

Прикладная эконометрика, 2017, т. 48, с. 63–84.
Applied Econometrics, 2017, v. 48, pp. 63–84.

А. Д. Аганин¹

Сравнение GARCH и HAR-RV моделей для прогноза реализованной волатильности на российском рынке

В работе выполняется множественное сравнение большого количества моделей GARCH, ARFIMA и HAR-RV семейств на данных по качеству одношагового прогноза реализованной волатильности на один день вперед. Сравнение проводится при помощи MCS теста на 10 российских биржевых активах, включая 8 акций и 2 биржевых индекса. Результаты говорят о явном преимуществе моделей семейства HAR-RV над семействами GARCH и ARFIMA и подтверждают обнаруженные в литературе результаты при сравнении отдельных представителей данных семейств.

Ключевые слова: GARCH; реализованная волатильность; HAR-RV; MCS тест.

JEL classification: C52; C53; C58.

Введение

Волатильность является важным финансовым показателем на финансовых рынках. Обычно под волатильностью принято понимать величину, описывающую изменчивость цены актива. Хотя волатильность не может быть рассчитана напрямую, существуют разнообразные методы ее оценки. Для описания свойств и предсказания волатильности существует множество различных моделей, разрабатываемых на протяжении уже более чем сорока лет, и много работ посвящено исследованию их свойств. Одним из наиболее популярных на протяжении долгого времени является семейство GARCH моделей. Основной особенностью этих моделей является то, что они рассматривают волатильность как ненаблюдаемую величину. Одновременно GARCH модели способны моделировать некоторые важные эмпирические особенности волатильности, чем и объясняется их популярность. Помимо семейства GARCH моделей, в литературе предлагается использование других семейств, таких как ARFIMA, MIDAS, модель Хестона и др. В качестве оценки волатильности и качества прогнозной силы в литературе и на практике широко используется «реализованная волатильность», которая получила широкое применение с появлением высокочастотных внутридневных данных котировок цен. Согласно свойствам этой оценки, реализованная волатильность сходится к ненаблюдаемой волатильности при высокой частоте данных, т. е. является состоятельной оценкой ненаблюдаемой волатильности.

¹ Аганин Артем Давидович — Национальный исследовательский университет «Высшая школа экономики», Москва; aaganin@hse.ru.

После того как Corsi (2009) предложил новое семейство HAR-RV моделей, появилось много работ с различными модификациями этой модели, а также сравнениями моделей этого семейства с уже существующими по их прогнозной силе. В литературе можно найти множественные сравнения большого числа GARCH моделей, однако сравнение GARCH и HAR-RV моделей не выполнялось. Среди работ, посвященных сравнению и выбору наилучших моделей с наибольшей прогнозной силой, основное внимание уделяется сравнению небольшого числа моделей из разных семейств, либо сравнению моделей из одного семейства, на основе информационных критериев или R^2 статистиках. Несмотря на то что сравнение GARCH моделей выполнялось многократно, большая прогнозная сила HAR-RV моделей по сравнению с GARCH моделями остается под вопросом. Некоторые исследования (например (Mastro, 2014)) обнаружили, что простейшая HAR-RV модель превосходит модели GARCH(1,1) и EGARCH(1,1).

Целью настоящей работы является сравнение разных семейств моделей волатильности и проверка большей прогнозной силы HAR-RV моделей по сравнению с GARCH и ARFIMA моделями. Сравняется большое число широко используемых GARCH моделей с распространенным в использовании набором HAR-RV моделей и несколькими ARFIMA моделями. В качестве критерия используется реализованная волатильность. Сравнение моделей выполняется на «out-of-sample» данных, т. е. по имеющимся данным до дня прогноза, и фактическое значение реализованной волатильности, как и другая информация о дне прогноза, являются неизвестными для моделей в момент построения прогноза. В качестве данных были выбраны 10 российских биржевых активов. Сравнение выполняется при помощи MCS теста, предложенного в (Hansen et al., 2011). В результате сравнения HAR-RV модели с логарифмическим преобразованием имеют статистически большую прогнозную силу и превосходят семейства GARCH и ARFIMA моделей по прогнозу значений реализованной волатильности.

1. Сравнимые модели

1.1. Семейство GARCH моделей

Волатильность — один из важнейших показателей, характеризующий разброс временного ряда. Известно, что финансовые ряды на практике имеют непостоянную волатильность, что нарушает условия моделей Блэка–Шоулза и ARIMA, которые часто используются для оценивания опционов и прогнозирования значений финансовых рядов. Поскольку эти и другие модели исходят из предположений о постоянной волатильности, которые нарушаются на практике, для изучения переменной волатильности необходимы специальные модели.

Эмпирические исследования свойств финансовых рядов показали, что волатильности свойственны такие свойства, как кластеризация («группирование» периодов с высокой и низкой волатильностью), а также так называемый эффект «рычага» — уменьшение волатильности при росте цены, и наоборот.

Одним из первых семейств моделей, предназначенных для изучения рядов волатильности, стало семейство GARCH моделей, первая из которых была представлена в (Engle, 1982) и получила название ARCH (AutoRegressive Conditional Heteroscedasticity). Основным достоинством этой модели был учет гетероскедастичности волатильности. В названии

модели «AR» означает авторегрессию в квадратах доходностей, а «CH» — моделирование волатильности как условной дисперсии в момент t при известной информации только до момента $t-1$ включительно. Фактически в ARCH модели предполагается, что дисперсия текущих «инноваций» (нововведений на финансовом рынке) является функцией от значений инноваций в предыдущие моменты времени. ARCH модель можно описать следующим образом. Стандартно для моделей GARCH семейства определяется два уравнения: условного математического ожидания доходности и процесс условной дисперсии. Обозначим через σ_t дисперсию актива в момент t , а его доходность в этот момент как r_t . Тогда доходность финансового актива можно представить в виде процесса

$$r_t = \mu + \varepsilon_t, \quad (1)$$

где $\varepsilon_t = \sigma_t z_t$ — инновации (остатки) и $z_t \sim \text{i.i.d. } \mathcal{N}(0,1)$. Условная дисперсия ARCH(1) модели может быть записана как

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2. \quad (2)$$

На коэффициенты ARCH модели накладываются следующие ограничения: $\omega > 0$, $\alpha_1 > 0$ для положительности значений дисперсии, а также $\alpha_1 < 1$ для стационарности ряда (2). Безусловная волатильность такого процесса может быть рассчитана из уравнения (2) взятием математического ожидания, и она равняется $\frac{\omega}{1 - \alpha_1}$.

Можно показать, что безусловное распределение остатков ε в ARCH модели имеет более толстые хвосты по сравнению с нормальным распределением, т.е. обладает большим куртозисом (показателем, отражающим остроту вершины и толщину хвостов одномерных распределений, и равным 3 для нормального распределения). Куртозис распределения ε определяется как

$$\text{Kurt}(\varepsilon_t) = \frac{\mathbb{E}(\varepsilon_t^4)}{(\mathbb{E}(\varepsilon_t^2))^2} = \frac{\mathbb{E}(\sigma_t^4)\mathbb{E}(z_t^4)}{(\mathbb{E}(\sigma_t^2))^2(\mathbb{E}(z_t^2))^2} = \frac{3\mathbb{E}(\sigma_t^4)}{(\mathbb{E}(\sigma_t^2))^2}. \quad (3)$$

Поскольку z_t нормально распределена (с куртозисом, равным 3), то из формулы (3) с учетом неравенства Йенсена ($\mathbb{E}(f(X)) \geq f(\mathbb{E}(X))$ для выпуклой функции $f(x)$) следует, что $\text{Kurt}(\varepsilon_t) \geq 3$, а это говорит о «толстых хвостах» распределения остатков.

Bollerslev (1986) предложил обобщенную GARCH(p, q) модель, которая предлагает более элегантный способ представления ARCH модели с большим числом лагов:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2. \quad (4)$$

На параметры модели (4) налагается ограничение $\omega > 0$, а также условие стационарности $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$.

