

ДМИТРИЙ БЕЛИКОВ

**Объединение прогнозов волатильности и
направленности рынка акций: макроэкономические
переменные как основные индикаторы**

Лондонская школа экономики
Бакалаврская работа, 2021

Аннотация

В этой статье я расширяю модель GARCH-MIDAS, чтобы включить переключение режимов на основе макроэкономических переменных. Я изучаю влияние макропеременных на прогнозируемость режима и волатильности фондового рынка. Прогнозирование рыночного режима дополняется нелинейной зависимостью от длительности режима. Я также предлагаю новый метод идентификации рыночного режима. Эмпирические результаты подтверждают выводы предыдущих работ и демонстрируют значительное улучшение точности прогнозов волатильности.

Содержание

1	Введение	3
2	Обзор литературы	4
2.1	Определение и прогнозирование направления	4
2.2	Прогнозирование волатильности	6
3	Эконометрическая постановка	7
3.1	Определение направления	8
3.2	Прогнозирование направления	8
3.3	Прогнозирование волатильности	9
3.4	Объединение двух подходов	10
4	Данные	12
5	Эмпирические результаты	14
6	Заключение	18
7	Приложение	22
7.1	Пики и впадины при ЕМА(500)	22
7.2	Тест ADF	22

1 Введение

Прогнозирование волатильности рынка давно занимает финансовую литературу — из-за применения в оценке опционов и в различных мерах риска. При этом большинство моделей волатильности эндогенны и не включают внешние переменные, особенно макроэкономические, из-за несовпадения частот: цены на рынке доступны практически с любой частотой, тогда как макроэкономические показатели в лучшем случае недельные, а чаще месячные или квартальные. В то же время финансовая волатильность, как отмечал ряд исследователей (см. раздел «Обзор литературы»), берёт начало в реальной экономике и в фундаментальной неопределённости. Приходя к таким выводам, эти авторы использовали рыночные данные месячной частоты, что делает их результаты малоприменимыми для практики.

Положение изменилось с появлением моделей типа GARCH-MIDAS, которые позволили включать в модели волатильности данные разной частоты. Такие модели оказались лучше стандартных, но им по-прежнему не хватало важного свойства, широко описанного в прежних работах, — переключения режимов волатильности. У бычьего и медвежьего состояний рынка разная безусловная волатильность. За единственной попыткой встроить переключение режимов в GARCH-MIDAS (Pan et al., 2017) вопрос так и не был разобран всерьёз. Более того, в той работе вероятность не учитывает ряд важных особенностей, в частности зависимость от длительности режима. В этой статье я представляю GARCH-MIDAS с переключением режимов, где режим рынка прогнозируется по макропеременным. Я сравниваю эту модель с моделью без переключения и обнаруживаю, что первая значительно превосходит вторую.

Вторая цель работы — проверить прогнозирование режима на новых данных. Весной 2020 года, когда все показатели реальной экономики резко падали, фондовый рынок словно не заметил глобального спада. Объяснений предлагалось много, но ни одна формальная модель это расхождение не объясняла. В общем виде цены акций опираются на реальную экономику напрямую: стоимость отдельной акции определяется будущими дивидендами (дисконтированными по подходящей ставке), дивиденды выплачиваются из прибыли, а прибыль отражает результат работы фирмы, то есть её вклад в добавленную стоимость экономики. Казалось бы, рынок обязан отражать изменения в реальной экономике. Однако данные показывают, что связь реального и финансового секторов скорее косвенная, — хотя нет сомнений, что в долгосрочной перспективе все финансовые активы окажутся в реальной экономике. Поэтому я хочу проверить, можно ли было предсказать падение рынка в марте 2020 года и последующее восстановление по макропеременным и согласуется ли оно с ними. Ещё один, меньший вклад этой работы — включение нелинейной зависимости от длительности в динамику состояния рынка.

Прогнозировать состояние рынка полезно по нескольким причинам. Во-первых, инвесторам и управляющим фондами это нужно для стратегии выбора момента входа — перехода из рискованных активов в безрисковые в зависимости от состояния рынка. Во-вторых, это пригодилось бы регуляторам при планировании мер по поддержанию финансовой устойчивости.

Работа построена так: в разделе 2 дан обзор литературы с упором на недавние продвижения в моделях прогнозирования состояния рынка и волатильности, в разделе 3 излагается

эконометрическая постановка, где эти две задачи объединяются, в разделе 4 приводятся результаты оценивания численной оптимизацией, раздел 5 подводит итоги.

2 Обзор литературы

Обзор разбит на две части. В первой рассматриваются работы о предсказуемости режима фондового рынка. Во второй — существующие модели прогнозирования рыночной волатильности.

2.1 Определение и прогнозирование направления

За последние три десятилетия согласия о том, что именно называть бычьим и медвежьим рынком, так и не сложилось. Обычно встречаются общие определения вроде *«продолжительных периодов снижения и роста рыночных цен»* (Chauvet and Potter, 2000). Большинство авторов пользуется тем или иным вариантом этого определения, но есть и более строгие формулировки. Например, Sperandeo (1990) определяет режимы рынка точнее: *«долгосрочное движение цены вверх/вниз, характеризующееся серией более высоких/низких промежуточных максимумов/минимумов, прерываемой серией более высоких/низких промежуточных минимумов/максимумов»*, что по существу соответствует определению *тренда* в техническом анализе. Беда таких определений, как я покажу дальше, в том, что они неявно предполагают определение текущего состояния с заглядыванием вперёд: нынешнее состояние зависит от того, окажется ли следующий максимум/минимум выше или ниже предыдущего. Поэтому прогноз на коротком горизонте оказывается невозможен. В этой работе под бычьим и медвежьим рынком понимается соответственно положительная и отрицательная средняя доходность рынка.

Хотя за последние два десятилетия работ об идентификации бычьих и медвежьих рынков стало больше, их по-прежнему немного. В литературе встречаются два основных подхода: параметрический марковский с переключением режимов и полупараметрический. У марковских моделей есть преимущество, отмеченное в том числе Kole and van Dijk (2017): такая модель задаёт единый порождающий процесс, пригодный и для идентификации режима, и для прогноза, тогда как полупараметрические модели требуют отдельного шага параметрического прогнозирования. Kole and van Dijk (2017) проводят сравнительное исследование марковских и полупараметрических моделей и приходят к выводу, что полупараметрические лучше работают внутри выборки, а марковские — вне её. Причина в том, что марковские модели учитывают волатильность, а полупараметрические — только среднюю доходность.

