Статьи: финансовая эконометрика

Бутстраповская инференция об интегрированной волатильности[⋆]

Андрей Рафальсон[†]

Barclays Capital, Лондон, Великобритания

В работе Gonçalves & Meddahi (2009) было предложено использовать iid и дикий бутстрап для реализованной волатильности вместо асимптотического подхода с целью оценивания интегрированной волатильности. Мы предлагаем использовать блочный бутстрап и GARCH остаточный бутстрап, мотивируя данные подходы присутствием внутридневной временной структуры в доходностях активов. Используя технику симуляций Монте–Карло, мы показываем, что блочный бутстрап является более точным подходом на низкочастотных внутридневных данных, более робастным и валидным. Другим результатом является факт, что GARCH бутстрап опережает другие подходы в случае высокой условной гетероскедастичности в данных. Также данный подход демонстрирует точную инференцию на симулированных данных при соответствующей модели с высокой частотой наблюдений. Однако GARCH бутстрап является более вычислительно емким и менее робастным, чем остальные.

Ключевые слова: интегрированная волатильность, реализованная волатильность, блочный бутстрап, GARCH бутстрап

Классификация JEL: C12, C22, C58

1 Введение

Моделирование доходностей активов на высокочастотных данных в области финансовых рынков стало важной проблемой в финансовой эконометрике с того момента, как этот подход был рассмотрен в Andersen & Bollerslev (1997). Возможность непараметрического оценивания процесса волатильности является основным преимуществом такого типа данных. Используя внутридневные наблюдения, мы можем вывести эмпирическое распределение дневных доходностей и получить информацию об их моментах. В частности, применение теории квадратичной вариации в моделировании распределения доходностей активов позволяет оценить волатильность распределения.

В стандартных моделях с непрерывным временем волатильность доходностей обычно представлена интегрированной волатильностью, которая по своей сути латентна. Учеными было показано, что при некоторых строгих предположениях реализованная волатильность является состоятельной и достаточно точной оценкой интегрированной волатильности (Andersen, Bollerslev, Diebold & Labys 2001a, далее ABDL). Но, к сожалению, присутствие микроструктурных эффектов, таких как дискретность цен, разрывы спроса-предложения, нерегулярные торги и т.д. (ABDL 2000), а также наличие прыжков, дрейфа цен активов и эффекта рычага, делают реализованную волатильность не совсем идеальной оценкой в реальной практике (Meddahi 2002, Barndorff-Nielsen & Shephard 2002a, 2002b).

[⋆]Цитировать как: Рафальсон, Андрей (2012) «Бутстраповская инференция об интегрированной волатильности», Квантиль, №.10, стр. 91–108. Citation: Rafalson, Andrey (2012) “Bootstrap inference about integrated volatility”, Quantile, No.10, pp. 91–108.

[†]Адрес: 5 North Colonnade, Canary Wharf, London E14 4BB, United Kingdom. Электронная почта: *rafalson.andrey@gmail.com*

Теория реализованной волатильности была усилена асимптотическими теоремами в работе Barndorff-Nielsen & Shephard (2002a) (далее BNS). В частности, авторы презентовали центральную предельную теорему (ЦПТ) для реализованной волатильности, когда частота наблюдений растет. Этот результат позволил строить доверительные интервалы для интегрированной волатильности. Более того, ЦПТ утверждает, что ошибки оценивания интегрированной волатильности с помощью реализованной являются асимптотически нормальными с нулевым средним. С того момента, как было показано, что реализованная волатильность, будучи смещенной оценкой, является асимптотически точной, асимптотическая теория начала активно использоваться в построении доверительных интервалов для интегрированной волатильности на реальных данных.

Предложив использовать метод бутстрапирования на высокочастотных данных, Gonçalves & Meddahi (2004) улучшили теорию инференции для интегрированной волатильности. Они проанализировали и предложили два бутстрап метода для реализованной волатильности: iid и дикий бутстрап. Метод iid бутстрапа использовал внутридневные доходности из фактического наблюдаемого множества доходностей. Метод был мотивирован моделями, в которых волатильность постоянна, и, следовательно, доходности внутри дня независимы и одинаково распределены (что и означает iid). Наблюдения в диком бутстрапе генерировались путем произведения каждой выборочной внутридневной доходности на независимую нормально распределенную случайную величину. Этот метод мотивировался в Wu (1986). Gonçalves & Meddahi (2009) также показали валидность этого подхода в моделях со стохастической волатильностью.

Данная работа ставит задачи расширить множество бутстрап-подходов для реализованной волатильности с целью улучшения существующих результатов и построить более точные асимптотические интервалы для интегрированной волатильности при невысокой частоте внутридневных наблюдений. Мотивацией использовать умеренную частотность наблюдений является подверженность данных микроструктурным эффектам. Блочный бутстрап (Hall 1995) кажется хорошим подходом для этих целей, так как в некоторой степени он сохраняет временную структуру временных рядов при построении подвыборок. Более того, как показали ABDL (2001c), очень важно сохранять временную структуру, так как реализованная волатильность склонна демонстрировать сильную временную зависимость, которую можно описать процессом с долгой памятью и латентными факторами. Использование блочного бутстрапа было предложено Gonçalves & Meddahi (2004) как возможное расширение их работы, но, насколько нам известно, ничего не было сделано в этом направлении. В этой статье мы предлагаем подход, как построить блочный бутстрап, валидный в первой аппроксимации.

Другим расширением работы является контроль внутридневной условной гетероскедастичности в данных. Для этой цели мы также предлагаем GARCH бутстрап, впервые введенный Maercker (1997). Наличие условной гетероскедастичности в финансовых данных (Engle 1982, Bollerslev 1986) мотивирует контроль данного эффекта в процессе бутстрапирования. Стандартные бутстрап-подходы в теории инференции полагают ошибки независимыми и одинаково распределенными. Таким образом, эти подходы теряют валидность при наличии условной гетероскедастичности. Присутствие данного эффекта особенно сильно проявляется на помесячных, недельных и дневных наблюдения. Однако, мы полагаем, что внутридневные наблюдения могут быть подвержены данному эффекту в такой же степени (ABDL 2001a).

Мы видим основной вклад работы в следующем. Во-первых, мы расширяем набор подходов для бутстрапирования реализованной волатильности. Предлагаемый блочный бутстрап улучшает асимптотику на низкой частоте данных, что может быть полезно при микроструктурных эффектах. Предложенный GARCH бутстрап позволяет контролировать условную гетероскедастичность во внутридневных данных. Во-вторых, мы демонстрируем свойство робастности для обоих методов бутстрапа при весьма общих предположениях, допускающих ненулевой дрейф и эффект рычага. В-третьих, мы используем симуляции Монте–Карло для сравнения точности построения бутстраповских доверительных интервалов для интегрированной волатильности между предложенными подходами, уже существующими разновидностями бутстрапа и асимптотическим подходом.

Полученные результаты заключаются в следующем. При высокой частоте данных все подходы дают относительно похожие результаты. При низкой частоте, асимптотический подход, основанный на ЦПТ, и iid бутстрап дают немного слишком узкий асимптотический интервал и склонны не накрывать интегрированную волатильность с заявленной точностью. Дикий бутстрап дает слишком широкий доверительный интервал. Неперекрывающийся блочный бутстрап дает достаточно точный доверительный интервал превосходя все остальные подходы. Перекрывающийся блочный бутстрап превосходит другие подходы на достаточно частотных данных и является более робастным, чем остальные подходы. GARCH бутстрап превосходит другие подходы, когда данные подвержены условной гетероскедастичности. Он также демонстрирует хорошую инференцию на симулированных данных при соответствующей GARCH модели с высокой частотностью. Однако этот подход менее робастен, чем остальные в смысле отклонения от условий соответствующей модели, в частности, от отсутствия дрейфа и эффекта рычага.

Структура данной статьи организована следующим образом. Раздел 2 вводит в курс базовых концепций реализованной и интегрированной волатильности и основных результатов, относящихся к данной тематике. Раздел 3 описывает блочный бутстрап. Раздел 4 описывает GARCH бутстрап на регрессионных остатках. В разделе 5 мы сравниваем результаты бутстрап-подходов, используя технику симуляций Монте–Карло, и обсуждаем их. Приложение содержит доказательства результатов из разделов 3 и 4.