Можно заметить, что GARCH модель фактически является ARMA моделью в квадратах остатков. Для этого запишем GARCH(1,1) модель как

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \quad (5)$$

Условный прогноз ошибки (разность между квадратом остатков доходностей ε_t и его условным математическим ожиданием) обозначим как $v_t = \varepsilon_t^2 - E_{t-1}(\varepsilon_t^2) = \varepsilon_t^2 - \sigma_t^2$, где $E_{t-1}(\varepsilon_t)$ является математическим ожиданием ε_t , основанным на доступной информации в момент $t-1$. Это позволяет переписать GARCH(1,1) модель как

$$\begin{aligned}\varepsilon_t^2 - v_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 (\alpha_{t-1}^2 - v_{t-1}), \\ \varepsilon_t^2 &= \alpha_0 + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + v_t - \beta_1 v_{t-1},\end{aligned}\quad (6)$$

что является уравнением ARMA(1,1) модели относительно квадратов остатков. Из этого следует, что GARCH модели обладают свойствами ARMA моделей.

После введения классической GARCH модели было предложено множество ее модификаций, которые позволяют учесть дополнительные эффекты, наблюдаемые на эмпирических данных, такие как асимметрическая реакция на прошлые изменения волатильности, эффект «рычага», замеченный в (Black, 1976) и др. Наиболее популярными моделями GARCH семейства являются следующие.

1. GARCH (иногда называемая SGARCH): $\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$.
2. IGARCH: $\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i (\varepsilon_{t-i}^2 - \varepsilon_{t-1}^2) + \sum_{j=1}^p \beta_j (\sigma_{t-j}^2 - \varepsilon_{t-1}^2)$.
3. NAGARCH: $\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i (\varepsilon_{t-i} + \gamma_i \sigma_{t-i})^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$.
4. Thr.-GARCH: $\sigma_t = \omega + \sum_{i=1}^q \alpha_i [(1 - \gamma_i) \varepsilon_{t-i}^+ - (1 + \gamma_i) \sigma_{t-i}^-]^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}$.
5. GJR-GARCH: $\sigma_t^2 = \omega + \sum_{i=1}^q (\alpha_i + \gamma_i I_{(\varepsilon_{t-i} > 0)}) \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$.
6. Taylor/Schwert: $\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i |\varepsilon_{t-i}| + \sum_{j=1}^p \beta_j \sigma_{t-j}$.
7. EGARCH: $\log(\sigma_t^2) = \omega + \sum_{i=1}^q [\alpha_i \varepsilon_{t-i} + \gamma_i (|\varepsilon_{t-j}| - E|\varepsilon_{t-i}|)] + \sum_{j=1}^p \beta_j \log(\sigma_{t-j}^2)$.
8. NGARCH: $\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i |\varepsilon_{t-i}|^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta$.
9. A-PARCH: $\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i [|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i}]^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta$.
10. CSGARCH: $q_t = \omega + \rho q_t - 1 + \varphi(\varepsilon_{t-1}^2 - \sigma_{t-1}^2)$,
 $\sigma_t^2 = q_t + \sum_{i=1}^n \alpha_i (\varepsilon_{t-i}^2 - q_{t-i}) + \sum_{j=1}^p \beta_j (\sigma_{t-j}^2 - q_{t-j})^2$.
11. H-GARCH или ALLGARCH: $\sigma_t^\delta = \omega + \sum_{i=1}^n \alpha_i \delta \sigma_{t-j}^\delta [|\varepsilon_t - k| - \tau |\varepsilon_t - k|]^\delta + \sum_{j=1}^n \beta_j \sigma_{t-j}^2$.

Множество исследований показывают негативный вклад увеличения количества лагов в качество прогноза GARCH моделей. Поэтому обычно используются спецификации с $p \leq 2$ и $q \leq 2$. Следует отметить, что GARCH(1,1) модель является наиболее популярной и часто используемой среди остальных.

Для оценивания параметров GARCH моделей по данным доходностей актива общепринятым является метод максимального правдоподобия, а также его модификация — метод

квази-правдоподобия. Несмотря на их повсеместное применение, эти методы обладают существенными недостатками: их сходимость не гарантирована, а оптимизация для поиска параметров может занимать длительное время. Основным достоинством этих процедур является состоятельность полученных оценок.

В качестве процесса доходностей актива (1) часто также используется ARMA процесс с малым количеством лагов. Выбору лаговой структуры этого ARMA процесса обычно не уделяется много времени, поскольку он слабо влияет на прогноз волатильности. Помимо общепринятого предположения о нормальности ошибок, в (1) было предложено множество других распределений, которые могут лучше описывать некоторые наблюдаемые особенности финансовых данных, такие как асимметрия и толстые хвосты распределения остатков. Среди наиболее используемых можно выделить следующие четыре.

1. Нормальное распределение с плотностью

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (7)$$

2. Скошенное t -распределение Стьюдента. Пусть x_t имеет t -распределение, ν — число степеней свободы. Обозначим

$$\varepsilon_t = \frac{x_t}{\sqrt{\nu / (\nu - 2)}} \quad (8)$$

и его стандартизированной функцию плотности как

$$f(\varepsilon | \nu) = \frac{\Gamma((\nu + 1)/2)}{\sqrt{(\nu - 2)\pi}\Gamma(\nu/2)} \left(1 + \frac{\varepsilon^2}{\nu - 2}\right)^{-(\nu+1)/2}, \quad (9)$$

где $\Gamma(x)$ — гамма-функция, ν — число степеней свободы t -распределения. Lambert, Laurent (2001) предложили использовать подход (Fernandez, Steel, 1998) к стандартизированному t -распределению и из него получить стандартизированное скошенное t -распределение Стьюдента с учетом как скошенности, так и эксцесса стандартизированных остатков:

$$g(\varepsilon | \xi, \nu) = \begin{cases} \frac{2}{\xi + 1/\xi} \rho f(\xi(\rho\varepsilon + \bar{\omega}) | \nu), & \text{если } \varepsilon < -\bar{\omega} / \rho, \\ \frac{2}{\xi + 1/\xi} \rho f((\rho\varepsilon + \bar{\omega}) / \xi | \nu), & \text{если } \varepsilon > -\bar{\omega} / \rho. \end{cases} \quad (10)$$

В этом представлении ξ есть параметр скошенности, а ρ и $\bar{\omega}$ можно рассчитать как

$$\bar{\omega} = \frac{\Gamma((\nu - 1)/2)\sqrt{\nu - 2}}{\sqrt{\pi}\Gamma(\nu/2)} (\xi - 1/\xi), \quad \rho^2 = (\xi^2 + 1/\xi^2 - 1) - (\bar{\omega})^2.$$

3. Нормальное обратное гауссовское распределение. Плотность вероятности случайной величины X с обратным гауссовским распределением и вектором параметров $\theta = (\alpha, \beta, \delta, \mu)$ можно записать как

$$f(x; \theta) = \frac{\alpha \delta}{\pi} e^{\delta \gamma + \beta(x - \mu)} \frac{\kappa_1 \left(\alpha \sqrt{\delta^2 + (x - \mu)^2} \right)}{\sqrt{\delta^2 + (x - \mu)^2}}, \quad (11)$$

где $x \in R$, $\alpha > 0$, $\beta \in (-\alpha; \alpha)$, $\delta \in (0, \infty)$, $\gamma = \sqrt{\alpha^2 - \beta^2}$, а κ_1 обозначает модифицированную функцию Бесселя третьего рода с индексом 1: $\kappa_1(\omega) = \frac{1}{2} \int_0^\infty \exp\{-(\omega(t + t^{-1}))/2\} dt$.

