

РОССИЙСКАЯ ЭКОНОМИЧЕСКАЯ ШКОЛА

NEW ECONOMIC SCHOOL

**Шакин Д.А.**

**ВЫСОКОЧАСТОТНЫЕ  
ДАННЫЕ НА РОССИЙСКОМ  
РЫНКЕ ЦЕННЫХ БУМАГ**

---

Россия, 117418, Москва, Нахимовский проспект, 47

Suite 1721, Nakhimovsky Prospekt, 47, 117418, Moscow, Russia

Tel: (7)(095) 129-3911 or 129-3722 fax: (7)(095)129-3722 E-mail [nes@nes.ru](mailto:nes@nes.ru)

<http://www.nes.ru/>

Шакин Д.А.

Высокочастотные данные  
на российском рынке ценных бумаг

Препринт # BSP/2003/063 R

Эта работа была написана на основе магистерских тезисов в РЭШ в 2003 году в рамках исследовательского проекта "Динамика российских финансовых данных" под руководством проф. С.А.Анатольева (РЭШ).

Проект осуществлен при поддержке Фонда Форда, Всемирного Банка и Фонда Джона и Кэтрин МакАртуров.

Москва  
2003

**Шакин Д.А.** Высоочастотные данные на российском рынке ценных бумаг. / Препринт # BSP/2003/063 R. - М.: Российская Экономическая Школа, 2003. - 37с. (Рус.)

Данная работа посвящена изучению высоочастотных данных на российском рынке ценных бумаг. Было выяснено, что традиционные модели для анализа высоочастотных данных не очень хорошо описывают российские данные. В настоящей статье предлагаются две новые модели для изучения промежутков времени между сделками и волатильности доходностей акций. В работе эти модели применяются для исследования поведения семи наиболее часто торгуемых акций российских компаний. Используя log-ARMA модель, предложенную в данной статье, можно разделить движения в условном среднем и условной дисперсии процесса для промежутков времени между сделками; оказалось, что устойчивость в этом процессе очень высока. Высоочастотная мера волатильности, введенная в работе, позволяет учесть информацию о времени между сделками в процессе для доходностей акций посредством GARCH-подобной модели. Влияние дополнительных переменных (например, объемов сделок) на промежутки времени между сделками и на волатильность доходностей также может быть исследовано с помощью этих моделей; результаты оценивания показывают, что российский рынок ценных бумаг во многом сходен с западными. Кроме того, полученные оценки интерпретируются с точки зрения нескольких теоретических моделей микроструктуры рынка.

**Ключевые слова:** высоочастотные данные; волатильность цен; ARMA-GARCH модели; микроструктура рынка

**Shakin Dmitry.** High Frequency Data in the Russian Stock Market./ Working Paper # BSP/2003/063 R. – Moscow, New Economic School, 2003. –37p. (Rus.)

This work is devoted to the analysis of high frequency data in the Russian stock market. It turns out that traditional high frequency models do not fit Russian transaction data very well. In the paper new models both for durations between trade events and for return volatility are proposed and applied to the investigation of trading patterns for seven most frequently traded Russian stocks. Using log-ARMA model, introduced in this paper, movements in conditional mean and conditional variance of the duration process can be separated and it is found that persistence in variance is very strong. A high frequency measure of volatility proposed in the paper makes it possible to incorporate duration information into the process for returns through a GARCH-like model. The influence of additional variables (e.g. volumes) on durations and return volatility is analyzed and the results of estimation reveal some similarities between Russian and foreign stock markets. Besides that, the obtained estimates are interpreted from the viewpoint of several theoretical models of market microstructure.

**Keywords:** high-frequency transaction data; intertrade durations; price volatility; ARMA-GARCH models; market microstructure

ISBN

©Шакин Д.А., 2003 г.

© Российская экономическая школа, 2003 г.

# Содержание

1.	Введение .....	4
2.	Высокочастотные данные .....	6
2.1.	Электронная торговая система ММВБ .....	6
2.2.	Данные .....	6
2.3.	Междневная сезонность .....	8
2.4.	Внутредневная сезонность .....	9
3.	Теории рыночной микроструктуры .....	10
4.	Модели для промежутков времени между сделками .....	11
4.1.	ACD модели .....	11
4.2.	Логарифмическая ARMA модель .....	13
4.3.	Диагностика модели и проверка спецификации .....	15
5.	Высокочастотный анализ волатильности .....	16
5.1.	UHF-GARCH модель .....	16
5.2.	GARCH модель с временным агрегированием .....	17
5.3.	Самокорректирующаяся GARCH модель .....	18
6.	Эмпирические результаты .....	19
6.1.	Оценивание моделей для промежутков времени между сделками .....	20
6.2.	Оценивание моделей для волатильности .....	23
7.	Выводы .....	25
8.	Благодарности .....	25
9.	Литература .....	26
10.	Приложение А: Информация о ценных бумагах .....	28
10.1.	Торговые режимы .....	28
10.2.	Ценные бумаги .....	28
11.	Приложение Б: Таблицы и рисунки .....	30

# 1. Введение

Высокочастотные данные можно определить как данные, в которых для каждого наблюдения известно его точное время. Главной особенностью таких данных является то, что длина промежутка времени между двумя последовательными наблюдениями является случайной. С одной стороны, это дает исследователю дополнительную информацию об изучаемом процессе. С другой стороны, однако, такая нерегулярность данных затрудняет анализ, так как традиционные методы в этом случае неприменимы.

В последнее появилось довольно много новых моделей для финансовых данных, наблюдаемых в случайные моменты времени. Одна из самых часто используемых из них — это ACD (Autoregressive Conditional Durations) модель, предложенная Энглем и Расселлом (Engle and Russell, 1998), а также ее многочисленные модификации и обобщения. Кроме того, были выдвинуты и другие модели, например, SVD (Stochastic Volatility Duration) модель (Ghysels, Gouriéroux and Jasiak, 2003), которая сложнее с точки зрения оценивания, но позволяет более гибкую зависимость между условным средним и условной дисперсией соответствующего авторегрессионного процесса. Дальнейшее развитие этой области привело к появлению моделей, которые учитывают информацию о промежутках времени между сделками при моделировании процесса для доходностей: такая модель была предложена Энглем (Engle, 2000). В своей UHF-GARCH (Ultra-High-Frequency GARCH) модели он описывал динамику доходностей, деленных на квадратный корень из промежутков времени между сделками с помощью GARCH процесса. Более сложная модель, сочетающая в себе ACD модель и метод временного агрегирования для GARCH процессов, была выдвинута в работе (Ghysels and Jasiak, 1998). Кроме того, некоторые исследователи рассматривали совместную динамику промежутков времени между сделками и соответствующих им значений цен, объемов сделок и т.д. (см., например, Engle and Lunde, 1998, или McCulloch and Tsay, 2001)

Все упомянутые выше модели используют информацию о промежутках времени между сделками для исследования микроструктуры финансовых рынков. Некоторые теории микроструктуры рынков (см. Admati and Pfleiderer, 1988; Easley and O'Hara, 1992) позволяют сделать выводы о поведении участников торгов (например, информированных и неинформированных трейдеров), которые могут быть статистически проверены с использованием высокочастотных данных.

В нескольких недавних исследованиях предпринимались попытки анализа микроструктуры российских финансовых рынков. Медведев и Колодяжный (Medvedev and Kolodyazhny, 2001) исследовали микроэкономические аспекты поведения нерезидентов на рынке ГКО и нашли некоторые подтверждения гипотезы “herding behavior”. Они же (Medvedev and Kolodyazhny, 2002) рассматривали возможности для неинформированной торговли на российском рынке акций. Баринов, Первозванский и Первозванская (Barinov, Pervozvansky and Pervozvanskaya, 1999) изучали поведение трейдеров и тестировали выводы портфельной теории для российского рынка акций. Гольдберг и Тенорио (Goldberg and Tenorio, 1997) применили теорию аукционов для исследования валютного рынка в России. Стоит отметить однако, что ни в одном из упомянутых выше исследований для

анализа не использовались высокочастотные данные.

В данном исследовании мы изучаем микроструктуру российского рынка акций, используя высокочастотные данные. Для нашего анализа мы взяли данные по сделкам с наиболее часто торгуемыми акциями российских компаний в электронной системе Московской Межбанковской Валютной Биржи (ММВБ). Подробное описание этих данных приведено в параграфе 2 и Приложении А. На предварительном этапе исследования было обнаружено, что кроме внутрдневной сезонной структуры, часто обнаруживаемой в высокочастотных данных, присутствует также и междневная сезонная компонента (в частности, в промежутках времени между сделками и в доходностях). Так как мы хотим сосредоточиться в основном на стохастической динамике, то эти детерминистические сезонности были удалены из данных (см. подпараграфы 2.3 и 2.4).

В параграфе 4 предлагается новая логарифмическая ARMA (log-ARMA) модель для промежутков времени между сделками. Основная черта модели состоит в том, что она разделяет динамику условного среднего и условной дисперсии этого процесса, и поэтому является более гибкой, чем ACD модель. С другой стороны, эта модель гораздо проще для оценивания чем SVD, которая также разделяет динамику условного среднего и условной дисперсии в процессе для промежутков времени между сделками.

В параграфе 5 рассматриваются достоинства и недостатки существующих высокочастотных моделей для волатильности доходностей сделок. Главная проблема при анализе доходностей с использованием высокочастотных данные — это правильный выбор меры волатильности. Один подход к решению этого вопроса был предложен Энглем (Engle, 2000): в его UHF-GARCH модели объектом изучения является дисперсия за единицу времени, измеряемая квадратом доходности, деленным на соответствующий промежуток времени. В данной работе мы предлагаем новую самокорректирующуюся (SA-GARCH) модель, в которой высокочастотная мера волатильности определяется эндогенно посредством GARCH-подобного уравнения. В этой модели условная дисперсия инновации в процессе для доходностей может быть записана как сумма предыдущих условных дисперсий и квадратов предыдущих инноваций с коэффициентами, которые зависят от промежутков времени между соответствующими сделками. Непостоянные коэффициенты делают эту модель похожей на ACD-GARCH модель (Ghysels and Jasiak, 1998), однако, в SA-GARCH модели текущий промежуток времени между сделками непосредственно входит в уравнение для условной дисперсии, в то время как в ACD-GARCH модели — нет.

Наконец, в параграфе 6 новые log-ARMA и SA-GARCH модели применяются для исследования акций семи российских компаний. Включая в модели разнообразные переменные, связанные с изучаемым процессом (например, объемы сделок), мы протестировали несколько теоретических моделей микроструктуры рынка и обнаружили некоторые подтверждения для моделей Адмати и Пфляйдерера (Admati and Pfleiderer, 1988) и Эсли и О'Хара (Easley and O'Hara, 1992). Однако, выводы модели Даймонда и Верречия (Diamond and Verrecchia, 1987) не нашли подтверждения на российских данных.

## 2. Высокочастотные данные

Высокочастотные данные стали доступными с внедрением электронных торговых систем. В России в настоящее время несколько бирж используют такие системы. В данном исследовании мы используем данные по торгам в электронной системе Московской Межбанковской Валютной Биржи (ММВБ), которая является лидером по среднесуточному объему сделок с акциями российских компаний.

### 2.1. Электронная торговая система ММВБ

Электронная система ММВБ устроена как рынок конкурирующих между собой заявок (order-driven market). Схема этой системы такова: в каждый момент времени все заявки на покупку и продажу по каждой конкретной ценной бумаге хранятся в так называемой книге заявок (order book); заявки содержат желаемые цены и количества. Как только в системе появляется заявка на продажу (покупку) по цене не меньшей (не большей) чем указанная в некоторой заявке на покупку (продажу) для той же ценной бумаги, эта сделка автоматически исполняется и происходит обновление книги заявок в соответствии с исполненной сделкой<sup>1</sup>. Однако наряду с этим “нормальным” торговым режимом с непрерывным сопоставлением заявок, в системе ММВБ есть несколько специальных режимов торговли (по той же ценной бумаге), например, режим “неполные лоты” или “режим переговорных сделок”<sup>2</sup>. Несомненно, сделки, производимые в этих режимах, могут наблюдаться участниками торгов и их результаты принимаются ими во внимание. Однако данные по торгам в этих режимах весьма “шумные” (особенно для режима переговорных сделок, в котором поставка ценных бумаг не обязана производиться немедленно, а лишь через несколько дней или недель). Поэтому мы удалили все такие сделки из нашей выборки и рассматривали только “нормальный” торговый режим, длящийся с 10:30 до 18:45 каждый будний день.

### 2.2. Данные

Для анализа мы избрали период с 12 августа по 27 октября 2002 года, насчитывающий 50 торговых дней. Семь самых часто торгуемых акций были отобраны для исследования; мы рассматривали обыкновенные акции<sup>3</sup> Лукойла, Юкоса, СургутНефтеГаза, Ростелекома, Сбербанка, Мосэнерго и РАО ЕЭС. Доля суммы сделок по акциям этих компаний превосходит 95 процентов от общей суммы сделок с акциями российских компаний в системе ММВБ.

Для каждой сделки мы имеем следующие данные:

1. Время ( $T_t$ ), когда эта сделка была совершена (с точностью до 1 секунды);

---

<sup>1</sup>Подробное описание рынка конкурирующих между собой заявок можно найти в статье Тюрин (Tyurin, 2002).

<sup>2</sup>Описание этих режимов приведено в Приложении А.

