (скажем, в куб при $n = 3$);
- использовать неединичный лаг τ . Например, при отображении ряда x_t на плоскость исходное наблюдение x_1 станет точкой с координатами $\{x_1; x_{1+\tau}\}$, а не $\{x_1; x_2\}$, как в IFS clumpiness test;
- неравномерность заполнения предоставляемого пространства (т. н. пространства вложения) определять количественно⁶.

Наиболее распространенным методом, реализующим все три идеи, является расчет корреляционного интеграла:

$$C_{(\varepsilon)} = \frac{1}{N^2} \cdot \sum_{\substack{i,j \neq 1 \\ i \neq j}}^N \theta(\varepsilon - |x_i - x_j|), \tag{5}$$

⁵ За аналогичный период был произведен расчет для GZP, LKOH, ROSN.

⁶ Хотя бы потому, что в пространстве больше трехмерного определить неоднородность его заполнения на глаз затруднительно.

$$\text{где } \theta(x) = \begin{cases} 0 & \text{при } x < 0 \\ 1 & \text{при } x > 0 \end{cases};$$

$|x_i - x_j|$ — расстояние между векторами в пространстве вложения.

Смысл данной статистики интуитивно понятен: $C_{(\varepsilon)}$ показывает вероятность того, что расстояние между двумя произвольными точками исходного ряда после их проецирования в n -мерное пространство будет меньше ε . Очевидно, что $C_{(\varepsilon)}$ растет с увеличением ε . Представляет интерес поведение функции $C_{(\varepsilon)} = f(\varepsilon)$. Для хаотических процессов она растет иначе, чем при стохастическом блуждании. Если ряд случаен, то он равномерно заполнит любое предоставляемое пространство (см. рис. 3б и 4б для $n = 2$). В общем случае будет присутствовать зависимость

$$\begin{cases} C_{(\varepsilon)} = \varepsilon^{D_2} \\ D_2 \sim n \end{cases}, \quad (6)$$

где n — размерность пространства вложения;

D_2 — размерность восстановленного аттрактора⁷.

Зависимость $D_2 \sim n$ будет сохраняться при любой комбинации n и τ , при погружении в куб $C_{(\varepsilon)} \sim \varepsilon^3$, в десятимерное пространство и т. д.

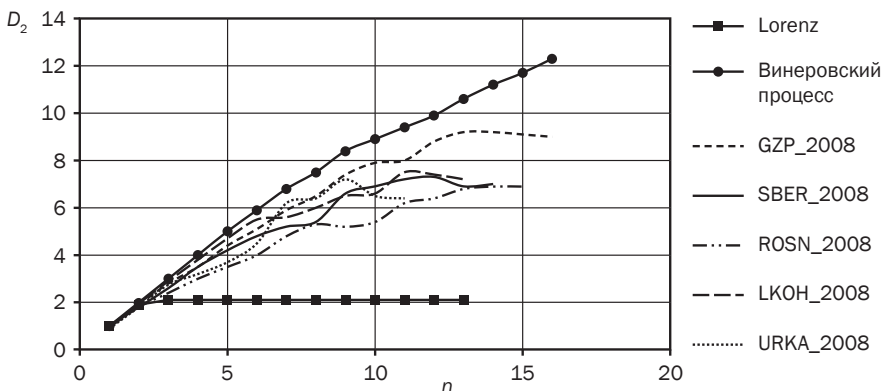
У хаотического процесса закономерность $D_2 \sim n$ нарушается:

$$\begin{cases} C_{(\varepsilon)} = \varepsilon^{D_2} \\ D_2 \leq n \end{cases}. \quad (7)$$

Для хаотической системы всегда можно подобрать такие n и τ , что начиная с некоторого n наступит насыщение D_2 . И дальнейшее погружение ряда в пространства больших размерностей не повлечет увеличения степени при ε . Поэтому характер зависимости $D_2 = f(n)$ принято использовать для строгого доказательства наличия во временном ряде детерминизма. С помощью специального программного обеспечения⁸ мы произвели расчет $D_2(n)$ для высокочастотных рядов приращений логарифмов цен некоторых акций, котирующихся на ММВБ (рис. 5).

Рисунок 5

Расчет корреляционной размерности D_2 для системы Лоренца, Винеровского процесса и рядов минутной логарифмической доходности некоторых акций ММВБ в 2008 г.



Источник: рассчитано автором.

⁷ Экономическая интерпретация этого термина будет приведена ниже.

⁸ Программа Fractan 4.4. / Институт математических проблем биологии РАН (<http://www.impb.ru/files.php>).