2 Базовые понятия

В данном разделе мы полагаем, что у нас есть ликвидный актив, чья цена S_t имеет непрерывную структуру, определенную с помощью следующего стохастического дифференциального уравнения:

$$d\log S_t = \mu_t dt + \sigma_t dW_t, \tag{1}$$

где W_t обозначает стандартное броуновское движение, μ_t обозначает дрейф, который имеет конечную вариацию, а σ_t обозначает процесс волатильности, такой что $\int_0^t \sigma_u^2 du < \infty$. Предположим, что σ_t и μ_t являются непрерывными процессами, и σ_t независима от W_t. Допустим, что время t измеряется в дневных единицах. В соответствии с решением стохастического дифференциального уравнения (1) мы можем определить однодневную непрерывную ставку доходности за день t:

$$r_t = \log S_t - \log S_{t-1} = \int_{t-1}^{t} \mu_u du + \int_{t-1}^{t} \sigma_u dW_u. \tag{2}$$

Из модели (1) мы можем определить и другие характеристики процесса. Так, например, интегрированная волатильность есть внутренняя мера вариации доходности, определяемая следующей формулой:

$$IV_t = \int_{t-1}^{t} \sigma_u^2 du.$$

Рассмотрим действительное число h, такое что $1/h$ есть положительное целое число, и определим $r_{t-1+ih}^{(h)}$ как доходность за период $[t-1+(i-1)h; t-1+ih]$. Тогда реализованная волатильность $RV_t(h)$ есть

$$RV_t(h) = \sum_{i=1}^{1/h} r_{t-1+ih}^{(h)2}.$$

В соответствии с теорией квадратичной вариации, состоятельность реализованной волатильности как оценки интегрированной волатильности оправдывается растущим количеством внутридневных наблюдений и уменьшающимся диаметром разбиения $h \to 0$ (BNS 2002a). Однако, на реальных данных микроструктурные эффекты, такие как дискретность цен, разрывы спроса-предложения и нерегулярные торги, ограничивают ультра-высокочастотность, т.к. они нарушают полумартингальные свойства доходностей (Andreou & Ghysels 2002). Имея в виду работу Andersen, Bollerslev, Diebold & Labys (2001b), с целью избежания микроструктурных эффектов для большинства активов лучшей практикой является использование внутридневных доходностей с частотой не выше, чем 30 минут. Таким образом, качество оценки с разумной частотой будет определяться случайной ошибкой

$$U_t^h = RV_t(h) - IV_t.$$

Основные свойства данной «шумовой» случайной величины были рассмотрены BNS (2002a) и Meddahi (2002a):

(a) *Среднее* U_t^h *в общем случае не равно нулю, когда дрейф* μ_t *отличен от нуля.*

(b) U_t^h *в общем случае гетероскадастична.*

(c) *При наличии эффекта рычага* U_t^h *коррелирует с интегрированной волатильностью* IV_t.

Асимптотические свойства ошибки были изначально описаны с помощью ЦПТ для реализованной волатильности (BNS 2002a). В частности, полагая, что процессы дрейфа и волатильности совместно независимы от W_u, было показано следующее асимптотическое свойство:

$$\sqrt{h^{-1}}\frac{RV_t(h) - IV_t}{\sqrt{2IQ_t}} \xrightarrow{d} \mathcal{N}(0,1), \tag{3}$$

где IQ обозначает интегрированую квартильность, которая определяется как

$$IQ_t = \int_{t-1}^{t} \sigma_u^4 du.$$

Заметим, что данный асимптотический результат не требует какого-либо знания о процессах дрейфа или диффузии, и асимптотика верна, даже если четвертый момент доходностей не существует. Аналогично, присутствие эффекта рычага не влияет на данное асимптотическое свойство, а ненулевое среднее «шумовой» случайной величины не противоречит асимптотическому свойству (3).

Используя теорию степенной вариации, мы можем вывести реализованную квартильность, которая будет состоятельной оценкой интегрированной квартильности при тех же предположениях (BNS 2002a, 2004b, 2006). Она определяется как

$$RQ_t = \frac{1}{h}\frac{1}{3}\sum_{i=1}^{1/h} r_{t-1+ih}^{(h)4}. \tag{4}$$

Пивотизацией статистики (3) мы можем построить доступный 95% асимптотический доверительный интервал для интегрированной волатильности:

$$CI_t^a = RV_t \pm 1.96\frac{1}{\sqrt{h^{-1}}}\sqrt{2RQ_t}. \tag{5}$$

Альтернативный метод построения доверительных интервалов был предложен в Gonçalves & Meddahi (2004), предложившими использовать бутстрап для инференции об интегрированной волатильности. Благодаря свойству асимптотического рафинирования бутстрап обычно дает более точную аппроксимацию распределения на конечной выборке, чем асимптотическое распределение. При условиях нулевого дрейфа $\mu_t = 0$ и отсутствия эффекта рычага

Gonçalves & Meddahi (2004) вывели бутстраповские статистики для дикого и iid бутстрапов, а также доказали их асимптотическую валидность первого порядка.

Метод iid бутстрапа мотивировался моделями, в которых волатильность постоянна, и следовательно, внутридневные доходности являются независимыми и одинаково распределенными. Таким образом, этот метод использует внутридневные доходности из начального множества доходностей:

$$r^*_{t-1+ih} \sim iid$$

из $\{r_{t-1+ih}\}$. Чтобы построить симметричный 95% t-процентильный iid бутстраповский доверительный интервал для интегрированной волатильности

$$CI_t^{\%t} = RV_t \pm q^*_{0.95} \frac{1}{\sqrt{h^{-1}}} \sqrt{2RQ_t}$$

нужно взять 95% квантиль $q^*_{0.95}$ из распределения множества

$$\left| \frac{\sqrt{h^{-1}}(RV_t^* - RV_t)}{u_t^{(h)*}} \right|,$$

где

$$u_t^{(h)*2} = h^{-1} \sum_{i=1}^{1/h} r^{*4}_{t-1+ih} - \left(\sum_{i=1}^{1/h} r^{*2}_{t-1+ih} \right)^2.$$

Метод дикого бутстрапа мотивировался моделями, в которых волатильность является стохастической. Наблюдения в диком бутстрапе генерируются путем умножения первоначальных внутридневных доходностей на независимую нормальную случайную величину:

$$r^*_{t-1+ih} = r_{t-1+ih}\eta_i, \quad \eta_i \sim \; iid\, \mathcal{N}(0,1).$$

Чтобы построить симметричный 95% бутстраповский t-процентильный доверительный интервал для интегрированной волатильности, нужно взять 95% квантиль $q^*_{0.95}$ из распределения множества

$$\left| \frac{\sqrt{h^{-1}}(RV_t^* - RV_t)}{\sqrt{2RQ_t^*}} \right|.$$

Более того, Gonçalves & Meddahi (2009) доказали, что бутстраповская аппроксимация лучше, чем асимптотическая, и продемонстрировали свойства робастности этих бутстрапов при ненулевом дрейфе и эффекте рычага. Они также продемонстрировали, что iid бутстрап может улучшить аппроксимацию даже в случае стохастической волатильности.

3 Блочный бутстрап

В этом разделе мы расширяем множество бутстрапов, предложенных Gonçalves & Meddahi (2009), и предлагаем блочное бутстрапирование реализованной волатильности. Цель данного метода — сохранить временную структуру доходностей при создании подвыборок. Использование блочного бутстрапа — популярный метод для улучшения качества бутстрапирования временных рядов (Hall 1995). Основной принцип данного метода лежит в разделении данных на несколько блоков, которые могут хранить в себе информацию о структуре начального временного ряда. Блоки могут быть перекрывающиеся и неперекрывающиеся (Lahiri 1999).