<sup>3</sup>Было бы интересно исследовать совместную динамику всех типов ценных бумаг, выпущенных одной компанией, однако, в настоящей работе мы сосредоточились только на обыкновенных акциях. Более того, для каждой компании мы рассматривали только один выпуск акций. Подробное описание тех ценных бумаг, данные по которым мы использовали, приведено в Приложении А.

2. Цена ( $P_t$ ), по которой она была совершена;
3. Число купленных/проданных акций ( $V_t$ ).

Необходимо заметить, что наша база данных не содержит информацию по ценам на покупку и продажу, установленным маркет-мейкером<sup>4</sup>. Используя описанные выше данные, мы построили ряды, которые будут интересовать нас в первую очередь:

1. Промежуток времени между последовательными сделками:  $d_t = T_t - T_{t-1}$ ;
2. Логарифмическая доходность:  $r_t = \ln P_t - \ln P_{t-1}$ ;
3. Логарифм объема:  $v_t = \ln V_t$ .

При построении этих рядов особое внимание было уделено одновременным сделкам, которые могут возникать по двум причинам. Во-первых, это случается, когда у новой заявки возникает пересечение по цене сразу с несколькими стоящими в очереди заявками и она имеет достаточный объем для исполнения всех их. Естественно рассматривать такие одновременные сделки как одну. Отличительной чертой таких сделок является то, что они происходят по неубывающим или невозрастающим ценам. Во-вторых, одновременные сделки записываются, когда односекундной точности не хватает для различения двух последовательных сделок, произошедших в разные моменты времени. Так как база данных не содержит информации о причине одновременных сделок, то мы использовали следующую мнемоническую процедуру: если в некоторый момент времени  $t$  записано несколько сделок, то они объединяются начиная с первой пока их цены не убывают (не возрастают), причем агрегированной сделке приписывается средневзвешенная цена и суммарный объем. Если после этого в момент  $t$  остаются несколько сделок, то все они, кроме первой удаляются (удалить оставшиеся сделки можно по двум причинам: во-первых, число таких ситуаций для всех акций, кроме РАО ЕЭС достаточно мало; во-вторых, альтернативный подход — объединение всех сделок в момент  $t$  в одну — искажает как процесс для промежутков времени между сделками, так и процесс для доходностей, в то время как подход, используемый нами, искажает только первый из них).

Таблица 1. Описательные статистики для акций Лукойла

Переменная	Среднее	Стандартное отклонение	Скошенность	Куртосис	$Q(15)$	Число наблюдений
$d_t$	43.19	75.43	5.07	44.63	7455.9	34316
$r_t$	$1.99 \cdot 10^{-6}$	$8.99 \cdot 10^{-4}$	0.09	8.68	2537.2	34316
$v_t$	4.79	2.19	-0.29	2.62	1439.2	34316

<sup>4</sup>В такой автоматической системе маркет-мейкер обязан поддерживать заявки на покупку и продажу для обеспечения ликвидности акции. Размер этих заявок, а также величина спреда регулируется правилами ММВБ.



В Таблице 1 представлены описательные статистики для данных по сделкам с акциями Лукойла. Стоит обратить внимание на некоторые характерные черты высокочастотных данных. Во-первых, ряды доходностей и промежутков времени между сделками сильно автокоррелированы, как показывает статистика Льюнга-Бокса (столбец  $Q(15)$  в Таблице 1) с 15 лагами (однопроцентное критическое значение равно 30.58). Другая характерная особенность — избыточная дисперсия в процессе для промежутков времени между сделками. Также стоит отметить большое значение кurtосиса для доходностей (8.69), что гораздо выше кurtосиса для нормального распределения.

В высокочастотных данных (в особенности, в доходностях) обычно отмечают дискретность, например, Расселл и Энгль (Russell and Engle, 1999) предложили специальную модель для моделирования этой черты доходностей. Однако в наших данных дискретность присутствует в значительно меньшей степени, чем в данных, которые использовали Расселл и Энгль. Это очень хорошо видно на Рисунке 1, на котором изображена гистограмма для ненулевых изменений цены акций Лукойла (примерно для одной трети сделок цена не изменялась).

### 2.3. Междневная сезонность

При рассмотрении свойств высокочастотных данных в предыдущем подпараграфе мы объединили сделки за разные дни в один длинный ряд. Однако, можно предположить, что активность на рынке (измеряемая средней длиной промежутка времени между сделками) и волатильность доходностей может варьироваться от одного дня к другому, причем причины этих изменений не связаны с эффектами асимметричной информации, которые мы хотим изучить. Наличие междневных вариаций можно без труда обнаружить на Рисунках 2 и 3 (см. Приложение Б), на которых показаны средние значения промежутков времени между сделками и волатильности доходностей для разных дней. Для объема сделок (переменная  $v_t$ ) этот эффект выражен не так ярко. Можно предложить несколько объяснений для существования таких междневных колебаний. Во-первых, возможно присутствие в данных детерминистической структуры, например, промежутки времени между сделками меньше во вторник, среду и четверг чем в понедельник и пятницу (в среднем). Во-вторых, цены акций российских компаний сильно подвержены влиянию внешних факторов, например, цены акций российских экспортеров нефти (Лукойл, Юкос, СургутНефтеГаз) тесно связаны с мировыми ценами на нефть, что может приводить к сильным вариациям волатильности их доходностей. Наконец, мы можем предполагать, что число участников торгов каждый день случайно, и причины этой случайности не связаны с асимметричной информацией.

Междневная сезонность должна быть принята во внимание, потому что в противном случае коэффициенты оцениваемых моделей могут быть неустойчивыми. Одно из возможных решений — оценивать модели отдельно для каждого дня. При таком подходе связь между отдельными днями теряется полностью. Более тонкий подход использовался в работе (McCulloch and Tsay, 2001), где было предложено рассматривать параметры процесса как случайные величины, принимающие новые значения каждый день. В этой модели связь между различными днями возникает из-за того, что

случайные параметры для разных дней выбираются из одного и того же распределения. Однако, такой вид зависимости все же довольно слаб; весьма вероятно, что параметры для разных дней связаны гораздо более тесно. Поэтому мы решили использовать следующую процедуру для устранения междневных вариаций из данных: были построены ряды:

$$\begin{aligned}\tilde{d}_t &= \frac{d_t}{\bar{d}_t}, \\ \tilde{r}_t &= \frac{r_t}{\sqrt{\bar{r}_t^2}}, \\ \tilde{v}_t &= v_t - \bar{v}_t,\end{aligned}$$

где  $\bar{d}_t$  — средний промежуток времени между сделками за тот день, которому принадлежит наблюдение  $t$ ,  $\bar{r}_t^2$  — средний квадрат доходностей за соответствующий день, и  $\bar{v}_t$  — средний логарифм объема за этот день. При этом преобразовании исходных рядов мы использовали мультипликативную форму корректировки для промежутков времени между сделками и для квадратов доходностей, а для логарифма объема — аддитивную (можно было бы использовать мультипликативную форму корректировки для объема, а потом взять логарифм — результаты были бы схожими).

#### 2.4. Внутредневная сезонность

Кроме междневных колебаний, описанных в предыдущем подпараграфе, в данных (промежутках времени между сделками и волатильности доходностей) также присутствует и специфическая внутредневная сезонная структура. Так как природа этой структуры детерминистическая, мы тоже хотим удалить ее из данных. Так как трудно подобрать гибкую параметризацию для моделирования этого вида сезонности, то на этой стадии обычно применяются непараметрические методы. В работах (Engle and Russell, 1998, Engle, 2000, Grammig and Wellner, 2002) для этой цели использовался метод сплайнов; а в работе (Zhang, Russell and Tsay, 2001) внутредневная структура оценивалась с помощью метода локальной линейной регрессии. Мы использовали похожий подход: следующие уравнения

$$\begin{aligned}E[\tilde{d}_t | T_{t-1}] &= f_1(T_{t-1}), \\ E[\tilde{r}_t^2 | T_{t-1}] &= f_2(T_{t-1}), \\ E[\tilde{v}_t | T_{t-1}] &= f_3(T_{t-1}),\end{aligned}$$

были оценены с помощью локальной линейной регрессии с гауссовским ядром и вручную выбранной шириной окна. Однако, так как логично предположить, что соответствующая сезонная структура может быть различной для разных дней недели, то для каждого дня недели использовалась своя функция  $f$ . Оцененная сезонная структура для промежутков времени между сделками и для квадратов доходностей показана на Рисунках 4 и 5 (см. Приложение Б). Полученные оценки весьма схожи с теми, которые были получены в предыдущих исследованиях для западных финансовых

данных: промежутки времени между сделками длиннее в середине дня и короче утром и вечером, волатильность доходностей утром выше чем на протяжении остального дня. Интересно отметить, что в данных не наблюдается увеличения волатильности доходностей к окончанию торговой сессии (18:45), что можно объяснить тем фактом, что после 18:45 торговля еще продолжается в других режимах<sup>5</sup>. Полученные оценки внутредневной сезонности были использованы для построения окончательных версий рядов:

$$\hat{d}_t = \frac{\tilde{d}_t}{\hat{f}_1(T_{t-1})},$$

$$\hat{r}_t = \frac{\tilde{r}_t}{\sqrt{\hat{f}_2(T_{t-1})}},$$

$$\hat{v}_t = \tilde{v}_t - \hat{f}_3(T_{t-1}).$$

Здесь, по тем же причинам, что и в предыдущем подпараграфе для промежутков времени между сделками и для квадратов доходностей использовалась мультипликативная форма корректировки (сходная процедура использовалась и в работе (Engle, 2000)), а для логарифма объема — аддитивная.

Таблица 2. Описательные статистики для акций Лукойла после удаления сезонности

Переменная	Среднее	Стандартное отклонение	Скошенность	Куртосис	Q(15)	Число наблюдений
$\hat{d}_t$	1.00	1.56	4.04	28.51	3397.6	34316
$\hat{r}_t$	0.00	0.99	0.06	6.10	2576.5	34316
$\hat{v}_t$	0.00	2.18	-0.28	2.60	918.46	34316

Описательные статистики для откорректированных данных приведены в Таблице 2. Снова все характерные особенности высокочастотных данных, такие как избыточная дисперсия для промежутков времени между сделками, избыточный куртосис в доходностях и сильная автокорреляция во всех рядах, легко заметны. Однако, соответствующие значения стали несколько меньше по сравнению с первоначальными данными (см. Таблицу 1).

В дальнейшем, чтобы не загромождать обозначения, мы будем писать  $d_t$ ,  $r_t$  и  $v_t$  вместо  $\hat{d}_t$ ,  $\hat{r}_t$  и  $\hat{v}_t$  соответственно.

### 3. Теории рыночной микроструктуры

В данном параграфе мы кратко обсудим те теоретические модели микроструктуры рынка, которые мы хотим протестировать, используя наши данные. Более подробный вариант этой дискуссии можно найти в работе (Engle, 2000).

<sup>5</sup>См. Приложение А.

Мы будем рассматривать три модели, в которых поведение трейдеров описывается на основе асимметричной информации. Во всех моделях предполагается наличие агентов двух типов: информированных и неинформированных. Первая модель — это модель Даймонда и Верречия (Diamond and Verrecchia, 1987). В этой модели рассматриваются ограничения на занятие коротких позиций по акциям, в силу которых те информированные трейдеры, которые получают плохие новости, не могут использовать эту информацию. Таким образом, эта модель предсказывает, что если наблюдаются длинные промежутки времени между сделками, то это может означать, что пришли плохие новости, следовательно, доходности должны падать и волатильность должна расти. В модели Эсли и О’Хара (Easley and O’Hara, 1992), напротив, предполагается отсутствие ограничений на занятие коротких позиций, поэтому в ней длинные промежутки времени между сделками интерпретируются как отсутствие новостей, и, следовательно, волатильность доходностей в таких ситуациях должна быть меньше. Последняя модель, которую мы хотим рассмотреть — это модель Адмати и Пфляйдера (Admati and Pfleiderer, 1988). В ней короткие промежутки времени между сделками объясняются как результат совместной торговли тех трейдеров (так называемые “ликвидные” трейдеры), которые торгуют по причинам, несвязанным напрямую с прибылью (ликвидность, диверсификация портфелей, и т.п.). Таким образом, короткие промежутки времени, сопровождаемые большими объемами, должны восприниматься как несущие дополнительной информации, и, следовательно, волатильность должна в таких ситуациях убывать. Кроме того, так как ликвидные трейдеры стремятся торговать в одно время, то промежутки времени между сделками должны кластеризоваться, и эта кластеризация должна возрастать при больших объемах сделок.

## 4. Модели для промежутков времени между сделками

### 4.1. ACD модели

Класс ACD (Autoregressive Conditional Duration) моделей был предложен Энглем и Расселлом (Engle and Russell, 1998), они использовали GARCH-подобную модель для моделирования высокой устойчивости в данных по промежуткам времени между последовательными сделками.