Расчет корреляционной размерности позволяет однозначно отличить динамику логарифмов цен акций от стохастического случайного блуждания. Для всех пяти рассмотренных бумаг D_2 растет медленнее, чем у сгенерированного белого шума. Это говорит о том, что ряды имеют **нестохастическую природу**.

Более того, анализ поведения функции $D_2 = f(n)$ позволяет оценить количество независимых переменных, определявших эволюцию хаотической системы. Например, для системы Лоренца с тремя переменными (3) D_2 достигает максимума при $n = 3$. При погружении в пространство больших размерностей корреляционная размерность множества не растет (см. рис. 5).

Для **любых** хаотических систем действует аналогичная логика. Если состояние системы описывается k независимыми сущностями, то D_2 входит в насыщение при пространстве вложения $n = k$. После чего увеличение n уже не будет находить отклик у D_2 . Вернемся к рис. 5. Как видно, для всех акций D_2 перестает расти при $n \sim 10-12$. Это значит, что:

- динамика акций была хаотической;
- подавляющую часть эволюции цен активов определяло действие десяти-двенадцати фундаментальных факторов.

Какую практическую роль может иметь это знание? Прежде всего, это дает ценную подсказку для построения экономических моделей. В них необходимо закладывать **обратные связи между переменными**. Это на порядок увеличивает сложность их конструирования и особенно настройки, но, по-видимому, является обязательным. Приведем ли построение подобных моделей к улучшению прогноза⁹ На первый взгляд — да. Ведь теоретически у нелинейных систем всегда существует т. н. горизонт прогноза, в пределах которого можно сравнительно точно — во всяком случае, лучше, чем $M(X_{t+1} | \Omega_t) = M(X_t)$ — предсказывать поведение процесса. Казалось бы, это открывает перспективы получения безрисковой доходности и обыгрывания рынка.

К сожалению (или, к счастью), это иллюзия. Даже 10 независимых переменных для нелинейной динамики — это очень много¹⁰. Горизонт прогноза в такой системе (как в любой хаотической), конечно, есть, но при таком числе степеней свободы он должен быть очень непродолжительным. Арбитражная доходность скорее всего не превзойдет и спреда. Поэтому экономическую прибыль получить не удастся. В этом смысле вероятностные прогнозы ценовой динамики остаются практически безальтернативными.

Библиография

1. Занг, В.-Б. Синергетическая экономика / Пер. с англ. — М.: Мир, 1999.
2. Канеман, Д., Тверски, А. Рациональный выбор, ценности и фреймы // Психологический журнал. — 2003. — Т. 24. — № 4. — С. 31-42.
3. Кроновер, Р. Фракталы и хаос в динамических системах / Пер. с англ. — М.: Постмаркет, 2000.
4. Мандельброт, Б., Хадсон, Р. (Не) послушные рынки / Пер. с англ. — М.: Вильямс, 2006.
5. Петерс, Э. Хаос и порядок на рынке капитала / Пер. с англ. — М.: Мир, 2000.
6. Петерс, Э. Фрактальный анализ на рынке капитала / Пер. с англ. — М.: Интернет-трейдинг, 2004.
7. Сорнетте, Д. Как предсказывать крахи финансовых рынков / Пер. с англ. — М.: Интернет-трейдинг, 2003.
8. Головки, В. Нейросетевые методы обработки хаотических процессов. — М.: МИФИ, 2005.
9. Ширяев, А. Основы стохастической финансовой математики. Т. 1, 2. — М.: Фазис, 1998.
10. Артюхин, И., Домбровский, Е. Оптимизация параметра функции полезности инвестиционного портфеля по статистическому ансамблю реализаций // Научная конференция по радиофизике, посвященная 100-летию М. Т. Греховой. — ННГУ, 2002.
11. Самойлов, Д. Факторы, оказывающие влияние на индекс РТС во время финансового кризиса 2008-2009 гг. и до него // Экономический журнал ВШЭ. — 2010. — № 2.
12. Смирнов, А. Макрофинансы I: методология моделирования пузырей и кризисов // Экономический журнал ВШЭ. — 2010. — № 3.
13. Стихова, О. Математическое моделирование статистик финансовых данных и их эмпирических распределений // Аудит и финансовый анализ. — 2008. — № 3.
14. Lempel, A., Ziv, J. On the complexity of finite sequences / IEEE Trans. Infor. Theory, IT-22. — 1976. — P. 75-81.

⁹ Сложность осуществления опустим.

¹⁰ Для сравнения: в ранних метеорологических моделях $D_2 = 3$; анализ ЭЭГ человека показывает $D_2 \approx 6-8$.