Если обозначить размер выборки как $T = h^{-1}$ и длину блока как l, то в неперекрывающемся бутстрапе мы разделим данные на T/l блоков. Для подхода с перекрывающимися блоками мы разделим данные на $T - l - 1$ блоков, так что блок 1 будет $\{r_1, r_2, ..., r_l\}$, блок 2 будет $\{r_2, r_3, ..., r_{l+1}\}$ и т.д. Подход к выбору длины блока был рассмотрен Andrews (2004). Автор продемонстрировал, что хорошие асимптотические свойства могут быть достигнуты, если длина блоков зависит от размера выборки таким образом, что $l(T) \to \infty$ при $T \to \infty$ и $l(T)/T \to 0$. Более того, длина блока должна быть асимптотически пропорциональна $T^{1/3}$. В данной работе мы демонстрируем, что хорошие асимптотические свойства блочного бутстрапа могут быть достигнуты, предположив, что длина блоков равна $l = \lfloor T^{1/3} \rfloor$, где $\lfloor \, \rfloor$ — оператор взятия нижней целочисленной границы.

Для простоты мы предположим, что в уравнении (1) $\mu_t = 0$ для любых t. На меленьком интервале времени порядок дрейфа есть dt, что меньше, чем порядок $\sqrt{dt}$, т.е. порядок диффузии в уравнении (1). Таким образом, дрейф-компонента достаточно мала на высокой частоте данных, чтобы ей пренебречь (Andersen, Bollerslev & Diebold 2009). Тогда наша модель будет задаваться следующим стохастическим дифференциальным уравнением

$$d \log S_t = \sigma_t dW_t, \tag{6}$$

где $\sigma_t > 0$ — в общем случае стохастический процесс от времени. На протяжении всей статьи мы будем полагать независимость между процессом стохастической волатильности σ_t и броуновским движением W_t, т.е. отсутствие эффекта рычага.

Gonçalves & Meddahi (2009) показали, что в модели (6) внутридневные доходности условно на реализовавшемся процессе волатильности являются независимыми, но гетероскедастичными:

$$r_{t-1+ih} \sim N(0, \sigma_i^2),$$

где

$$\sigma_i^2 = \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du.$$

Однако это может не выполняться, если данные подвергнуты микро-структурным эффектам (ABDL 2001a).

Допустим, множество $\{r^*_{ih,h} : i = 1, ...1/h\}$ будет бутстраповской выборкой из начального множества внутридневных доходностей. Рассмотрим, например, неперекрывающийся блочный бутстрап. Тогда неперекрывающиеся блоки будут

$$\{r_{t-1+ih}\}_{i=1}^{l}, \{r_{t-1+ih}\}_{i=l+1}^{2l}, ..., \{r_{t-1+ih}\}_{i=1/h-l+1}^{1/h}.$$

База бутстрапа формируется из случайно выбранных блоков с возвращением из всего набора блоков до тех пор, пока не сформируется бутстраповская выборка длины $T = h^{-1}$.

$$\{r^*_{t-1+ih}\}_{i=1}^{1/h} = \left\{ \{r_{t-1+ih}\}_{i=l\zeta_k+1}^{l\zeta_k+l} \right\}_{k=1}^{T/l},$$

где ζ_k есть случайное число из множества $\{0, 1, 2, ..., T/l - 1\}$. Бутстраповская реализованная волатильность есть обычная реализованная волатильность, но оцененная на бутстраповских внутридневных доходностях

$$RV_t^* = \sum_{i=1}^{1/h} r^{*2}_{t-1+ih}.$$

Для блочного бутстрапа 95% процентильный интервал определяется как

$$CI_t^{\%} = RV_t \pm Q^*_{0.95},$$

где $Q^*_{0.95}$ есть 95% процентиль распределения нестьюдентизованной бутстраповской статистики $|\sqrt{h^{-1}}(RV_t - IV_t)\,|$. Для блочного бутстрапа 95% t-процентильный интервал равен

$$CI_t^{\%t} = RV_t \pm q^*_{0.95}\frac{1}{\sqrt{h^{-1}}}\sqrt{2RQ_t},$$

где $q^*_{0.95}$ есть 95% процентильный распределения уже стьюдентизованной бутстраповской статистики.

На протяжении всей статьи будем полагать, что $\mathbb{P}^*$ означает вероятностную меру, индуцированную бутстрапом условно на первоначальной выборке. Аналогично, положим $\mathbb{E}^*$ и $\mathbb{V}^*$ обозначать ожидание и дисперсию в бутстраповской мере, условно на первоначальной выборке. Стоит заметить, что раз мы действуем условно на первоначальной выборке внутридневных доходностей, добавление процесса волатильности в информационное множество не изменяет бутстраповской вероятностной меры. Таким образом, $\mathbb{P}^*$ можно также интерпретировать как вероятностную меру, индуцированную бутстрапом, условно на первоначальной выборке внутридневных доходностей и волатильности.

Gonçalves & Meddahi (2009) доказали, что для iid бутстрапа и дикого бутстрапа

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}^*\left\{\frac{\sqrt{h^{-1}}(RV_t^* - RV_t)}{v_t^{(h)*}} \le x\right\} - \Phi(x)\right| \to 0, \tag{7}$$

при условиях схожих с ЦПТ, и $h \to 0$, где $v_t^{(h)*2} = \mathbb{V}^*(\sqrt{h^{-1}}RV_t^*)$, и $\Phi(x) = \mathbb{P}\{Z \le x\}$ с $Z \sim \mathcal{N}(0,1)$. В данной статье мы полагаем, что данный результат верен для блочного и GARCH бутстрапов.

Далее характеризуем первые два выборочных момента бутстраповской реализованной волатильности RV^*.

Предложение 1. *При перекрывающемся и неперекрывающемся блочном бутстрапе для реализованной волатильности верны следующие свойства:*
(i) $\mathbb{E}^*(RV_t^* - RV_t) = 0$ *для любых* h *и* t.
(ii) Дисперсия бутстраповской статистики

$$v_t^{(h)*2} = \mathbb{V}^*\left(\sqrt{h^{-1}}RV_t^*\right) = h^{-1}\left((RV_t^{(1)})^2 + (RV_t^{(2)})^2 + \ldots + (RV_t^{(T/l)})^2\right) - (RV_t)^2 l,$$

такая, что $v_t^{(h)*2} - v_t^{*2} \to 0$ *по вероятности при* $h \to 0$, *где* $RV_t^{(k)}$ *обозначает реализованную волатильность в блоке* k *в день* t

$$RV_t^{(k)} = \sum_{i=1}^{l} r^2_{t-1+(k-1)l+ih},$$

и

$$v_t^{*2} = h^{-1}\left(\left(\int_{t-1}^{t-1+l}\sigma_u^2du\right)^2 + \ldots + \left(\int_{t-l+1}^{t}\sigma_u^2du\right)^2\right) - l\left(\int_{t-1}^{t}\sigma_u^2du\right)^2 \ne 2\int_{t-1}^{t}\sigma_u^4du.$$

Приложение содержит доказательство этого и остальных результатов этого и следующего разделов. Первый результат показывает, что блочный бутстрап дает несмещенную оценку реализованной волатильности для любой частотности h. Второй результат выводит предел по вероятности бутстраповской дисперсии при растущей частоте выборки $h \to 0$. Заметим, что, если длина блока будет равна единице, этот подход будет в точности таким же, как и iid бутстрап, и первые два момента бутстраповской статистики будут равны в данных подходах. Более формально, если $l = 1$, тогда

$$v_t^{(h)*2} = \mathbb{V}^* \left(\sqrt{h^{-1}} RV_t^* \right) = h^{-1} \sum_{i=1}^{1/h} r_{t-1+ih}^4 - \left(\sum_{i=1}^{1/h} r_{t-1+ih}^2 \right)^2$$

и

$$v_t^{(h)*2} \to 3 \int_{t-1}^{t} \sigma_u^4 du - \left(\int_{t-1}^{t} \sigma_u^2 du \right)^2 \neq 2 \int_{t-1}^{t} \sigma_u^4 du \quad \text{при } h \to 0.$$

Как мы можем видеть, когда волатильность гетероскедастична дисперсия бутстраповской статистики не сходится к $2IQ_t = 2 \int_{t-1}^{t} \sigma_u^4 du$ по вероятности. Следовательно, блок-бутстраповский процентильный интервал, построенный на квантилях распределения

$$\left| \sqrt{h^{-1}} (RV_t^* - RV_t) \right|,$$

не будет валидным. Однако, это не мешает блочному бутстрапу быть асимптотически валидным, когда волатильность имеет временную структуру и является стохастической. Действительно, если мы центрируем и стьюдентизируем бутстраповскую реализованную волатильность нужным образом, блочный бутстрап будет асимптотически валидным, в смысле, что бутстраповская статистика реализованной волатильности сходится к нормальной случайной величине при $h \to 0$ по вероятности. Мы можем определить бутстраповский t-процентильный интервал, основанный на квантилях распределения

$$\left| \frac{\sqrt{h^{-1}} (RV_t^* - RV_t)}{u_t^{(h)*}} \right|, \tag{8}$$

где $u_t^{(h)*} \to v_t^{(h)*}$ при $h \to 0$. Если мы положим

$$u_t^{(h)*2} = h^{-1} \left((RV_t^{(1)*})^2 + (RV_t^{(2)*})^2 + ... + (RV_t^{(T/l)*})^2 \right) - (RV_t^*)^2 l,$$

то бутстраповский t-процентильный интервал, основанный на квантилях распределения выражения (8), будет валидным, так как мы предположили результат (7).