Пусть  $d_t$  — промежуток времени между сделками<sup>6</sup>, произошедшими в моменты  $T_t$  и  $T_{t-1}$ , то есть,  $d_t = T_t - T_{t-1}$ . В ACD модели величина  $d_t$  разлагается в следующее произведение:

$$d_t = E[d_t | I_{t-1}] \varepsilon_t = \psi_t \varepsilon_t, \quad (1)$$

где  $I_{t-1}$  — множество, содержащее предыдущую информацию, которое включает в себя значения предыдущих промежутков времени между сделками, и, возможно, какие-то другие переменные, а  $\varepsilon_t$  — инновация, независимая от переменных, входящих в  $I_{t-1}$ . Поэтому,  $E[\varepsilon_t | I_{t-1}] = 1$ . Кроме того, так как  $d_t > 0$ , то мы получаем, что  $\varepsilon_t > 0$  почти наверное. Энгль и Расселл использовали следующую

---

<sup>6</sup>Предполагается, что  $d_t > 0$ .

функциональную форму для условного математического ожидания  $\psi_t$ :

$$\psi_t = \alpha + \sum_{i=1}^p \beta_i d_{t-i} + \sum_{i=1}^q \gamma_i \psi_{t-i}. \quad (2)$$

Такая спецификация называется ACD(p,q). Для завершения этой параметрической модели требуется специфицировать распределение для инновации  $\varepsilon$ . Так как инновация неотрицательна (с вероятностью единица), то естественно рассмотреть в качестве возможных вариантов экспоненциальное, вейбулловское, а также обобщенное гамма распределения<sup>7</sup>. После выбора распределения, модель может быть оценена с помощью метода максимального правдоподобия.

Однако, часто сообщается, что перечисленные выше распределения не всегда хорошо подходят к данным. В такой ситуации есть возможность рассматривать ACD модель как полупараметрическую, то есть, не специфицировать распределение для инноваций  $\varepsilon$ . Используя метод максимального квази-правдоподобия (состоящий в данном случае в максимизации функции правдоподобия, соответствующая экспоненциальному распределению для инноваций) можно получить оценки, которые являются состоятельными для псевдо-истинных значений параметров (при условии правильной спецификации  $\psi$ ). Однако, при этом теряется эффективность, так как нужно использовать стандартные ошибки в робастной форме (см. Bollerslev and Wooldridge, 1992).

ACD модель позволяет исследовать влияние других переменных на процесс для промежутков времени между сделками с помощью включения их в уравнение (2). Кроме того, эта модель дает возможность получить оценку для функции риска (hazard function). Однако главное достоинство ACD модели заключается в том, что с помощью нее можно смоделировать высокую устойчивость, наблюдаемую в процессе для промежутков времени между сделками, используя лишь небольшое число параметров. Более того, модель может быть расширена разными способами по аналогии с GARCH моделями. Например, часто оказывается, что сумма оценок авторегрессионных коэффициентов в уравнении (2) очень близка к единице, что дает возможность подозревать нестационарность процесса. Эта проблема может быть решена с помощью ACD модели с дробной интеграцией (fractionally integrated ACD, FIACD), предложенной в работе (Jasiak, 1999). Другое возможное решение — ACD модель с порогами (threshold ACD, TACD), использовавшаяся в работе (Zhang, Russell and Tsay, 2001).

Другое направление для модификации ACD модели — поиск более подходящей параметризации для функции  $\psi$ . Например, в логарифмической ACD модели (Bauwens and Giot, 2000) эта функция представлена в следующем виде:

$$\psi_t = \exp \left( \alpha + \sum_{i=1}^p \beta_i \ln d_{t-i} + \sum_{i=1}^q \gamma_i \ln \psi_{t-i} \right), \quad (3)$$

Эта спецификация удобна тем, что условное среднее  $\psi$  всегда неотрицательно.

---

<sup>7</sup>В литературе встречаются и некоторые другие варианты, например, распределение Бурра было использовано в работе (Grammig and Maurer, 2000).

Однако, несмотря на гибкость ACD модели при моделировании высокой устойчивости в процессе для промежутков времени между последовательными сделками, основным недостатком<sup>8</sup> этого класса моделей является то, что условное среднее и условная дисперсия очень жестко связаны соотношением

$$V[d_t|I_{t-1}] = \psi_t^2 V[\varepsilon_t]. \quad (4)$$

Здесь  $V[\varepsilon_t]$  — некоторая константа, определяемая выбором распределения для инноваций  $\varepsilon$ . Следовательно, можно предположить, что авторегрессионные коэффициенты в уравнении (2) отражают не только устойчивость в условном среднем процесса, но и в его условной дисперсии. Для того, чтобы разделить эти два источника устойчивости, авторы работы (Ghysels, Gouriéroux and Jasiak, 2003) предложили модель SVD (Stochastic Volatility Duration). Результаты их исследования говорят о том, что эта модель успешно моделирует оба источника устойчивости. Стоит заметить однако, что SVD модель достаточно сложна для оценивания: в работе (Ghysels, Gouriéroux and Jasiak, 2003) предлагаются два метода для ее оценивания, оба с использованием симуляций. Из-за сложности этих методов, модель должна быть очень компактной, вследствие чего влияние некоторых переменных весьма непросто учесть в рамках этой спецификации.

В следующем подпараграфе будет предложена модель, старающаяся сочетать простоту ACD модели и гибкость SVD модели через ARMA процесс для логарифмов промежутков времени между сделками.

## 4.2. Логарифмическая ARMA модель

Пусть  $d_t$  — промежуток времени между сделками. Мы предположим, что справедливо следующее разложение:

$$\ln d_t = \alpha + \sum_{i=1}^p \beta_i \ln d_{t-i} + \sum_{i=1}^q \gamma_i \varepsilon_{t-i} + \varepsilon_t, \quad (5)$$

где  $\varepsilon_t = \sigma_t \eta_t$  и

$$\begin{aligned} \eta_t &\sim iid(0,1), \\ \sigma_t^2 &= \mu + \sum_{i=1}^r \lambda_i \varepsilon_{t-i}^2 + \sum_{i=1}^s \nu_i \sigma_{t-i}^2. \end{aligned}$$

Здесь  $\alpha, \beta_i, \gamma_i, \mu, \lambda_i$  и  $\nu_i$  — неизвестные параметры. То есть, в этой модели  $\ln d_t$  следует ARMA( $p, q$ ) процессу с ошибками типа GARCH( $r, s$ ). Аналогично ACD модели, имеются две возможности для завершения описания: либо специфицировать распределение для  $\eta$  (подходящими вариантами здесь являются нормальное и, быть может,  $t$ -распределение) и использовать метод максимального правдоподобия для оценивания, либо рассматривать модель как полупараметрическую, и,

<sup>8</sup>См. обсуждение этого вопроса в книге (Gouriéroux and Jasiak, 2000, стр. 462).

следовательно, использовать метод максимального квази-правдоподобия (с нормальной функцией плотности) для ее оценивания.

Для того, чтобы увидеть как log-ARMA модель расширяет логарифмическую ACD модель, мы рассмотрим случай логарифмической ACD(1,1) модели. Используя уравнения (1) и (3) мы получим:

$$\begin{aligned}\ln d_t &= \alpha + \beta \ln d_{t-1} + \gamma \ln \psi_{t-1} + \ln \varepsilon_t \\ &= \alpha + (\beta + \gamma) \ln d_{t-1} - \gamma \ln \varepsilon_{t-1} + \ln \varepsilon_t \\ &= (\alpha + \gamma m - m) + (\beta + \gamma) \ln d_{t-1} - \gamma \zeta_{t-1} + \zeta_t,\end{aligned}$$

где  $m = E(\ln \varepsilon_t)$  и  $\zeta_t = \ln \varepsilon_t - m$ , то есть,  $\ln d_t$  следует ARMA(1,1) процессу. Если специфицировать лог-нормальное распределение<sup>9</sup> для  $\varepsilon_t$  с параметрами<sup>10</sup>  $(-\rho^2/2, \rho^2)$ , то мы получим, что  $\zeta_t$  имеет нормальное распределение с параметрами  $(0, \rho^2)$ , то есть, случай логарифмической ACD(1,1) с лог-нормальным распределением для инноваций сводится к гомоскедастичной log-ARMA модель с нормальным распределением.

Простой пример, рассмотренный выше, показывает сходства логарифмической ACD и log-ARMA моделей, поэтому мы можем ожидать, что log-ARMA модель способна моделировать высокую устойчивость в условном среднем процесса для промежутков времени между сделками. Однако, бóльшая гибкость log-ARMA модели, обусловленная наличием дополнительного уравнения для условной дисперсии, позволяет разделить два источника устойчивости, следовательно, мы можем рассматривать эту модель как простую (с точки зрения методов оценивания) альтернативу SVD модели.

Еще одной интересной чертой log-ARMA модели является функция риска, которую можно получить при параметрической спецификации. Чтобы проанализировать этот вопрос, рассмотрим случай гомоскедастичной log-ARMA модель с нормальным распределением. В этой ситуации условное распределение промежутков времени между сделками является лог-нормальным, и, следовательно, функция риска является немонотонной (см. Рисунок 6 в Приложении Б). Это дает log-ARMA модели преимущество над ACD моделями с экспоненциальным, вейбулловским и обобщенным гамма распределением для шоков, так как зачастую оказывается, что функция риска для промежутков времени между сделками немонотонна. Более того, в гетероскедастичной log-ARMA модели функция риска еще и меняется во времени.

Зависимость условной дисперсии от времени может помочь также и при моделировании избыточной дисперсии в высокочастотных данных. В книге (Gourieroux and Jasiak, 2000, стр. 462) утверждается, что избыточная дисперсия меняется во времени, в то время как в ACD модели она

<sup>9</sup>Такое распределение может рассматриваться как предельный случай обобщенного гамма распределения.

<sup>10</sup>Напомним, что если случайная величина  $x$  имеет лог-нормальное распределение с параметрами  $(\mu, \sigma^2)$ , то  $\ln x$  имеет нормальное распределение с параметрами  $(\mu, \sigma^2)$ ; кроме того,  $E x = \exp(\mu + \sigma^2/2)$ . Следовательно, так как  $E \varepsilon = 1$ , то  $\mu$  и  $\sigma$  должны быть связаны соотношением  $\mu = -\sigma^2/2$ .

постоянна (см. уравнение (4)). В log-ARMA модели, напротив, избыточная дисперсия меняется в зависимости от предыдущих значений переменных, и, более того, авторегрессионные коэффициенты в уравнении для дисперсии показывают степень кластеризации наблюдений с избыточной и недостаточной дисперсией.

### 4.3. Диагностика модели и проверка спецификации

Как ACD так и log-ARMA модели могут быть протестированы на правильности спецификации с помощью стандартизованных остатков. Для ACD модели из уравнения (1) следует, что  $\varepsilon_t = d_t / \psi_t$  есть последовательность независимых одинаково распределенных случайных величин. Следовательно, диагностика модели может быть проведена на стандартизованных остатках, определенных равенством

$$e_t = \frac{d_t}{\hat{\psi}_t},$$

где  $\hat{\psi}_t$  вычисляется с использованием полученных оценок неизвестных параметров. Здесь можно воспользоваться, например, тестом Льюнга-Бокса на отсутствие автокорреляции (для остатков и их квадратов). Кроме того, параметрическую спецификацию можно тестировать на соответствие остатков конкретному распределению, например, для тестирования гипотезы об экспоненциальном распределении возможно использовать тест на избыточную дисперсию, предложенный Энгле и Расселлом (Engle and Russell, 1998).

Аналогичный набор тестов применим и к log-ARMA модели. Здесь стандартизованные остатки определяются как

$$e_t = \frac{w_t}{\hat{\sigma}_t},$$

где  $w_t$  есть остатки ARMA модели:

$$w_t = \ln d_t - \left( \hat{\alpha} + \sum_{i=1}^p \hat{\beta}_i \ln d_{t-i} + \sum_{i=1}^q \hat{\gamma}_i w_{t-i} \right)$$

и  $\hat{\sigma}_t^2$  — оцененная условная дисперсия. Так как последовательность  $\eta_t$ , по предположению, состоит из независимых одинаково распределенных случайных величин, то можно применить тест Льюнга-Бокса для рядов  $e_t$  и  $e_t^2$ . Кроме того, log-ARMA спецификацию с нормальными ошибками можно проверить с помощью тестов на нормальность для  $e_t$  (тесты Колмогорова-Смирнова, Жарке-Бера и др.).

Еще одним важным вопросом при тестировании адекватности модели является стабильность коэффициентов во времени. Часто оказывается, что параметры в высокочастотных моделях очень неустойчивы, например, в работе (Zhang, Russell and Tsay, 2001) с помощью тестов SupLM, AveLM и ExpLM (Andrews, 1993; Andrews and Ploberger, 1994) было найдено шесть структурных сдвигов (при



периоде наблюдений — три месяца). Для проверки стабильности параметров во времени мы используем тест Ниблома (Nyblom, 1989) для индивидуальных коэффициентов и для модели в целом.

## 5. Высокочастотный анализ волатильности

При анализе волатильности доходностей акций с помощью высокочастотных данных есть два очень важных момента. Во-первых — нерегулярность процесса, которая, в частности, не позволяет использовать обычную меру волатильности (средний квадрат доходности); частота торговли должна быть обязательно учтена при анализе. Второй важный вопрос — это направление причинно-следственной связи между различными переменными, то есть, между ценами, объемами и промежутками времени между сделками. Любая модель, предназначенная для высокочастотного исследования волатильности доходностей, должна как-либо решить эти вопросы.