4. Johnson's reparametrized SU распределение. Первоначальное Johnson's SU распределение было предложено Johnson (1949). Репараметризованная версия этого распределения имеет больший, по сравнению с нормальным, куртозис и выводится из Johnson's SU распределения с теми же математическим ожиданием и дисперсией, как и первоначальное распределение. Функцию плотности вероятности такого распределения можно записать в виде следующих выражений:

$$f(y | \mu, \sigma, \nu, \tau) = \frac{\tau}{c\sigma} \cdot \frac{1}{\sqrt{r^2 + 1}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad (12)$$

$$z = -\nu + \tau / \sinh r, \quad (13)$$

$$r = \frac{y - (\mu + c\sigma\sqrt{\omega} \sinh \Omega)}{c\omega}, \quad (14)$$

$$c = \sqrt{2/((\omega - 1)[\omega \cosh(2\Omega) + 1])}. \quad (15)$$

Здесь $\tau > 0$ — куртозис этого распределения, ν — коэффициент скошенности, ω и Ω определяются как $\omega = \exp\{r^{-2}\}$ и $\Omega = \nu/\tau$ соответственно. Если $\tau \rightarrow \infty$, то это распределение совпадает с нормальным.

Как можно заметить, GARCH модели рассматривают волатильность как ненаблюдаемую величину и поэтому требуют полностью определенного условного математического ожидания и условной вероятности для расчета латентной волатильности. Несмотря на такое представление, GARCH модели неспособны воспроизвести важные наблюдаемые особенности волатильности финансовых данных. В частности, GARCH модели не обладают долгой памятью (автокорреляция уменьшается с экспоненциальной скоростью), а использование сложных оптимизационных процедур оценивания требует больших затрат времени.

1.2. Семейство HAR-RV моделей

Первая HAR-RV (The Heterogeneous Autoregressive model of the Realized Volatility) модель была предложена F. Corsi в 2009 году (Corsi, 2009). Эта модель основана на концепции поведения агентов на финансовых рынках, согласно которой они отличаются своим восприятием волатильности в зависимости от их инвестиционных горизонтов и делятся на краткосрочных, среднесрочных и долгосрочных. Гипотеза о существовании таких гетерогенных структур на финансовых рынках (гипотеза о гетерогенности рынков) была сформулирована в (Müller et al., 1997) следующим образом.

1. Разные агенты на гетерогенном рынке обладают разными сроками инвестирования и участвуют в торговле на бирже с разной частотой. Например, к участникам высокочастотных торгов (краткосрочным агентам) относятся Fogex дилеры и маркетмейкеры (чья позиция открывается и закрывается в течение дня), к долгосрочным агентам можно отнести центральные банки и некоторые фонды. Каждая отдельная группа таких агентов (компонент гетерогенного рынка) реагирует на новости, поступающие на рынок, с разной скоростью в зависимости от ее горизонта инвестирования. Если предположить, что память каждого компонента (т. е. скорость убывания значений автокорреляционной функции с ростом лага) является экспоненциально убывающей с определенной временной константой (что верно для GARCH(1,1) модели), то память всего рынка будет состоять из множества таких экспоненциально убывающих составляющих с разнообразными значениями временных констант. При этом их суперпозиция будет вести себя как процесс с гиперболически убывающей автокорреляцией.

2. В то же время, чем больше агентов присутствует на гомогенном рынке, тем быстрее оценки стоимости некоторого актива разных агентов сходятся к «реальной рыночной цене», относительно которой все агенты имеют рациональные ожидания. То есть в случае однородного рынка должна наблюдаться отрицательная корреляция волатильности с рыночной активностью, что не наблюдается в действительности. Однако на гетерогенном рынке агенты будут надеяться на разные уровни цены и будут принимать разные решения о покупке/продаже активов в различных ситуациях, что приведет к росту волатильности с увеличением числа агентов на рынке.

3. Рынки также отличаются по географическому положению своих агентов. Так, была замечена низкая корреляция между внутрисуточными ценами с лагом 1.5 суток, что соответствует агентам в противоположных полушариях, и высокая корреляция между ценами с лагом в 1–2 суток, что соответствует одним и тем же агентам.

Согласно гипотезе о гетерогенности рынка, рыночные агенты отличаются также по другим характеристикам, таким как склонность к риску, институциональные ограничения и транзакционные издержки.

В результате агенты с разными временными горизонтами будут анализировать рынок с разной частотой и поэтому создавать (либо оказывать влияние на) разные виды волатильности. Экономическая интерпретация долгосрочной волатильности σ_m заключается в наличии неопределенности фундаментальных макроэкономических факторов, среднесрочная компонента волатильности σ_w отражает текущую неопределенность на рынке, связанную с обработкой поступающих на рынок новостей, а краткосрочная волатильность σ_d включает в себя неопределенность, связанную с возникновением спекулятивных моментов.

Семейство HAR-RV моделей основывается на гипотезе о гетерогенности рынка. Модели этого семейства способны воспроизводить наблюдаемые свойства волатильности, такие как: «длинная память» и устойчивость волатильности, «каскад волатильности» (иерархичная структура волатильности, при которой долгосрочная волатильность оказывает сильное влияние на краткосрочную волатильность, однако краткосрочная волатильность почти не влияет на долгосрочную), «толстые хвосты», а также имеют экономическую интерпретацию. Хотя сами по себе они не являются моделями с длинной памятью, а представляют собой сумму AR процессов, однако в работе (LeBaron, 2001) было показано наличие эффекта долгой памяти даже для суммы трех AR(1) процессов, что также подтверждается эмпирическими результатами расчета корреляционной функции для HAR-RV моделей. Еще одной

особенностью HAR-RV модели является рассмотрение волатильности как ненаблюдаемой величины, для которой существует наблюдаемая оценка, которую можно оценить при помощи высокочастотных данных. Эта оценка называется реализованной волатильностью и обозначается RV .

Простейшую HAR-RV модель можно сформулировать следующим образом. Пусть $p_{t,j}$ — цена актива в день t на конец внутрисуточного интервала j длины Δ (секунд), $j = 1, \dots, N$ с общим числом таких интервалов за один день, равным N . Тогда $r_{t,j} = \log(p_{t,j}) - \log(p_{t,j-1})$ — внутрисуточная доходность актива на временном интервале длины Δ .

Реализованную волатильность в день t можно записать как

$$RV_t = \sum_{j=1}^N r_{t,j}^2. \quad (16)$$

Такая форма RV является оценкой дневной волатильности. Использование высокочастотных данных о ценах позволяет моделировать волатильность как некоторую виртуальную, но обозримую величину, в отличие от моделирования ненаблюдаемой латентной переменной, как это делается в ранее рассмотренных GARCH моделях (Engle, 1982; Bollerslev, 1986) или латентной стохастической модели волатильности (например (Taylor, 1994)). Другой важной особенностью HAR-RV модели является ее гетерогенная структура, которая позволяет учесть рассмотренную выше гетерогенность рынка. HAR-RV модель представляет волатильность как сумму дневного, недельного и месячного компонентов. Обычно HAR-RV модель рассчитывается с параметрами (1,5,21), что соответствует лагам дневной, недельной и месячной RV :

$$RV_{t+1} = \beta_0 + \beta_d RV_t^d + \beta_w RV_t^w + \beta_m RV_t^m + \varepsilon_{t+1}, \quad (17)$$

где ε_{t+1} — случайный шок, RV_t^d — реализованная волатильность в день t , RV_t^w — лагированная недельная волатильность за предыдущую неделю, определяемая как $RV_t^w = \frac{1}{4} \sum_{i=1}^4 RV_{t-i}$, и RV_t^m — месячная волатильность, определяемая похожим образом как $RV_t^m = \frac{1}{17} \sum_{i=5}^{21} RV_{t-i}$.

Вместо параметров (1,5,21) можно использовать любой другой набор лагов в правой части уравнения (17). Однако в работе (Andersen et al., 2007) в результате сравнения моделей с разными параметрами было показано, что дневная, недельная и месячная компоненты наиболее точно описывают гетерогенную структуру и являются оптимальным набором параметров. В последнее время во многих сравнительных работах авторы рассматривали HAR-RV модель как возможно наилучшую модель при анализе и моделировании RV . Такими, например, статья (Souček, Todorova, 2014), посвященная исследованию с помощью HAR-RV модели роли скачков в эффекте увеличения волатильности одного актива при увеличении волатильности другого для индекса S&P500, а также работа (Lahaie, Shaw, 2014), в которой для моделирования индекса S&P500 оказалось достаточно линейной спецификации HAR-RV модели.