Заметим, что все вышесказанное верно и для перекрывающегося, и для неперекрывающегося подходов к блочному бутстрапированию, так как доказательство Предложения 1 не использует свойства перекрытия блоков. Таким образом, эти два подхода являются эквивалентными в асимптотике первого порядка. Однако, они обладают неодинаковыми свойствами — распределение выражения (8) имеет разные моменты высокого порядка. В дальнейшем мы рассмотрим эти два подхода отдельно друг от друга.

4 GARCH бутстрап

Присутствие условной гетероскедастичности в моделях для финансовых данных (Engle 1982, Bollerslev 1986) мотивирует контроль этих эффектов во время бутстрапирования. Стандартные бутстраповские методы инференции полагают ошибки независимыми и одинаково распределенными, а также не подверженными условной гетероскедастичности. Присутствие условной гетероскедастичности особо сильно проявляется в регрессиях на помесячных, недельных и дневных данных. В данном случае мы полагаем, что внутридневные данные подвергнуты этому эффекту точно таким же образом.

С того момента, как Bollerslev (1986) ввел GARCH модель для контроля условной гетероскедастичности, асимптотические теории для моделей GARCH-типа были представлены многими учеными. Также появилось несколько работ, описывающих метод бутстрапирования для GARCH моделей. Kreiß & Franke (1992) доказали асимптотическую валидность метода бутстрапирования авторегрессионых моделей ARMA(p, q). Maercker (1997) показал слабую состоятельность дикого бутстрапа для GARCH(1, 1) модели, основываясь на асимптотических свойствах QML оценки, которые были изучены Lee & Hansen (1994). Более того, существует некоторое количество эмпирических приложений метода бутстрапирования GARCH моделей; так, например, появились работы Reeves (2005), Pascual, Romo & Ruiz (2005) и Robio (1999). В данной секции мы применяем GARCH бутстрап к реализованной волатильности с целью получения инференции об интегрированной волатильности.

Для простоты мы также будем рассматривать упрощенную модель (6) без дрейфа и эффекта рычага. Так как ABDL (2001a, 2001c) мотивировали сильную гетероскедастичность внутридневных доходностей для обменных курсов и сильную временную зависимость реализованной волатильности для индексов акций, имеет смысл рассматривать модель (6). Положим, процесс волатильности σ_t подчиняется такой модели, что наблюдаемые внутридневные доходности имеют GARCH(1,1) структуру.

$$\begin{aligned} r_{t-1+ih} &= \varepsilon_{t-1+ih} \\ \varepsilon_{t-1+ih} &= \psi^{\frac{1}{2}}_{t-1+ih}\xi_{t-1+ih} \\ \psi_{t-1+ih} &= \beta_0 + \beta_1\psi_{t-1+(i-1)h} + \beta_2\left(\varepsilon_{t-1+(i-1)h}\right)^2, \end{aligned} \tag{9}$$

где ψ_{t-1+ih} обозначает гетероскедастичность, а ξ_{t-1+ih} обозначает модельные шоки. Неотъемлемым свойством модели (6) является тот факт, что доходности нормально распределены с нулевым средним. Таким образом, мы можем сделать стандартное QML-предположение, что модельные шоки нормально распределены с нулевым средним. Также мы ограничиваем параметры модели (9) с целью получения положительности процесса гетероскедастичности и существования долгосрочного значения гетероскедастичности: $\beta_0, \beta_1, \beta_2 > 0$, $\beta_0 + \beta_1 + \beta_2 < 1$.

Методика GARCH бутстрапа заключается в следующих шагах:

- **Шаг 1:** Имея в виду состоятельность QML оценки при условиях $\xi_{t-1+ih} \sim N(0,1)$, вычисляем QML оценку для $\beta_0, \beta_1, \beta_2$. Таким образом, имея $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$, вычисляем остатки и полагаем

$$\psi_{t-1+ih} = \hat{\beta}_0 + \hat{\beta}_1\psi_{t-1+(i-1)h} + \hat{\beta}_2\left(\psi^{\frac{1}{2}}_{t-1+(i-1)h}\hat{\xi}_{t-1+(i-1)h}\right)^2.$$

- **Шаг 2:** Организуем процедуру бутстрапирования, набирая остатки ξ^*_{t-1+ih} из $N(0,1)$ распределения или из начального множества остатков $\{\hat{\xi}_{t-1+ih}\}$ с возвращением, т.е. из эмпирического распределения.

- **Шаг 3:** Полагая начальное значение гетероскедастичности равным долгосрочному значению

$$\psi^*_{t-1+0h} = \frac{\hat{\beta}_0}{1-\hat{\beta}_1-\hat{\beta}_2},$$

восстанавливаем процесс гетероскедастичности ψ^*_{t-1+ih} для каждой бутстрап-итерации:

$$\psi^*_{t-1+ih} = \hat{\beta}_0 + \hat{\beta}_1\psi^*_{t-1+(i-1)h} + \hat{\beta}_2\left(\psi^*_{t-1+(i-1)h}\xi^{*2}_{t-1+(i-1)h}\right).$$

- **Шаг 4:** Для каждой бутстрап-итерации рассчитываем множество бутстраповских доходностей:

$$r^*_{t-1+ih} = \psi^{*1/2}_{t-1+ih}\xi^*_{t-1+ih}$$

- **Шаг 5:** Для каждой бутстрап-итерации считаем бутстраповскую реализованную волатильность:

$$RV^* = \sum_{i=1}^{1/h} r^{*2}_{t-1+ih} = \sum_{i=1}^{1/h} \left(\psi^*_{t-1+ih}\xi^{*2}_{t-1+ih}\right)$$

Аналогично случаю блочного бутстрапа мы определим 95% t-процентильный интервал для GARCH бутстрапа как

$$CI_t^{\%t} = RV_t \pm q^*_{0.95}\frac{1}{\sqrt{h^{-1}}}\sqrt{2RQ_t}, \tag{10}$$

где $q^*_{0.95}$ есть 95% процентильный распределения стьюдентизованной бутстраповской статистики. Чтобы стьюдентизировать статистику соответствующим образом, мы должны найти первые два бутстраповских выборочных момента для RV^*. Обозначим их $\mathbb{E}^*(RV_t^*)$ и $v_t^{(h)*2} = \mathbb{V}^*(\sqrt{h^{-1}}RV_t^*)$ соответственно.

Предложение 2. *Для GARCH бутстрапа с набором остатков из начальной выборки или из нормального распределения мы можем численно подсчитать бутстраповское математическое ожидание $\mathbb{E}^*(RV_t^*)$ и дисперсию $v_t^{(h)*2}$ для реализованной волатильности.*

Приложение 2 содержит доказательство данного результата и в точности описывает метод, как получить бутстраповское ожидание и дисперсию реализованной волатильности. Следуя данному алгоритму, мы можем организовать расчет стьюдентизованной статистики

$$\left|\frac{\sqrt{h^{-1}}(RV_t^* - IV_t) - \mathbb{E}^*\left(\sqrt{h^{-1}}(RV_t^* - IV_t)\right)}{\sqrt{\mathbb{V}^*\left(\sqrt{h^{-1}}(RV_t^* - IV_t)\right)}}\right| = \left|\frac{\sqrt{h^{-1}}\left(RV_t^* - \mathbb{E}^*(RV_t^*)\right)}{v_t^{(h)*}}\right| \tag{11}$$

и для каждой бутстрап-итерации вычислить 95%-ный процентиль $q^*_{0.95}$ полученного распределения статистики (11). В конечном итоге, используя выражение (10), легко получить инференцию об интегрированной волатильности.