Марковская модель с переключением режимов широко применялась для деления рынка на состояния. Отправная точка — работа Hamilton (1989), предложившего спецификацию марковской модели переключения, которая эндогенно разделяет рынок на режимы и допускает изменение условного среднего и дисперсии во времени. Maheu and McCurdy (2000) вводят в модель зависимость от длительности: вероятность смены режима зависит от того, сколько режим уже длится. Они сообщают, что волатильность в медвежьем состоянии растёт с длительностью. В бычьем состоянии с длительностью растёт доходность,

то есть вероятность увидеть конец серии положительных доходностей с длительностью падает.

Полупараметрический подход отличается от марковского тем, что идентификация и прогнозирование в нём разделены. Отсюда и название — *полупараметрический*. Сначала выполняется непараметрическая идентификация, затем строится модель бинарного выбора для прогноза. Поскольку шаги не связаны, рассмотрим их по отдельности.

Есть несколько основных методик непараметрического разделения рынка на режимы. Изначально это была задача моделирования делового цикла — выделение рецессий и спадов в реальной экономике. Позже часть этих приёмов применили к режимам рынка. Например, известный алгоритм Bru and Boschian (1971), использованный и у Nyberg (2013). Подход сочетает две скользящие средние, кривую Спенсера¹ и различные процедуры фильтрации. Nyberg (2013) показывает, что состояние рынка предсказуемо лучше, чем доходность, на горизонте от 1 до 12 месяцев. Он использует типичный для таких исследований набор данных: SP500 как прокси рынка, доходности казначейского рынка (спреды временной структуры, ставки по федеральным фондам, доходности постоянной срочности), уровень безработицы, промышленное производство и инфляцию. И здесь спред временной структуры оказался самым значимым предиктором. Главный вклад его работы — введение динамики в прежде статические пробит-модели. Это заметно повысило точность прогноза, а наиболее значимой предсказывающей переменной оказался временной спред.

Два похожих и широко используемых подхода к идентификации режимов предложены Pagan and Sossounov (2003) и Lunde and Timmerman (2004); далее — PS и LT соответственно. Правило PS требует, чтобы состояние (бычье или медвежье) имело определённую длительность и величину изменения цен, тогда как LT учитывает только амплитуду. Как формулируют LT,

«...рынок переключается из бычьего состояния в медвежье, если цены акций снизились на определённый процент от предыдущего (локального) пика внутри бычьего состояния».

Недостаток этого подхода в том, что он опирается на параметр, заглядывающий вперёд². Отсюда временной лаг, обычно немалый, и малая пригодность для прогноза вне выборки. Определение цикла с заглядыванием вперёд надёжнее и вполне полезно в анализе делового цикла, но в определении состояния рынка мы себе этого позволить не можем: смена тренда происходит за дни. Например, индекс S&P500 упал с 3393 пунктов 16.02.2020 до 2280 пунктов 15.03.2020 — потеря 32% за месяц. Скорость реакции здесь принципиальна.

Любопытно, что Lunde and Timmerman (2004) упоминают применение технических правил для определения трендов³, замечая, что горизонты в таких правилах довольно коротки.

¹Скользящая средняя со сложной схемой весов: наибольший вес в центре и отрицательные веса по краям. Такая кривая близко следует за данными.

²Строго говоря, любой подход, основанный на выделении пиков и впадин, заглядывает вперёд: текущее состояние рынка зависит от того, окажется ли следующий локальный минимум/максимум выше или ниже предыдущего.

³Из LT: «Технические торговые правила ищут закономерности в ценах при обычно довольно коротком горизонте. Например, значение 100-дневной скользящей средней цен можно сравнить со значением 25-дневной. Мы же не обуславливаем анализ временем конкретного движения, а прямо рассматриваем его как случайную величину, распределение которой нас и интересует».

Как я уже говорил, тренды и правда бывают короткими — как в феврале-марте 2020 года, — и всё же это был полноценный, экономически обоснованный тренд. Это существенно, поскольку многие авторы исходят из некоторой минимальной длительности фазы рынка. Например, и Nyberg (2013), и Chen (2009), как и Candelon et al. (2008), берут минимальную длительность состояния в 6 месяцев. Понятно, что такое ограничение лишает нас возможности выделять движения рынка на более высоких частотах — например падения 2016 или 2018 года.

В качестве процедуры идентификации Chen (2009) рассматривает то, что называет наивным подходом. Он берёт простую скользящую среднюю доходностей с фиксированным периодом сглаживания⁴. Когда средняя доходность за $N \geq 1$ периодов положительна, состояние рынка в данном периоде считается бычьим, когда отрицательна — медвежьим. Он не находит значимой разницы в точности прогноза при переходе от марковской модели к полупараметрической с наивной идентификацией. По его результатам инфляция и спред процентных ставок — самые значимые предикторы состояния рынка, а денежные агрегаты и валютные курсы оказались слабыми предикторами.

В целом все авторы используют схожие наборы данных и, применяя разные методики и подходы, приходят к близким выводам: состояние рынка предсказуемо лучше, чем доходность, а самым полезным предиктором оказывается спред процентных ставок.

2.2 Прогнозирование волатильности

Если в случае режима рынка самым сложным было однозначно и последовательно его определить, то рыночная волатильность (реализованная волатильность) наблюдаема напрямую — это квадраты доходностей, — и основное внимание уделяется процедурам прогнозирования. Обзор литературы по теме можно начать с Schwert (1989), хотя были и более ранние работы, например Officer (1973). Schwert (1989) изучает влияние различных макроэкономических переменных — инфляции, базовой ставки, финансового рычага, денежных агрегатов, — а также их дисперсий на волатильность фондового рынка. Большинство коэффициентов оказались положительными, часть — надёжно выше нуля. Финансовый рычаг влиял слабо. Schwert (1989) показывает, что доходности финансовых активов более волатильны во время рецессий в реальной экономике, и объясняет это эффектом операционного рычага. Финансовая волатильность берёт начало в реальной экономике, но свидетельств того, что макроэкономическая волатильность способна предсказывать рыночную, немного. Позднейшие авторы, впрочем, продвинулись в сложных прогнозных моделях и показали обратное.