### 5.1. UHF-GARCH модель

В работе (Engle, 2000) была предложена модель для анализа волатильности доходностей с помощью высокочастотных данных; в ней объектом изучения является волатильность за соответствующий промежуток времени:

$$\sigma_t^2 = \frac{V[r_t | I_{t-1}, d_t]}{d_t} = V \left[ \frac{r_t}{\sqrt{d_t}} \middle| I_{t-1}, d_t \right].$$

Кроме этого, в модели предполагается, что цены (и, следовательно, доходности) не являются причиной (в смысле Гранжера) для промежутков времени между сделками<sup>11</sup>. Опираясь на это предположение, можно сначала оценить модель для промежутков времени между сделками, а затем — модель для доходностей, используя функцию правдоподобия, условную на предыдущих значениях переменных, а также на текущее значение промежутка времени между сделками. Получившуюся модель можно записать следующим образом:

$$\begin{aligned} \frac{r_t}{\sqrt{d_t}} &= \rho \frac{r_{t-1}}{\sqrt{d_{t-1}}} + \varepsilon_t + \phi \varepsilon_{t-1}, \\ \varepsilon_t &= \sigma_t \eta_t, \\ \eta_t &\sim iid(0,1), \\ \sigma_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2. \end{aligned}$$

Эта простая ARMA(1,1)-GARCH(1,1) спецификация называется UHF-GARCH (ultra-high-frequency GARCH) из-за весов  $\sqrt{d}$ . Кроме этого, модель может быть расширена включением дополнительных переменных, принадлежащих информационному множеству  $I_{t-1} \cup \{d_t\}$ , как в уравнение для среднего, так и в уравнение для дисперсии. Например, следующая спецификация, в которой текущее значение промежутка времени между сделками является более информативным, была рассмотрена в

<sup>11</sup>См. Engle, 2000, стр. 11.

(Engle, 2000):

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma \frac{1}{d_t}. \quad (6)$$

Однако, стоит отметить три достаточно важных момента. Во-первых, хотя измерение волатильности за период времени кажется естественным, выбор весов в UHF-GARCH все-таки является достаточно произвольным. Во-вторых, даже если такое взвешивание является корректным, то интерпретация самих доходностей, деленных на корень из  $d_t$ , весьма затруднительна; не совсем понятно, почему выводы моделей микроструктуры рынка должны применяться именно к этим взвешенным доходностям (как это было сделано в работе (Engle, 2000)). Наконец, выбор функциональной формы, посредством которой текущий промежуток времени оказывает влияние на волатильность (уравнение б), кажется немного неестественным.

## 5.2. GARCH модель с временным агрегированием

Альтернативная модель для моделирования волатильности доходностей, которая опирается на метод временного агрегирования для GARCH процессов (Drost and Nijman, 1993; Drost and Werker, 1996), была предложена в работе (Ghysels and Jasiak, 1998). А именно, рассматривается следующая система уравнений для промежутков времени между сделками и для доходностей:

$$\begin{aligned} d_t &= \psi_t \varepsilon_t, \\ \psi_t &= \alpha + \beta d_{t-1} + \gamma \psi_{t-1}, \\ \sigma_t^2 &= \omega_t + \alpha_t r_{t-1}^2 + \beta_t \sigma_{t-1}^2, \end{aligned}$$

где переменные коэффициенты  $\omega_t, \alpha_t$  и  $\beta_t$  являются (известными) функциями<sup>12</sup> некоторых параметров и переменных, принадлежащих информационному множеству  $I_{t-1}$ . Эта модель называется ACD-GARCH. Первые два уравнения системы — это ACD(1,1) модель для промежутков времени между сделками. Третье уравнение специфицирует процесс для условной дисперсии доходностей (предполагается, что условное среднее доходностей равно нулю). Одним из основных достоинств этой модели является то, что она не опирается на конкретный выбор весов для доходностей, и, следовательно, является более предпочтительной по сравнению с UHF-GARCH моделью с этой точки зрения. Однако, такая гибкая спецификация, с меняющимися во времени коэффициентами требует достаточно сложных методов оценивания<sup>13</sup>. Еще одним важным моментом является то, что текущее значение промежутка времени между сделками не оказывает влияния на условную дисперсию доходностей, однако, этот эффект является весьма значительным (см. работу (Meddahi, Renault and Werker, 2003)).

---

<sup>12</sup>Точные выражения для этих функций приведены в (Ghysels and Jasiak, 1998).

<sup>13</sup>Например, в работе (Ghysels and Jasiak, 1998) сообщается, что стандартные ошибки некоторых коэффициентов не были вычислены по причине сложности вычислений.

### 5.3. Самокорректирующаяся GARCH модель

В этом параграфе предлагается новая модель, которая исправляет недостатки UHF-GARCH и ACD-GARCH моделей. Мы начнем с разложения условной функции правдоподобия:

$$l(d_t, r_t, v_t | I_{t-1}, \theta) = l_d(d_t | I_{t-1}, \theta_d) l_r(r_t | I_{t-1}, d_t, \theta_r) l_v(v_t | I_{t-1}, d_t, r_t, \theta_v).$$

Наше ключевое предположение состоит в том, что векторы параметров  $\theta_d$ ,  $\theta_r$  и  $\theta_v$  не пересекаются. Следовательно, сначала можно оценить модель для промежутков времени между сделками используя только функцию  $l_d$ , затем оценить модель для доходностей и наконец — для объемов. Аналогичный подход использовался в работе (Engle, 2000). Однако, необходимо оговориться, что это лишь один из возможных способов разложения функции правдоподобия и наложения причинно-следственных ограничений. В книге (Gourieroux and Jasiak, 2001, стр. 404) приводятся аргументы в пользу других ограничений на причинную связь между переменными, например, тех, которые были рассмотрены в статье (Darolles, Gourieroux and Le Fol, 2000). Так как из трех переменных нас более всего интересуют доходности и промежутки времени между сделками, то будем рассматривать только первые две модели. Для промежутков времени между сделками можно использовать ACD или log-ARMA модели из предыдущего параграфа. Для доходностей мы предположим, что:

$$r_t = \alpha + \sum_{i=1}^p \beta_i r_{t-i} + \sum_{i=1}^q \gamma_i \varepsilon_{t-i} + \varepsilon_t$$

где для  $\varepsilon_t$  выполняется следующее:

$$\begin{aligned} \varepsilon_t &= \sigma_t \eta_t, \\ \eta_t &\sim iid(0,1), \\ \frac{\sigma_t^2}{d_t^\rho} &= \mu + \sum_{i=1}^r \lambda_i \frac{\varepsilon_{t-i}^2}{d_{t-i}^\rho} + \sum_{i=1}^s \nu_i \frac{\sigma_{t-i}^2}{d_{t-i}^\rho}. \end{aligned}$$

Здесь  $\alpha, \beta_i, \gamma_i, \mu, \lambda_i, \nu_i$  и  $\rho$  — неизвестные параметры. Кроме того, можно специфицировать конкретное семейство распределений для  $\eta$ , однако, предпочтительнее не делать этого и использовать метод максимального квази-правдоподобия для оценивания<sup>14</sup>. Мы будем называть описанную выше спецификацию самокорректирующейся GARCH (self-adjusting GARCH, SA-GARCH) моделью. Конечно, базовые уравнения могут быть расширены с помощью включения в них дополнительных переменных, например, лагированных значений объема и текущих значений промежутков времени между сделками; это позволяет тестировать выводы теорий микроструктуры рынков (см. параграф 3). Диагностика модели может быть проведена с помощью теста Льюнга-Бокса для стандартизованных остатков и их квадратов. Тест Ниблома может быть использован для

<sup>14</sup>Это действительно предпочтительнее, так как распределение доходностей имеет атом в нуле (см. подпараграф 2.2), следовательно, предположения о нормальном распределении или  $t$ -распределении для шоков выглядит нереалистичным. С другой стороны, распределение доходностей достаточно симметрично, и поэтому можно предположить, что метод максимального квази-правдоподобия будет работать достаточно хорошо.

тестирования стабильности коэффициентов во времени.

В SA-GARCH модели предполагается, что некоторая высокочастотная мера волатильности,  $S_t^2 = \sigma_t^2 / d_t^\rho$ , определяется эндогенно посредством параметра “самокоррекции”  $\rho$ . Мы будем говорить, что  $S_t^2$  есть *внутренняя волатильность* акции. Таким образом, параметр  $\rho$  показывает подходящие веса для обычной меры волатильности. При ограничении  $\rho = 1$  SA-GARCH модель становится похожей (хотя и не полностью) на UHF-GARCH. Мы можем вывести выражение для “естественной” меры волатильности, предложенной Энглем (для простоты рассмотрим SA-GARCH(1,1)):

$$\frac{\sigma_t^2}{d_t} = d_t^{\rho-1} \left[ \mu + \lambda d_{t-1}^{1-\rho} \frac{\varepsilon_{t-1}^2}{d_{t-1}} + \nu d_{t-1}^{1-\rho} \frac{\sigma_{t-1}^2}{d_{t-1}} \right]. \quad (7)$$

Если  $\rho$  равно единице, как в UHF-GARCH модели, то естественная мера волатильности постоянна (при заданных предыдущих значения переменных), или, другими словами, условная дисперсия шока  $\varepsilon_t$  прямо пропорциональна  $d_t$ . Другая спецификация, в которой дисперсия за единицу времени сходится в ненулевому пределу при  $d_t$  стремящемся к бесконечности, предложена в работе (Meddahi, Renault and Werker, 2003). Однако, такое поведение условной дисперсии неочевидно: чем длиннее период  $d_t$ , тем больше времени проходит без новой информации об акции, поэтому естественно считать, что дисперсия за единицу времени должна стремиться к нулю, иначе говоря,  $\sigma_t^2$  является вогнутой функцией  $d_t$ , что соответствует случаю  $0 < \rho < 1$ . Частичную поддержку этому можно найти в работе (Engle, 2000): там сообщается, что коэффициент при  $1/d_t$  очень значим в уравнении для условной дисперсии UHF-GARCH модели (уравнение (6)).

Вернемся теперь опять к уравнению для условной дисперсии в SA-GARCH модели, мы можем переписать его в следующем виде:

$$\sigma_t^2 = d_t^\rho \left[ \mu + \lambda d_{t-1}^{-\rho} \varepsilon_{t-1}^2 + \nu d_{t-1}^{-\rho} \sigma_{t-1}^2 \right]$$

Из этого уравнения вытекает, что условная дисперсия шока  $\varepsilon_t$  выражается через предыдущие значения квадратов шоков и их дисперсий посредством авторегрессионного процесса с коэффициентами, которые зависят от промежутков времени между сделками. Таким образом, мы можем относиться к SA-GARCH модели как к модификации UHF-GARCH в направлении ACD-GARCH. Стоит отметить, однако, что в SA-GARCH модели текущее значение промежутка времени между сделками непосредственно входит в уравнение для дисперсии, и, следовательно, SA-GARCH исправляет этот недостаток ACD-GARCH модели.

## 6. Эмпирические результаты

В этом параграфе представлены результаты применения описанных выше моделей к российским

высокочастотным данным. При оценивании всех моделей использовались ряды  $d_t$ ,  $r_t$  и  $v_t$  с устраненной меж- и внутредневной сезонностями. Для оценивания моделей использовался пакет GAUSS, версия 4.0.23, и библиотека MAXLIK, версия 4.0.24<sup>15</sup>. Кроме этого, так как маловероятно, что последняя сделка предыдущего дня как-либо связана с первой сделкой следующего, то все авторегрессионные процессы реинициализировались в начале каждого дня (походный подход использовался, например, в работах (Engle and Russell, 1998) и (Zhang, Russell and Tsay, 2001)).

Везде далее, все результаты приводятся для акций Лукойла (если явно не указано обратное).

### 6.1. Оценивание моделей для промежутков времени между сделками

Вначале была оценена полупараметрическая ACD(2,2) модель:

$$\begin{aligned}d_t &= \psi_t \varepsilon_t, \\ \psi_t &= \alpha + \beta_1 d_{t-1} + \beta_2 d_{t-2} + \gamma_1 \psi_{t-1} + \gamma_2 \psi_{t-2} + \delta v_{t-1}.\end{aligned}$$

Здесь лагированное значение  $v_{t-1}$  введено для учета зависимости от объема предыдущих сделок. Результаты оценивания показывают, что процесс очень устойчив (сумма коэффициентов  $\beta$  и  $\gamma$  превосходит 0.95). Значение статистики Льюнга-Бокса для тестирования нулевой гипотезы об отсутствии автокорреляции в стандартизованных остатках значимо отличается от нуля на однопроцентном уровне ( $Q(15) = 32.67$ , в то время как однопроцентное критическое значение равно 30.58). Еще более значимая корреляция наблюдается в квадратах стандартизованных остатков (значение статистики Льюнга-Бокса равно 44.90). Кроме этого, тест на избыточную дисперсию уверенно отвергает гипотезу об ее отсутствии (значение статистики равно 13.50, однако при нулевой гипотезе экспоненциального распределения ее асимптотическое распределение является стандартным нормальным). Аналогично, параметрические ACD(2,2) модели с вейбулловским и обобщенным гамма распределениями не смогли пройти тесты на отсутствие автокорреляции в стандартизованных остатках и их квадратах.

После этого была оценена полупараметрическая логарифмическая ACD(2,2) модель, и оказалось, что и она не может пройти тесты на автокорреляцию (соответствующие значения статистик Льюнга-Бокса — 108.57 и 137.92 для стандартизованных остатков и их квадратов). Таким образом, мы можем заключить, что традиционные модели не очень подходят для российских данных и мы имеем все основания для применения log-ARMA спецификации, предложенной в параграфе 4.