Спецификация лагов HAR-RV модели была протестирована в работе (Craioveanu, Hillebrand, 2012), где также подтвердилось преимущество HAR-RV(1,5,21) модели. Относительно частоты внутрисуточных данных, используемых при построении RV , Liu et al. (2013) обратили особое внимание на оценку RV , построенную по данным пятиминутных интервалов.

Это важное для настоящей статьи исследование показало, что из 400 оценок волатильности, построенных для 31 актива из пяти разных классов активов, ни одна оценка не может превзойти RV , построенную по пятиминутным интервалам, при прогнозировании будущей волатильности при помощи HAR-RV модели. В соответствии с этим результатом в настоящей статье реализованная волатильность также рассчитывалась по данным цен закрытия пятиминутных интервалов.

Поскольку реализованная волатильность имеет логнормальное распределение, то естественно перейти к HAR-RV модели с логарифмами:

$$\log(RV_{t+1}) = \beta_0 + \beta_d \log(RV_t^d) + \beta_w \log(RV_t^w) + \beta_m \log(RV_t^m) + \varepsilon_{t+1}. \quad (18)$$

Другой популярный подход заключается в моделировании стандартных отклонений, а не самих реализованных волатильностей, в результате чего получается нелинейная модель вида

$$\sqrt{RV_{t+1}} = \beta_0 + \beta_d \sqrt{RV_t^d} + \beta_w \sqrt{RV_t^w} + \beta_m \sqrt{RV_t^m} + \varepsilon_{t+1}. \quad (19)$$

Позже было предложено множество модификаций базовой HAR-RV модели, учитывающих дополнительные свойства RV . Одной из первых стала LHAR-RV-J модель, отдельно рассматривающая скачки. Рассмотрим построение этого компонента.

Общим подходом к расчету компоненты скачков является подход, основанный на работе (Barndorff-Nielsen, Shephard, 2004), который формулируется следующим образом.

Пусть, как и прежде Δ — период (длина внутрисуточного интервала) внутрисуточной доходности. Как было сказано выше, реализованная волатильность получается суммированием соответствующих N внутрисуточных доходностей:

$$RV_{t+1}(\Delta) = \sum_{j=1}^N r_{t+1,j}^2. \quad (20)$$

Из свойств квадратической вариации (см., например, (Andersen et al., 2003)) следует, что реализованная волатильность равномерно (по t из любого конечного интервала $[0, T]$) сходится по вероятности к приращению квадратической вариации процесса доходности с уменьшением длины внутрисуточного интервала Δ , используемого для расчета RV . Это одна из главных причин использования реализованной волатильности в качестве критерия прогнозирования при сравнении моделей волатильности. Так, если предположить наличие дискретных скачков в RV , то получим, что при $\Delta \rightarrow 0$

$$RV_{t+1}(\Delta) \rightarrow \int_t^{t+1} \sigma^2(s) ds + \sum_{t < s \leq t+1} \kappa(s), \quad (21)$$

где $\kappa(s)$ — величина скачка, а суммирование идет по скачкам в $t+1$ день (по всем интервалам внутри дня, на которых были скачки). Если бы скачки в рядах RV отсутствовали, то реализованная волатильность была бы состоятельной оценкой интегрированной волатильности, широко используемой в научной литературе по стохастической волатильности и оцениванию опционов. Однако реализованная волатильность часто включает в себя и непрерывный процесс, и процесс скачков. Из этого можно сделать вывод, что модель с наибольшей прогнозной силой должна моделировать эти компоненты по отдельности, в отличие от базовой

HAR-RV модели (17), в пользу чего говорит исследование (Corsi et al., 2010). На практике было предложено использовать оценку bipower variation, рассчитанную как

$$BV_{t+1}(\Delta) = \frac{\pi}{2} \sum_{j=2}^N |r_{t,j}| |r_{t,j-1}|, \quad (22)$$

для которой можно показать, что при $\Delta \rightarrow 0$ она будет сходиться к интегрированной волатильности:

$$BV_{t+1}(\Delta) \rightarrow \int_t^{t+1} \sigma^2(s) ds. \quad (23)$$

Объединяя результаты (21) и (23), можно сделать вывод, что вклад скачков, вносимый в процесс квадратической вариации, может быть состоятельно оценен (при $\Delta \rightarrow 0$) как

$$RV_{t+1}(\Delta) - BV_{t+1}(\Delta) \rightarrow \sum_{t < s \leq t+1} \kappa^2(s). \quad (24)$$

Это основная идея расчета скачков. Barndorff-Nielsen, Shephard (2004) предложили при расчетах ограничивать разность в левой части (24) нулем, чтобы обойти ситуации с отрицательным значением для квадрата скачка. Формально это можно записать так:

$$J_{t+1} = \max \{ RV_{t+1}(\Delta) - BV_{t+1}(\Delta), 0 \}. \quad (25)$$

Первой модификацией, рассматривающей скачки как отдельный компонент, стала HAR-RV-J модель, включающая значение скачка за предыдущий день:

$$RV_{t+1} = \beta_0 + \beta_D RV_t^d + \beta_W RV_t^w + \beta_M RV_t^m + \beta_J J_t^d + \varepsilon_{t+1}, \quad (26)$$

а также ее лог-модификация

$$\log RV_{t+1} = \beta_0 + \beta_D \log RV_t^d + \beta_W \log RV_t^w + \beta_M \log RV_t^m + \beta_J \log(J_t^d + 1) + \varepsilon_{t+1}. \quad (27)$$

Рассчитав скачок, его можно включать не только как значение в момент t (как это делается в HAR-RV-J модели), но и представлять в виде той же гетерогенной структуры, которая используется при разбиении RV на гетерогенные компоненты, аналогично поступая и с непрерывной частью реализованной волатильности. Полученная модель получила название HAR-RV-CJ модели и может быть записана в виде

$$RV_{t+1} = \beta_0 + \beta_{CD} C_t^d + \beta_{CW} C_t^w + \beta_{CM} C_t^m + \beta_{JD} J_t^d + \beta_{JW} J_t^w + \beta_{JM} J_t^m + \varepsilon_{t+1}, \quad (28)$$

где C соответствует непрерывному компоненту RV и оценивается как bipower variation по формуле (22).

Помимо скачка, в литературе было предложено включать компонент «рычага» для учета соответствующего эффекта, объясненного ранее. В качестве оценки для дня t предлагается использовать минимум между дневной доходностью и нулем: $r_t^- = \min(r_t, 0)$. Это позволяет отдельно оценить эффект от увеличения волатильности в случае отрицательных доходностей.

В результате для учета эффекта «рычага» были предложены LHAR-RV-J модель и ее лог-версия:

$$RV_{t+1} = \beta_0 + \beta_D RV_t^d + \beta_W RV_t^w + \beta_M RV_t^m + \beta_J J_t^d + \gamma^d r_t^- + \varepsilon_{t+1}, \quad (29)$$

$$\log RV_{t+1} = \beta_0 + \beta_D \log RV_t^d + \beta_W \log RV_t^w + \beta_M \log RV_t^m + \beta_J + \log(J_t^d + 1) + \gamma^d r_t^- + \varepsilon_{t+1}, \quad (30)$$

а также LHAR-RV-CJ модель и ее лог-модификации:

$$RV_{t+1} = \beta_0 + \beta_{CD} C_t^d + \beta_{CW} C_t^w + \beta_{CM} C_t^m + \beta_{JD} J_t^d + \beta_{JW} J_t^w + \beta_{JM} J_t^m + \gamma^d r_t^{d-} + \gamma^w r_t^{w-} + \gamma^m r_t^{m-} + \varepsilon_{t+1}, \quad (31)$$

$$\log RV_{t+1} = \beta_0 + \beta_{CD} \log C_t^d + \beta_{CW} \log C_t^w + \beta_{CM} \log C_t^m + \beta_{JD} \log(J_t^d + 1) + \beta_{JW} \log(J_t^w + 1) + \beta_{JM} \log(J_t^m + 1) + \gamma^{d-} r_t^{d-} + \gamma^{w-} r_t^{w-} + \gamma^{m-} r_t^{m-} + \varepsilon_{t+1}. \quad (32)$$

Эти и другие варианты модели с эффектом рычага были рассмотрены в (Corsi, Reno, 2012). В целом, семейство HAR-RV моделей показывает хорошие результаты при моделировании волатильности (по информационным критериям и R^2), а простота применения этих моделей и понятная экономическая интерпретация коэффициентов служат катализаторами роста их популярности.