Engle and Rangel (2008) предложили модель Spline-GARCH, где дисперсия состоит из двух компонент: детерминированной и возвращающейся к среднему дневной GARCH. Такое разделение позволило безусловной дисперсии меняться во времени. Engle and Rangel обнаруживают, что волатильность макроэкономических факторов (они берут рост ВВП, инфляцию и краткосрочные процентные ставки) — важные объясняющие переменные для волатильности фондового рынка. Engle et al. (2013) идут дальше и предлагают развитие

⁴Скользящая средняя (MA) в смысле технического анализа. Это не то же самое, что эконометрическое скользящее среднее.

spline-GARCH, позволяющее учесть экономические корни финансовой волатильности. Они возвращаются к вопросу Шверта: *почему волатильность меняется во времени?*. Они изучают долгосрочную волатильность рынка США за период 1884–2010 годов с помощью более совершенного инструмента — модели GARCH-MIDAS, которая позволяет долгосрочной компоненте волатильности быть функцией макроэкономических переменных. Engle et al. (2013) ограничивают выбор макропеременных инфляцией и ростом промышленного производства. Они подтверждают выводы Engle and Rangel (2008) о влиянии макропеременных: *«промышленное производство и инфляция объясняют примерно от 10% до 35% ожидаемой волатильности на день вперёд»*. Что до прогнозной способности нового класса моделей, GARCH-MIDAS *«примерно на уровне временных моделей волатильности на квартальном горизонте и на уровне или лучше них на полугодовом»*.

Самое привлекательное свойство GARCH-MIDAS — допустимость разных частот временных рядов. Оно очень полезно в макрофинансовом анализе, потому что большинство макропеременных месячные или квартальные, тогда как ценовые данные доступны с дневной частотой. Подход MIDAS (Mixed Data Sampling), предложенный Ghysels et al. (2006), позволяет разложить дисперсию на долгосрочную и краткосрочную компоненты, где за долгосрочную волатильность отвечают макроэкономические переменные, а за краткосрочную — возвращающийся к среднему дневной GARCH.

Asgharian et al. (2013) немного изменили постановку Engle et al. (2013), включив в долгосрочную компоненту волатильности и сглаженную реализованную волатильность доходностей, и макроэкономические данные (как уровень, так и дисперсию). Они сообщают о более высокой точности прогноза по сравнению со стандартной GARCH-MIDAS. Они используют данные рынка США за 1991–2008 годы и находят, что самая значимая объясняющая переменная — краткосрочная процентная ставка.

Расширение GARCH-MIDAS предложили Pan et al. (2017). Они допустили переключение режимов волатильности, управляемое марковской матрицей переходов. Классические GARCH-модели без структурных сдвигов могут давать ложную устойчивость волатильности. Поэтому переключение режимов — принципиальная деталь для точных прогнозов, тем более при значительных эмпирических свидетельствах существенно разной волатильности в бычьем и медвежьем состояниях. В этой работе я использую переключение режимов волатильности, как у Pan et al. (2017), но ввожу и переключение средней доходности — естественное, но у Pan (2017) опущенное. Вероятности я к тому же прогнозирую экзогенно, а не внутри модели.

3 Эконометрическая постановка

Здесь я излагаю формальную модель, подлежащую оцениванию. В определении и прогнозировании направления я объединяю подходы Chen (2009), Nyberg (2013) и Maheu and McCurdy (2000). В прогнозировании волатильности я по существу использую постановку Asgharian et al. (2013). Наконец, я соединяю их так же, как Pan et al. (2017), но с другим процессом вероятности режима и рядом иных расширений.

3.1 Определение направления

Проще всего разделить режимы рынка *простой скользящей средней* доходностей с периодом сглаживания N :

$$\bar{r}_t(N) = \frac{1}{N} \sum_{i=0}^{N-1} r_{t-i}$$

Когда средняя доходность за $N \geq 1$ периодов положительна, состояние рынка в момент t считается бычьим, когда отрицательна — медвежьим. Здесь можно применить и схему весов, и комбинацию скользящих средних. Например, бычий рынок можно определять условием $\bar{r}_t(N_{low}) > \bar{r}_t(N_{high})$. Это классические инструменты технического анализа, в научной литературе используемые редко.

Достоинство таких индикаторов в том, что они, во-первых, не заглядывают вперёд (а значит, пригодны для прогноза вне выборки) и, во-вторых, просто калибруются. Здесь я беру экспоненциальную скользящую среднюю (ЕМА), которая придаёт больший вес недавним значениям и потому чувствительнее к развороту тренда:

$$EMA_t(\gamma, N) = r_t \frac{\gamma}{N+1} + EMA_{t-1} \left(1 - \frac{\gamma}{N+1}\right)$$

, где γ — постоянная сглаживания (обычно 2), а N — число периодов сглаживания. С ростом N ЕМА улавливает более крупные движения цены, пренебрегая мелкими. Меньшее значение N отвечает большему числу выделенных режимов, и они будут короче.

3.2 Прогнозирование направления

Дальше состояние рынка $s_t \in \{0; 1\}$ считается наблюдаемым: $s_t = 0$ отвечает медвежьему состоянию, $s_t = 1$ — бычьему. Начнём с базовой постановки бинарного выбора, где прогнозная вероятность того, что через k периодов рынок окажется в бычьем состоянии, задаётся как

$$p_t \equiv P(s_{t+k}) = \Phi(\beta_0 + \beta_1 x_t) \quad (1)$$

где x_t — макроэкономическая переменная, а Φ — функция распределения. Для оценки прогнозов вне выборки я использую квадратичную оценку вероятности:

$$QPS = \frac{1}{T} \sum_t 2(p_{t+k} - s_{t+k})^2 \quad (2)$$

при этом значение 0 отвечает 100% точности, а максимальное значение 2 означает полную несостоятельность прогноза.

Такая спецификация включает лишь одну объясняющую переменную и лишена ряда других важных свойств. Введём сначала линейную функцию объясняющих переменных для состояния рынка:

$$\pi_t = \alpha + x_t^T \beta \quad (3)$$

где x^T — вектор объясняющих переменных, а β — вектор коэффициентов. Тогда

$$p_t \equiv P_t(s_{t+k}) = E_t(s_{t+k}) = \Phi(\pi_t) \quad (4)$$

Можно также включить лагированное значение состояния рынка:

$$\pi_t = \alpha_0 + \alpha_1 s_{t-1} + x_t^T \beta \quad (5)$$

Maheu and McCurdy (2000) и другие авторы обращают внимание на значительную разницу в длительности бычьего и медвежьего рынков. Действительно, бычьи рынки обычно продолжительнее и менее волатильны, чем медвежьи. Поэтому длительность разумно включить в уравнение. Maheu and McCurdy (2000) определяют её как

$$D_t \equiv D(s_t) = \begin{cases} D_{t-1} + 1 & s_{t-1} = s_t \\ 1 & \text{otherwise} \end{cases} \quad (6)$$

