

Федеральное государственное автономное образовательное учреждение высшего образования

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет экономических наук

КУРСОВАЯ РАБОТА

Оценка риска российских ценных бумаг

по направлению подготовки «Экономика»

образовательные программы: «Экономика», «Экономика и статистика»

Выполнили:

Студентка группы БЭК192

Кодзаева Валерия Витальевна

Студент группы БЭК193

Краевский Артем Андреевич

Студентка группы БСТ191

Елисеева Екатерина Игоревна

Руководитель:

профессор, кандидат физико-
математических наук,

Канторович Григорий

Гельмутович

Москва 2021

Оглавление

Введение.....	3
Постановка проблемы.....	3
Актуальность.....	3
Предмет и объект исследования.....	4
Цель исследования.....	5
Задачи и этапы	5
Литературный и теоретический обзор	6
САРМ Шарпа-Литнера (Классическая модель)	7
Трехфакторная модель Фамы-Френча.....	12
Пятифакторная модель Фамы-Френча	14
Тестирование и проверка модели	17
Выбор рыночной и безрисковой ставки	18
Аналогичные исследования для российского рынка	19
Работа с данными.....	20
Расчет доходности за рассматриваемый период	20
Конструирование факторов	21
Выбор композиции разбиения	24
Методология	25
Построение регрессий.....	25
Формулирование и проверка гипотез	27
Результаты.....	29
Общее описание результатов.....	29
Лучшая описательная модель	30
Коэффициент детерминации R^2	30
P-value для коэффициента α	33
Наилучшая модель для российского рынка	35
Лучшая прогностическая модель	36
MAPE.....	36
Качество работы моделей для разных активов: причины и следствия.....	38
Выводы.....	41
Список используемой литературы	43
Приложения	44

Введение

Постановка проблемы

В рамках нашей работы мы рассмотрим применение модели Capital Asset Pricing Model (далее – CAPM) и её вариаций для оценки рисков российских ценных бумаг. Главной задачей проекта является установить, какая из предложенных вариаций модели CAPM наиболее точно описывает и предсказывает динамику российского рынка акций. Для анализа будут использованы реальные дневные котировки всех ценных бумаг, входивших в индекс Московской биржи с 2014 по 2019 годы, на основе которых в дальнейшем предполагается строить регрессионные модели и сравнивать точность их предсказаний между собой. Также в рамках работы будут описаны основные трудности, с которым столкнулись авторы и пути их решения, перспективы исследуемой темы и применения её результатов для анализа рынка российских ценных бумаг, возможные пути развития предложенной траектории исследования.

Актуальность

На современном этапе экономического развития одним из ключевых факторов для долгосрочного экономического роста как отдельной отрасли в экономике, так и всей экономической системы в целом является успешное функционирование финансовых рынков. Вопросом описания и предсказания изменений котировок активов на фондовом рынке занимались ещё экономисты в конце первой половины XX века, с течением времени по мере становления экономической теории и более глубинной интеграции методов

математической статистики в прикладную экономику всё более актуальной становилась проблема математического моделирования процессов, происходящих на фондовых рынках, тем самым породив сразу несколько подходов к изучению данного феномена.

На сегодняшний день моделирование процессов, происходящих на фондовом рынке, является одной из главных задач специалистов в области финансов, эконометрики, математической статистики, теоретической экономики и ряда других дисциплин, её решение позволило бы в первую очередь решить проблемы асимметрии информации и рисков в финансовом секторе.

В рамках работы мы используем один из наиболее известных подходов к описанию динамики на рынках ценных бумаг, специфицируя его для российского рынка. Предполагается, что на основе полученных в ходе работы результатов можно делать состоятельные прогнозы об изменениях на фондовом рынке, дополнять и обобщать рассматриваемую модель на произвольный финансовый рынок, создавать инвестиционные стратегии и осуществлять портфельное управление с минимальными финансовыми рисками.

Предмет и объект исследования

Основным предметом исследования в рамках нашей работы является модель CAPM в представлениях Шарпа – Литнера (далее – классическая модель CAPM) и Фамы-Френча (трёхфакторная и пятифакторная). В ходе работы будут так же построены факторы¹, которые в дальнейшем будут использованы для построения регрессионных моделей.

¹ О процессе построения факторов для трёхфакторной и пятифакторной моделей Фамы-Френча будет подробнее рассказано далее – прим. авторов.

Модель Оценки Долгосрочных Активов используется для того, чтобы определить необходимый и достаточный уровень доходности актива, который предполагается включить в относительно диверсифицированный портфель в предпосылках существования ненулевого системного рыночного риска, описанием и прогнозированием которого мы и будем заниматься в рамках дальнейшей работы.

Объектом исследования выступает российский рынок ценных бумаг, входивших в индекс Московской Биржи с 2014 по 2019 год². Все данные, которые будут использованы в рамках работы, а именно котировки, тикеры, даты и прочие были получены из надёжных источников, все ссылки на которые так же будут приведены в списке используемой литературы.

Цель исследования

Главной целью нашего исследования является сравнение результатов применения рассматриваемых вариаций модели CAPM на реальных данных, самостоятельное построение факторов для использования трехфакторной и пятифакторной моделей CAPM для российского рынка и установление модели, способной наиболее точно предсказывать динамику котировок российских ценных бумаг на фондовом рынке.

Задачи и этапы

В рамках работы были выделены следующие этапы и соответствующие им цели:

² См. Приложения, Таблица 8, стр. 52

- 1) Изучение работ по исследуемой теме, ознакомление с теоретическими предпосылками создания каждой из рассматриваемых моделей.
- 2) Выбор трёх моделей, которые по нашему предположению наиболее точно должны описывать динамику на российском рынке ценных бумаг, одной из которых должна являться классическая модель CAPM.
- 3) Сбор рыночных данных Московской биржи из наиболее надежных и авторитетных источников в период с 2014 по 2019 годы.
- 4) Первичная обработка данных и форматирование.
- 5) Построение факторов, необходимых для построения выбранных нами моделей для каждой из представленных в базе данных компаний в соответствии с моделями Фамы-Френча.
- 6) Вторичная обработка данных, подготовка базы данных для построения дневной, месячной доходности и скользящего окна для каждого тиккера, добавление построенных ранее факторов.
- 7) Построение регрессионных моделей с использованием подготовленной базы данных, визуализация численной информации.
- 8) Проверка гипотез о точности рассмотренных нами моделей, сравнение точности моделей между собой.
- 9) Подведение итогов о проделанной работе, определение модели, наиболее подходящей для описания динамики российского рынка ценных бумаг и перспектив для будущих исследований.

Литературный и теоретический обзор

САРМ Шарпа-Литнера (Классическая модель)

Основополагающим трудом, сформулировавшим предпосылки классической модели САРМ, выступает работа Уильяма Шарпа «Equilibrium in a Capital Asset Market»³. Сам Шарп, ссылаясь на работы Марковица, Тобина, Фон Неймана и Morgenstern, Гордона и Ганголли, заключает, что все модели, разработанные в рамках портфельной теории Марковица, опираются на единую концепцию оптимального поведения инвесторов, но никто из них не пытался обобщить её, тем самым детерминировав теорию общего равновесия.

Шарп начинает свои рассуждения из предположения, что функция полезности любого индивида в предположении существования некоторой функции распределения доходов субъектов рынка представима в виде:

$$U = f(E_w, \sigma_w)$$

Где за E_w принимается ожидаемый доход, а за σ_w – отклонения фактического дохода от E_w соответственно. Так же Шарп вводит некоторые ограничения для данной функции:

$\frac{\partial U}{\partial E_w} > 0$, то есть максимизация доходов инвестора и $\frac{\partial U}{\partial \sigma_w} < 0$, то есть неприязнь к рискам и предпочтение более надежных финансовых инструментов. Таким образом мы предъявляем требования к кривым безразличия, чтобы они удовлетворяли условиям рациональности потребления (а именно – наклон вверх).

Теперь определим норму прибыли от инвестирования как $R \equiv \frac{W_t - W_i}{W_i}$, где W_t – его доход от инвестирования, а W_i – объём исходных вложений

³ Sharpe, William F. (1964). "Capital asset prices: A theory of market equilibrium under conditions of risk". *Journal of Finance*.