1.3. ARFIMA модели

Еще одной моделью, получившей распространение для моделирования волатильности после появления оценок реализованной волатильности, является ARFIMA(p, d, q) модель, которая является обобщением ARIMA модели. Ее главным отличием является возможность использования действительного показателя интегрированности, принимающего значения от -0.5 до 0.5 . Основным ее преимуществом является возможность моделирования долгосрочных зависимостей, т. е. она обладает долгой памятью, в отличие от обычной ARIMA модели с короткой памятью.

ARFIMA(p, d, q) модель можно записать как

$$\Phi(L)(1-L)^d (RV_t - \mu) = \Theta(L)\varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } (0, \sigma_\varepsilon^2),$$

где L — лаговый оператор, $\Phi(L) = 1 - \varphi_1 L - \dots - \varphi_p L^p$, $\Theta(L) = 1 + \nu_1 L + \dots + \nu_q L^q$ и $(1-L)^d$ — оператор дробного интегрирования, определяемый как

$$(1-L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)L^k}{\Gamma(-d)\Gamma(k+1)}.$$

В литературе по примеру Andersen et al. (2003) в качестве параметров p, d, q принято использовать наборы $(1, 0.401, 0)$ и $(5, 0.401, 0)$. Использование данной модели позволяет напрямую моделировать реализованную волатильность и учитывать долгосрочные зависимости, которые не способны учесть GARCH и HAR-RV модели.

2. Методика сравнения

К настоящему времени в литературе предложено множество тестов для сравнения набора конкурирующих моделей. Основным недостатком всех этих тестов является зависимость результата от особенностей природы используемых данных, что на практике может привести к разным результатам сравнения для разных активов. Большинство существующих тестов предназначены для сравнения небольшого числа моделей, поэтому неуместны при сравнении двух больших семейств.

Для сравнения больших наборов моделей можно выделить два подхода. Первый из них опирается на использование тестов, предложенных Hansen. Первым тестом, получившим большую популярность при сравнении большого набора моделей, стал тест из (Hansen, Lunde, 2004) под названием Superior Predictive Ability (SPA). Будучи модификацией теста White (2000), он использовался для сравнения представителей GARCH моделей, обладая при этом большей робастностью и мощностью, что и было продемонстрировано в (Hansen, Lunde, 2004). SPA тест оставался популярным некоторое время, поскольку мог использоваться для сравнения наборов моделей и даже стратегий, не связанных с волатильностью. Основное достоинство и одновременно недостаток этого теста заключается в возможности сравнить одну модель (бенчмарк) с набором конкурирующих и выявить среди них наличие либо отсутствие моделей, превосходящих бенчмарк. При этом возможность одновременно существования нескольких одинаковых по прогнозной силе моделей не рассматривается, и выявить множество таких моделей нельзя.

Позже Hansen et al. (2011) предложили тест Model Comparison Set (MCS), имеющий более широкое применение, допускающий существование набора нескольких наилучших моделей с одинаковой прогнозной силой (или с одинаковой производительностью по какому-то другому критерию). Другим часто используемым подходом являются тесты Value at Risk (VaR), основанные на понятии VaR. Эти тесты требовательны к размеру периода прогнозирования, размеру базового периода и производительности компьютера, чаще всего применяются для оптимизации портфелей, поэтому проведение VaR тестов для данной задачи сравнения моделей можно считать нецелесообразным и излишне времязатратным.

Тест Model Confidence Set

Для описанных ранее целей Hansen et al. (2011) предложили новый тест — Model Confidence Set (MCS). Этот тест не требует предположения о превосходстве какой-либо одной модели над другими (выбора бенчмарка), допускает возможность существования множества моделей (а не единственной, как в SPA тесте) с одинаковым превосходством над остальными, а также позволяет учитывать несовершенство данных. Под этим учетом понимается расширение искомого множества моделей с включением некоторых дополнительных, не являющихся равными им по заданному критерию, в случае, если имеющейся информации в данных недостаточно для определения истинного набора «равнопроизводительных» моделей.

Перейдем к построению самого теста. Пусть M_0 — множество, включающее конечное число моделей (проиндексированных $i = 1, \dots, m_0$). Выполнение теста начинается с оценивания всех моделей из множества по ряду данных за R дней, предшествующих дню,

для которого необходимо выполнить прогноз. После получения оценок коэффициентов модели делается прогноз на следующий день k , где $k = 1, \dots, n$, а n — общая длина горизонта прогнозирования. Полученный прогноз обозначим через h_k , а фактическую реализованную волатильность в этот день — RV_k . После прогнозирования модель оценивается по данным за R дней, но уже со сдвигом на один день вперед, включая информацию дня k , на который перед этим сделан прогноз. После оценивания моделей по новым данным снова делается прогноз на день $k + 1$ и так далее, пока для каждой модели не будет получен ряд прогнозов для всех k .

Для каждого дня t и каждой модели i обозначим функцию «потерь» через $L_{i,t}$. Эта функция является тем самым критерием, относительно которого выполняется сравнение моделей. Поскольку данная работа посвящена сравнению моделей волатильности, то в качестве критерия можно использовать, например, функции потерь (33). Очевидно, чем ближе прогнозное значение к фактическому, тем лучше выполнен прогноз.

$$\begin{aligned} MSE &= \frac{1}{n} \sum_{t=1}^n (RV_t - h_t)^2, \\ MAE &= \frac{1}{n} \sum_{t=1}^n |RV_t - h_t|, \\ R2LOG &= \frac{1}{n} \sum_{t=1}^n (\log(RV_t/h_t))^2. \end{aligned} \quad (33)$$

Относительную производительность модели i относительно модели j в день t определим как $d_{ij,t} = L_{i,t} - L_{j,t}$ для всех $i, j \in \mathcal{M}_0$. Производительность модели i в день t относительно всех остальных моделей можно определить как

$$d_{i,t} = \frac{1}{m-1} \sum_{j \in \mathcal{M}} d_{ij,t}, \quad i = 1, \dots, m. \quad (34)$$

Поскольку в дальнейшем нас будет интересовать прогнозная сила моделей на всем интервале, то индекс t будет опускаться. Множество моделей с наилучшей производительностью (финальное множество) на всем временном интервале определим как

$$\mathcal{M}^* = \{i \in \mathcal{M}_0 : E(d_{ij}) \leq 0 \text{ для всех } j \in \mathcal{M}_0\}. \quad (35)$$

Формально нулевую и альтернативную гипотезы теста на эквивалентность можно сформулировать двумя способами:

$$\begin{aligned} \mathbf{H}_{0,\mathcal{M}} &: E(d_{ij}) = 0 \text{ для всех } i, j \in \mathcal{M}_0, \\ \mathbf{H}_{A,\mathcal{M}} &: E(d_{ij}) \neq 0 \text{ для некоторых } i, j \in \mathcal{M}_0 \end{aligned} \quad (36)$$

или

$$\begin{aligned} \mathbf{H}_{0,\mathcal{M}} &: E(d_{i,\cdot}) = 0 \text{ для всех } i \in \mathcal{M}_0, \\ \mathbf{H}_{A,\mathcal{M}} &: E(d_{i,\cdot}) \neq 0 \text{ для некоторых } i \in \mathcal{M}_0, \end{aligned} \quad (37)$$

где $E(d_{ij})$ и $E(d_{i,\cdot})$ предполагаются конечными и не зависящими от времени. Hansen et al. (2011) использовали следующие две статистики для тестирования нулевой гипотезы \mathbf{H}_0 против \mathbf{H}_A :

$$t_{ij} = \bar{d}_{ij} / \sqrt{\widehat{\text{Var}}(\bar{d}_{ij})}, \quad t_i = \bar{d}_{i,\cdot} / \sqrt{\widehat{\text{Var}}(\bar{d}_{i,\cdot})} \quad \text{для } i, j \in \mathcal{M}, \quad (38)$$

где \mathcal{M} — тестируемое множество моделей, \bar{d}_{ij} и $\bar{d}_{i,\cdot}$ — относительные потери, рассчитанные по имеющимся данным, $\widehat{\text{Var}}(\bar{d}_{ij})$ и $\widehat{\text{Var}}(\bar{d}_{i,\cdot})$ являются оценками дисперсии истинных дисперсий $\text{Var}(\bar{d}_{ij})$ и $\text{Var}(\bar{d}_{i,\cdot})$ соответственно, полученными при помощи процедуры бутстрапа. Hansen et al. (2011) использовали процедуру блочного бутстрапа с построением 5000 «повторных» выборок, которые состоят из значений имеющихся потерь моделей, случайным образом переставленных блоками на произвольные места. Такие блоки позволяют сохранить временную зависимость, имеющуюся в исходных данных.