Мы также хотим уловить нелинейность зависимости от длительности. Естественно ожидать, что длительность тренда сначала увеличивает π_t , а с продолжением тренда уменьшает. Значит, нужна нелинейная структура; берём квадратичную. Более того, поскольку зависимость должна различаться для бычьего и медвежьего рынков, мы их разделяем.

$$\begin{aligned} \pi_t = & \alpha_0 + \alpha_1 s_{t-1} + \alpha_2 s_t D_t + \\ & \alpha_3 s_t D_t^2 + \alpha_4 (1 - s_t) D_t + \alpha_5 (1 - s_t) D_t^2 + x_t^T \beta \end{aligned} \quad (7)$$

Разумно ожидать, что длительность каждого состояния в долгосрочной перспективе сходится к некоторому среднему значению. Это и есть устойчивость состояния рынка. Тогда вероятность того, что следующее состояние совпадёт с предыдущим, должна вогнуто-параболически зависеть от его длительности. Отсюда я ожидаю положительного знака у α_2 и отрицательного у α_3 . У α_4 и α_5 знаки должны быть противоположными (поскольку мы по существу прогнозируем вероятность бычьего рынка) — соответственно отрицательным и положительным.

3.3 Прогнозирование волатильности

Перейдём к формальной постановке GARCH-MIDAS (см., например, Engle et al. (2013) или Asgharian et al. (2013)). Доходность $r_{i,t}$ в месяце i , день t подчиняется процессу:

$$r_{i,t} = \mu + \sqrt{\tau_i g_{i,t}} \epsilon_{i,t} \quad (8)$$

, где $g_{i,t}$ — краткосрочная компонента волатильности, а τ_i — долгосрочная. $g_{i,t}$ следует дневному процессу GARCH(1,1):

$$g_{i,t} = (1 - \alpha - \beta) + \alpha \frac{(r_{i-1,t} - \mu)^2}{\tau_t} + \beta g_{i-1,t} \quad (9)$$

Теперь я ввожу переключение режимов и в среднем, и в дисперсии. Во-первых, средняя

доходность становится переменной во времени и зависит от текущего режима рынка: $\mu_{i,t} = \mu(s_{i,t})$. Аналогично переопределяем $g_{i,t}$

$$g_{i,t} = \omega(s_{i,t}) + \alpha \frac{(r_{i,t} - \mu(s_{i,t}))^2}{\tau_t} + \beta g_{i-1,t} \quad (10)$$

Компонента τ (без изменений) задаётся в духе MIDAS-регрессии:

$$\tau_i = m + \theta \sum_{k=1}^K \phi_k(w) RV_{t-k} \quad (11)$$

, где RV — реализованная волатильность

$$RV_t = \sum_{i=1}^{N_t} r_{i,t}^2 \quad (12)$$

, $\phi_k(w)$ — экспоненциальная схема весов⁵

$$\phi(w) = \frac{w^K}{\sum_{j=1}^K w^j} \quad (13)$$

а K — число периодов сглаживания. Полная условная дисперсия определяется как

$$\sigma_{i,t}^2 = \tau_t g_{i,t} \quad (14)$$

Уравнения (9)–(13) составляют модель GARCH-MIDAS с пространством параметров $\Theta = \{\mu(0), \mu(1), \omega(0), \omega(1), \alpha, \beta, m, \theta_0, \theta_1, \theta_2, w\}$.

Asgharian et al. (2013) предлагают также включать макроэкономические переменные и их волатильности прямо в уравнение (11).

$$\tau_i = \exp(m + \theta_0 \sum_{k=1}^K \phi_k(w) RV_{t-k} + \theta_1 \sum_{k=1}^K \phi_k(w) X_{t-k}^l + \theta_2 \sum_{k=1}^K \phi_k(w) X_{t-k}^v) \quad (15)$$

, где X^l — уровень макроэкономической переменной, а X^v — её дисперсия. Asgharian et al. (2013) экспоненту не используют, но Engle et al. (2013) доказывают, что она нужна для положительности долгосрочной волатильности.

По аналогии с волатильностью доходностей рынка определим X^v как

$$X_t^v = \left(\frac{X_t^l}{X_{t-1}^l} - 1 \right)^2 \quad (16)$$

Вслед за Asgharian et al. (2013) я беру окно оценивания в 10 лет и провожу проверку вне выборки с постоянными параметрами на следующий год.

3.4 Объединение двух подходов

Есть весомые свидетельства того, что волатильность зависит от состояния рынка (см., например, Kole and van Dijk (2017), Maheu and McCurdy (2000), Lunde and Timmermann

⁵Возможны и другие схемы весов, например бета-структура лагов, разобранный у Ghysels et al. (2006).

(2004), Schwert (1989), Marcucci (2005)). Бычьи рынки обычно длинные и малволатильные, медвежьи — короткие и высоковолатильные.

Как и Pan et al. (2017), введём вектор условного распределения $\eta_{i,t}$:

$$\eta_{i,t} = \begin{bmatrix} f(r_{i,t}|s_{i,t} = 1; \Theta) \\ f(r_{i,t}|s_{i,t} = 0; \Theta) \end{bmatrix} = \begin{bmatrix} \Phi(\mu(1), \tau_t g_{i,t}(1)) \\ \Phi(\mu(0), \tau_t g_{i,t}(0)) \end{bmatrix} \quad (17)$$

, где Φ — функция нормального распределения.

Тогда, безотносительно состояния рынка,

$$f(r_{i,t}|\Theta) = \begin{bmatrix} p \\ 1-p \end{bmatrix}^T \eta_{i,t} \quad (18)$$

, где p — вероятность бычьего состояния в следующем периоде — предполагается заданной экзогенно. Её прогнозирование описано в разделе 4.2. Здесь виден важный вклад этой работы: если обычно модели GARCH-MIDAS используют данные низкой частоты только для прогноза долгосрочной компоненты волатильности, а краткосрочную динамику моделируют дневным GARCH(1,1), то я включаю данные дневной частоты и в прогноз вероятности следующего состояния.