инвестора. Таким образом мы можем выразить наш доход от инвестирования через норму прибыли и начальные вложения инвестора следующим образом:

$$W_t = RW_i + W_i$$

Далее Шарп делает ссылку на работу Фаррара⁴, в которой тот принимает в качестве исходной функции ожидаемой доходности, ссылаясь на функцию полезности Фон-Неймана – Моргенштерна:

$$E(U) = a + bE_R - cE_R^2 - c\sigma_R^2,$$

то есть для нашей задачи:

$$E(U) = k_1E_R - k_2\sigma_R^2.$$

Здесь, в отличие от первой функции, зависимость нашей функции от E_R и σ_R^2 линейна, то есть полученные Фарраром выводы не согласуются с аксиомами рационального потребителя (в частности если мы возьмём три точки на двух кривых, то на исходной они не будут связаны линейно, в то время как в предложенной Фарраром модели мы получим обратный вывод, то есть противоречие к исходным предположениям⁵

⁴ Donald E. Farrar «The Investment Decision Under Uncertainty» (Prentice-Hall, 1962)

⁵ См. Рисунок 1, стр. 9

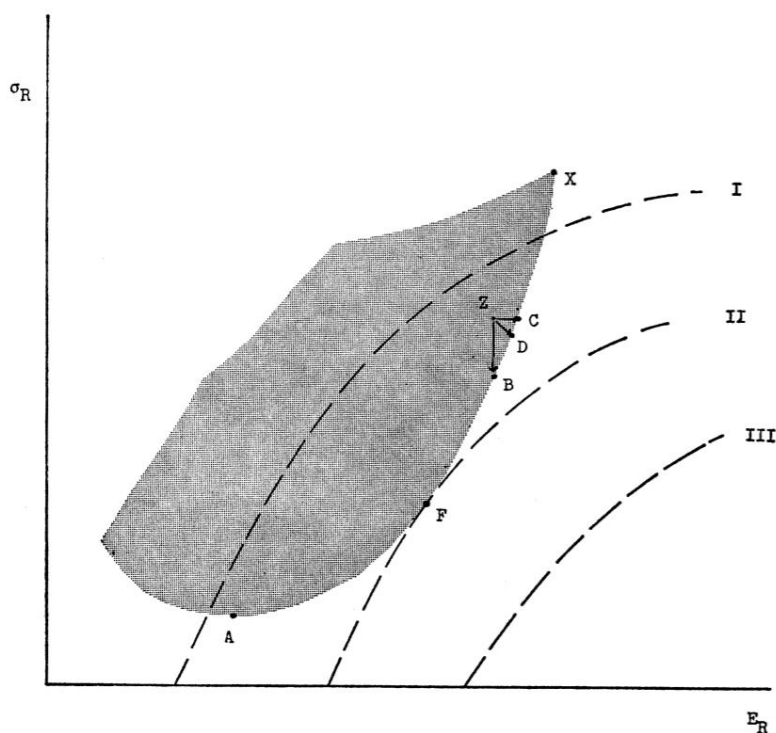


Рисунок 1. The investment opportunity curve.

Сам же Шарп предлагает рассматривать следующую предпосылку: давайте рассмотрим две инвестиционные стратегии – А и В, которые индивид принимает с вероятностями α и $(1 - \alpha)$ соответственно. В таком случае ожидаемая доходность представима в виде следующей линейной комбинации:

$$E_{R_e} = \alpha E_{R_a} + (1 - \alpha) E_{R_b}$$

Тогда стандартное отклонение реальной доходности от ожидаемой мы можем посчитать по формуле как:

$$\sigma_{R_e} = \sqrt{\alpha^2 \sigma_{R_a}^2 + (1 - \alpha)^2 \sigma_{R_b}^2 + 2r_{ab}\alpha(1 - \alpha)\sigma_{R_a}\sigma_{R_b}}$$

Где r_{ab} – коэффициент корреляции между нормами прибыли рассматриваемых нами инвестиционных стратегий.

Далее описывается сущность так называемого «безрискового» актива (обозначим его за P):

$$\exists P: \sigma_{R_P} = 0, E(U_P) = E_{R_P}, r_{ap} = \text{corr}(A, P) = 0 \forall A,$$

где A – произвольная инвестиционная стратегия

Тогда, полагая B стратегию вложения в «безрисковый» актив, а A – произвольной имеем:

$$E_{R_e} = \alpha E_{R_P} + (1 - \alpha) E_{R_a}$$

$$\sigma_{R_e} = \sqrt{\alpha^2 \sigma_{R_P}^2 + (1 - \alpha)^2 \sigma_{R_a}^2 + 2 r_{ap} \alpha (1 - \alpha) \sigma_{R_P} \sigma_{R_a}}$$

В нашем предположении $\sigma_{R_P} = 0$, $r_{ap} = 0$, тогда имеем:

$$\sigma_{R_e} = (1 - \alpha) \sigma_{R_a}$$

Далее Шарп обосновывает дополнительное построение линии рынка капитала (Security market line), необходимая для лучшей экономической интерпретации результатов CAPM. Мы пропустим некоторую часть обоснования использования данной модели, так как к нашей работе она не будет иметь в дальнейшем непосредственного отношения.

Возвращаясь к модели CAPM, Шарп связывает линию рынка капитала и функции полезности отдельных индивидов следующими соотношениями (полагая, что они пересекаются в одной точке):

$$\frac{\partial E_{R_e}}{\partial \alpha} = -[E_{R_a} - E_{R_P}]$$

$$\frac{\partial \sigma_{R_e}}{\partial E_{R_e}} = \frac{\sigma_{R_a} - r_{ap} \sigma_{R_P}}{E_{R_a} - E_{R_P}}$$

Если уровень $\alpha = 0$ (т.е. мы придерживаемся единой стратегии инвестирования в «безрисковый» актив)

Тогда, полагая, что наша линия рынка капитала задана уравнением:

$$\sigma_{R_e} = s(E_{R_e} - P)$$

Где P – чистая процентная ставка по «безрисковым» активам

Имеем справедливым следующее тождество:

$$\frac{r_{ap}\sigma_{R_P}}{\sigma_{R_a}} = -\left[\frac{P}{E_{R_a} - P}\right] + \left[\frac{1}{E_{R_a} - P}\right]E_{R_P}$$

Наконец мы добрались до основном результате исследования – Шарп вводит так называемый им «предсказываемый ответ» (predicted response)⁶ изменения инвестиционного портфеля безрискового актива на изменение второго портфеля А, обозначая его как B_{ap} . Тогда предполагается, что при прогнозируемой σ_{R_a} объём систематических предсказываемых рисков может быть однозначно определён следующим соотношением:

$$r_{ap} = \sqrt{\frac{B_{ap}^2 \sigma_{R_a}^2}{\sigma_{R_P}^2}} = \frac{B_{ap} \sigma_{R_a}}{\sigma_{R_P}}$$

$$B_{ap} = \frac{r_{ap} \sigma_{R_P}}{\sigma_{R_a}}$$

Таким образом, этот «предсказываемый ответ», связывающий теоретическую и прогнозируемую доходность портфеля детерминирован как:

$$B_{ap} = -\left[\frac{P}{E_{R_a} - P}\right] + \left[\frac{1}{E_{R_a} - P}\right]E_{R_P}$$

Этот фундаментальный результат, полученный Уильямом Шарпом в 1964 и положил начало работы над моделью CAPM. Схожие с ним результаты, получили и Моссин, и Литнер, и Трейнор, однако именно в этой работе дано опровержение существовавшим ранее взглядом о целесообразности применения функции полезности Фон-Неймана-Моргенштерна для моделирования рынков ценных бумаг и поведения на них рациональных инвесторов. Предложенный всеми ими «предсказываемый ответ», который сейчас принято называть β -коэффициентом, давал возможность связать портфельную теорию Марковица со статистическими данными реальных

⁶ Прим. авторов

рынков, тем самым создав базовый инструментарий для оценки финансовых рисков.

В рамках обзора работы Шарпа, мы смогли проследить весь путь математического моделирования предложенного коэффициента, что, безусловно, даёт более состоятельное представление о рассматриваемых далее модификациях Фамы и Френча, где суждения строятся в более жестких предпосылках, чем предъявляли экономисты того времени. О них далее и пойдёт речь в нашей работе.

Трехфакторная модель Фамы-Френча

Несмотря на своевременность и востребованность модели Г. Марковица, Дж. Тобина и У. Шарпа, было проведено достаточно большое количество исследований, подтверждавших несостоятельность полученных ранее выводов. По мере их проведения выявлялись проблемы, которые данная модель не может описать. Одно из наиболее значимых исследований было проведено Юджином Фамой и Кеннетом Френчем. Оно основывалось на аномальном систематическом отклонении от модели CAPM.⁷

В ходе своей работы ими были выявлены зависимости: обратно пропорциональная зависимость доходности акции от размера компании, которая определяется рыночной капитализацией, и прямо пропорциональная зависимость доходности акции от соотношения балансовой стоимости акционерного капитала к рыночной стоимости.