От нулевых гипотез теста на эквивалентность, сформулированных в (36) и (37), логично перейти к статистикам

$$T_{R,\mathcal{M}} = \max_{i,j \in \mathcal{M}} |t_{ij}|, \quad T_{\max,\mathcal{M}} = \max_{i,j \in \mathcal{M}} t_{i,\cdot}, \quad (39)$$

где t_{ij} и t_i определены в (38). Асимптотическое распределение этих статистик не является стандартным, поэтому это распределение при выполнении гипотезы \mathbf{H}_0 тоже необходимо оценивать при помощи процедуры бутстрапа.

Выбор наилучшей модели на исключение из множества \mathcal{M} выполняется при помощи исключающего правила, которое основывается на тесте, основанном на статистике (38), и формулируется как

$$e_{R,\mathcal{M}} = \operatorname{argmax}_{i \in \mathcal{M}} \left\{ \sup_{j \in \mathcal{M}} \left(\bar{d}_{ij} / \sqrt{\widehat{\text{Var}}(\bar{d}_{ij})} \right) \right\}, \quad (40)$$

$$e_{\max,\mathcal{M}} = \operatorname{argmax}_{i \in \mathcal{M}} \left\{ \bar{d}_{i,\cdot} / \sqrt{\widehat{\text{Var}}(\bar{d}_{i,\cdot})} \right\}.$$

Фактически, MCS тест состоит из двух повторяющихся шагов, применяемых к множеству \mathcal{M} — проведение теста на эквивалентность $\delta_{\mathcal{M}}$ и проверка исключающего правила $e_{\mathcal{M}}$. Если тест $\delta_{\mathcal{M}}$ отклоняется, то не все модели в множестве \mathcal{M} одинаково «хороши» по заданному критерию, и в этом случае применяется исключающее правило, чтобы исключить модель с наименьшей производительностью на используемых данных в множестве \mathcal{M} . После исключения такой модели процедура повторяется заново до тех пор, пока нулевая гипотеза теста $\delta_{\mathcal{M}}$ не будет отклонена, что будет означать, что \mathcal{M} состоит только из моделей с одинаковой производительностью. При проведении всех тестов используется один и тот же уровень значимости α , что гарантирует $\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{M}^* \subset \mathcal{M}_{1-\alpha}^*) \geq 1 - \alpha$. Процедура MCS теста рассчитывает P -значения для каждой из моделей. Для модели $i \in \mathcal{M}_0$ MCS P -значение \hat{p}_i является порогом, при котором модель $i \in \mathcal{M}_{1-\alpha}^*$ тогда и только тогда, когда $\hat{p}_i \geq \alpha$. В результате модели с маленькими P -значениями, скорее всего, имеют плохую производительность относительно других моделей в \mathcal{M}_0 .

Алгоритм выполнения MCS теста состоит из следующих шагов.

1. Изначально считаем все модели равносильными: $\mathcal{M} = \mathcal{M}_0$.
2. Тестируем $\mathbf{H}_{0,\mathcal{M}}$, используя $\delta_{\mathcal{M}}$ с доверительным уровнем α .
3. Если $\mathbf{H}_{0,\mathcal{M}}$ не отвергается, то алгоритм завершается, и искомым множеством моделей с вероятностью $1 - \alpha$ является текущее: $\widehat{\mathcal{M}}_{1-\alpha} = \mathcal{M}$. Если же \mathbf{H}_0 отвергается, то не все

модели статистически равны по производительности, и используется исключающее правило e_M , чтобы определить модель с наихудшей производительностью и исключить ее из текущего множества M , после чего алгоритм возвращается к шагу 2.

В лучшем случае после выполнения алгоритма MCS теста в финальном множестве остается единственная модель. Однако если в данных содержится недостаточно информации для теста на эквивалентность, то ему не будет хватать мощности для разделения моделей по их производительности на хорошие и плохие. Недостаток мощности приведет к тому, что алгоритм MCS теста будет в среднем завершаться преждевременно, а финальное множество будет состоять из большого числа моделей, включая некоторые с худшей производительностью.

Если же имеющейся информации в данных достаточно, то тест отвергнет все ложные нулевые гипотезы, и, более того, исключающее правило не будет ставить под сомнения действительно превосходящие модели, пока в текущем множестве остаются модели с плохой производительностью. При этом вероятность того, что одна или несколько наилучших моделей будут исключены, асимптотически ограничена доверительным уровнем теста.

В силу последовательного применения тестов при выполнении алгоритма MCS логично задуматься о проблеме накопления ошибок первого рода с нежелательными последствиями. Однако последовательное применение тестов прекращается, как только нулевая гипотеза отвергается.

В целом, в результате выполнения MCS теста определяется множество моделей с одинаковой производительностью согласно заданного критерия, что может быть, в частности, полезно для прогнозирования волатильности.

3. Результаты эмпирического сравнения

Для эмпирического сравнения были взяты 88 GARCH моделей со структурой, рассмотренной в разделе 1.1, десять HAR-RV моделей, описанных в разделе 1.2, а также четыре ARFIMA модели — две с фиксированными параметрами: $(1, 0.401, 0)$, $(5, 0.401, 0)$, и две с оцениваемым параметром d : $(1, d, 0)$, $(5, d, 0)$. Сравнение выполнялось на данных котировок десяти российских активов: акциях алмазодобывающей компании «АЛРОСА» (ALRS), акциях энергетической корпорации «Газпром» (GAZP), горно-металлургической компании «Норникель» (GNWK), нефтяной компании «ЛУКОЙЛ» (LKOH), композитного фондового индекса ММББ (MICEX), телекоммуникационной компании «МТС» (MTSS), нефтегазовой компании «Роснефть» (ROSN), биржевого индекса Russian Trading System Cash Index (RTSI), акциях Сбербанка России (SBER) и акциях ВТБ банка (VTBR) с 09 сентября 2013 г. по 12 мая 2016 г., что в целом составляет 654 торговых дня.

Для сравнения использовались данные котировок на закрытии торговых сессий, а также внутрисессионные данные — цены закрытия внутрисессионных пятиминутных интервалов. При этом данные первых 504 дней использовались исключительно для оценивания моделей, а «out-of-sample» периодом служили следующие 150 дней, на которых происходило сравнение прогнозной силы моделей.

На рисунке 1 темные линии соответствуют логарифму из корня реализованной волатильности, рассчитанной согласно (16), а более светлые линии соответствуют логарифму из корня дневной волатильности, рассчитанному как $\log \sigma_t = \log \sqrt{(\log p_t - \log p_{t-1})^2}$.