Выпишем теперь функцию распределения доходности явно:

$$\begin{aligned} f(r_{i,t}|\Theta) &= \\ &= \mathcal{N}(p\mu(1) + (1-p)\mu(0), p^2\tau_t g_t(1) + (1-p)^2\tau_t g_t(0)) \\ &= \frac{1}{\sqrt{2\pi p^2\tau_t g_t(1) + (1-p)^2\tau_t g_t(0)}} \cdot \\ &\cdot \exp\left(-\frac{(r_{i,t} - (p\mu(1) + (1-p)\mu(0)))^2}{2p^2\tau_t g_t(1) + (1-p)^2\tau_t g_t(0)}\right) \end{aligned} \quad (19)$$

Отсюда строится оценка методом максимального правдоподобия

$$\mathcal{L}(\Theta) = \sum_{t=1}^T \sum_{i=1}^{N_t} \log f(r_{i,t}|\Theta) \rightarrow \max_{\Theta} \quad (20)$$

, откуда получаем $\hat{\Theta}$. Здесь T — число месяцев, а N_t — число дней в месяце t .

Мы также хотим оценить стандартные ошибки. Стандартная процедура (Pan et al., 2017) состоит в оценке ковариационной матрицы по гессияну. Обозначив гессиян H , ковариацию вкладов логарифмического правдоподобия J , скалярный шаг s_i и вектор e_i , все элементы которого, кроме i -го, нулевые, ковариационную матрицу $H^{-1}JH^{-1}$ можно оценить так:

$$\hat{H}_{i,j} \approx \frac{1}{Ts_i s_j} (\mathcal{L}(\hat{\Theta} + e_i s_i + e_j s_j) - \mathcal{L}(\hat{\Theta} + e_i s_i) - \mathcal{L}(\hat{\Theta} + e_j s_j) + \mathcal{L}(\hat{\Theta})) \quad (21)$$

Derivative is approximated as

$$\frac{\partial \mathcal{L}(\hat{\Theta})}{\partial \Theta} \approx \frac{\mathcal{L}(\hat{\Theta} + e_i s_i) - \mathcal{L}(\hat{\Theta})}{s_i} \quad (22)$$

and J is given as

$$\hat{J} = \frac{1}{T} \left(\frac{\partial \mathcal{L}(\hat{\Theta})}{\partial \Theta} \right)^T \left(\frac{\partial \mathcal{L}(\hat{\Theta})}{\partial \Theta} \right) \quad (23)$$

После этого стандартные ошибки строятся обычным образом:

$$s.e.(\hat{\Theta}) = \sqrt{\frac{Tr(H^{-1}JH^{-1})}{T}} \quad (24)$$

4 Данные

Все данные взяты из FRED ⁶ за период с января 2003 по апрель 2021 года. Инфляция построена как месячное процентное изменение ИПЦ. Так же получена и доходность рынка. Здесь я использую индекс Wilshire 5000 вместо привычных S&P500 или NASDAQ, поскольку Wilshire 5000 включает *все* торгуемые в США акции (на сегодня это 3141 бумага) и потому полнее отражает экономическую активность, хотя и почти идеально коррелирует с S&P500. Я также разделяю дневные и месячные данные: дневные пойдут на прогноз направления рынка, месячные — в долгосрочную компоненту процесса волатильности.

К данным дневной частоты относятся

- **BRENT**⁷: цена нефти Brent (Европа)
- **TBY2**: доходность двухлетних казначейских облигаций постоянной срочности
- **TBY10**: доходность десятилетних казначейских облигаций постоянной срочности
- **TBY30**: доходность тридцатилетних казначейских облигаций постоянной срочности
- **BEINF**: вменённая инфляция⁸
- **MARKET**: доходность рыночного индекса

Стоит отметить, что временную структуру процентных ставок я учитываю через сами ставки, а не через спред, как делает большинство авторов. Как будет видно дальше, разницы это не составляет.

⁶<https://fred.stlouisfed.org>

⁷Здесь и далее я изменил обозначения по сравнению с исходными названиями FRED для краткости. Например, в обозначениях FRED BRENT — это DCOILBRENTU.

⁸Из FRED: «Вменённая инфляция — мера ожидаемой инфляции, выводимая из десятилетних казначейских облигаций постоянной срочности и десятилетних казначейских облигаций постоянной срочности, индексируемых по инфляции. Последнее значение показывает, какой в среднем участники рынка ожидают инфляцию в ближайшие 10 лет».

Таблица 1: Описательные статистики дневных данных

Показатель	Min	Mean	Median	St. Dev.	Max
BRENT	9.12	69.60	64.31	27.38	143.95
TBY2	0.09	1.64	1.15	1.43	5.29
TBY10	0.52	2.96	2.79	1.13	5.26
TBY30	0.99	3.66	3.56	1.06	5.61
BEINF	0.04	2.04	2.12	0.41	2.76
MARKET	-12.24	0.05	0.06	1.15	11.40

Есть также переменные месячной частоты. Это макроэкономические показатели, используемые в долгосрочной компоненте волатильности модели GARCH-MIDAS. По общепринятому правилу берутся первые разности.

- **INF:** месячное процентное изменение ИПЦ.
- **INDPROD:** индекс промышленного производства — показатель реального выпуска всех расположенных в США предприятий обрабатывающей и добывающей промышленности, электро- и газоснабжения.
- **RETSALES:** совокупные продажи розничной торговли и общественного питания, млн долларов
- **UNRATE:** уровень безработицы.

Таблица 2: Описательные статистики месячных данных

Показатель	Min	Mean	Median	St. Dev.	Max
Δ INF	-1.77	0.16	0.19	0.31	1.38
Δ INDPROD	87.07	100.82	101.84	5.44	110.55
Δ RETSALES	-45217	924	1446	4656.493	9926
Δ UNRATE	3.50	6.11	5.40	2.03	14.80

Тесты ADF на стационарность приведены в приложении.

Теперь рассмотрим статистику по бычьим и медвежьим состояниям. Например, изучим длительности каждого состояния, разделяя их по ЕМА(30).

Таблица 3: Описательные статистики: длительность режимов рынка

Показатель	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Бычий	280	15.604	20.191	1	2	21	127
Медвежий	280	8.171	9.455	1	2	10.2	48

Уже отсюда видно, что бычьи состояния в среднем вдвое длиннее медвежьих. Но проведём и стандартный t-тест на различие средних:

Таблица 4: Проверка равенства длительностей бычьего и медвежьего режимов

t	5.5779
df	395.75
p-значение	4.519e-08

Рассмотрим также характеристики доходности бычьего и медвежьего рынка:

Таблица 5: Описательные статистики: доходности в разных режимах рынка

Показатель	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Бычий	4,406	0.272	0.867	-9.120	-0.110	0.630	9.000
Медвежий	2,288	-0.386	1.469	-12.240	-1.050	0.290	11.400

Нас интересует прежде всего различие дисперсий, хотя различие средних тоже следует проверить — этого требует само определение бычьего и медвежьего состояний. Но уже сейчас видно, что стандартное отклонение доходностей в медвежьем состоянии в среднем вдвое больше, чем в бычьем.