5 Симуляции Монте–Карло

В данной секции мы сравниваем асимптотическую теорию, дикий бутстрап и iid бутстрап с предложенными выше блочным и GARCH методами бутстрапирования на конечной частоте наблюдений. С целью достичь чистых и репрезентативных результатов мы следуем технике, используемой в статье Gonçalves & Meddahi (2009). Используя симуляции Монте–Карло, предложенные в Andersen, Bollerslev & Meddahi (2004), мы симулируем следующий процесс цен:

$$d\log S_t = \mu_t dt + \sigma_t\left(\rho_1 dW_{1t} + \rho_2 dW_{2t} + \sqrt{1-\rho_1^2-\rho_2^2}dW_{3t}\right),$$

где W_{1t}, W_{2t} и W_{3t} являются тремя независимыми броуновскими движениями. С целью достичь разнородной репрезентации результатов, мы рассмотрим 4 разных порождающих процесса волатильности σ_t.

Первый — это $GARCH(1,1)$ диффузия, впервые примененная в Andersen & Bollerslev (1998):

$$d\sigma_t^2 = 0{,}035(0{,}636 - \sigma_t^2)dt + 0{,}144\sigma_t^2 dW_{1t}.$$

Второй — логнормальная диффузия, впервые рассмотренная в Andersen, Benzoni & Lund (2002):

$$d\log\sigma_t^2 = -0{,}0136(0{,}8382 + \log\sigma_t^2)dt + 0{,}1148dW_{1t}.$$

Третий процесс — двухфакторная аффинная диффузия, рассмотренная в Bollerslev & Zhou (2002):

$$\sigma_t^2 = \sigma_{1,t}^2 + \sigma_{2,t}^2,$$
$$d\sigma_{1,t}^2 = 0{,}5708(0{,}3257 - \sigma_{1,t}^2)dt + 0{,}2286\sigma_{1,t}^2 dW_{1t},$$
$$d\sigma_{2,t}^2 = 0{,}0757(0{,}1786 - \sigma_{2,t}^2)dt + 0{,}1096\sigma_{2,t}^2 dW_{2t}.$$

Четвертый процесс — двухфакторная диффузия, примененная в Chernov, Gallant, Ghysels & Tauchen (2003), а также в Huang & Tauchen (2003):

$$\sigma_t = exp(-1{,}2 + 0{,}04\sigma_{1,t}^2 + 1{,}5\sigma_{2,t}^2),$$
$$d\sigma_{1,t}^2 = -0{,}00137\sigma_{1,t}^2 dt + dW_{1t},$$
$$d\sigma_{2,t}^2 = -1{,}386\sigma_{2,t}^2 dt + (1 + 0{,}25\sigma_{2,t}^2)dW_{2t}.$$

Наша базовая модель из уравнения (6) фиксирует $\mu_t = 0$ и $\rho_1 = \rho_2 = 0$, подразумевая, что $d\log S_t = \sigma_t dW_{3t}$ и ни дрейф, ни эффект рычага не присутствуют. Однако мы также проверяем свойства робастности, нарушая данные предположения и рассматривая присутствие дрейфа и эффекта рычага.

Следуя работе Gonçalves & Meddahi (2009), возьмем следующие значения параметров. Для однофакторных моделей диффузии положим $\mu = 0{,}0314$, $\rho_1 = -0{,}576$ и $\rho_2 = 0$. Для двухфакторных моделей диффузии мы положим $\mu = 0{,}03$, $\rho_1 = -0{,}03$ и $\rho_2 = -0{,}03$.

Наше внимание будет сосредоточено на двухсторонних симметричных 95% t-процентильных доверительных интервалах. Для построения таких интервалов мы будем использовать доступную асимптотическую теорию (CLT), iid бутстрап (iidB) и дикий бутстрап (WB), предложенные Gonçalves & Meddahi (2009). Также мы включим два типа блочного бутстрапа — неперекрывающийся (BB1) и перекрывающийся (BB2), и GARCH бутстрап с выбором остатков из нормального распределения.

Мы создали 10000 репликаций для четырех разных размеров выборки: $1/h = 1152$, 288, 48 и 12, соответствующие 1,25-минутным, 5-минутным, получасовым, и двухчасовым доходностям. Методы бутстрапирования основываются на 1000 бутстраповских повторах для каждой Монте–Карло репликации. Таблица 1 показывает долю покрытия для разных моделей диффузий и частот наблюдений. Доли покрытия меньшие, чем 95%, подразумевают свойство недостатка покрытия, т.е. модель строит слишком узкий доверительный интервал. Свойство излишнего покрытия подразумевает слишком широкий доверительный интервал.

Мы обобщаем результаты следующим образом: (1) При высокой частоте данных все подходы приблизительно эквивалентны. (2) При низкой частоте, асимптотический подход основанный на ЦПТ и iid бутстрап дают слегка узкие доверительные интервалы, WB дает слегка широкий доверительный интервал, а неперекрывающийся блочный бутстрап дает достаточно точный доверительный интервал, превосходя все остальные подходы. (3) Перекрывающийся блочный бутстрап превосходит остальные подходы на достаточно частотных данных и он более робастен, чем остальные подходы. (4) GARCH бутстрап демонстрирует хорошие результаты на симулированных данных при соответствующей модели с высокой частотностью. (5) GARCH бутстрап менее робастный, чем остальные подходы, в смысле отклонения от соответствующей модели, в частности, допуская дрейф и эффект рычага.

Таблица 1: Проценты покрытия номинального 95% доверительного интервала для интегрированной волатильности

	Нет рычага, нет дрейфа						Модель с рычагом и дрейфом					
$1/h$	CLT	iidB	WB	BB1	BB2	GB	CLT	iidB	WB	BB1	BB2	GB
$GARCH(1,1)$ диффузия												
12	85,0	92,8	98,4	95,6	94,6	95,4	80,4	90,5	98,0	94,5	93,2	96,0
48	92,1	94,9	98,4	96,3	95,3	95,4	91.3	93.5	98,0	95,3	94,6	95,8
288	94,3	95,2	97,2	95,8	95,8	95,2	94,4	95,1	96,5	95,8	95,7	95,5
1152	94,5	94,9	95,4	95,3	95,2	95,1	94,8	95,0	96,1	95,3	95,5	95,1
Логнормальная диффузия												
12	84,6	93,4	98,4	96,2	94,5	96,4	81,0	94,4	98,8	96,9	95,3	96,8
48	91,7	94,9	98,6	95,9	95,3	96,2	93,6	96,3	99,0	97,0	96,2	96,5
288	94,4	95,3	97,0	96,0	95,4	96,0	93,9	95,1	97,2	95,5	95,3	95,2
1152	95,2	95,3	96,1	95,4	95,6	95,2	94,8	95,0	96,2	95,3	95,4	95,2
Двухфакторная аффинная диффузия												
12	83,4	92,0	97,7	95,4	94,0	96,4	80,3	90,6	98,5	93,5	92,7	97,5
48	90,9	94,7	97,9	95,4	94,7	96,3	92,3	94,4	98,1	95,1	94,7	96,4
288	94,1	95,2	97,6	96,1	95,8	95,7	94,4	95,2	96,7	95,1	95,4	96,0
1152	94,7	95,2	96,1	95,7	95,3	95,3	95,8	95,5	95,7	95,4	95,2	95,4
Двухфакторная диффузия												
12	79,6	90,6	96,9	94,6	92,0	97,0	80,2	90,6	97,2	93,5	92,3	98,1
48	90,8	94,5	98,3	95,0	95,1	96,6	91,4	94,4	98,1	95,1	94,7	97,6
288	94,8	95,2	97,1	95,8	95,6	95,5	94,4	95,2	97,0	96,0	95,5	96,5
1152	95,1	95,4	95,6	95,2	95,0	95,1	95,7	95,0	95,8	95,5	95,4	95,5

Замечание: Таблица показывает доли покрытия для разных моделей диффузии и частот данных. Доля покрытия меньшая, чем 95%, подразумевает свойство недостатка покрытия, т.е. подход дает слишком узкий доверительный интервал. Свойство излишнего покрытия подразумевает слишком широкий доверительный интервал. Все методы бутстрапа построены на 1000 бутстраповских повторах для каждой из 10000 симуляций Монте–Карло.