Мы остановились на следующей log-ARMA модели:

$$\begin{aligned}\ln d_t &= \alpha + \beta \ln d_{t-1} + \gamma_1 \varepsilon_{t-1} + \gamma_2 \varepsilon_{t-2} + \delta v_{t-1} + \varepsilon_t, \\ \varepsilon_t &= \sigma_t \eta_t, \\ \sigma_t^2 &= \mu + \lambda \varepsilon_{t-1}^2 + \nu \sigma_{t-1}^2.\end{aligned}$$

Результаты оценивания этой модели для семи акций приведены в Таблице 6 (см. Приложение Б).

<sup>15</sup>Применялись алгоритмы BFGS и ВННН, для вычисления стандартных ошибок использовались численные частные производные.

Легко видеть, что все модели проходят тесты Льюнга-Бокса для стандартизованных остатков и их квадратов на однопроцентном уровне. Также можно заметить, что модели для акций Лукойла, РАО ЕЭС, Сбербанка, Юкоса, СургутНефтеГаза и Ростелекома проходят тесты Ниблома и для индивидуальных коэффициентов и для модели в целом на однопроцентном уровне. Нестабильность модели для МосЭнерго может быть объяснена очень нестабильным коэффициентом при лагированном объеме. Подводя итог, можно считать, что рассмотренная log-ARMA спецификация вполне адекватно описывает данные.

Дальнейшая диагностика модели может быть проведена с помощью тестов на соответствие стандартизованных остатков какому либо конкретному распределению. Непараметрическая оценка плотности стандартизованных остатков для акций Лукойла приведена на Рисунке 7 (метод Надарайя-Уотсона, гауссовское ядро, правило Силвермана (Silverman, 1986, уравнение 3.31) для выбора ширины окна). Оцененная плотность весьма похожа на нормальную, поэтому естественно протестировать гипотезу о нормальности остатков. Однако, тесты Колмогорова-Смирнова, Жарке-Бера, Крамера-фон Мизеса и Андерсона-Дарлинга отвергают эту гипотезу. Возможной причиной этого является тот факт, что распределение остатков для всех акций, кроме РАО ЕЭС немного скошено и имеет недостаточный кurtosis (для Лукойла значения скошенности и кurtosis равны, соответственно, -0.174 и 2.678). Вероятно, что можно учесть эту скошенность распределения остатков с помощью введения несимметричных членов в ARMA уравнение, однако, мы не рассматриваем эту возможность в данной работе.

Таблица 3. Средние значения функции правдоподобия

	Сбербанк	Юкос	РАО ЕЭС	СургутНефтеГаз
Log-ARMA	-0.7712	-0.8687	-0.7809	-0.8524
ACD	-0.9188	-0.9504	-0.9177	-0.9301
Log-ACD	-0.9169	-0.9502	-0.9197	-0.9272
	Ростелеком	МосЭнерго	Лукойл	–
Log-ARMA	-0.8458	-0.8558	-0.8398	–
ACD	-0.9259	-0.9335	-0.9229	–
Log-ACD	-0.9261	-0.9326	-0.9226	–

Сравнение полупараметрических ACD, log-ACD и log-ARMA моделей может быть проведено также и с помощью сравнения значений функций квази-правдоподобия<sup>16</sup>. Средние значения функций правдоподобия (в оцененных параметрах) для ACD, log-ACD и log-ARMA моделей, описанные в начале этого подпараграфа, приведены в Таблице 3. Легко видеть, что log-ARMA модель более предпочтительна для всех рассмотренных акций. Так как log-ARMA модель использует больше

<sup>16</sup>Так как в log-ARMA модели максимизируется функция квази-правдоподобия для логарифмов промежутков времени между сделками, мы домножили ее на соответствующий якобиан чтобы получить значение функции для самих промежутков времени между сделками.

параметров чем ACD и log-ACD, то мы также вычислили информационные критерии Акаике и Шварца (результаты здесь не приводятся) и они также предпочитают log-ARMA спецификацию. Из Таблицы 3 можно заметить, что log-ACD и ACD модели имеют почти одинаковые значения функций правдоподобия, поэтому мы можем заключить, что log-ARMA модель выигрывает по одной из двух причин: либо нормальное распределение значительно ближе к истинному распределению логарифмов промежутков времени между сделками, чем экспоненциальное — к истинному распределению самих промежутков времени между сделками; либо динамика условного среднего и условной дисперсии процесса для промежутков времени между сделками не связаны таким жестким соотношением, какое предполагается в ACD модели.

Перейдем теперь к интерпретации полученных оценок. Очень значимые AR и MA коэффициенты показывают, что условное среднее процесса очень устойчиво, что совпадает с результатами предыдущих исследований (см., например, Engle and Russell, 1998; Engle, 2000; или Zhang, Russell, and Tsay, 2001). Лагированное значение объема имеет отрицательный коэффициент, очень значимый для всех акций, таким образом, мы заключаем, что когда объемы сделок велики, промежутки между сделками становятся меньше. Это может рассматриваться как проявление эффекта совместной торговли “ликвидных” трейдеров, который предсказывается моделью Адматти и Пфляйдера (см. параграф 3). Кластеризация торговых событий, наличие которой видно из высоких значений AR и MA коэффициентов также интерпретируется в пользу этой модели микроструктуры рынка.

Дальнейший анализ процесса для промежутков времени между сделками возможно осуществить с помощью уравнения для дисперсии. Устойчивость в этом уравнении даже больше чем в уравнении для условного среднего: сумма ARCH и GARCH коэффициентов близка к 0.99 для всех рассмотренных акций. Эти коэффициенты могут быть интерпретированы с помощью функций риска, которые выведены при предположении о нормальном распределении остатков. Если значения  $\varepsilon_{t-1}^2$  и  $\sigma_{t-1}$  были велики (то есть значение предыдущего промежутка времени между сделками было далеко от его условного среднего), то  $\sigma_t$  тоже будет велико и, следовательно, функция риска будет довольно крутой (см. Рисунок 6 для  $\sigma = 1.3$  в Приложении Б), то есть, ожидается, что следующий промежуток времени будет либо очень коротким, либо значительно более длинным. С другой стороны, если предыдущие значения  $\varepsilon_{t-1}^2$  и  $\sigma_{t-1}$  были небольшими, то функция риска будет гораздо более пологой (см. Рисунок 6 для  $\sigma = 0.7$  в Приложении Б), поэтому ожидания о следующем промежутке времени будут более однородными.

Кроме этого, оцененные коэффициенты дают повод утверждать, что избыточная дисперсия в этом процессе на самом деле является непостоянной во времени, что противоречит предположениям ACD моделей. Более того, высокая устойчивость в уравнении для условной дисперсии показывает, что наблюдения с избыточной (недостаточной) дисперсией будут кластеризоваться.

## 6.2. Оценивание моделей для волатильности

Сперва была оценена расширенная UHF-GARCH модель (уравнение (6)), с ARMA(1,1) процессом для условного среднего (кроме того, текущее значение промежутка времени было включено в уравнение для среднего). Эта спецификация оказалась неадекватной, так как статистики Льюнга-Бокса для стандартизованных остатков и их квадратов были очень велики (например, для для Лукойла значение статистики Льюнга-Бокса с 15 лагами для остатков было равно 1109.13, что весьма далеко от однопроцентного критического значения 30.58). Поэтому мы полностью перешли к рассмотрению SA-GARCH моделей. Более точно, оценивалась следующая спецификация:

$$r_t = \alpha + \beta r_{t-1} + \gamma \varepsilon_{t-1} + \delta_1 d_{t-1} + \delta_2 i_{t-1} + \varepsilon_t,$$
$$\frac{\sigma_t^2}{d_t^\rho} = \mu + \lambda_1 \frac{\varepsilon_{t-1}^2}{d_{t-1}^\rho} + \lambda_2 \frac{\varepsilon_{t-2}^2}{d_{t-2}^\rho} + \nu_1 \frac{\sigma_{t-1}^2}{d_{t-1}^\rho} + \nu_2 \frac{\sigma_{t-2}^2}{d_{t-2}^\rho} + \delta_3 i_{t-1}.$$

Здесь переменная  $i_t$  определяется как  $i_t = \nu_t - \ln d_t$  и характеризует интенсивность потока акций<sup>17</sup>.

Результаты оценивания представлены в Таблице 7. Во-первых, нужно отметить, что все модели, кроме РАО ЕЭС, прошли тест Льюнга-Бокса на однопроцентном уровне<sup>18</sup> (для Лукойла значение статистики маржинально значимо на этом уровне). Более того, для Сбербанка и Юкоса оказалось достаточным использовать SA-GARCH(1,1) (то есть,  $\lambda_2 = \nu_2 = 0$ ). Во-вторых, все модели и их коэффициенты (кроме коэффициента  $\rho$  для Лукойла) являются стабильными на однопроцентном уровне, как показывают статистики Ниблома (для модели с 10 параметрами критическое однопроцентное значение равно 3.04, для модели с 12 параметрами — около 3.40; критические значения для индивидуальных статистик приведены в примечаниях ниже Таблицы 6).

Мы начнем анализ оцененных моделей с уравнения для условного среднего. Очень значимый MA(1) коэффициент указывает на отрицательную автокорреляцию доходностей, похожие результаты были получены Энглем при использовании UHF-GARCH моделей. Такая отрицательная автокорреляция часто присутствует в высокочастотных данных и может быть объяснена структурой электронной торговой системы; другие объяснения можно найти в книге (Gourieroux and Jasiak, 2001, параграф 14.2), но стоит подчеркнуть, что такая предсказуемость доходностей на основе предыдущей информации не противоречит эффективности рынка. Незначимость AR(1) коэффициента (в работе (Engle, 2000) этот коэффициент был значим во всех моделях) для пяти акций из семи можно объяснить тем фактом, что MA член объясняет значительную часть отрицательной автокорреляции; однако, если оценивать модель без MA членов, то первые два AR коэффициента будут значимыми (на пятипроцентном уровне) и отрицательными. Несмотря на значимость MA коэффициентов, устойчивость в уравнении для условного среднего не очень высока. Удивительно, что константа в

---

<sup>17</sup>Напомним, что  $\nu_t$  есть скорректированное значение  $\ln V_t$ , логарифма числа акций, купленных в момент времени  $T_t$ .

<sup>18</sup>См. критические значения ниже Таблицы 6.



этом уравнении оказалась значимой и отрицательной для трех акций (Лукойл, МосЭнерго и Юкос). Далее, коэффициент при текущем промежутке времени между сделками значим и положителен для трех акций и незначим для остальных, что можно рассматривать как противоречие с моделью Даймонда и Верречия, по которой этот коэффициент должен быть значим и отрицателен. Это отличается от результатов в (Engle, 2000), где были найдены подтверждения для этой модели в данных по сделкам с акциями IBM. Однако, приведенные там оценки довольно сложно интерпретировать: в модели Энгля используются не сами доходности, а доходности, деленные на корень из соответствующего промежутка времени между сделками, таким образом, наблюдая отрицательный коэффициент в его модели мы можем лишь утверждать, что доходность растет медленнее чем корень из  $d_t$ . Значимые положительные коэффициенты при переменной  $i_{t-1}$  для большинства акций можно интерпретировать следующим образом: если предположить, что большие сделки инициируются продавцами, и следовательно, дают отрицательные доходности (так как при таких сделках вероятнее всего будут исполнены несколько стоящих в очереди заявок и поэтому средняя цена для этой сделки будет меньше), то следующая сделка с большей вероятностью будет иметь положительную доходность в силу отрицательной автокорреляции в процессе для доходностей; таким образом, мы будем наблюдать положительное влияние лагированных значений интенсивности потока акций на доходность. Такой эффект не наблюдается только для ПАО ЕЭС, что, скорее всего, вызвано искажениями, внесенными в процесс при удалении одновременных сделок (напомним, что акции ПАО ЕЭС торгуются примерно в 10 раз чаще чем остальные).

Перейдем теперь к анализу уравнения для условной дисперсии. Во-первых, для всех акций хорошо заметна кластеризация волатильности: уравнение для условной дисперсии довольно устойчиво, как видно из большой суммы ARCH и GARCH коэффициентов. Кроме этого, коэффициент при переменной, характеризующей интенсивность торговли значим и положителен, что можно воспринимать как свидетельство против модели Адматти и Пфляйдера, в которой большие объемы сделок ассоциируются с присутствием на рынке большой доли “ликвидных” трейдеров, и, следовательно, волатильность должна снижаться в этой ситуации, так как эти сделки не несут новой информации об акции.

Наконец, мы получили высокозначимые коэффициенты  $\rho$ , которые, тем не менее, далеки от 1, они варьируются от 0.1 для ПАО ЕЭС до 0.28 для Сбербанка. Этот результат непосредственно подтверждает нашу гипотезу о том, что условная дисперсия шока в процессе для доходностей,  $\sigma_t^2$ , является вогнутой, а вовсе не линейной, функцией текущего промежутка времени между сделками.