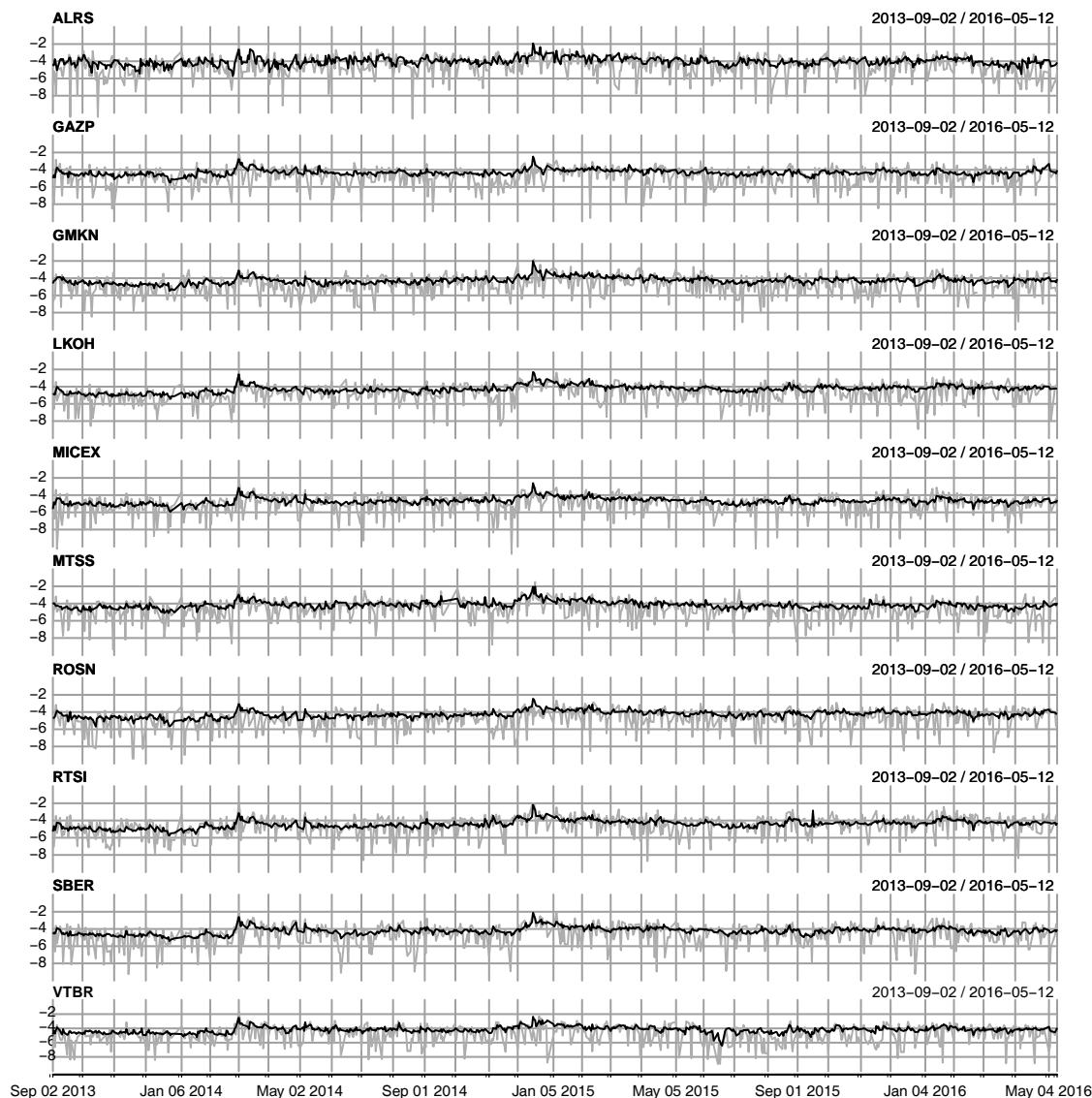


Рис. 1. Реализованные и дневные лог-волатильности

По графику легко убедиться в наличии большого количества шумов в дневной оценке волатильности (глядя на высоту вершин дневных оценок реализованной волатильности и оценок дневной волатильности, построенной по ценам закрытия), в то время как реализованная волатильность выглядит более сглаженной. Это подтверждает сказанное выше о преимуществе использования реализованной волатильности в качестве критерия при прогнозировании волатильности.

Как видно, у всех активов наблюдаются два пика: в марте и в декабре 2014 г., что соответствует падениям рынка, связанным с политическими событиями. Программа для сравнения моделей была написана на языке программирования *R* с использованием библиотеки *rugarch*, реализованной Ghalanos (2015), которая способна оценивать большое число

GARCH моделей. Эта библиотека написана большей частью на языке *C*, что позволяет оценивать GARCH модели гораздо быстрее по сравнению с кодом, написанным на *R*, что существенно для данной работы из-за необходимости многократной переоценки большого числа моделей.

Из GARCH семейства были взяты 11 моделей, описанных в разделе 1.1, с двумя структурами лагов ((1,1) и (2,1)), а также четыре разными условными распределениями случайной ошибки (нормальное, скошенное *t*-распределение Стьюдента, обратное гауссовское распределение и Johnson's reparametrized SU распределение, описанные в (7)–(15)), что в сумме дает 88 моделей. В качестве процесса математического ожидания доходностей использовался ARMA(1,1) процесс.

Из семейства линейных HAR-RV моделей в сравнении участвовали HAR-RV, HAR-RV-J, HAR-RV-CJ, LHAR-RV-J и LHAR-RV-CJ модели, а также их логарифмические нелинейные модификации, описанные в уравнениях (17)–(28). Все модели оценивались со стандартными параметрами гетерогенности (1, 5, 22).

Использовались данные «движущегося окна» из 504 торговых дней с дальнейшим расчетом «out-of-sample» одношагового прогноза на один день вперед. Это означает, что в каждый момент времени *t* модель оценивалась по данным дней $[t - 503 : t]$, а затем выполнялся прогноз на день *t* + 1. Затем модель заново оценивалась в день *t* + 1 и строился прогноз для дня *t* + 2.

В итоге для каждой из 102 моделей было получено 150 одношаговых прогнозов волатильности на один день вперед. Затем они сравнивались с фактическими реализованными волатильностями в эти дни при помощи теста MCS для выявления множества моделей, значительно превосходящих остальные по прогнозной силе, а также разницы в прогнозах между GARCH, ARFIMA и HAR-RV семействами. Сравнение большого набора GARCH, ARFIMA и HAR-RV моделей, а также использование теста MCS позволяет ответить на вопрос «Действительно ли HAR-RV модели превосходят семейство GARCH моделей при прогнозе реализованной волатильности?», о чем говорят сравнения отдельных представителей этих семейств. Включение ARFIMA моделей должно не только проверить обнаруженное в литературе преимущество HAR-RV моделей за счет включения реализованной волатильности, но и ответить на вопрос «Действительно ли гетерогенная структура HAR-RV необходима для прогнозирования реализованной волатильности?».

В таблице 1 приведены результаты выполнения теста MCS для 10 биржевых активов. В качестве функции потерь были рассмотрены три варианта, приведенные в (33), при этом каждый раз строилось 5000 ресэмплов для тестовых статистик процедуры MCS теста. В результате сравнения можно сделать вывод, что семейство HAR-RV моделей значительно превосходит все рассмотренные GARCH и ARFIMA модели при прогнозировании реализованной волатильности. Согласно проведенным тестам, при трех разных функциях потерь в финальное множество моделей, равных по предсказательной силе и превосходящих остальные, вошли только HAR-RV модели.

Более того, практически всегда финальное множество состояло из нелинейных LOG-HAR-RV моделей, что согласуется с логнормальным распределением реализованной волатильности и предыдущими исследованиями. Поскольку на практике реализованная волатильность чаще всего используется в качестве наиболее точной оценки волатильности, то логарифмические модификации HAR-RV моделей являются предпочтительными для использования при работе с дневной волатильностью.

Таблица 1. Результаты тестирования при помощи процедуры MCS

Модели	ALROSA			GAZP			GMKN			LKOH			MICEX			MTSS			ROSN			RTSI			SBER			VTBR		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
HAR-RV	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LOG-HAR-RV	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
HAR-RV-J	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LOG-HAR-RV-J	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
HAR-RV-CJ	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LOG-HAR-RV-CJ	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LHAR-RV-J	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LOG-LHAR-RV-J	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LHAR-RV-CJ	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LOG-LHAR-RV-CJ	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Примечание. Цифры 1,2,3 обозначают функции потерь MSE, MAE, R2LOG соответственно. Звездочками отмечены модели, попавшие в финальное множество MCS теста.