6 Заключение

В данной статье мы предлагаем два бутстрап подхода для реализованной волатильности с целью построения доверительных интервалов для интегрированной волатильности. Один из них это блочный бутстрап, другой — это GARCH бутстрап. В работе показано, что применение блочного бутстрапа для реализованной волатильности позволяет достичь более точной инференции интегрированной волатильности. При низкой частоте данных неперекрывающийся блочный бутстрап превосходит асимптотическую теорию, iid бутстрап и дикий бутстрап, описанные Gonçalves & Meddahi (2004). Более того, перекрывающийся бутстрап превосходит другие подходы на достаточно частотных данных и он более робастен, чем остальные подходы. Более сложный GARCH бутстрап основан на предположении присутствия условной гетероскедастичности в данных в форме GARCH. Он демонстрирует хорошую инференцию на симулированных данных при соответствующей модели и превосходит другие подходы. Однако, он менее робастен, вычислительно сложен и требует высокой частотности для достижения хороших результатов.

Благодарности

Автор выражает благодарность своему научному руководителю, профессору РЭШ Игорю Хейфецу за ценные замечания, а также руководителю исследовательского проекта РЭШ профессору Станиславу Анатольеву за организацию научной работы. Автор также благодарен своим коллегам по исследованиям в РЭШ — Ренату Хабибуллину, Людмиле Кауровой, Олегу Грошеву и Игорю Верещетину — за плодотворные обсуждения. Автор несет ответственность за все возможные ошибки и неточности.

Приложение: доказательства результатов разделов 3 и 4

Доказательство Предложения 1.

(i) Благодаря независимости бутстраповских блоков и дискретной структуре бутстраповской реализованной волатильности по вероятностной мере, индуцированной бутстрапом, мы имеем:

$$\mathbb{E}^* \left(RV_t^* - RV_t\right) = \mathbb{E}^* \left(\sum_{k=1}^{T/l} RV_t^{(k)*} - RV_t\right) = \sum_{k=1}^{T/l} \left(\mathbb{E}^* \left(RV_t^{(k)*}\right)\right) - RV_t$$

$$= \sum_{k=1}^{T/l} \sum_{i=1}^{T/l} \left(\left(\frac{T}{l}\right)^{-1} RV_t^{(i)}\right) - RV_t = \sum_{i=1}^{T/l} RV_t^{(i)} - RV_t = RV_t - RV_t = 0$$

для любых h и t.

(ii) Благодаря независимости бутстраповских блоков,

$$v_t^{(h)*2} = \mathbb{V}^* \left(\sqrt{h^{-1}} RV_t^*\right) = h^{-1} \mathbb{V}^* \left(\sum_{i=1}^{T/l} RV_t^{(i)*}\right) = h^{-1} \sum_{i=1}^{T/l} \mathbb{V}^* \left(RV_t^{(i)*}\right)$$

$$= h^{-1} \left(\frac{T}{l}\right) \mathbb{V}^* \left(RV_t^{(i)*}\right). \tag{12}$$

Дисперсия в бутстраповской мере

$$\mathbb{V}^* \left(RV_t^{(i)*}\right) = \mathbb{E}^* \left(RV_t^{(i)*2}\right) - \mathbb{E}^* \left(RV_t^{(i)*}\right)^2$$

$$= \left((RV_t^{(1)})^2 + (RV_t^{(2)})^2 + ... + (RV_t^{(i)})^2\right) \left(\frac{T}{l}\right)^{-1} - (RV_t)^2 \left(\frac{T}{l}\right)^{-2}.$$

Подставляя ее в уравнение (12), мы получаем желаемый результат. ■

Доказательство Предложения 2.

(i) Для простоты мы рассмотрим GARCH бутстрап с выборкой остатков из нормального распределения. Таким образом, легко подсчитать

$$\mathbb{E}^*(\xi^{*2}_{t-1+ih}) = 1,$$

и

$$\mathbb{E}^*(\xi^{*4}_{t-1+ih}) = 3.$$

Так как волатильность в определенный момент времени зависит только от предыдущих шоков, а все шоки независимы, в любой момент времени ψ^*_{t-1+ih} и ξ^*_{t-1+ih} будут независимы. Следовательно,

$$\mathbb{E}^*(RV_t^*) = \mathbb{E}^*\left(\sum_{i=1}^{1/h} r^{*2}_{t-1+ih}\right) = \mathbb{E}^*\left(\sum_{i=1}^{1/h} \psi^*_{t-1+ih}\xi^{*2}_{t-1+ih}\right) = \sum_{i=1}^{1/h} \mathbb{E}^*(\psi^*_{t-1+ih}). \tag{13}$$

Используя метод прямой индукции, мы можем с легкостью оценить все компоненты в уравнении (13):

$$\begin{aligned}
\mathbb{E}^*(\psi^*_{t-1+h}) &= \hat\beta_0 + \hat\beta_1\psi^*_{t-1+0h} + \hat\beta_2\mathbb{E}^*(\xi^{*2}_{t-1+0h})\psi^*_{t-1+0h} = \hat\beta_0 + \psi^*_{t-1+0h}(\hat\beta_1+\hat\beta_2),\\
\mathbb{E}^*(\psi^*_{t-1+2h}) &= \hat\beta_0 + \hat\beta_1\mathbb{E}^*(\psi^*_{t-1+h}) + \hat\beta_2\mathbb{E}^*(\xi^{*2}_{t-1+0h})\mathbb{E}^*(\psi^*_{t-1+h})\\
&= \hat\beta_0 + \mathbb{E}^*(\psi^*_{t-1+h})(\hat\beta_1+\hat\beta_2),\\
\mathbb{E}^*(\psi^*_{t-1+3h}) &= \hat\beta_0 + \mathbb{E}^*(\psi^*_{t-1+2h})(\hat\beta_1+\hat\beta_2),\\
&\dots\\
\mathbb{E}^*(\psi^*_{t-1+ih}) &= \hat\beta_0 + \mathbb{E}^*(\psi^*_{t-1+(i-1)h})(\hat\beta_1+\hat\beta_2).
\end{aligned}$$

Шаг за шагом легко вычислить каждую компоненту в уравнении (13). В конечном итоге, складывая все вышеуказанные выражения, бутстраповское ожидание реализованной волатильности можно рассчитать численно.

(ii) В данной части мы рассмотрим 2 возможных случая отдельно: когда бутстраповские доходности независимы условно на процессе волатильности, и более общий случай, когда бутстраповские доходности зависимы.