Таблица 6: Проверка равенства средней доходности в разных режимах

t	19.708
df	3135.8
p-значение	<2.2e-16

Проверим также доходности бычьего и медвежьего рынка на равенство дисперсий с двусторонней альтернативой:

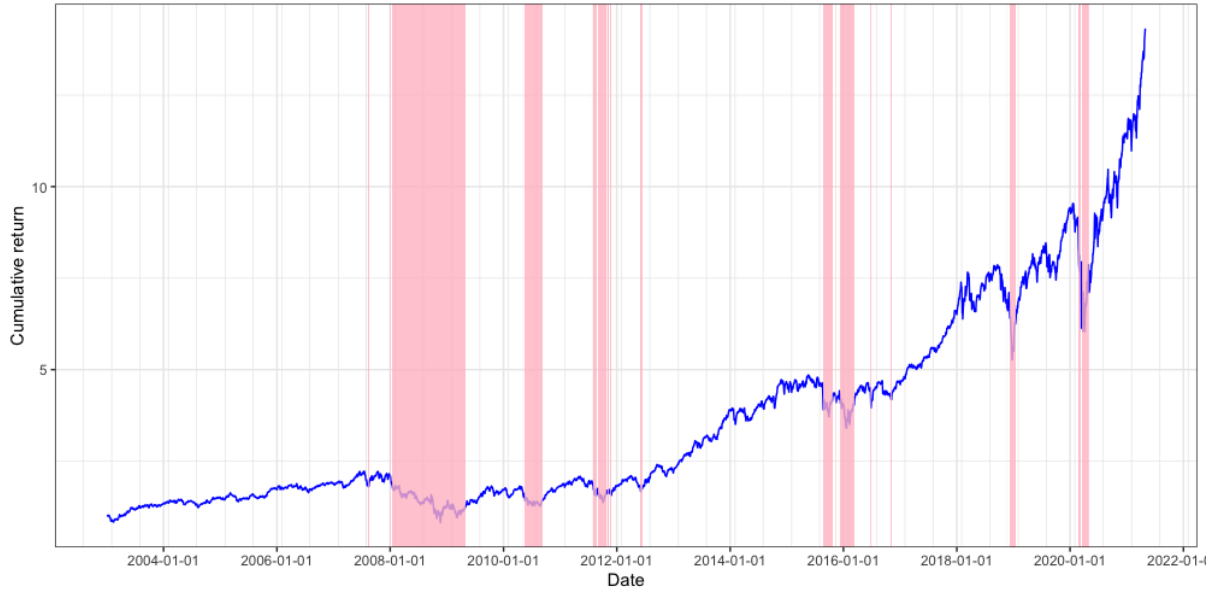
Таблица 7: Проверка равенства дисперсий доходностей в разных режимах

F	0.34828
df1	4405
df2	2287
p-значение	<2.2e-16

5 Эмпирические результаты

Начнём с разделения рынка на состояния. На рисунке 1 периоды медвежьего рынка отмечены розовым. Видно, что основные медвежьи периоды отвечают мировому финансовому кризису 2008 года, европейскому долговому кризису 2010 года, распродаже 2016 года, повышению ставок ФРС в 2018 году и ковидному шоку 2020 года.

Рис. 1: Непараметрическая идентификация режима рынка при ЕМА(500)



Полный список пиков и впадин приведён в приложении.

Сначала рассмотрим результаты оценивания модели прогнозирования направления.

Таблица 8: Прогноз направления: пробит-оценки по горизонтам k (дни)

k	5	30	90	180	360
(Intercept)	-2.248*** (0.262)	-1.126*** (0.168)	0.551*** (0.150)	1.006*** (0.165)	0.917*** (0.210)
state	2.274*** (0.118)	1.925*** (0.100)	0.575*** (0.098)	-0.258* (0.111)	-1.694*** (0.161)
BRENT	-0.016*** (0.003)	-0.017*** (0.002)	-0.018*** (0.001)	-0.011*** (0.001)	-0.023*** (0.002)
TBY2	0.219+ (0.124)	0.214** (0.081)	0.454*** (0.070)	1.241*** (0.074)	0.803*** (0.075)
TBY10	-1.189** (0.395)	-1.271*** (0.259)	-2.239*** (0.225)	-5.055*** (0.253)	-3.791*** (0.261)
TBY30	0.920** (0.350)	0.870*** (0.226)	1.662*** (0.198)	4.249*** (0.225)	2.925*** (0.231)
BEINF	1.290*** (0.222)	1.022*** (0.152)	0.521*** (0.130)	-0.888*** (0.138)	1.033*** (0.147)
D1	0.003*** (0.000)	0.001*** (0.000)	0.001** (0.000)	0.003*** (0.000)	0.005*** (0.000)
D2	0.000*** (0.000)	0.000* (0.000)	0.000 (0.000)	0.000*** (0.000)	0.000*** (0.000)
D3	0.017*** (0.003)	-0.003 (0.002)	0.001 (0.002)	0.016*** (0.002)	0.043*** (0.005)
D4	0.000*** (0.000)	0.000 (0.000)	0.000+ (0.000)	0.000*** (0.000)	0.000*** (0.000)
Num.Obs.	6689	6664	6604	6514	6334
AIC	1287.7	2929.8	4027.9	4208.4	3987.2
BIC	1362.6	3004.7	4102.7	4283.0	4061.5
Log.Lik.	-632.866	-1453.913	-2002.950	-2093.179	-1982.594
QPS	0.0226	0.0586	0.0864	0.0920	0.0934

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Для краткости здесь введены обозначения: $D1 = s_t D_t$, $D2 = s_t D_t^2$, $D3 = (1 - s_t) D_t$ и $D4 = (1 - s_t) D_t^2$. Отсюда видно, что модель работает лучше всего на коротких горизонтах. Приведённые результаты подтверждают значимую предсказательную способность данных кривой доходности. По оценке QPS точнее всего краткосрочный прогноз в рамках рассматриваемой модели. Нелинейная зависимость от длительности также оказалась значимой на всех горизонтах, кроме 30 и 90 дней. Как и ожидалось, цена нефти отрицательно влияет на вероятность бычьего рынка. Любопытно, что коэффициенты кривой доходности в сумме дают около нуля: это означает, что по существу мы используем два временных спреда — (TBY10 TBY2) и (TBY30 TBY10).