Выводы

В работе выполнено сравнение более 100 представителей трех семейств моделей по их прогнозной силе при прогнозировании реализованной волатильности на один день вперед с использованием котировок 10 российских биржевых активов, в том числе по ценам 8 акций российских компаний и двух биржевых индексов. Поскольку в абсолютном большинстве работ реализованная волатильность используется в качестве критерия при прогнозировании волатильности и как наилучшая оценка дневной волатильности, то на основе выполненного сравнения можно сделать вывод в пользу HAR-RV моделей. Полученные результаты уверенно говорят в пользу выбора модели из HAR-RV семейства при построении одношагового прогноза волатильности.

Главной причиной таких результатов является использование реализованной волатильности при оценивании HAR-RV моделей. GARCH модели, оцениваемые по ценам закрытия торговых дней, неспособны воспроизводить особенности реализованной волатильности, в результате чего ни одна из 88 моделей не попала в множество незначимо отличающихся по прогнозной силе моделей. Хотя ARFIMA модели напрямую моделируют реализованную волатильность, им также не удалось попасть в финальное множество моделей. Это говорит о том, что представление реализованной волатильности в виде суммы трех компонент в модели HAR-RV позволяет значительно улучшить прогноз.

Полученные результаты согласуются с проведенными в литературе исследованиями, содержащими сравнения отдельных представителей семейств GARCH, ARFIMA и HAR-RV моделей, и подтверждают превосходство HAR-RV семейства над остальными на данных российских биржевых активов.

Список литературы

Andersen T., Bollerslev T., Diebold F. X. (2007). Roughing it up: Including jump components in measuring, modeling and forecasting asset return volatility. *Review of Economics and Statistics*, 89 (4), 701–720.

Andersen T., Bollerslev T., Diebold F. X., Labys P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71 (2), 579–625.

Barndorff-Nielsen O., Shephard N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2 (1), 1–37.

Black F. (1976). Studies of stock price volatility changes. In: *Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economic Statistics Section*, 177–181.

Bollerslev T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307–327.

Corsi F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7 (2), 174–196.

Corsi F., Pirino D., Reno R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159 (2), 276–288.

Corsi F., Reno R. (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business and Economic Statistics*, 30 (3), 368–380.

Craioveanu M., Hillebrand E. (2012). Why it is OK to use the HAR-RV(1,5,21) model. *Working Papers No 1201*, University of Central Missouri, Department of Economics & Finance.

- Engle R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50 (4), 987–1008.
- Fernandez C., Steel M. (1998). On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association*, 93, 359–371.
- Ghahlanos A. (2015). *Rugarch: Univariate GARCH models*. <https://cran.r-project.org/web/packages/rugarch/index.html>.
- Hansen P., Lunde A. (2004). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20 (7), 873–889.
- Hansen P., Lunde A., Nason J. (2011). The model confidence set. *Econometrica*, 79 (2), 453–497.
- Johnson N. (1949). System of frequency curves generated by method of translation. *Biometrika*, 36, 149–176.
- Lahaye J., Shaw P. (2014). Can we reject linearity in an HAR-RV model for the S&P 500? Insights from a nonparametric HAR-RV. *Economics Letters*, 125 (1), 43–46.
- Lambert P., Laurent S. (2001). Modelling skewness dynamics in series of financial data using skewed location-scale distributions. *Stat Discussion Paper*, 0119. Université Catholique de Louvain.
- LeBaron B. (2001). Stochastic volatility as a simple generator of financial power-laws and long memory. *Quantitative Finance*, 1 (6), 621–631.
- Liu L., Patton A. J., Sheppard K. (2013). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 187 (1), 293–311.
- Mastro D. (2014). Forecasting realized volatility: ARCH-type models vs. the HAR-RV model. *Dissertation*, Kingston University, London.
- Müller U. A., Dacorogna M. M., Davé R. D., Olsen R. B., Pictet O. V., von Weizsäcker J. E. (1997). Volatilities of different time resolutions — analyzing the dynamics of market components. *Journal of Empirical Finance*, 4 (2-3), 213–239.
- Souček M., Todorova N., (2014). Realized volatility transmission: The role of jumps and leverage effects. *Economics Letters*, 122 (2), 111–115.
- Taylor S. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical Finance*, 4, 183–204.
- White H. (2000). A reality check for data snooping. *Econometrica*, 68 (5), 1097–1126.

Поступила в редакцию 11.04.2017;
принята в печать 20.09.2017.

Aganin A. Forecast comparison of volatility models on Russian stock market. *Applied Econometrics*, 2017, v. 48, pp. 63–84.

Artem Aganin

National Research University Higher School of Economics, Moscow, Russian Federation;
aaganin@hse.ru

Forecast comparison of volatility models on Russian stock market

This article is dedicated to multivariate comparison of big number of GARCH, ARFIMA and HAR-RV families' models considering their one-day ahead realized volatility, which is known to be a consistent measure of daily volatility. A total of 102 models from three families were included in comparison.

Comparison was completed with the help of Model Confidence Set test using 3 different loss functions on 10 Russian stock assets, including eight stock assets and two stock market indices. Received results strongly suggest HAR-RV superior performance to other two families of volatility models on Russian stock market and confirm local findings of previous studies.

Keywords: GARCH; realized volatility; HAR-RV; MCS.

JEL classification: C52; C53; C58.

References

Andersen T., Bollerslev T., Diebold F. X. (2007). Roughing it up: Including jump components in measuring, modeling and forecasting asset return volatility. *Review of Economics and Statistics*, 89 (4), 701–720.

Andersen T., Bollerslev T., Diebold F. X., Labys P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71 (2), 579–625.

Barndorff-Nielsen O., Shephard N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2 (1), 1–37.

Black F. (1976). Studies of stock price volatility changes. In: *Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economic Statistics Section*, 177–181.

Bollerslev T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307–327.

Corsi F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7 (2), 174–196.

Corsi F., Pirino D., Reno R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159 (2), 276–288.

Corsi F., Reno R. (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business and Economic Statistics*, 30 (3), 368–380.

Craioveanu M., Hillebrand E. (2012). Why it is OK to use the HAR-RV(1,5,21) model. *Working Papers No 1201*, University of Central Missouri, Department of Economics & Finance.

Engle R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50 (4), 987–1008.

Fernandez C., Steel M. (1998). On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association*, 93, 359–371.

Ghalanos A. (2015). *Rugarch: Univariate GARCH models*. <https://cran.r-project.org/web/packages/rugarch/index.html>.

Hansen P., Lunde A. (2004). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20 (7), 873–889.

Hansen P., Lunde A., Nason J. (2011). The model confidence set. *Econometrica*, 79 (2), 453–497.

Johnson N. (1949). System of frequency curves generated by method of translation. *Biometrika*, 36, 149–176.

Lahaye J., Shaw P. (2014). Can we reject linearity in an HAR-RV model for the S&P 500? Insights from a nonparametric HAR-RV. *Economics Letters*, 125 (1), 43–46.

Lambert P., Laurent S. (2001). Modelling skewness dynamics in series of financial data using skewed location-scale distributions. *Stat Discussion Paper*, 0119. Université Catholique de Louvain.

LeBaron B. (2001). Stochastic volatility as a simple generator of financial power-laws and long memory. *Quantitative Finance*, 1 (6), 621–631.

Liu L., Patton A. J., Sheppard K. (2013). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 187 (1), 293–311.

Mastro D. (2014). Forecasting realized volatility: ARCH-type models vs. the HAR-RV model. *Dissertation*, Kingston University, London.

Müller U. A., Dacorogna M. M., Davé R. D., Olsen R. B., Pictet O. V., von Weizsäcker J. E. (1997). Volatilities of different time resolutions — analyzing the dynamics of market components. *Journal of Empirical Finance*, 4 (2-3), 213–239.

Souček M., Todorova N., (2014). Realized volatility transmission: The role of jumps and leverage effects. *Economics Letters*, 122 (2), 111–115.

Taylor S. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical Finance*, 4, 183–204.

White H. (2000). A reality check for data snooping. *Econometrica*, 68 (5), 1097–1126.

Received 11.04.2017; accepted 20.09.2017.