Так Ю. Фама и К. Френч разработали трехфакторную модель, которая помимо рыночной премии за риск подразумевала наличие ещё двух факторов:

⁷ Fama E. F., French K. R. The cross section of expected stock returns //Journal of Finance. 1992. Vol. 47. P. 427–465.

- SMB – это фактор, определяющийся мерой доходности компании на основе размера, или рыночной капитализацией компании.
- HML – это фактор, основанный на отношении балансовой и рыночной стоимости компании.

В своем исследовании они использовали исторические данные по котировкам акций с 1963–1990 год, торгуемых на бирже NYSE. Для расчета факторов использовались комбинации портфелей с ранжированными акциями. По результатам проведенных исследований они предложили использовать следующую модель:

$$r_i = \alpha_i + \beta_i(r_m + r_f) + s_i * SMB_t + h_i * HML_t$$

где:

r_i – доходность актива

r_f – безрисковая доходность

r_m – доходность рыночного портфеля (среднерыночная доходность)

β_i, s_i, h_i – коэффициенты, которые говорят о влиянии рыночного фактора, факторов HML и SMB на доходность i-го актива

α_i – свободный член линейной регрессии.

Некоторые ученые в своих работах оспаривали данные выводы, например авторы А. Шлейфер и Р. Вишну, получившие результаты, что высокая доходность акций с высоким показателем балансовой стоимости к рыночной стоимости акции, создана искусственно инвесторами⁸, которым свойственно быть более оптимистично настроенными к акциям компании, публикующей полноценную и детализированную финансовую отчетность. Таким образом

⁸ Josef Lakonishok, Andrei Shleifer, Robert W. Vishny Contrarian Investment, Extrapolation, and Risk//Journal of Finance. 1994 . P. 1541–1578.

существует слишком сильная зависимость доходности с факторами стоимости и размера и незначительная зависимость с макроэкономическими факторами в модели Фамы и Френча.

Однако при расширении числового ряда ученый Р. Фафф, исследуя месячные данные по изменениям котировок на фондовом рынке Австралии с 1991 по 1999 гг., подтвердил состоятельность предложенной модели, используя для этого метод обобщенных моментов.⁹ В своем исследовании он обнаружил положительную зависимость между премией за риск, рыночной доходностью и отношением балансовой и рыночной стоимостей компании. Так же он показал отрицательную взаимосвязь между доходностью и размером.

Пятифакторная модель Фамы-Френча

Исследования Novy-Marx (2013), Titman, Wei and Xie (2004) показывают, что многие изменения в средней доходности, связанные с прибыльностью и инвестированием, были упущены трехфакторной моделью и остались необъясненными. Так, были обрисованы направления уточнения модели оценки теоретической основой для построения 5-факторной модели Фамы и Френча стала модель дисконтирования дивидендов (Миллер и Модильяни, 1961), из нее удалось вывести факторы прибыльности и инвестирования. Модель дисконтирования дивидендов говорит, что приведенная стоимость акции в момент t связана с дисконтированными, ожидаемыми дивидендами.

$$m_t = \sum_{\tau=1}^{\infty} \frac{E(d_{t+\tau})}{(1+r)^\tau}$$

⁹ Robert Faff, An Examination of the Fama and French Three-Factor Model Using Commercially Available Factors //Journal of Finance. 2001. P. 1-17;83-92

Здесь m_t – цена акции в момент t , $E(d_{t+\tau})$ – ожидаемые дивидендные выплаты на акцию в период $t + \tau$, а r – ожидаемая в долгосрочной перспективе внутренняя норма доходности на инвестиции в эти акции. Из этого уравнения Миллер и Модильяни выводят общую рыночную стоимость акции:

$$M_t = \sum_{\tau=1}^{\infty} E(Y_{t+\tau} - dB_{t+\tau}) / (1 + r)^\tau$$

Здесь $Y_{t+\tau}$ – общий заработок капитала за период $t + \tau$, а $dB_{t+\tau} = B_{t+\tau} - dB_{t+\tau-1}$ есть изменение в балансовой стоимости капитала. Разделяя полученное уравнение на балансовую стоимость капитала в момент t , Фама и Френч получают следующее:

$$\frac{M_t}{B_t} = \frac{\sum_{\tau=1}^{\infty} E(Y_{t+\tau} - dB_{t+\tau}) / (1 + r)^\tau}{B_t}$$

Из этого уравнения ученые делают несколько наблюдений. Во-первых, зафиксировав всё кроме текущей стоимости акции M_t и ожидаемой доходности r , можно получить, что с более низкими показателями M_t или что эквивалентно, с более высоким отношением балансовой стоимости к рыночной (Book-to-Market ratio) более высокой будет и ожидаемая доходность. Во-вторых, при прочих равных более высокий ожидаемая прибыль соответствует более высокой ожидаемой доходности. Наконец, высокий ожидаемый рост в балансовой стоимости, т.е. инвестициях – подразумевает более низкую ожидаемую доходность. Так, выводятся теоретические обоснования введения двух новых факторов риска, влияющих на ожидаемую доходность: фактор прибыльности и инвестирования в активы.

Так, например, была разработана четырехфакторная модель с фактором «моментум» (Carhart, 1997), который указывает на влияние доходности за последние периоды на ожидаемую доходность. Трехфакторная модель не смогла объяснить «momentum anomaly» - разницу в доходности портфелей, отсортированных по доходности прошлого года. Так, трехфакторная модель

показала сильно отрицательные ошибки на портфелях, проигравших по доходности и сильно положительные на победивших портфелях. Потому были созданы предпосылки для добавления четвертого фактора риска «моментум». Однако в пятифакторной модели Фама и Френч от него отказываются, обращая внимание на иные факторы. Они считают, что необходимо в первую очередь учитывать те факторы, которые непосредственно следуют из декомпозиции денежных потоков, описанной выше. А те показатели, которые явно не выводятся из модели дисконтирования потоков, но помогают прогнозировать доходность, они должны быть учтены в предсказаниях по прибыльности и инвестированию. Таким образом, Фама и Френч указывают на отсутствие теоретических предпосылок для включения этого фактора «моментум» в модель.

Итоговая 5-факторная модель выглядит так:

$$R_{it} - R_{ft} = \alpha_i + b_i(R_{Mt} - R_{ft}) + s_iSMB_t + h_iHML_t + r_iRMW_t + c_iCMA_t + \varepsilon_{it}$$

Здесь добавляется фактор прибыльности RMW (Robust minus Weak): разница между доходностями диверсифицированных портфелей из акций компаний с большой и малой прибыльностью - и фактор инвестирования CMA (Conservative Minus Aggressive): разница между доходностями акций компаний с консервативной и агрессивной инвестиционной политикой.

В пятифакторной модели возникает предположение об избыточности фактора HML, характеризующего оценку Book-to-Market Ratio, предполагается что эффект от него объясняется факторам прибыльности и инвестирования, которые влияют на балансовую стоимость компании. Однако Фама и Френч показывают, что модель без фактора HML показывает менее точные результаты на данных США 1963 – 2013 гг., чем пятифакторная. Именно поэтому в своей работе рассуждать о возможной проблеме мультиколлинеарности признаков не планируется, однако такой способ

проверки может рассматриваться в качестве перспективы для будущих исследований.

Тестирование и проверка модели

Опыт тестирования моделей

Для оценки качества моделей CAPM сложилась традиция проверять гипотезу о равенстве нулю коэффициента пересечения, что соответствует сильной описательной силе модели.

Для тестирования моделей доходности активов, в том числе CAPM и её многофакторных вариаций, учеными Фама и Макбет¹⁰. была разработана двухэтапная регрессия: на первом этапе оценивались коэффициенты при факторах риска на всем периоде обучающей выборки, в том числе и коэффициент пересечения α . На втором этапе предлагалось для каждого актива использовать оценки коэффициентов как зависимые переменные и оценить риск-премии γ на всех периодах. После повторения процедуры для всех активов и получения оценок риск-премий, проверялись две гипотезы:

- *о положительных риск-премиях*: естественным предположением о работе финансовых рынков является повышенная доходность при увеличении рисков, соответственно, если отвергается гипотеза о положительности риск-премий, то модель, очевидно, не может объективно описывать работу рынков
- *о равенстве риск-премий факторам*: отвержение этой гипотезы говорит о недостаточном качестве оценки коэффициентов регрессии

В этой работе оценка описательной силы модели будет проводиться лишь по гипотезе о равенстве коэффициента пересечения нулю ввиду

¹⁰ Fama, Eugene F.; MacBeth, James D. (1973). "Risk, Return, and Equilibrium: Empirical Tests". *Journal of Political Economy*. **81** (3): 607–636

недостаточного опыта тестирования моделей на российском рынке и фактического отсутствия в этой работе предобработки данных и работы с временными лагами, которые повлияют на качество оценки риск-премий и, возможно, необоснованно приведут к негативной оценке качества работы моделей на российском рынке.