Для исследования различий между UHF-GARCH уравнением ( $\rho = 1$ ) и SA-GARCH уравнением (мы положим  $\rho = 0.2$ , что близко к полученным нами оценкам), предположим, что предыдущий промежуток времени между сделками ( $d_{t-1}$ ) равен 1 (это выборочное среднее для скорректированных промежутков времени между сделками), следовательно, для UHF-GARCH модели мы имеем  $\sigma_t^2 = d_t s_{t-1}^2$ , где  $s_{t-1}^2 = \mu + \lambda \varepsilon_{t-1}^2 + \nu \sigma_{t-1}^2$  (для простоты мы рассматриваем GARCH(1,1)), в то время

как для SA-GARCH мы получаем  $\sigma_t^2 = d_t^{0.2} s_{t-1}^2$ . Таким образом, мы видим, что при малых значениях  $d_t$  условная дисперсия в SA-GARCH модели сильнее зависит от значения  $d_t$  и больше по величине, чем в UHF-GARCH модели. Однако, для больших значений  $d_t$ , наблюдается обратная ситуация. Используя уравнение (7), мы получаем, что естественная мера волатильности экспоненциально убывает (с показателем -0.8) с ростом  $d_t$ . Это вполне согласуется с результатами в работе (Engle, 2000), где сообщается, что  $d_t^{-1}$  имеет очень значимый положительный коэффициент в уравнении для условной дисперсии (уравнение (6)). Отрицательная зависимость естественной меры волатильности от текущего значения промежутка времени между сделками можно интерпретировать как подтверждение модели Эсли и О'Хара, в которой длинные промежутки времени между сделками свидетельствуют об отсутствии новостей. Кроме этого, такая отрицательная зависимость противоречит модели Даймонда и Верречия, в соответствии с которой длинные промежутки времени между сделками должны рассматриваться как индикатор плохих новостей, и, следовательно, должны увеличивать волатильность.

## 7. Выводы

В данной работе мы предложили новую модель для анализа промежутков времени между сделками, которая является более гибкой по сравнению с ACD моделью, и, с другой стороны, такой же простой с точки зрения оценивания. Еще одна новая модель предложена для исследования волатильности доходностей; эта модель гораздо лучше работает для российских данных чем традиционная UHF-GARCH модель. Обе новые модели применяются для анализа поведения акций семи российских компаний. С помощью них было установлено, что поведение участников торгов в России весьма сходно с поведением трейдеров на западных рынках, кроме того, результаты оценивания моделей дают частичное подтверждения выводам моделей Эсли и О'Хара, и Адмати и Пфляйдера. Однако, не найдено никаких подтверждений для модели Даймонда и Верречия, в которой предполагаются ограничения на занятие коротких позиций.

## 8. Благодарности

Автор хотел бы поблагодарить профессора Станислава Анатольева, под руководством которого это исследование было выполнено. Также, автор очень признателен Георгию Колодяжному (Центральный Банк), который любезно предоставил данные для этого проекта. Кроме того, автор выражает свою благодарность ассистенту проекта, Ольге Крюковской, за ее помощь в подготовке данных, а также всем остальным участникам проекта "Динамика российских финансовых данных" за их многочисленные ценные предложения и замечания.

## 9. Литература

1. Admati, A. R., Pfleiderer, P. (1988). *A theory of intraday patterns: volume and price variability*. The Review of Financial Studies 1, 3-40.
2. Andrews, D. W. K. (1993). *Tests for parameter instability and structural change with unknown change point*. Econometrica 61, 821-856.
3. Andrews, D. W. K., Ploberger, W. (1994). *Optimal tests when a nuisance parameter is present only under the alternative*. Econometrica 62, 1383-1414.
4. Barinov, V., Pervozvansky, A., Pervozvanskaya, T. (1999). *State debt policy and state bond market behavior*. EERC Working paper.
5. Bauwens, L., Giot, P. (2000). *The logarithmic ACD model: an application to the bid/ask quote process of three NYSE stocks*. Annales d'Economie et de Statistique 60, 117-149.
6. Bollerslev, T., Wooldridge, J. (1992). *Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances*. Econometric Reviews 11, 143-172.
7. Darolles, S., Gouriéroux, S., Le Fol, G. (2000). *Intraday transaction price dynamics*. Annales d'Economie et de Statistique 60, 239-270.
8. Diamond, D. W., Verrecchia, R. E. (1987). *Constraints on short-selling and asset price adjustments to private information*. Journal of Financial Economics 18, 277-311.
9. Drost, F. C., Nijman, T. E. (1993). *Temporal aggregation of GARCH processes*. Econometrica 61, 907-927.
10. Drost, F. C., Werker, B. J. M. (1996). *Closing the GARCH gap: continuous time GARCH modeling*. Journal of Econometrics 74, 31-57.
11. Easley, D., O'Hara, M. (1992). *Time and the process of security price adjustment*. Journal of Finance 47, 577-606.
12. Engle, R. F. (2000). *The econometrics of ultra-high frequency data*. Econometrica 68, 1-22.
13. Engle, R. F., Lunde, A. (1998). *Trades and quotes: a bivariate point process*. Discussion paper, University of California, San Diego.
14. Engle, R. F., Russell, J. R. (1998). *Autoregressive conditional duration: a new model for irregularly spaced transaction data*. Econometrica 66, 1127-1162.
15. Ghysels, E., Gouriéroux, C., Jasiak, J. (2003). *Stochastic volatility duration models*. Journal of Econometrics, to appear.
16. Ghysels, E., Jasiak, J. (1998). *GARCH for irregularly spaced financial data: the ACD-GARCH model*. Studies in Nonlinear Dynamics and Econometrics 2, 133-149.
17. Goldberg, L., Tenorio, R. (1997). *Strategic trading in a two-sided foreign exchange auction*. Journal of International Economics 42, 299-326.
18. Gouriéroux, C., Jasiak, J. (2000). *Durations*, in "A Companion to Theoretical Econometrics", B. Baltagi (ed.), Blackwell.
19. Gouriéroux, C., Jasiak, J. (2001). *Financial Econometrics: Problems, Models, and Methods*. Princeton

University Press.

20. Grammig, J., Maurer, K.-O. (2000). *Non-monotonic hazard functions and the autoregressive conditional duration model*. The Econometrics Journal 3, 16-38.
21. Grammig, J., Wellner, M. (2002). *Modeling the interdependence of volatility and inter-transaction duration processes*. Journal of Econometrics 106, 369-400.
22. Jasiak, J. (1999). *Persistence in intertrade durations*. Finance 19, 166-195.
23. McCulloch, R. E., Tsay, R. S. (2001). *Nonlinearity in high frequency financial data and hierarchical models*. Studies in Nonlinear Dynamics and Econometrics 5.
24. Meddahi, N., Renault, E., Werker, B. (2003). *GARCH and irregularly spaced data*. Preprint.
25. Medvedev, A. N., Kolodyazhny, G. V. (2001). *Financial crisis in Russia: the behavior of non-residents*. EERC Working paper.
26. Medvedev, A. N., Kolodyazhny, G. V. (2002). *Russian stock market: participants and their strategies*. EERC Working paper.
27. Nyblom, J. (1989). *Testing the constancy of parameters over time*. Journal of the American Statistical Association 84, 223-230.
28. Russell, J. R., Engle, R. F. (1999). *Econometric analysis of discrete-valued, irregularly-spaced financial transactions data*. Working Paper, University of Chicago and University of San Diego.
29. Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. Chapman & Hall, London.
30. Tyurin, K. (2002). *Modeling high-frequency dynamics of financial markets in continuous time: an event-based approach*. Preprint.
31. Zhang, M. Y., Russell, J. R., Tsay, R. S. (2001). *A nonlinear autoregressive conditional duration model with applications to financial transaction data*. Journal of Econometrics 104, 179-207.

## 10. Приложение А: Информация о ценных бумагах

### 10.1. Торговые режимы

В электронной системе ММВБ сделки могут проводиться в нескольких режимах. Эти режимы и их краткие описания представлены в Таблице 4.

Таблица 4. Торговые режимы в системе ММВБ

Торговый режим	Расписание	Описание
Предторговый период	10:00-10:30	Определение цены открытия.
Торговая сессия	10:30-18:45	Основной режим торгов.
Послеторговый период	18:45-19:00	В этом режиме сделки происходят по средневзвешенной цене за последние полчаса торговой сессии.
Режим переговорных сделок	10:00-19:05	Этот режим предназначен для крупных сделок; поставка не всегда выполняется сразу, она может быть отложена на срок до 30 дней.
Режим переговорных сделок (с ограничениями на коды расчетов)	19:05-22:00	Тоже, что предыдущее, но сделки производятся не для всех кодов расчетов.
Режим РЕПО	10:15-19:05	Заключение сделок РЕПО.
Неполные лоты	17:00-19:05	Заключения сделок с нестандартным размером лота.

В данном исследовании мы рассматривали только сделки, заключенные во время торговой сессии, все остальные наблюдения были удалены из выборки, а именно:

1. Сделки, заключенные не во время торговой сессии (10: 30-18: 45);
2. Сделки, отмеченные строкой “РПС:” (режим переговорных сделок);
3. Сделки, отмеченные строкой “РЕПО:” (режим РЕПО);
4. Сделки, отмеченные строкой “Неполные лоты” (режим “неполные лоты”).

### 10.2. Ценные бумаги

Мы использовали данные о сделках с обыкновенными акциями семи российских компаний, краткое описание этих ценных бумаг приведено в Таблице 5. Среди компаний, которые мы рассматриваем, три являются нефтедобывающими (Лукойл, СургутНефтеГаз и Юкос), РАО ЕЭС и Мосэнерго — это крупные производители электроэнергии, Ростелеком является самой крупной российской телекоммуникационной компанией и Сбербанк — самый крупный российский банк. Для некоторых из этих компаний было несколько выпусков обыкновенных акций, наряду с привилегированными

акциями и облигациями, однако, для каждой компании мы рассматриваем только один выпуск обыкновенных акций (для всех компаний есть один выпуск акций, торгуемый гораздо чаще остальных). Столбец “Доля в сумме сделок” показывает процент сделок по конкретной акции в общей сумме сделок по всем акциям за период с 19 по 23 августа 2002 года. Несложно подсчитать, что все семь акций занимают почти 97 процентов в общей сумме сделок. Из рассмотренных ценных бумаг, акции РАО ЕЭС торгуются наиболее часто (в среднем, 6000 сделок в день, одновременные сделки удалены), по акциям Лукойла, СургутНефтеГаза, Ростелекома и Мосэнерго совершается по 600-750 сделок в день; по акциям Юкоса — 400 и по акциям Сбербанка — 250. Период наблюдений — с 12 августа по 27 октября 2002 года, то есть, 50 торговых дней. Однако, так как для РАО ЕЭС мы имеем гораздо больше сделок чем для других компаний, то по этой акции мы используем только 4 недели для оценки внутрдневной структуры и одну неделю для оценки моделей. Также, для акций СургутНефтеГаза, данные за 14 августа 2002 года были неполными, поэтому этот день для этой акции пропущен.

Таблица 5. Описание акций

Компания	Код бумаги	Шаг цены	Лот	Период наблюдений	Доля в сумме сделок	Наблюдений в конечной выборке
Сбербанк	RU0009029540	1	1	12.08.02- 27.10.02	0.75	13156
Юкос	RU0009054449	1	1	12.08.02- 27.10.02	1.99	20125
Лукойл	RU0009024277	1	1	12.08.02- 27.10.02	15.71	34316
СургутНефтеГаз	RU0008926258	0.1	100	12.08.02- 27.10.02	5.27	38275
Ростелеком	RU0008943394	1	1	12.08.02- 27.10.02	2.75	38246
MosEnergo	RU14MSNG3008	0.1	100	12.08.02- 27.10.02	1.88	32452
РАО ЕЭС	RU0008926621	0.1	100	19.08.02-23.08.02	68.75	30537

Шаг цены в копейках, лот — в акциях.

## 11. Приложение Б: Таблицы и рисунки

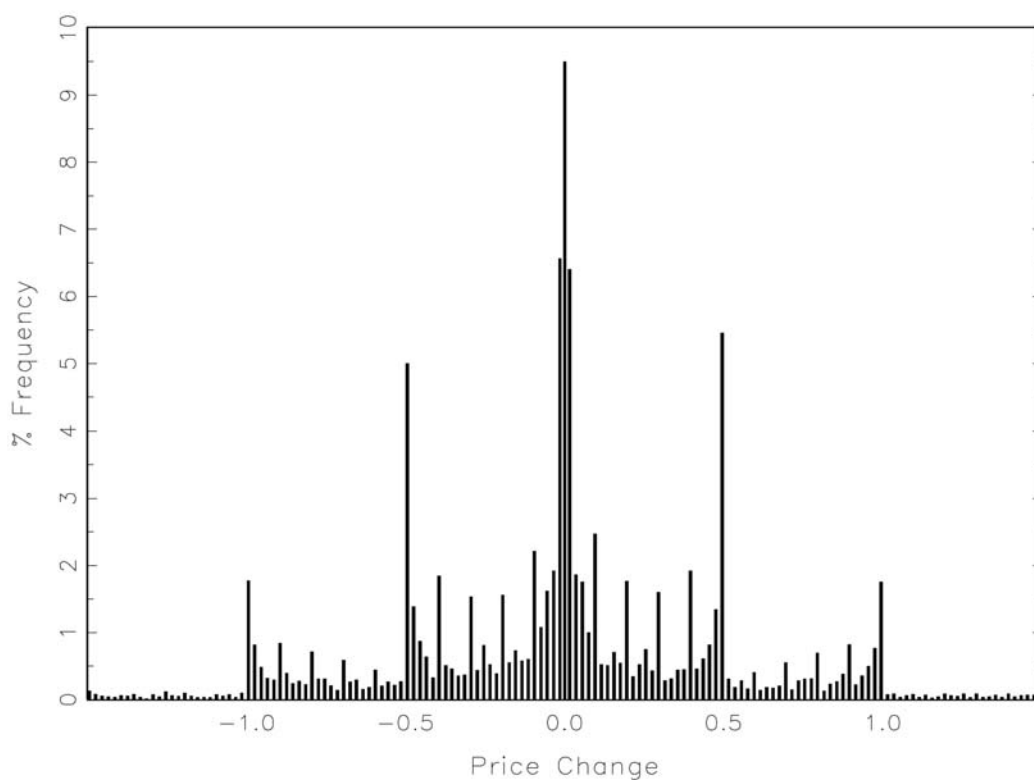


Рисунок 1. Гистограмма для изменений цены  
(Лукойл, нулевые изменения цены исключены).