Независимые доходности. В силу независимости бутстраповских доходностей, бутстраповская дисперсия может быть представлена следующим образом:

$$\begin{aligned}
\mathbb{V}^*\left(\sqrt{h^{-1}}RV_t^*\right) &= h^{-1}\mathbb{V}^*\left(\sum_{i=1}^{1/h} r^{*2}_{t-1+ih}\right) = h^{-1}\sum_{i=1}^{1/h}\mathbb{V}^*\left(\psi^*_{t-1+ih}\xi^{*2}_{t-1+ih}\right)\\
&= h^{-1}\sum_{i=1}^{1/h}\left(\mathbb{E}^*(\psi^{*2}_{t-1+ih})\mathbb{E}^*(\xi^{*4}_{t-1+ih}) - (\mathbb{E}^*\psi^*_{t-1+ih})^2(\mathbb{E}^*\xi^{*2}_{t-1+ih})^2\right)\\
&= h^{-1}\sum_{i=1}^{1/h}\left(3\mathbb{E}^*\left(\psi^{*2}_{t-1+ih}\right) - \left(\mathbb{E}^*\psi^*_{t-1+ih}\right)^2\right)\\
&= 3h^{-1}\sum_{i=1}^{1/h}\mathbb{E}^*\left(\psi^{*2}_{t-1+ih}\right) - h^{-1}\sum_{i=1}^{1/h}\left(\mathbb{E}^*\psi^*_{t-1+ih}\right)^2.
\end{aligned}$$

Каждая компонента во второй сумме может быть рассчитана по аналогии с уравнением (13). Первую сумму можно рассчитать, используя следующую технику.

$$\begin{aligned}
\psi^{*2}_{t-1+ih} &= \hat\beta_0^2 + \left(\hat\beta_1\psi^*_{t-1+(i-1)h}\right)^2 + \left(\hat\beta_2\psi^*_{t-1+(i-1)h}\xi^{*2}_{t-1+(i-1)h}\right)^2 + 2\hat\beta_0\hat\beta_1\psi^*_{t-1+(i-1)h}\\
&\quad + 2\hat\beta_0\hat\beta_2\psi^*_{t-1+(i-1)h}\xi^{*2}_{t-1+(i-1)h} + 2\hat\beta_1\hat\beta_2\psi^{*2}_{t-1+(i-1)h}\xi^{*2}_{t-1+(i-1)h}\\
&= \hat\beta_0^2 + \psi^{*2}_{t-1+(i-1)h}\left(\hat\beta_1 + \hat\beta_2\xi^{*2}_{t-1+(i-1)h}\right)^2 + \psi^*_{t-1+(i-1)h}\left(2\hat\beta_0\hat\beta_1 + 2\hat\beta_0\hat\beta_2\xi^{*2}_{t-1+(i-1)h}\right),
\end{aligned}$$

$$\begin{aligned}\mathbb{E}^*\psi^{*2}_{t-1+h} &= \hat\beta_0^2 + \psi^{*2}_{t-1+0h}\left(\hat\beta_1^2 + 3\hat\beta_2^2 + 2\hat\beta_1\hat\beta_2\right) + \psi^*_{t-1+0h}\left(2\hat\beta_0\hat\beta_1 + 2\hat\beta_0\hat\beta_2\right)\\ &= \hat\beta_0^2 + \psi^{*2}_{t-1+0h}\Omega + 2\psi^*_{t-1+0h}\hat\beta_0\left(\hat\beta_1 + \hat\beta_2\right),\end{aligned}$$

где

$$\Omega = \hat\beta_1^2 + 3\hat\beta_2^2 + 2\hat\beta_1\hat\beta_2.$$

Далее,

$$\begin{aligned}&\mathbb{E}^*\psi^*_{t-1+h} = \hat\beta_0 + \psi^*_{t-1+0h}\left(\hat\beta_1 + \hat\beta_2\right),\\ &\mathbb{E}^*\psi^{*2}_{t-1+2h} = \hat\beta_0^2 + \mathbb{E}^*\psi^{*2}_{t-1+h}\Omega + 2\mathbb{E}^*\psi^*_{t-1+h}\hat\beta_0\left(\hat\beta_1 + \hat\beta_2\right),\\ &\mathbb{E}^*\psi^*_{t-1+2h} = \hat\beta_0 + \mathbb{E}^*\psi^*_{t-1+h}\left(\hat\beta_1 + \hat\beta_2\right),\\ &\mathbb{E}^*\psi^{*2}_{t-1+3h} = \hat\beta_0^2 + \mathbb{E}^*\psi^{*2}_{t-1+2h}\Omega + 2\mathbb{E}^*\psi^*_{t-1+2h}\hat\beta_0\left(\hat\beta_1 + \hat\beta_2\right),\\ &\mathbb{E}^*\psi^*_{t-1+3h} = \hat\beta_0 + \mathbb{E}^*\psi^*_{t-1+2h}\left(\hat\beta_1 + \hat\beta_2\right),\\ &\ldots\\ &\mathbb{E}^*\psi^{*2}_{t-1+ih} = \hat\beta_0^2 + \mathbb{E}^*\psi^{*2}_{t-1+(i-1)h}\Omega + 2\mathbb{E}^*\psi^*_{t-1+(i-1)h}\hat\beta_0\left(\hat\beta_1 + \hat\beta_2\right),\\ &\mathbb{E}^*\psi^*_{t-1+ih} = \hat\beta_0 + \mathbb{E}^*\psi^*_{t-1+(i-1)h}\left(\hat\beta_1 + \hat\beta_2\right).\end{aligned}$$

Шаг за шагом мы можем подсчитать каждую компоненту в вышеуказанной сумме. Таким образом, бутстраповская дисперсия реализованной волатильности может быть подсчитана численно.

Зависимые доходности. Так как ψ^*_{t-1+ih} зависимы между собой, бутстраповскую дисперсию можно представить как:

$$\begin{aligned}\mathbb{V}^*\left(\sqrt{h^{-1}}RV_t^*\right) &= h^{-1}\mathbb{V}^*\left(\sum_{i=1}^{1/h} r^{*2}_{t-1+ih}\right) = h^{-1}\mathbb{V}^*\left(\sum_{i=1}^{1/h}\psi^*_{t-1+ih}\xi^{*2}_{t-1+ih}\right)\\ &\neq h^{-1}\sum_{i=1}^{1/h}\mathbb{V}^*\left(\psi^*_{t-1+ih}\xi^{*2}_{t-1+ih}\right).\end{aligned}$$

Следовательно, мы не можем применить метод прямой индукции в данном случае, как мы делали ранее. Альтернативный подход — подсчитать бутстраповскую дисперсию с помощью рекурсивного алгоритма. Введем функцию

$$f_k = \mathbb{V}^*\left(\sum_{i=1}^{k} r^{*2}_{t-1+ih}\right) = \mathbb{V}^*\left(\sum_{i=1}^{k}\psi^*_{t-1+ih}\xi^{*2}_{t-1+ih}\right).$$

Цель алгоритма — это найти f_{k+1} как функцию от f_k:

$$\begin{aligned}f_{k+1} &= \mathbb{V}^*\left(\sum_{i=1}^{k} r^{*2}_{t-1+ih} + \hat\beta_0 + \hat\beta_1\psi^*_{t-1+kh} + \hat\beta_2\psi^*_{t-1+kh}\xi^{*2}_{t-1+kh}\right)\\ &= f_k + \mathbb{V}^*\left(\psi^*_{t-1+kh}(\hat\beta_1 + \hat\beta_2\xi^{*2}_{t-1+kh})\right)\end{aligned}$$

$$+\,2\,\mathbb{C}^*\left(\sum_{i=1}^{k}\psi^*_{t-1+ih}\xi^{*2}_{t-1+ih},\psi^*_{t-1+kh}(\hat\beta_1+\hat\beta_2\xi^{*2}_{t-1+kh})\right). \tag{14}$$

Подсчитаем каждую компоненту в уравнении (14) отдельно. Из GARCH модели мы знаем представление функции гетероскедастичности

$$\psi^*_{t-1+kh}=\hat\beta_0+\psi^*_{t-1+(k-1)h}(\hat\beta_1+\hat\beta_2\xi^{*2}_{t-1+(k-1)h}).$$

Возведя в квадрат обе части, мы с легкостью можем вывести следующее выражение:

$$\psi^{*2}_{t-1+kh}=\hat\beta_0^2+\psi^{*2}_{t-1+(k-1)h}\left(\hat\beta_1^2+2\hat\beta_1\hat\beta_2\xi^{*2}_{t-1+(k-1)h}+\hat\beta_2^2\xi^{*4}_{t-1+(k-1)h}\right).$$

Из условия стационарности

$$\mathbb{E}^*\psi^{*2}_{t-1+kh}\left(1-\hat\beta_1^2-2\hat\beta_1\hat\beta_2-3\hat\beta_2^2\right)=\hat\beta_0^2+2\hat\beta_0\mathbb{E}^*\psi^*_{t-1+kh}(\hat\beta_1+\hat\beta_2).$$

Полагая $\mathbb{E}^*\psi^*_{t-1+kh}$ значением долгосрочной дисперсии гетероскедастичности

$$\mathbb{E}^*\psi^*_{t-1+kh}=\frac{\hat\beta_0}{1-\hat\beta_1-\hat\beta_2},$$

мы получаем выражение для $\mathbb{E}^*\psi^{*2}_{t-1+kh}$:

$$\mathbb{E}^*\psi^{*2}_{t-1+kh}=\hat\beta_0^2+2\hat\beta_0\frac{\hat\beta_0}{1-\hat\beta_1-\hat\beta_2}(\hat\beta_1+\hat\beta_2). \tag{15}$$