Выбор рыночной и безрисковой ставки

В качестве рыночной ставки для всех активов были использованы данные по котировкам на Московской бирже, информация о которых находится в открытом доступе, таким образом после получения дневных доходностей месячная доходность соответственно определялась как сумма дневных.

При определении ставки доходности безрискового актива было предложено протестировать сразу несколько вариантов:

- 1) Ставка по облигациям Федерального Займа Министерства Финансов Российской Федерации
- 2) Ставка по межбанковским кредитам
- 3) Ключевая ставка, устанавливаемая Банком России
- 4) Ставка по Еврооблигациям или облигациям Казначейства США

После проведенного анализа максимальной коррелированности безрисковой ставки с динамикой российского рынка акций было принято решение, что наиболее актуальным и показательным для нашего исследования будет являться выбор в качестве безрисковой ставки именно процента по ОФЗ МинФина РФ в силу оптимального сочетания коррелированности, минимизации рисков и актуальности в качестве безрискового актива для российского рынка.

Аналогичные исследования для российского рынка

Наиболее схожей по тематике работой с данной является исследование, проведенное Е.А. Фёдоровой и А.Р. Сивак в 2012 году и опубликованное в журнале «Финансы и кредит»¹¹. В нём авторы сравнивают между собой однофакторную и трехфакторную модели CAPM для российского рынка акций, в рамках своего исследования они собирают инвестиционный портфель из обычных и привилегированных акций, после чего разделяют компании на 6 категорий по параметрам Book equity/ Market equity и размерам фирмы соответственно.

В рамках работы не указывается, как именно рассчитывается доходность предложенного авторами «хорошо диверсифицированного» портфеля, так как не сказано о долях различных активов и не приведены данные о коррелированности акций между собой. После построения основных регрессоров на основе GARCH-моделей авторы рассчитывают условные вариации для оценки «выбросов» данных и лучшей описательной способности модели в вопросах волатильности, получая таким образом уравнение для трехфакторной модели и условную вариацию, используемую в дальнейшем для тестирования гипотез. Сами авторы сравнивают эффективность моделей на основе коэффициента детерминации (R^2), критерия Фишера (F-test), ARCH-LM и Jarque-Bera тестов.

В качестве выводов в своей работы авторы говорят о том, что так как трехфакторная модель является лишь расширением однофакторной, то результаты, полученные в обоих случаях, хорошо дополняют друг друга – можно сделать вывод о том, что значимость β -коэффициента для трехфакторной модели больше, чем двух оставшихся, однако их добавление

¹¹ Е.А. Федорова, А.Р. Сивак – «Сравнение моделей CAPM и Фамы-Френча на Российском фондовом рынке», журнал «Финансы и кредит», 42 (522) – 2012

позволяет сильно повысить точность предсказаний. Так же в ходе тестирования отклоняется гипотеза о незначимости двух новых факторов и принимается гипотеза о нормальности распределения нормированных остатков.

Завершением работы является выявленные авторами наиболее высоколиквидные безрисковые активы и обращение к участникам фондового рынка протестировать полученные в ходе работы выводы практически.

Работа с данными.

Расчет доходности за рассматриваемый период

Первым этапом работы был сбор данных за 2014–2019 годы и построение факторов для многофакторных моделей. Для проверки моделей был взят именно этот период, так как предполагается, что он однороден в терминах риска, то есть в течение него подверженность к систематическому риску для каждой из бумаг не менялась: оставались неизменными коэффициенты бета и альфа, а также другие коэффициенты регрессии, характеризующие влияние анализируемых факторов риска на доходность отдельной бумаги. Для обучения моделей будет взят период 2014-2018 гг., а для тестирования качества – 2019 г. Период в 5 лет показался достаточным для оценки коэффициентов регрессии, как для дневной доходности, так и для месячной, потому что он позволяет сохранить баланс между достаточным количеством точек для построения регрессий (1256 для дневной и 60 для месячной) и устойчивостью коэффициентов модели. Были взяты ежедневные цены закрытия для 65 бумаг фондового рынка, котирующихся на Московской бирже, которые в рассматриваемый период хотя бы единожды включались в индекс МосБиржи. На основе этих данных для каждой их бумаг была рассчитана ежедневная и месячная доходность: $\frac{P_t - P_{t-1}}{P_{t-1}}$. В число бумаг, которые затем подвергались оценке тремя моделями и на которых эти модели

тестировались, были взяты 38 тех, по которым не было пропущенных данных с 2014 по 2019 гг.

Конструирование факторов

Затем были сконструированы факторы риска SMB и HML для трехфакторной модели и SMB, HML, RMW, CMA для пятифакторной описанным в работе Фамы и Френча¹² способами. Для этого были собраны данные по количеству акций в обращении, балансовой стоимости, выручке, операционным и процентным расходам, активам за 2012–2018 гг. из годовой отчетности по стандартам МСФО для упомянутых 65 компаний. Затем из списка были исключены финансовые организации, так как высокий уровень долга, нормальный для этой отрасли, для других отраслей был бы крайне рискованным, потому показатели таких компаний неактуальны для построения риск-факторов на всем рынке. Подобный подход использован и в работе Фамы и Френча. Все показатели кроме количества акций были взяты в млрд российских рублей, те из них, что в отчетности были представлены в долларах были переведены в рубли по курсу от Центрального банка на конец отчетного года. Состав бумаг, которые входили в расчет доходности факторов, пересматривался ежегодно исходя из обновлений отчетности и того, входила ли бумага в этот год в индекс Московской биржи. Для каждой из бумаг в год t были посчитаны капитализация, отношение балансовой стоимости к рыночной (book-to-market ratio), прибыльность и инвестирование по следующим формулам из таблицы. Для формирования факторов на период t брались данные из отчетности за период $t-1$, так учитывалось влияние бухгалтерских показателей компании на факторы риска и на решения инвесторов в следующем периоде.

¹² См. Таблица 1. Стр 22

Таблица 1. Конструирование факторов.

Рыночная капитализация _t	$\text{Цена}_{t-1} \cdot \text{Количество акций в обращении}_{t-1}$
Book-to-market Ratio _t	$\frac{\text{Собственный капитал}_{t-1}}{\text{Рыночная капитализация}_{t-1}}$
Прибыльность _t	$\frac{\text{Выручка} - \text{Операционные расходы}_{t-1} - \text{Процентные расходы}_{t-1}}{\text{Собственный капитал}_{t-1}}$
Инвестирование _t	$\frac{\text{Активы}_{t-1} - \text{Активы}_{t-2}}{\text{Активы}_{t-2}}$

По рыночной капитализации рассчитывался фактор SMB, по book-to-market ratio HML, по прибыльности RMW, а по инвестированию степень агрессии CMA. Процесс формирования портфелей аналогичен описанным у Фамы и Френча, исходя из разбиения 2х3, первоначально по капитализации, а затем по каждому из других показателей. Однако на рынках с малой капитализацией и относительно небольшим присутствием компаний, каким можно назвать российский фондовый рынок, представляется логичным производить разбиение 2х2, так как есть риск, что при разбиении 2х3 в какую-то из групп попадут немного компаний, а это негативно скажется на степени диверсификации портфеля и ухудшит оценку систематического риска, который призваны оценивать каждая из трех моделей. Из-за недостаточного использования моделей на российском рынке и отсутствия сформировавшейся практики по конструированию факторов риска было решено строить регрессии как для факторов, построенных на основе портфелей 2х3, так и на основе разбиения 2х2, а затем сравнить результаты. Каждая из бумаг в год t

попадала в одну из групп¹³. Сначала все бумаги по медиане капитализации попадали либо в категорию Small, либо в Big, затем для разбиения 2х3 среди Small и среди Big формировался список из компаний по 30% и 70% квантилям. Так строился состав портфелей. Каждая бумага входила в портфель, взвешенная по капитализации, то есть занимая долю в портфеле, пропорционально отношению капитализации компании, эмитировавшей эту бумагу, к сумме капитализаций других компаний, входящих в портфель. Окончательно факторы были посчитаны следующим образом (при разбиении 2х2 для подсчета SMB не было нейтральных факторов SN и BN).¹⁴

Трехфакторная модель

$$SMB = (SH + SN + SL)/3 - (BH + BN + BL)/3$$

$$HML = (SH + BH)/2 - (SL + BL)/2$$

Пятифакторная модель

$$SMB_{B/M} = (SH + SN + SL)/3 - (BH + BN + BL)/3$$

$$SMB_{OP} = (SR + SN + SW)/3 - (BR + BN + BW)/3$$

$$SMB_{Inv} = (SC + SN + SA)/3 - (BC + BN + BA)/3$$

$$SMB = (SMB_{B/M} + SMB_{OP} + SMB_{Inv})/3$$

$$HML = (SH + BH)/2 - (SL + BL)/2$$

$$RMW = (SR + BR)/2 - (SW + BW)/2$$

$$CMA = (SC + BC)/2 - (SA + BA)/2$$

¹³ См. Приложения Таблица 2, Таблица 3, стр. 45-46

¹⁴ См. Трёхфакторная модель, Пятифакторная модель, стр. 23-24

Благодаря такому методу и усреднению портфелей для подсчета фактора SMB можно добиться большей диверсификации, потому что иначе портфели для разных факторов могли бы сильно коррелировать, если бы двухэтапного разбиения не происходило: так как капитализация, используемая для подсчета фактора SMB, стоит в знаменателе фактора HML и немало говорит о прибыльности бизнеса, учитывая, что количество акций в обращении ни для одной рассматриваемой компании не изменилось за 2014–2019 гг.