К интерпретации оценённых параметров зависимости от длительности следует подходить с большой осторожностью. Во-первых, мои исходные ожидания относительно знаков не подтвердились: все значения оказались положительными. Это, впрочем, согласуется с выводами Maheu and McCurdy (2000) об устойчивости состояния рынка: чем дольше длится бычий рынок, тем менее вероятно его завершение. Но для медвежьего рынка толкование иное: чем дольше он длится (параметры D3 и D4), тем выше вероятность перехода к бычьему. Таким образом, у бычьего рынка зависимость от длительности положительная, а у медвежьего — по существу отрицательная.

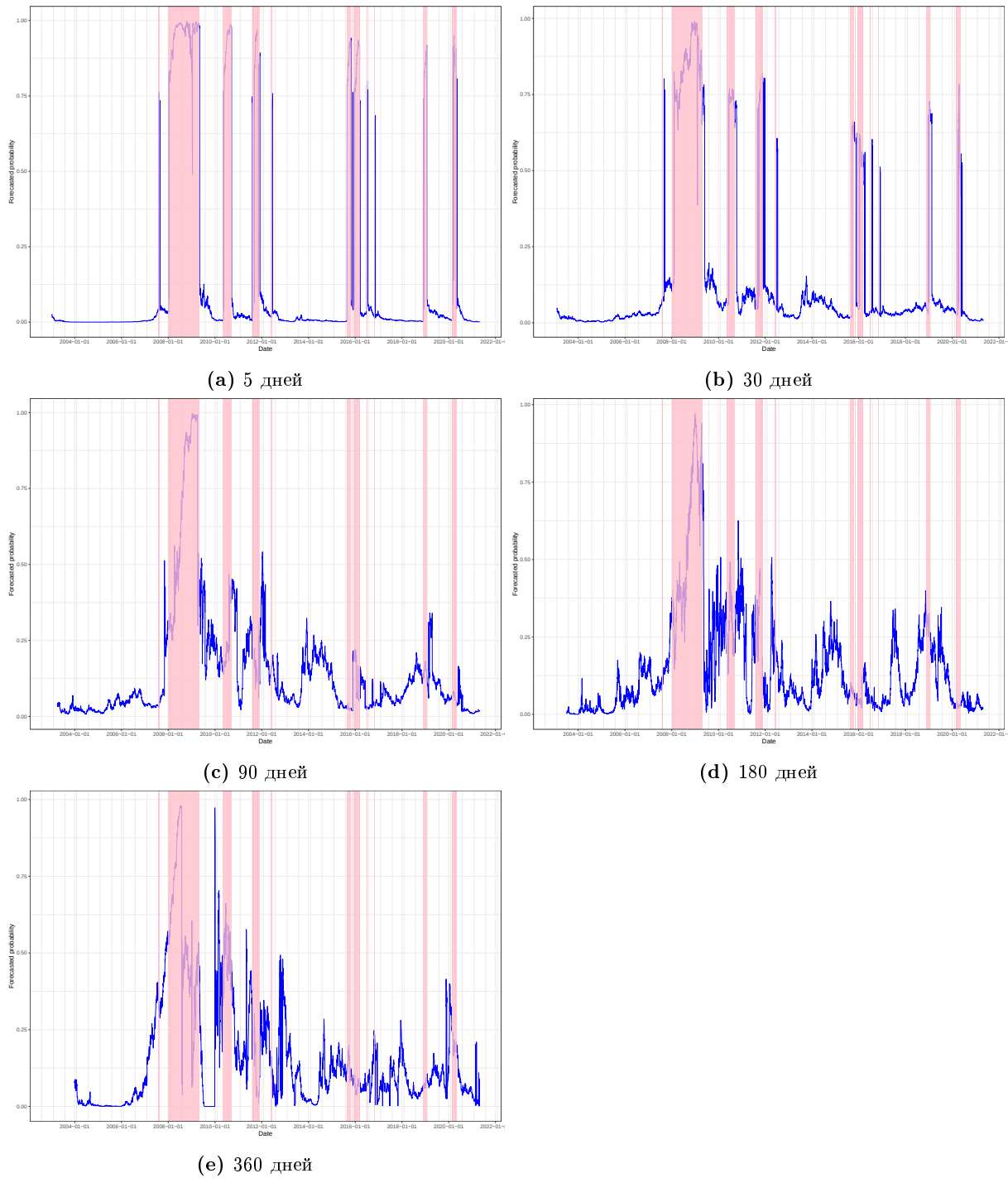


Рис. 2: Прогнозная вероятность медвежьего рынка на разных горизонтах

Если считать медвежий рынок предсказанным при прогнозной вероятности выше 0,5 (хотя можно взять и другие пороги), то применительно к обвалу 2020 года точными оказались только прогнозы на 5 и 30 дней вперёд. На более длинных горизонтах падение в рассмотренных моделях было непредсказуемо.

Перейдём к оцениванию методом максимального правдоподобия GARCH-MIDAS с переклещением режимов вместе со стандартным эталоном.

Таблица 9: Оценки максимального правдоподобия: GARCH-MIDAS с переключением режимов против стандартного эталона

model	RS-GARCH-MIDAS	GARCH-MIDAS
m	−0.0754*** (0.0001)	0.2240*** (0.0003)
μ_0	0.0431*** (0.0071)	
μ_1	0.0692*** (0.0001)	0.0880 (0.0067)
ω_0	0.2224*** (0.0104)	
ω_1	0.0444*** (0.0001)	0.0152*** (0.0036)
α	0.1117*** (0.0001)	0.0727*** (0.0035)
β	0.1408*** (0.0001)	0.7203*** (0.0088)
θ_0	0.1126*** (0.0001)	0.0101*** (0.0001)
θ_1	−0.0520*** (0.0000)	−0.0025 (0.9304)
θ_2	0.0791*** (0.0001)	0.7841 (7.3922)
w	0.0941*** (0.0002)	0.4276*** (0.0281)
Num.Obs.	6689	6689
AIC	11865.35	17050.57
Log.Lik.	−5921.677	−8516.284

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

По информационному критерию Акаике (AIC) модель с переключением режимов значительно превосходит стандартную GARCH-MIDAS. Компоненты θ_1 и θ_2 , отвечающие по существу за макроэкономическую информацию, в эталонной модели незначимы, но становятся значимыми в модели с переключением режимов. Стоит отметить знаки параметров: волатильность рынка и макроэкономическая волатильность (соответственно θ_0 и θ_2) положительно влияют на прогнозную волатильность рынка, что естественно, тогда как уровень макропеременной (первая разность розничных продаж) связан с ней отрицательно. Как и следовало ожидать, безусловная долгосрочная компонента волатильности ω больше для медвежьих рынков, чем для бычьих, а безусловная средняя доходность μ больше для бычьих. Тем самым модель отвечает разделению бычьего и медвежьего рынков, обсуждавшемуся вначале, и подтверждает выводы предыдущих работ.