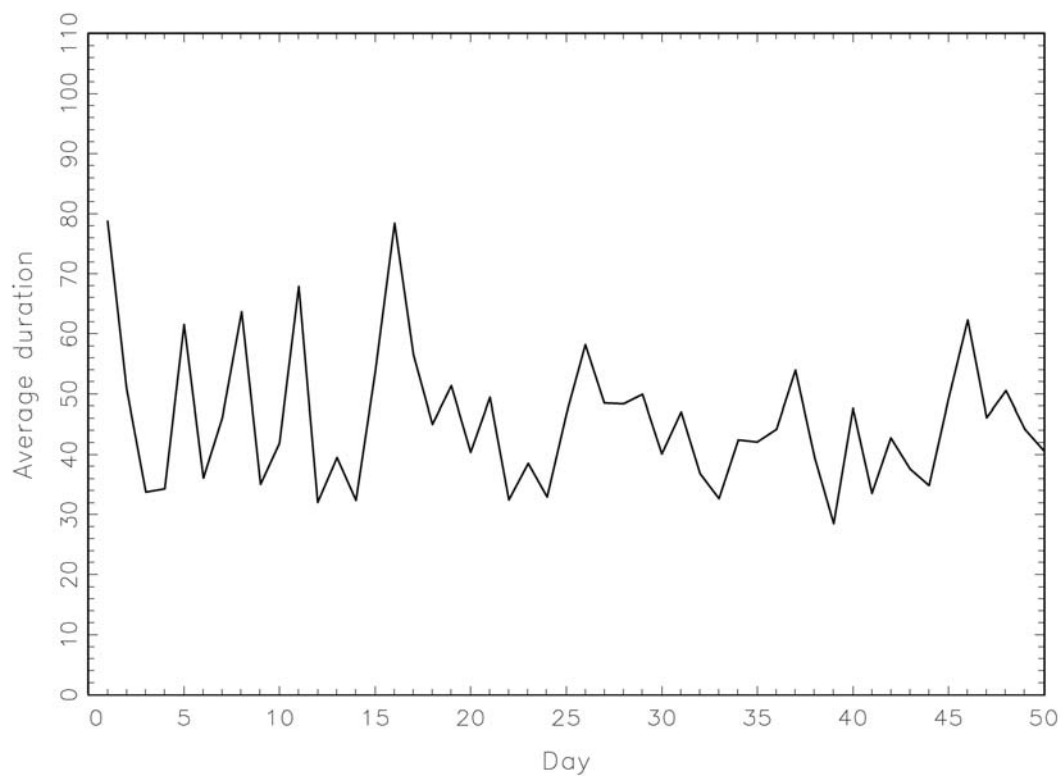


Рисунок 2. Средний промежуток времени между сделками для различных дней (Лукойл).

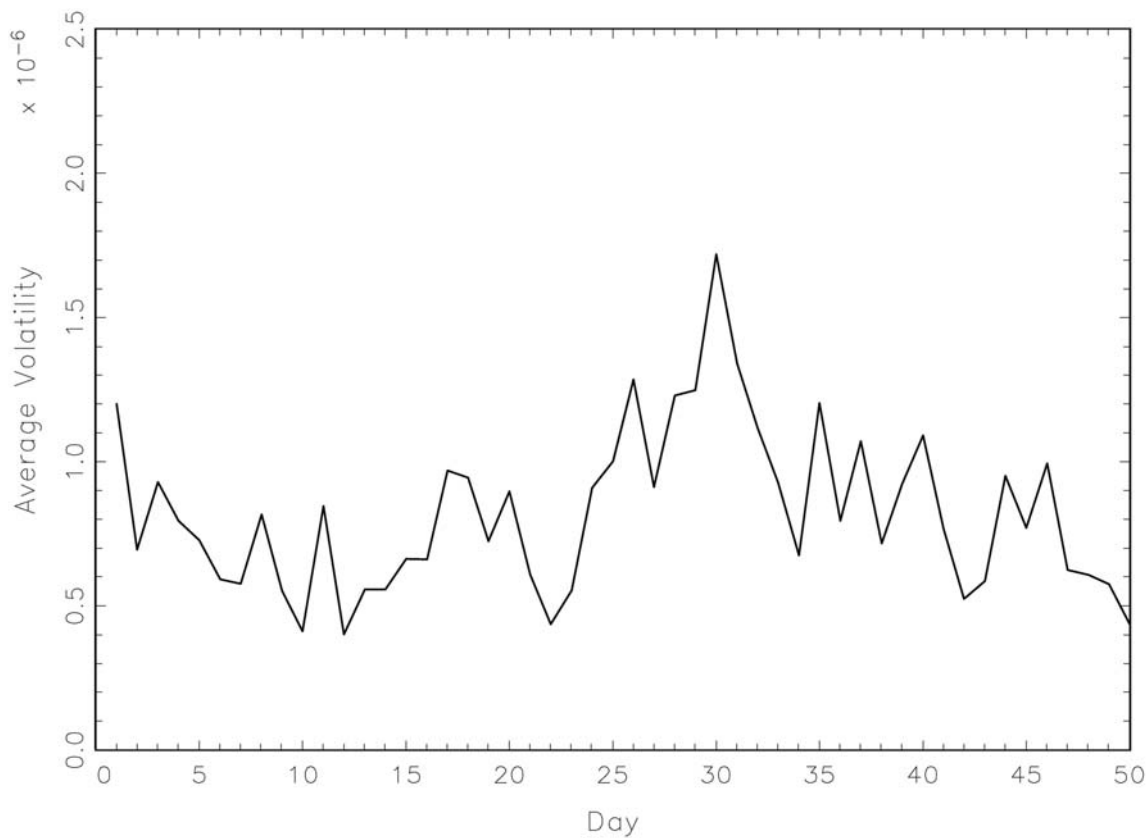


Рисунок 3. Средняя волатильность доходности для различных дней (Лукойл).

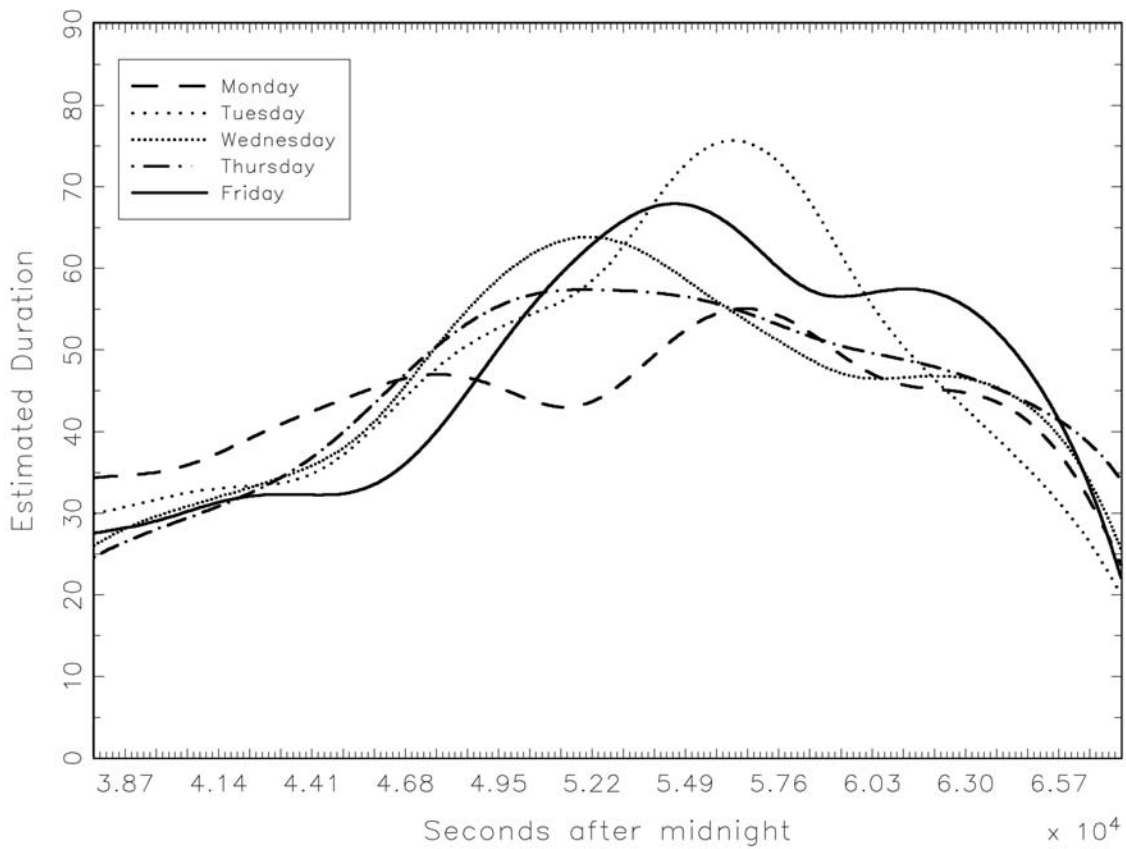


Рисунок 4. Оценка внутрдневной структуры промежутков времени между сделками (Лукойл, метод локальной линейной регрессии, гауссовское ядро, ширина окна = 2000).



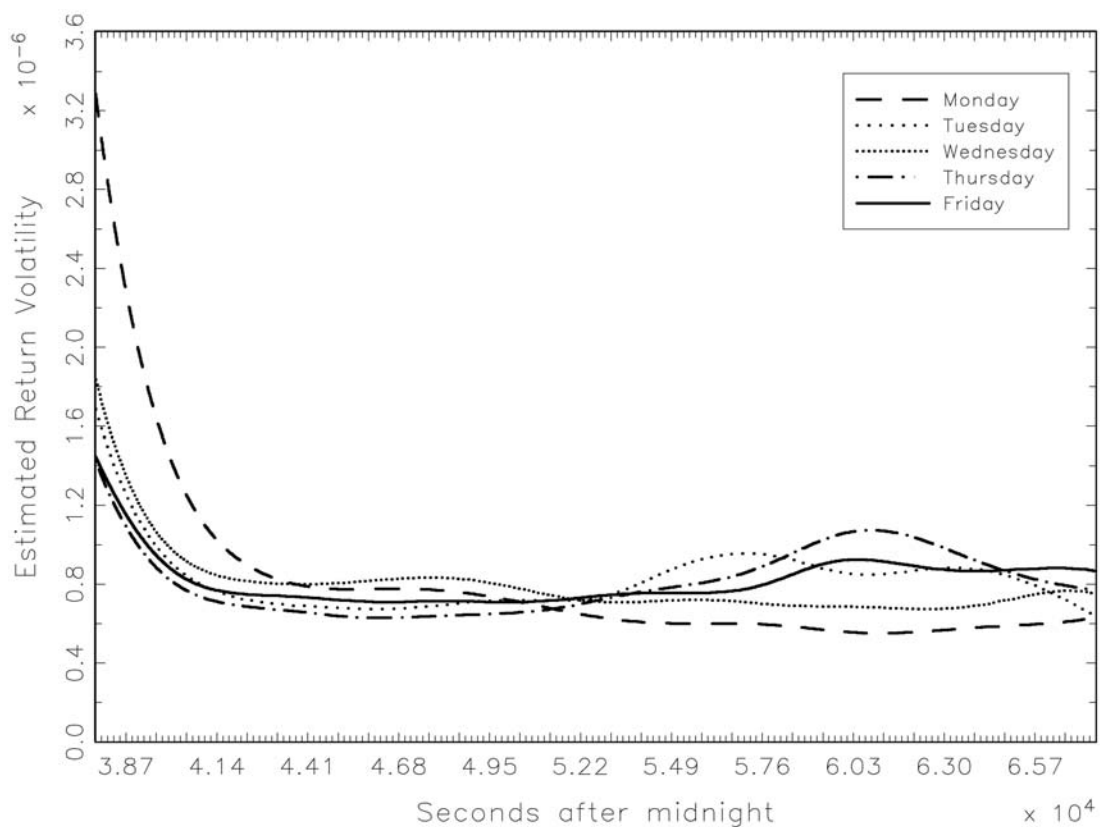


Рисунок 5. Оценка внутредневной структуры волатильности

(Лукойл, метод локальной линейной регрессии, гауссовское ядро, ширина окна = 2000).