В качестве рыночной доходности была взята доходность индекса МосБиржи (IMOEX), а в качестве безрисковой ставки – доходность индекса совокупного дохода государственных облигаций 3–5 лет (RUGBITR5Y). И если вопрос выбора рыночного индикатора для российского фондового рынка в литературе определен довольно точно: причем показано, что неважно, брать индекс полной доходности с дивидендами или нетто, из-за высокой корреляции, то выбор безрисковой ставки не так ясен. Главная черта безрисковой ставки – определенность инвестора по поводу будущих выплат: отсутствие риска заключается в гарантии выплат и в отсутствии риска реинвестирования. Такими характеристиками в наибольшей степени обладают именно долгосрочные государственные ценные бумаги, потому было решено остановиться на RUGBITR5Y.

Выбор композиции разбиения

Из-за низкой капитализации российского фондового рынка и относительно небольшого количества компаний, эмитирующих ценные бумаги, построение портфелей для подсчета риск-факторов методом разбиения 2x3 может обернуться наличием портфелей в составе двух бумаг, что, очевидно, влияет на степень диверсификации портфеля, а значит и на релевантность риск-метрики в негативную сторону. Именно поэтому факторы

были посчитаны двумя способами: с разбиением портфелей по квантилям, т. е. 2х3, и по медиане, т. е. 2х2. В Приложениях видно¹⁵ распределение бумаг по портфелям в двух случаях разбиения. При малой диверсификации в разбиении 2х3 в крайние категории, которые и влияют на расчет факторов, входят одни и те же бумаги, что должно увеличивать корреляцию между различными факторами и делать один из скоррелированных факторов избыточным для включения в модель.

С целью определения оптимального способа разбиения портфелей для подсчета факторов риска в специфике российского фондового рынка были построены корреляционные матрицы для факторов. Из результатов¹⁶ видно, что нельзя однозначно сказать, что при разбиении 2х3 корреляция факторов более высока, напротив, заметно больше корреляция между фактором RMW и HML, CMA для разбиения 2х2. Таким образом, на этом этапе нельзя определить, какой из способов построения даст более эффективный результат, поэтому тестирование многофакторных моделей будет проводиться как по факторам 2х2, так и 2х3.

Методология

Построение регрессий

В классической модели CAPM вычисление β -коэффициента осуществляется посредством построения линейной регрессии на основе данных дневных доходностей, в рамках этой работы были рассмотрены разные подходы к реформированию ортодоксального подхода – во-первых, была предложена идея перехода от наиболее простой с точки зрения реализации линейной регрессии к более сложным моделям, а именно квантильной или медианной регрессиям, полиномиальной или гребневой (ridge-regression)

¹⁵ См. Приложения Таблица 2, Таблица 3, стр. 45-46

¹⁶ См. Приложения Рисунок 2. Рисунок 3 стр. 47

регрессионным моделям, однако ни одна из предложенных регрессий не способна разрешить ключевую проблему работы с рыночными данными – гетероскедастичность, а именно большое скопление данных в окрестности нуля (то есть околонулевая дневная доходность акции в среднем) при наличии редких выбросов, которые сильно снижают описательную точность любой из рассматриваемых нами регрессий, не имеющие необходимого инструментария для предсказания несистемных рисков (а при попытке импликации такого подхода возникает проблема переобучения, снижающая прогнозную точность для всех случаев в целом).

Другим направлением, в котором так же велась работа, являлась постобработка полученных данных, в частности рассматривалась идея создания нормированной безразмерной базы доходностей для лучшего сравнения предположительно однородных данных для единого временного периода (например, при декомпозиции портфелей в пятифакторной модели различными подходами).

Одной из наиболее простых и часто используемых корректировок данных, получаемых при построении регрессии, является использование поправок к результатам регрессии, для российского рынка¹⁷ наиболее релевантными и уместными поправками являются:

1) Вариант М. Блюма

$$\beta_{Blume} = 0.67 * \beta_{OSL} + 0,33 * 1$$

Где β_{OSL} – показатель, полученный при построении линейно регрессии.

2) Вариант Шоулза-Виллимса:

¹⁷ Т.В.Теплова, Н.В. Селиванова – «Эмпирическое исследование применимости модели DСАРМ на развивающихся рынках», журнал «Корпоративные финансы», №3 2007, стр. 5 - 23

$$\beta_{sw} = \frac{\beta^{-1} + \beta + \beta^{+1}}{(1 + 2\rho_m)},$$

Где β – показатель, полученный при построении регрессии для i – го периода, β^{-1} – для $(i - 1)$ – го периода, а β^{+1} – для $(i + 1)$ – го периода соответственно. ρ_m – коэффициент автокорреляции рыночной доходности.

Первый вариант используется Bloomberg и ValueLine, второй вариант чаще всего используют для построения вариации модели MСРМ.

В рамках нашей работы было принято решение строить линейные регрессии для дневных и месячных данных, а также дневных данных с плавающим окном, так как для проверки сформулированных нами гипотез было достаточно установлении точности посредством работы с факторами, нежели экспериментами с новаторством в области оптимизации, которая могла бы нивелировать грань прогностической точности для различных моделей.

Каждая из вариаций моделей обучалась на периоде 2014–2018 гг., так, для модели с дневными доходностями получилось 1257 точек, с месячными – 60. Для скользящего окна проводилась оценка коэффициентов регрессии каждые 3 месяца, потому из рабочих дней получалось 60 точек, затем окно смещалось на одну точку и проводилась новая оценка.

Для моделей с дневными и месячными данными затем строился пост-прогноз дневных и месячных доходностей соответственно на 2019 г. Для модели со скользящим окном для каждого из окон строился пост-прогноз дневных доходностей на следующий месяц.

Формулирование и проверка гипотез

Модель CAPM и ее вариации статичны, т. е. ориентированы, в первую очередь, на описание рынка и объяснение доходности актива, а не на его прогнозирование. Однако в рамках работы будет также сравниваться качество пост-прогноза, сделанного моделями с учетом оценок коэффициентов регрессии на предыдущих периодах.

Обучение во всех трех модификациях модели проводилось на данных 2014–2018 гг., а по рыночным и факторным доходностям 2019 г. была получена оценка прогноза.

Построение трех регрессий разными способами должно было объяснить:

- 1) Улучшается ли качество модели с её усложнением, или с увеличением количества параметров
- 2) Какой из способов построения регрессии позволит достичь наивысшего качества
- 3) На всех ли активах работает модель, наблюдается ли устойчивость
- 4) Что объединяет те активы, на которых модель работает хорошо, и те, на которых не работает

Для ответов на эти вопросы отдельно проводилась оценка моделей по их описательной силе на данных тренировочной выборки и по прогнозной на доходностях из тестовой. Основные факторы, влияющие на сравнение моделей: коэффициент детерминации R^2 и р-значение коэффициента пересечения, а также их динамика с усложнением моделей. Как известно, каждая из моделей претендует на то, чтобы учитывать все риск-факторы, влияющие на доходность актива. И если для оригинальной модели CAPM это только рыночный фактор, то каждая следующая была нацелена на то, чтобы охватить все больше параметров, влияющих на оценку риска на рынке. Во всех этих моделях отсутствует коэффициент пересечения α . Соответственно, в полученных в ходе построения регрессий оценках для α может содержаться

необъясненная моделями информация о доходностях активов. Именно поэтому важно проверить, что

- α уменьшаются с усложнением модели
- $\alpha = 0$ для отдельного актива

Для этого будет рассматриваться р-значение коэффициента α . Если оно окажется слишком малым и гипотеза о равенстве α нулю будет отвергаться, то это будет означать, что модель не объясняет все факторы риска при верных предположениях. Но отвержение гипотезы также может быть связано с издержками оценки моделями отдельных активов, а не портфелей, а значит улавливанием коэффициентом α несистематических рисков, связанных с отсутствием диверсификации, которые факторы модели CAPM и её вариаций не призваны учитывать.