Таким образом, дисперсию в уравнении (14) можно представить как

$$\mathbb{V}^*\left(\psi^*_{t-1+kh}(\hat\beta_1+\hat\beta_2\xi^{*2}_{t-1+kh})\right)=\left(\hat\beta_1^2+2\hat\beta_1\hat\beta_2+3\hat\beta_2^2\right)\mathbb{E}^*\psi^{*2}_{t-1+kh}-(\hat\beta_1+\hat\beta_2)\frac{\hat\beta_0}{1-\hat\beta_1-\hat\beta_2}$$

$$=\left(\hat\beta_1^2+2\hat\beta_1\hat\beta_2+3\hat\beta_2^2\right)\left(\hat\beta_0^2+2\hat\beta_0\frac{\hat\beta_0}{1-\hat\beta_1-\hat\beta_2}(\hat\beta_1+\hat\beta_2)\right)-(\hat\beta_1+\hat\beta_2)\frac{\hat\beta_0}{1-\hat\beta_1-\hat\beta_2}.$$

Распишем ковариационный член в уравнении (14), используя определение ковариации:

$$\mathbb{C}^*\left(\sum_{i=1}^{k}\psi^*_{t-1+ih}\xi^{*2}_{t-1+ih},\psi^*_{t-1+kh}(\hat\beta_1+\hat\beta_2\xi^{*2}_{t-1+kh})\right)=(\hat\beta_1+\hat\beta_2)\mathbb{C}^*\left(\sum_{i=1}^{k}\psi^*_{t-1+ih}\xi^{*2}_{t-1+ih},\psi^*_{t-1+kh}\right)$$

$$=(\hat\beta_1+\hat\beta_2)\left(\mathbb{E}^*\psi^{*2}_{t-1+kh}+\mathbb{E}^*\left(\psi^*_{t-1+kh}\sum_{i=1}^{k-1}\psi^*_{t-1+ih}\xi^{*2}_{t-1+ih}\right)-k\left(\frac{\hat\beta_0}{1-\hat\beta_1-\hat\beta_2}\right)^2\right).$$

Используя уравнение (15) и численную оценку $\mathbb{E}^*\left(\psi^*_{t-1+kh}\sum_{i=1}^{k-1}\psi^*_{t-1+ih}\xi^{*2}_{t-1+ih}\right)$, мы можем вывести выражение для ковариационного члена и, следовательно, функцию рекурсии. Если прогнать рекурсивный алгоритм начиная с $k=1$ до $k=1/h$, мы сможем численно получить бутстраповскую дисперсию реализованной волатильности. ■

Список литературы

Andersen, T.G., L. Benzoni & J. Lund (2002). An empirical investigation of continuous-time equity return models. *Journal of Finance* 57, 1239–1284.

Andersen, T.G. & T. Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns. *Journal of Finance* 52, 975–1005.

Andersen, T.G. & T. Bollerslev (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39, 885–905.

Andersen, T.G., T. Bollerslev & F.X. Diebold (2009). Parametric and nonparametric measurements of volatility. Глава 2 в *Handbook of Financial Econometrics* под редакцией Y.Aït-Sahalia & L.P. Hansen, том 1, North Holland: Amsterdam.

Andersen, T.G., T. Bollerslev, F.X. Diebold & P. Labys (2000). Exchange rate returns standardized by realized volatility are (nearly) gaussian. *Multinational Finance Journal* 4, 159–79.

Andersen, T.G., T. Bollerslev, F.X. Diebold & P. Labys (2001a). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96, 42–55.

Andersen, T.G., T. Bollerslev, F.X. Diebold & P. Labys (2001b). Modeling and forecasting realized volatility. *NBER Working Paper 8160*, Cambridge, MA.

Andersen, T.G., T. Bollerslev, F.X. Diebold & P. Labys (2001c). The distribution of realized stock return volatility. *Journal of Financial Economics* 61, 43–76.

Andersen, T.G., T. Bollerslev, F.X. Diebold & P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica* 71, 599–625.

Andersen, T.G., T. Bollerslev & N. Meddahi (2004). Correcting the errors: volatility forecast evaluation using high-frequency data and realized volatilities. *Econometrica* 37, 279–296.

Andreou E. & E. Ghysels (2002). Rolling-sampling volatility estimators: some new theoretical, simulation and empirical results. *Journal of Business & Economic Statistics* 20, 363–376.

Andrews W.K. (2004). The block-block bootstrap: improved asymptotic refinements. *Econometrica* 72, 673–700.

Barndorff-Nielsen, O.E. & N. Shephard (2002a). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society* 64, 253–280.

Barndorff-Nielsen, O.E. & N. Shephard (2002b). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics* 17, 457–477.

Barndorff-Nielsen, O.E. & N. Shephard (2003). Realised power variation and stochastic volatility models. *Bernoulli* 9, 243–265.

Barndorff-Nielsen, O.E. & N. Shephard (2004a). Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics. *Econometrica* 72, 885–925.

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

Barndorff-Nielsen, O.E. & N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4, 1–30.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 37, 307–327.

Bollerslev, T. & H. Zhou (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics* 109, 33–65.

Chernov, M., R. Gallant, E. Ghysels & G. Tauchen (2003). Alternative models for stock price dynamics. *Journal of Econometrics* 116, 225–257.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987–1007.

Gonçalves, S. & L. Kilian (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics* 123, 89–120.

Gonçalves, S. & N. Meddahi (2004). Supplement to bootstrapping realized volatility. Manuscript, Université de Montréal.

Gonçalves, S. & N. Meddahi (2009). Bootstrapping realized volatility. *Econometrica* 77, 283–306.

Hall, P. (1992). *The bootstrap and Edgeworth expansion.* Springer-Verlag: New York.

Hall, P., J.L. Horowitz & B. Jing (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* 82, 561–574.

Huang, X. & G. Tauchen (2004). The relative contribution of jumps to total price variance. Manuscript, Duke University.

Kreiß, J.-P. & J. Franke (1992). Bootstrapping stationary autoregressive moving-average models. *Journal of Time Series Analysis* 13, 297–317.

Lahiri S.N. (1999). Theoretical comparisons of block bootstrap methods. *The Annals of Statistics* 27, 386–404.

Lee, S.-W. & B.E. Hansen (1994). Asymptotic theory for the GARCH(1,1) quasimaximum likelihood estimator. *Econometric Theory* 10, 29–52.

Maercker, G. (1997). Statistical inference in conditional heteroskedastic autoregressive models. *PhD Dissertation*, Technische University Braunschweig.

Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics* 17, 475–508.

Pascual, L., J. Romo & E. Ruiz (2005). Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics and Data Analysis* 50, 2293–2312.

Reeves, J.J. (2005). Bootstrap prediction intervals for ARCH models. *International Journal of Forecasting* 21, 237–248.

Robio, P.O. (1999). Forecast intervals in ARCH models: bootstrap versus parametric methods. *Applied Economics Letters* 6, 323–327.

Wu, C.F.J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics* 14, 1261–1295.

Bootstrap inference about integrated volatility

Andrey Rafalson

Barclays, Equity Derivatives Group, London, UK

We extend the work of Gonçalves & Meddahi (2009) who suggest using the iid and wild bootstrap for realized volatility instead of the asymptotic approach in order to estimate integrated volatility. We propose the block bootstrap and GARCH residual bootstrap approaches motivated by the persistence of the intraday term structure of returns. Using Monte Carlo simulations we show that the block bootstrap is more accurate for a low intraday frequency, more robust and valid. Another result is that the GARCH bootstrap outperforms others when the data imply strong persistence in conditional heteroskedasticity. It also demonstrates good inference on simulated data along the baseline model with a high frequency. However, the GARCH bootstrap is more computationally costly and less robust than the others.

Keywords: integrated volatility, realized volatility, block bootstrap, GARCH bootstrap
JEL Classification: C12, C22, C58