6 Заключение

В этой работе я разобрал существующие модели прогнозирования направления и волатильности доходностей фондового рынка. Опираясь на предшествующие исследования, я предложил расширение стандартной модели GARCH-MIDAS, дополнив её прогнозом состояния рынка на макроэкономической основе. Эмпирические результаты подтверждают выводы прежних работ о значимости данных кривой доходности для прогноза направле-

ния рынка. Более того, нелинейная зависимость от длительности в прогнозе направления оказалась высокосignимой на большинстве горизонтов. У бычьих рынков зависимость от длительности положительная, у медвежьих — отрицательная, что согласуется с выводами предыдущих исследований. Точнее всего оказался краткосрочный прогноз (на 5 дней вперёд). Полученные результаты говорят о том, что обвал 2020 года был предсказуем за месяц до самого падения, но не предсказуем на более длинных горизонтах. Включение макроэкономического переключения режимов в GARCH-MIDAS заметно повысило точность прогноза.

Основной вклад работы — включение в модель GARCH-MIDAS переключения режимов на макроэкономической основе, проверка прогнозирования режима на новых данных, введение нелинейной зависимости от длительности в прогноз состояния рынка и новый способ идентификации рыночного режима.

Список литературы

- [1] Asgharian, H., Hou, A. J., & Javed, F. (2013). The importance of the macroeconomic variables in forecasting stock return variance: A GARCH-MIDAS approach. *Journal of Forecasting*, 32(7), 600-612. <https://doi.org/10.1002/for.2256>
- [2] Chen S., 2009, Predicting the bear stock market: Macroeconomic variables as leading indicators, *Journal of Banking & Finance*
- [3] Cochrane J. H., 1991, Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations, *The Journal of Finance*
- [4] Demir C., 2019, Macroeconomic Determinants of Stock Market Fluctuations: The Case of BIST-100, *Economies*
- [5] van Dijk D. and Cakmakl C. , 2016, Getting the most out of macroeconomic information for predicting excess stock returns, *International Journal of Forecasting*
- [6] Engle R. F., Ghysels E., and Sohn B., 2013, STOCK MARKET VOLATILITY AND MACROECONOMIC FUNDAMENTALS, *The Review of Economics and Statistics*
- [7] Ibbotson R. G.& Chen P., 2019, Long-Run Stock Returns: Participating in the Real Economy, *Financial Analysts Journal*
- [8] Kole, E., and van Dijk, D. (2017) How to Identify and Forecast Bull and Bear Markets?. *J. Appl. Econ.*, 32: 120– 139. doi: 10.1002/jae.2511.
- [9] Lunde A. & Timmermann A. (2004) Duration Dependence in Stock Prices, *Journal of Business & Economic Statistics*, 22:3, 253-273, DOI: 10.1198/073500104000000136
- [10] Maheu J. M. & McCurdy T.H. (2000) Identifying Bull and Bear Markets in Stock Returns, *Journal of Business & Economic Statistics*, 18:1, 100-112, DOI: 10.1080/07350015.2000.10524851
- [11] Marcucci, J. (2005). Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics & Econometrics*, 9(4), 1-53.
- [12] Naik P.K., 2013, Does Stock Market Respond to Economic Fundamentals? Time- series Analysis from Indian Data, *Journal of Applied Economics and Business Research*
- [13] Nyberg H., Predicting bear and bull stock markets with dynamic binary time series models, *Journal of Banking & Finance*, Volume 37, Issue 9, 2013, Pages 3351-3363, ISSN 0378-4266, <https://doi.org/10.1016/j.jbankfin.2013.05.008>.
- [14] Pan Z., Wang Y., Wu C., Yin L. (2017), Oil price volatility and macroeconomic fundamentals: A regime switching GARCH-MIDAS model, *Journal of Empirical Finance*, Volume 43, Pages 130-142, ISSN 0927-5398, <https://doi.org/10.1016/j.jempfin.2017.06.005>.
- [15] Rapach D. E., Strauss J. K., Zhou G., 2009, Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy, *Oxford University Press*

- [16] SCHWERT, G.W. (1989), Why Does Stock Market Volatility Change Over Time?. The Journal of Finance, 44: 1115-1153. <https://doi.org/10.1111/j.1540-6261.1989.tb02647.x>
- [17] Hlavac M (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.2. <https://CRAN.R-project.org/package=stargazer>
- [18] Tangjitprom N., 2012, The Review of Macroeconomic Factors and Stock Returns, *International Business Research*

7 Приложение

7.1 Пики и впадины при ЕМА(500)

Таблица 10: Список пиков и впадин, выделенных по ЕМА(500)

Peak	Trough
05.08.2007	04.08.2007
17.08.2007	13.08.2007
09.01.2008	05.01.2008
02.01.2009	10.01.2008
30.04.2009	08.01.2009
09.05.2010	08.05.2010
12.09.2010	15.05.2010
14.08.2011	03.08.2011
28.08.2011	15.08.2011
23.10.2011	01.09.2011
26.10.2011	24.10.2011
02.11.2011	31.10.2011
07.11.2011	05.11.2011
10.11.2011	08.11.2011
29.11.2011	20.11.2011
14.06.2012	31.05.2012
21.10.2015	21.08.2015
17.11.2015	13.11.2015
15.12.2015	11.12.2015
05.03.2016	16.12.2015
10.03.2016	07.03.2016
01.07.2016	25.06.2016
05.07.2016	04.07.2016
06.11.2016	31.10.2016
25.11.2018	24.11.2018
18.01.2019	07.12.2018
01.03.2020	26.02.2020
03.03.2020	02.03.2020
14.03.2020	04.03.2020
18.04.2020	15.03.2020
23.04.2020	19.04.2020
07.05.2020	01.05.2020

7.2 Тест ADF

	DF	Lag oreder	p-значение
ΔINF	-9.2188	5	0.01
$\Delta RETSALES$	-4.6387	5	0.01
$\Delta INDPROD$	-1.4954	5	0.7875
$\Delta UNEMP$	-0.26413	5	0.99