Таким образом, по совокупности показателей будет выбрана лучшая модель для статичного описания доходности активов и сделаны выводы о возможных причинах плохого качества работы моделей на некоторых активах.

Для тестирования прогностической силы модели будут получены значения функции потерь MAPE (Mean Absolute Percentage Error). Оценка качества прогноза будет производиться именно ей, так как она измеряется в процентах и лишена разницы из-за размерности дневной и месячной доходности, которая присуща MSE (Mean Squared Error).

Результаты

Общее описание результатов

Главным итогом нашей работы стало построение трех моделей, которые могли бы с разной точностью описывать и предсказывать динамику рынка, исходя из располагаемых данных.

После построения пяти рыночных факторов на основе представляемой компаниями-эмитентами финансовой отчетности мы обработали массив

данных по дневным котировкам акций за пять лет, после чего создав так же базу данных для месячных доходностей. На основе полученных данных были построены регрессионные модели для одно-, трёх- и пятифакторной интерпретации CAPM, при том для любой из них функция задавалась в виде линейной комбинации многих переменных (в частности, факторов).

После обучения регрессий на данных 2014–2018 годов, были получены соответствующие уравнения, для которых проверялись различные метрики (R^2 , р-значение, MAPE и прочие) и гипотезы (О равенстве свободного члена линейной регрессии нулю, о равенстве выборочной и прогнозируемой дисперсий и прочие). Так же посредством обученных регрессий был дан прогноз на 2019 год, согласованность которого с реальными рыночными данными так же являлась одним из ключевых предметов нашей работы.

Итоговым результатом нашей работы стало сравнение моделей и способов их построения на различных форматах исследуемых баз данных (для дневных, месячных котировок и при построении скользящего окна), выявление активов, описательная точность моделей для которых была оптимальной, выявление возможных перспектив дальнейшего развития подходов к построению моделей и метрик для оценки их точности.

Лучшая описательная модель

Коэффициент детерминации R^2

Коэффициент детерминации является одной из наиболее популярных метрик оценки качества моделей, однако имеет свою специфику, в частности, он может увеличиваться при добавлении в модель новых данных. Таким образом не совсем корректно сравнивать между собой модели, в которых имеет место разное количество обучаемых переменных или же объём данных, однако в силу совпадения исследуемых данных и логического «дополнения» модели при её эволюции от однофакторной до пятифакторной мы можем

сделать состоятельный вывод об улучшении\ухудшении точности модели на одинаковых данных.

Таблица 4 Среднее значение коэффициента детерминации R^2

R2	Дневная	Скользящее окно	Месячная
CAPM	0,244	0,240	0,186
3 Factor 2x3	0,269	0,319	0,263
3 Factor 2x2	0,276	0,317	0,266
5 Factor 2x3	0,291	0,376	0,347
5 Factor 2x2	0,293	0,376	0,301

Из представленных данных в таблице¹⁸ хорошо видно, что наибольшей точностью обладают модели, в которых, во-первых, имеет место дополнительное конструирование факторов (трёхфакторная и пятифакторная соответственно), а, во-вторых, использование в качестве методики оценивания скользящее окно. Также можно заметить, что декомпозиция факторов 2x3 и 2x2 не сильно оказывает влияние на итоговый результат, в том числе и учитывая, что дисперсии для распределения коэффициентов детерминации по каждой акции в отдельности почти совпадают для каждой из рассматриваемых моделей.

В соответствии с полученными результатами, наибольшая описательная точность для всех рассматриваемых моделей была получена при исследовании акций крупных корпораций, в особенности связанных с нефтедобывающей и нефтеперерабатывающей отраслью (GAZP, LKOH, ROSN)¹⁹, для всех этих компаний расчетный коэффициент детерминации был максимальным в ряду,

¹⁸ См Таблица 4, стр. 31

¹⁹ См. Приложения Рисунок 3, Рисунок 4, Рисунок 5, стр. 48

в частности для пятифакторной модели 2×2 и 2×3 он превышал во всех случаях 0,7 (максимальное значение $R^2 = 0,757$ было получено при описании GAZP).

Компании, для которых точность была наименьшей, не могут быть классифицированы столь же общим признаком, как наиболее точно описываемые в рамках модели, здесь представлены крупные ритейлеры, строительные и торговые компании, а также некоторые фирмы, занимающиеся разработкой и добычей драгоценных металлов (PLZL, NMTP, SVAV, MVID, MSTT и прочие)²⁰, для них коэффициент детерминации не превышал показателя 0,25 даже на моделях с наибольшей точностью (а именно пятифакторная с обеими декомпозициями и построением скользящего окна).

Стоит отметить, что по мере роста описательной точности моделей не было замечено значительных «выбросов» в описательной точности, а именно при построении на однородных данных все более и более продвинутых моделей наблюдался пропорциональный рост коэффициента детерминации, не противоречащий предположениям, сделанным в начале работы.

Принимая во внимание тот факт, что коэффициент детерминации не может выступать как абсолютно верный показатель качества используемых моделей, нельзя сделать однозначный вывод о целесообразности применения какого-либо подхода для работы с данными, однако на основании полученных результатов можно заметить, что в отличие от однофакторной модели, для которой точность существенно повышалась при работе с предсказаниями для дневных котировок, пятифакторная модель давала более точные результаты при использовании сгруппированных месячных данных, тогда как для скользящего окна качество модели планомерно возрастало при добавлении большего количества факторов. Декомпозиция портфелей 2×2 и 2×3 не сильно сказалась на полученных результатах, так как коэффициент детерминации в среднем не сильно различался, а выборочные дисперсии в свою очередь

²⁰ См. Приложения Рисунок 6, Рисунок 7, Рисунок 8, стр. 49

практически совпадали как для трехфакторной, так и для пятифакторной моделей.

P-value для коэффициента α

Одним из показателей качества работы модели и ее описательной силы было распределение оценок коэффициента пересечения. По предположению оригинальной модели CAPM и её вариаций, $E(\alpha) = 0$. Эта гипотеза проверялась при построении оценок для коэффициента. Для каждого из активов было получено р-значение гипотезы о равенстве математического ожидания α нулю. Чем больше это значение, тем менее значим для описания доходности рассчитанный коэффициент. Для каждой вариации модели р-значения по всем активам были усреднены и в среднем ни для одной из моделей гипотеза о $E(\alpha) = 0$ не отвергалась.

Таблица 5 Среднее р-значение по коэффициенту пересечения

α P-value	Дневная	Скользящее окно	Месячная
CAPM	0.379	0.512	0.341
3 Factor 2x3	0.378	0.522	0.365
3 Factor 2x2	0.379	0.526	0.361
5 Factor 2x3	0.38	0.53	0.363
5 Factor 2x2	0.37	0.525	0.365

Во всех вариациях модели при построении регрессии по дневным и по месячным доходностям на пятилетнем периоде находились активы, для которых гипотеза отвергалась на 5%-ом уровне значимости, но в

то же время находились акции, по которым p -значение для константы оказывалось больше 0.9, что говорит о близости $E(\alpha)$ к нулю. В среднем более высокие p -значения получились при оценке коэффициентов методом скользящего окна с периодом в 3 месяца. Однако для этого метода ни в одной модели не нашлось акции, по которой гипотеза бы отвергалась даже на 10%-ом уровне значимости. При этом также не нашлось актива с p -значением выше 0.8.

Также важно отметить, что наблюдается устойчивость в том, какие акции оказываются среди тех, где гипотеза о равенстве $E(\alpha)$ нулю отвергается и где она, напротив, оказывается высоковероятной²¹. Интересно, что для некоторых бумаг устойчивым оказывается результат, рассчитанный только по месячной доходности.

Рассмотрим распределение оценок α бумаги MAGN и выявим закономерности, из-за которых гипотеза о равенстве α нулю отвергается.