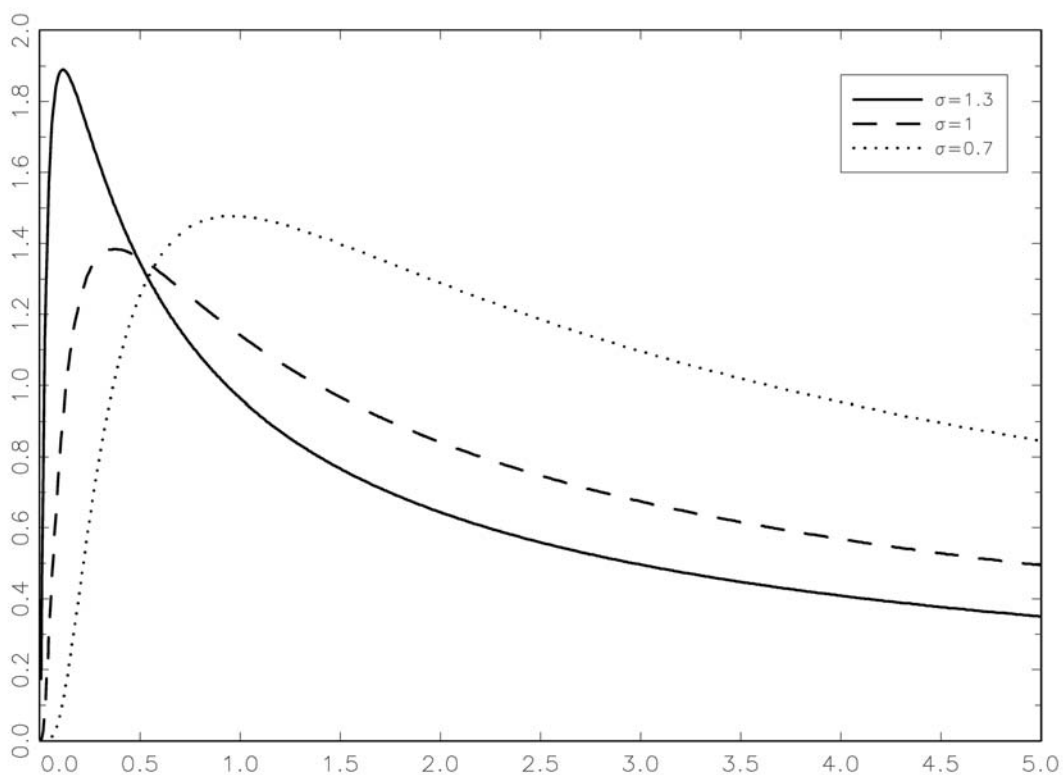


Рисунок 6. Функции риска для лог-нормального распределения

(плотность:  $f(x) = \frac{1}{\sqrt{2\pi\sigma x}} \exp\left(-\frac{(\ln x + \sigma^2/2)^2}{2\sigma^2}\right)$  при  $x > 0$ ).

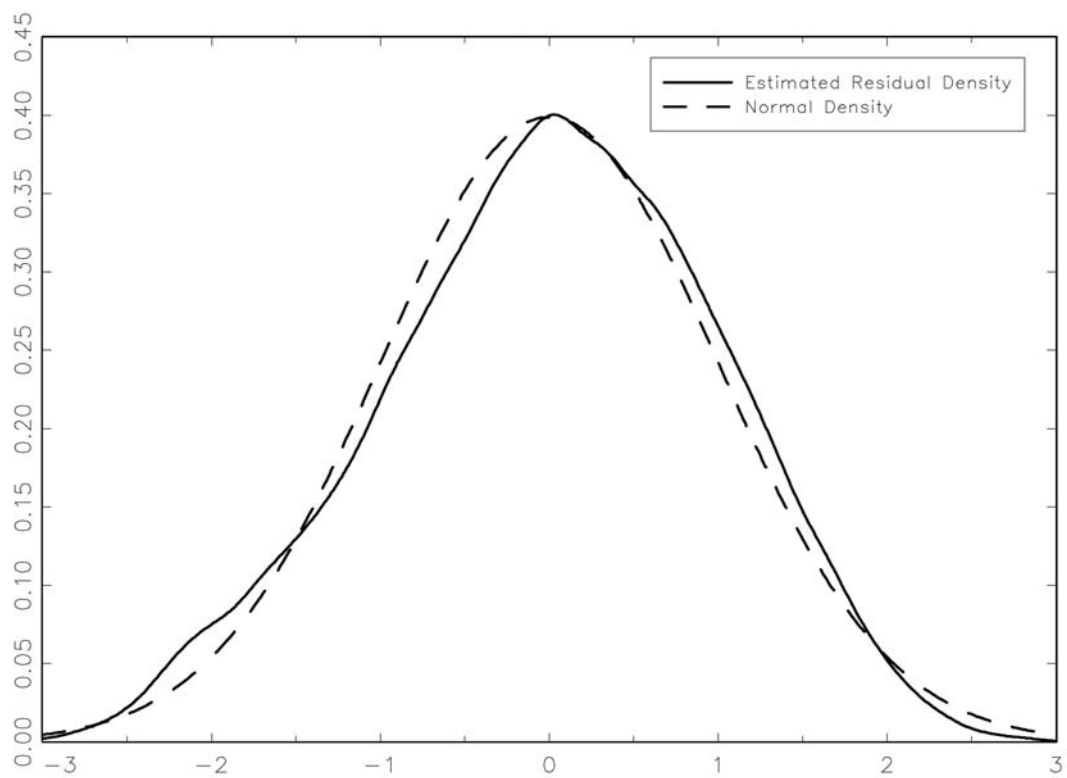


Рисунок 7. Оценка плотности распределения остатков для log-ARMA модели (Лукойл, метод Надарья-Уотсона, гауссовское ядро, ширина окна по методу Силвермана).

Таблица 6. Log-ARMA Модели								
	РАО ЕЭС		Сбербанк		Мосэнерго		Лукойл	
	Коэф.	Нуb	Коэф.	Нуb	Коэф.	Нуb	Коэф.	Нуb
Среднее								
$\alpha$	-0.025 (-11.81)	0.14	-0.186 (-8.70)	0.15	-0.085 (-12.01)	0.13	-0.109 (-10.54)	0.06
$\beta$	0.950 (237.09)	0.10	0.812 (39.19)	0.10	0.897 (107.73)	0.15	0.875 (74.67)	0.22
$\gamma_1$	-0.819 (-115.58)	0.19	-0.605 (-26.35)	0.11	-0.732 (-71.51)	0.17	-0.691 (-52.06)	0.07
$\gamma_2$	-0.028 (-4.38)	0.24	-0.029 (-2.57)	0.13	-0.033 (-4.97)	0.50	-0.049 (-6.71)	0.19
$\delta$	-0.013 (-11.37)	0.27	-0.061 (-9.28)	0.12	-0.014 (-7.85)	1.40	-0.028 (-11.34)	0.31
Дисперсия								
$\mu$	0.011 (6.64)	0.37	0.019 (2.86)	0.25	0.012 (4.47)	0.43	0.046 (4.84)	0.05
$\lambda$	0.044 (13.38)	0.38	0.020 (4.85)	0.12	0.013 (9.19)	0.37	0.020 (8.93)	0.05
$\nu$	0.941 (185.68)	0.43	0.970 (140.07)	0.16	0.979 (363.89)	0.39	0.954 (133.11)	0.05
Диагностические статистики								
$Q(15)$	18.47		27.46		18.04		22.15	
$Q^2(15)$	12.14		19.24		11.34		15.12	
Nyblom	1.96		2.04		3.86		1.52	

**Примечания:** Робастные t-статистики приведены в скобках. Нуb обозначает индивидуальную статистику Ниблома, 5- и 1-процентные критические значения которой равны 0.46 и 0.74 соответственно. Nyblom — это статистика Ниблома для модели в целом, 5- и 1-процентные критические значения равны 2.12 и 2.58.  $Q(15)$  и  $Q^2(15)$  — статистики Льюнга-Бокса с 15 лагами для стандартизованных остатков и их квадратов, 5- и 1-процентные критические значения равны 25 и 30.58 соответственно.

Таблица 6. Log-ARMA Модели (Продолжение)

	СургутНефтеГаз		Юкос		Ростелеком	
	Коэф.	Nyb	Коэф.	Nyb	Коэф.	Nyb
Среднее						
$\alpha$	-0.101 (-12.87)	0.13	-0.170 (-8.31)	0.11	-0.078 (-11.36)	0.16
$\beta$	0.876 (93.27)	0.20	0.808 (35.43)	0.20	0.906 (111.68)	0.20
$\gamma_1$	-0.701 (-63.74)	0.11	-0.639 (-26.24)	0.53	-0.736 (-73.92)	0.10
$\gamma_2$	-0.039 (-6.18)	0.12	-0.023 (-2.35)	0.61	-0.048 (-7.66)	0.15
$\delta$	-0.028 (-11.78)	0.12	-0.032 (-7.85)	0.35	-0.017 (-9.66)	0.17
Дисперсия						
$\mu$	0.022 (4.52)	0.25	0.028 (2.28)	0.35	0.015 (4.76)	0.19
$\lambda$	0.014 (8.76)	0.25	0.010 (3.21)	0.39	0.012 (9.06)	0.16
$\nu$	0.972 (227.44)	0.21	0.975 (106.87)	0.38	0.978 (334.83)	0.20
Диагностические статистики						
$Q(15)$	24.13		29.88		19.94	
$Q^2(15)$	27.60		20.81		20.96	
Nyblom	2.15		1.93		1.37	

Таблица 7. SA-GARCH Модели

	Лукойл		Сбербанк		Мосэнерго		СургутНефтеГаз	
	Коэф.	Нуb	Коэф.	Нуb	Коэф.	Нуb	Коэф.	Нуb
Среднее								
$\alpha$	-0.016 (-3.29)	0.40	0.014 (1.69)	0.07	-0.010 (-2.11)	0.08	-0.004 (-0.71)	0.12
$\beta$	-0.002 (-0.07)	0.37	-0.012 (-0.29)	0.11	0.044 (1.99)	0.31	0.005 (0.16)	0.07
$\gamma$	-0.270 (-12.55)	0.23	-0.226 (-5.44)	0.10	-0.311 (-14.59)	0.33	-0.268 (-8.28)	0.13
$\delta_1$	0.012 (3.05)	0.13	-0.008 (-1.26)	0.13	0.009 (2.27)	0.16	0.0005 (0.13)	0.08
$\delta_2$	0.005 (2.56)	0.40	0.010 (3.05)	0.09	0.003 (1.69)	0.25	0.005 (2.70)	0.33
Дисперсия								
$\mu$	0.060 (4.54)	0.22	0.139 (5.72)	0.54	0.099 (3.60)	0.20	0.125 (3.50)	0.06
$\lambda_1$	0.212 (24.04)	0.02	0.184 (11.31)	0.33	0.193 (24.12)	0.03	0.204 (21.20)	0.03
$\nu_1$	1.006 (28.24)	0.10	0.676 (19.38)	0.54	0.977 (10.01)	0.11	0.683 (3.77)	0.05
$\lambda_2$	-0.143 (-10.98)	0.02			-0.112 (-5.55)	0.04	-0.064 (-1.69)	0.03
$\nu_2$	-0.135 (-5.35)	0.12			-0.156 (-2.71)	0.11	0.051 (0.44)	0.06
$\delta_3$	0.005 (3.83)	0.24	0.013 (3.04)	0.21	0.002 (1.42)	0.08	0.008 (3.29)	0.07
$\rho$	0.223 (24.30)	1.07	0.277 (20.16)	0.07	0.202 (24.22)	0.42	0.216 (22.83)	0.46
Диагностические статистики								
$Q(15)$	30.99		19.95		15.10		24.93	
$Q^2(15)$	25.69		26.26		27.13		29.32	
Nyblom	3.14		1.85		2.18		2.22	

Таблица 7. SA-GARCH Модели (продолжение)

	Юкос		Ростелеком		РАО ЕЭС	
	Коэф.	Nyb	Коэф.	Nyb	Коэф.	Nyb
Среднее						
$\alpha$	-0.018 (-2.95)	0.08	-0.003 (-0.72)	0.09	-0.006 (-1.15)	0.24
$\beta$	0.018 (0.60)	0.04	0.026 (0.84)	0.06	-0.033 (-2.25)	0.17
$\gamma$	-0.302 (-10.13)	0.04	-0.260 (-8.50)	0.09	-0.342 (-24.50)	0.15
$\delta_1$	0.013 (2.78)	0.04	0.001 (0.28)	0.06	0.005 (1.32)	0.27
$\delta_2$	0.004 (1.74)	0.18	0.004 (2.89)	0.62	-0.004 (-2.29)	0.17
Дисперсия						
$\mu$	0.224 (8.62)	0.15	0.048 (3.53)	0.09	0.010 (6.28)	0.18
$\lambda_1$	0.189 (14.55)	0.04	0.201 (23.45)	0.03	0.191 (24.61)	0.07
$\nu_1$	0.586 (17.29)	0.15	1.126 (13.86)	0.05	1.447 (73.75)	0.10
$\lambda_2$			-0.140 (-8.77)	0.03	-0.170 (-19.88)	0.07
$\nu_2$			-0.236 (-4.13)	0.05	-0.481 (-23.84)	0.10
$\delta_3$	0.013 (3.74)	0.30	0.004 (3.06)	0.18	0.001 (2.86)	0.10
$\rho$	0.203 (16.92)	0.17	0.238 (24.90)	0.03	0.102 (7.60)	0.03
Диагностические статистики						
$Q(15)$	19.80		15.86		62.74	
$Q^2(15)$	23.07		13.89		26.38	
Nyblom	1.63		2.10		2.00	