Из рис.²² видно, что оценка α на периоде в 5 лет смещена вправо, в положительную сторону. Совмещенный график регрессии и месячной доходности рынка²³ показывает, что из-за выбросов, изображенных в верхней правой части графика, что соответствует более низким доходностям бумаги к рыночным по сравнению с типичным состоянием, прямая оказалась очень пологой. Это повлекло за собой увеличение коэффициента пересечения. Видно, что на рис.²⁴ благодаря более узкому периоду построения регрессии $\alpha = 0$ входит практически во все доверительные интервалы, что позволяет достичь усредненного p -значения в 0.487. Соответственно, выбросы обусловлены короткими периодами изменения рыночной конъюнктуры и соотношениям риска на рынке, бумага на какой-то период стала относительно безопасной и её коэффициент β стал ниже. Когда модель, построенная на

²¹ См. Приложения Таблица 6 стр. 50

²² См. Приложения Рисунок 10, Рисунок 11 стр. 51

²³ См. Приложения Рисунок 12 стр. 51

²⁴ См. Приложения Рисунок 13 стр. 51

длинном периоде в 5 лет, пытается уловить эти изменения, минимизируя функцию потерь, получается высокий коэффициент пересечения. Все это не позволяет модели выявить наиболее объективную зависимость доходности от риск-факторов. Видно, что аналогичные выводы можно сделать и по другим бумагам, для которых р-значение оказывалось стабильно малым.²⁵

Таким образом, анализ коэффициентов α показал, что более точные оценки подверженности риск-факторам можно получить на данных более узких периодов, чем 5 лет.

Наилучшая модель для российского рынка

Исходя из совокупности рассуждений, приведённых выше, нельзя сделать однозначный вывод о том, какая модель лучше или хуже описывает динамику российского рынка, результат оказался неоднородным по активам, отчего неопределенность усилилась.

Во-первых, при достаточно низкой волатильности и высокой ликвидности активов почти любая модель показывает себя достаточно состоятельной в качестве описания динамики рынка, однако наилучшие результаты мы наблюдались для трехфакторной модели с декомпозицией портфелей 2x2, в то же время если работать с данными по неликвидной акции, то любая из моделей показывает весьма пессимистичные результаты, а причины таких результатов связаны как с особенностями динамики доходностей отдельных бумаг, так и с особенностями формирования риск-факторов, связанными с низкой капитализацией российского фондового рынка.

Во-вторых, в случае работы с данными по дням точность моделей по мере увеличения числа переменных возрастает гораздо меньше, чем для месячных данных и скользящего окна, однако во всех трёх случаях

²⁵ См. Приложения Таблица 6 стр. 50

наблюдается положительная зависимость точности модели от количества риск-факторов (что, очевидно, согласовывается с концепцией о «надстройке» более сложных моделей над однофакторной CAPM). Таким образом при работе со скользящим окном и месячными данными выигрыш более чем в 1,4 и 1,8 раз соответственно будет иметь пятифакторная (любая из двух декомпозиций) модель при сравнении её с однофакторной.

В-третьих, стоит заметить, что разные способы построения данных дают разные результаты, однако их интерпретация зависит от конкретной задачи, которая ставится в рамках проекта, поэтому не совсем корректным является сравнение их даже с использованием безразмерных коэффициентов оценки точности предсказаний, таким образом для каждой задачи следует выбирать свою модель, однако при любой конфигурации данных пятифакторная модель показывает лучшие результаты для любого критерия.

Лучшая прогностическая модель

MAPE

Таблица 7. Средние показатели MAPE по прогнозам

	Дневная	Скользящее окно	Месячная
CAPM	1193	428	311
3 Factor 2x3	483	566	346
3 Factor 2x2	753	586	367
5 Factor 2x3	1988	576	497
5 Factor 2x2	878	581	575

Для сравнения прогностической силы моделей были рассчитаны коэффициенты MAPE. Из табл. выше²⁶ видна устойчивость результатов, полученных с помощью скользящего окна и по месячным данным, однако для дневных данных эта стабильность не была обнаружена. Важно также отметить, что внутри каждого варианта построения моделей (по дневным данным, по месячным и скользящим окном) не найдено закономерности в динамике коэффициента MAPE при увеличении числа факторов. Таким образом, невозможно сделать однозначный вывод о том, какая из вариаций модели SARIM лучше подходит для пост-прогнозов.

Вероятнее всего, такая глобальная неустойчивость прогностических результатов связана с особенностями динамики рынка и подверженности рискам. Наиболее устойчивыми получились результаты моделей, построенных методом скользящего окна, потому как оценки коэффициентов регрессии проводились по узкому периоду в 60 дней. Разница между различными способами построения регрессий в коэффициентах MAPE объясняется степенью обобщения модели. При использовании месячных данных на 5-летнем периоде переобучения практически не наступает, благодаря тому что шумов почти не остается, этим же объясняются наилучшие показатели по MAPE в сравнении с регрессиями, обученными на дневных данных. Использование дневных доходностей на периоде в 5 лет из-за обилия шумов и отсутствия баланса между точностью и прогнозным периодом показало наихудшие и неустойчивые коэффициенты MAPE. Метод скользящего окна продемонстрировал устойчивость в оценке, однако из-за узости периода обучения модели шириной в 3 месяца, по-видимому, наступало переобучение, и модель не могла качественно построить прогноз на следующий месяц.

Таким образом, нельзя с уверенностью говорить о превосходстве одной из вариаций моделей в прогнозной силе над другой, так как полученные

²⁶ См. Таблица 7 стр. 37

оценки коэффициента качества прогноза MAPE не продемонстрировали устойчивой динамики. Однако однозначно можно утверждать о преимуществах использования месячных доходностей для построения прогнозных оценок из-за снижения вероятности переобучения. На данном этапе метод скользящего окна показал меньшее прогнозное качество, но это может быть связано и с необходимостью поиска баланса между периодом обучения и периодом прогнозирования. Не исключено, что при ином соотношении длины данных для обучения и для прогноза, метод скользящего окна показал бы наилучшие результаты.

Качество работы моделей для разных активов: причины и следствия

Стоит заметить, что не для всех акций, которые подвергались оценке, модели работают одинаково, иначе говоря, результат неоднороден и на каких-то ценных бумагах результат стабильно плохой, на других – стабильно хороший. Интересно разобраться в причинах такого разделения, это может дать идеи для дальнейшей работы над моделями в рамках российского фондового рынка.

Среди лидеров по показателям R2: акции Газпрома, Лукойла, Роснефти, Новатэка и Норникеля. Наихудшие показатели устойчиво показывают: М-Видео, Россети, Группа ЛСР.

Из графиков прогнозного и реального индекса цен акций²⁷ по отношению к 6 января 2014 г. (первая точка данных) видно, что по акциям с высокой описательной силой предсказания, основная ошибка состоит в большом смещении. Верхние и нижние экстремумы модель определяет довольно неплохо. При этом с усложнением модели и с увеличением количества факторов наблюдается увеличение коэффициента детерминации за счёт снижения смещения. Также видно, что в первой трети периода 2014–2018 гг. на всех трёх моделях получается качественный прогноз относительно

²⁷ См. Приложения Рисунок 20, Рисунок 21, Рисунок 22 стр. 54

остальных данных. Основное ухудшение качества наблюдается позднее и улучшение каждой модели происходит, в первую очередь, благодаря снижению смещения именно на этой части данных. Отсюда логичное предположение, что рыночный фактор являлся основным риск-фактором в первой трети периода, затем зависимость от рыночного фактора уменьшалась в сторону увеличения зависимости от других, используемых в многофакторных моделях. Естественный вывод заключается в том, что предположение об однородности тренировочного периода неверно, а значит зависимость от факторов во времени не является статичной.

Похожая картина складывается среди бумаг, на которых модель показала среднюю²⁸ описательную силу. В первой трети рассматриваемого периода прогностическая доходность очень близка к реальной, затем наблюдается увеличение смещения. И так же с каждым усложнением модели смещение уменьшается и пики устремляются к совпадению. Потому аналогичные выводы относительно зависимости от риск-факторов можно сделать и по этим бумагам.

Для бумаг с плохим²⁹ коэффициентом детерминации R^2 усложнение моделей также давало улучшение описательной силы, однако не всегда такое заметное, как для акций, на которых модели работали хорошо. И если для рассмотренных ранее активов модель хорошо улавливала локальные экстремумы, однако на некоторых участках была смещена, то для этой группы бумаг модель явно не обратила внимание на многие реальные яркие подъемы, падения и тренды. Для бумаг, которые попали в эту категорию, по-видимому, работают иные риск-факторы.

Из полученных результатов: графиков прогнозных и реальных цен, распределения бумаг по группам – следуют предположения о возможных

²⁸ См. Приложения Рисунок 23, Рисунок 24, Рисунок 25, стр. 55

²⁹ См. Приложения Рисунок 26, Рисунок 27, Рисунок 28 стр. 56

причинах такой устойчивой разницы в силах моделей, которые можно будет учесть в последующих работах о моделях CAPM на российском рынке.

Первое предположение связано с построением риск-факторов. Все портфели, используемые для конструирования факторов, были сформированы с учетом рыночной капитализации на начало года каждой из компаний. К тому же и вес бумаги, с которым она входит в Индекс Мосбиржи, зависит от размеров капитализации. Получается, что доходности компаний с высокой стоимостью больше коррелированы с доходностями риск-факторов, чем те компании, которые вошли в портфели с малым весом из-за своей невысокой стоимости. Этим объясняется тот факт, что наилучшее качество показывают модели, которые вошли в категорию ‘Big’ при построении факторов, т. е. с высокой капитализацией.

Второе предположение основано на том, что для бумаг с плохим качеством более ярко выражены временные лаги: прогнозные падения случаются несколько раньше реальных. Иными словами, из-за специфики бумаг, особенностей информированности участников рынка, падение риск-факторов в одном периоде влияет на доходность в другом. Однако оригинальная версия CAPM не предполагает такой динамической зависимости. Это становится причиной горизонтального, а не вертикального смещения локальных экстремумов.

Третье предположение связано с более высокой волатильностью бумаг из плохой категории. Резкие скачки и отсутствие ярко выраженного стабильного долгосрочного тренда говорит о возможно более частых изменениях подверженности риск-факторам. Из этого следует вывод о необходимости оценивать коэффициенты регрессии для таких активов на более коротком промежутке данных, чтобы улавливать эти более частые изменения.

Выводы

Исходя из совокупности рассуждений, приведённых выше, нельзя сделать однозначный вывод о том, какая модель лучше или хуже описывает динамику российского рынка, результат оказался неоднородным по активам, отчего неопределенность усилилась.

Основная идея состоит в том, что полученные результаты гораздо сильнее разнятся в зависимости от описываемой акции, чем в зависимости от используемой модели. Для бумаг с высокой рыночной капитализацией наилучшие метрики чаще всего получались при использовании 5-факторной модели с портфелями 2x2. И гораздо более пессимистичные результаты демонстрируют акции с высокой волатильностью и низкой капитализацией, а причины таких результатов связаны как с особенностями динамики доходностей отдельных бумаг, так и с особенностями формирования риск-факторов, связанными с низкой капитализацией российского фондового рынка.

Также в случае работы с данными по дням точность моделей по мере увеличения числа переменных возрастает гораздо меньше, чем для месячных данных и скользящего окна, однако во всех трёх случаях наблюдается положительная зависимость точности модели от количества риск-факторов (что, очевидно, согласовывается с концепцией о «надстройке» более сложных моделей над однофакторной CAPM). Таким образом при работе со скользящим окном и месячными данными выигрыш более чем в 1,4 и 1,8 раз соответственно будет иметь пятифакторная (любая из двух декомпозиций) модель при сравнении её с однофакторной.

Чаще наилучшие результаты среди многофакторных моделей демонстрировали те, что были обучены на факторах, построенных по портфелям 2x2, однако этот результат нельзя назвать устойчивым, для разных бумаг он отличался и для разных метрик тоже.

Важно смотреть и на динамику метрик: в основном все из них улучшаются при увеличении количества риск-факторов в модели. Это свидетельствует о том, что модель CAPM и её вариации имеют основание быть использованными для описания доходностей на российском фондовом рынке, а её предпосылки не быть отвергнутыми.

Стоит также заметить, что разные способы построения данных дают разные результаты, однако их интерпретация зависит от конкретной задачи, которая ставится в рамках проекта, поэтому не совсем корректным является сравнение их даже с использованием безразмерных коэффициентов оценки точности предсказаний, таким образом для каждой задачи следует выбирать свою модель, однако при любой конфигурации данных пятифакторная модель показывает лучшие результаты для любого критерия.

Ещё одним важным итогом стало то, что для получения достойных метрик важно правильно подбирать однородный период для оценки статичных коэффициентов, причем, весьма вероятно, что для разных бумаг этот период также будет разным, в зависимости от внутренних процессов компании, не учтённых в риск-факторах.

Описательная точность по дневным данным продемонстрировала себя гораздо лучше, чем по месячным, однако для построения пост-прогнозов более качественный результат показывали модели, построенные на месячных данных, ввиду большей устойчивости оценок.

Перспективы дальнейшей работы с моделями в рамках российского рынка были обрисованы в ходе работы: это предобработка данных и избавление от шумов, использование иных способов оценки прогнозов, анализ специфики каждой из бумаг для нивелирования влияния ее особенностей на результаты, введение динамики для некоторых активов с запоздалым реагированием на рыночные потрясения и другое. Возможно, с этими корректировками модели покажут себя гораздо лучше и более уверенно войдут в инструментарий инвесторов.

Список используемой литературы

- Т.В.Теплова, Н.В. Селиванова – «Эмпирическое исследование применимости модели DСАРМ на развивающихся рынках», журнал «Корпоративные финансы», №3 2007, стр. 5 – 23
- Е.А. Федорова, А.Р. Сивак – «Сравнение моделей САРМ и Фамы-Френча на Российском фондовом рынке», журнал «Финансы и кредит», 42 (522) – 2012
- Robert Faff, «An Examination of the Fama and French Three-Factor Model Using Commercially Available Factors» //Journal of Finance. 2001. P. 1-17;83-92
- Fama E. F., French K. R. «The cross section of expected stock returns» //Journal of Finance. 1992. Vol. 47. P. 427–465.
- Fama, Eugene F.; MacBeth, James D. (1973). «Risk, Return, and Equilibrium: Empirical Tests». *Journal of Political Economy*. **81** (3): 607–636
- Donald E. Farrar «The Investment Decision Under Uncertainty» (Prentice-Hall, 1962)
- Josef Lakonishok, Andrei Shleifer, Robert W. Vishny «Contrarian Investment, Extrapolation, and Risk»//Journal of Finance. 1994 . P. 1541–1578.
- Sharpe, William F. (1964). "Capital asset prices: A theory of market equilibrium under conditions of risk". *Journal of Finance*.

Приложения

Таблица 2, Конструирование факторов с разбиением портфелей 2х3 на 2014 год

Small	Book-to-Market (BM)	High (SH)	FEES, MAGN, IRAO, MTLR, MSRS, TATNP, RTKMP, BANEP
		Neutral (SN)	VSMO, TRMK, AFLT, LSRG, SVAV, GCHE, NMTP, RUALR, PHST
		Low (SL)	POLY, PHOR, PIKK, DIXY, MVID, MSTT
	Operating Profitability	Robust (SR)	TATNP, BANEP, AFLT, MVID, MTLR, PHST
		Neutral (SN)	RTKMP, VSMO, PHOR, PIKK, DIXY, MSTT, SVAV, MSRS
		Weak (SW)	POLY, TRMK, FEES, LSRG, RUALR, IRAO, MAGN, GCHE, NMTP
	Investment	Conservative (SC)	RTKMP, BANEP, FEES, RUALR, IRAO, MAGN, MTLR, SVAV
		Neutral (SN)	TATNP, TRMK, PHOR, AFLT, PIKK, DIXY, NMTP, MSRS, GCHE
		Aggressive (SA)	POLY, VSMO, LSRG, MVID, MSTT, PHST
Big	Book-to-Market (BM)	High (BH)	GAZP, SNGS, SNGSP, TRNFP, HYDR, RSTI
		Neutral (BN)	LKOH, ROSN, TATN, AFKS, BANE, ALRS, CHMF, NLMK, UPRO
		Low (BL)	MGNT, NVTK, MTSS, GMKN, URKA, RTKM, MFON, MOEX
	Operating Profitability	Robust (BR)	LKOH, ROSN, TATN, MGNT, MTSS, GMKN, MFON, BANE
		Neutral (BN)	GAZP, NVTK, URKA, RTKM, AFKS, TRNFP, ALRS, UPRO
		Weak (BW)	SNGS, SNGSP, HYDR, CHMF, NLMK, RSTI
	Investment	Conservative (BC)	GMKN, URKA, RTKM, BANE, NLMK, UPRO, RSTI
		Neutral (BN)	GAZP, MTSS, TATN, AFKS, TRNFP, HYDR, CHMF
		Aggressive (BA)	LKOH, MGNT, SNGS, SNGSP, NVTK, ROSN, MFON, ALRS, MOEX

Small	Book-to-Market (BM)	High (SH)	FEES, MAGN, IRAO, MTLR, MSRS, TATNP, RTKMP, BANEP, LSRG, GCHE, RUALR
		Low (SL)	POLY, PHOR, PIKK, DIXY, MVID, MSTT, VSMO, TRMK, AFLT, SVAV, NMTP, PHST
	Operating Profitability	Robust (SR)	TATNP, BANEP, AFLT, MVID, MTLR, PHST, VSMO, PHOR, PIKK, SVAV
		Weak (SW)	POLY, TRMK, FEES, LSRG, RUALR, IRAO, MAGN, GCHE, NMTP, RTKMP, DIXY, MSTT, MSRS
	Investment	Conservative (SC)	RTKMP, BANEP, FEES, RUALR, IRAO, MAGN, MTLR, SVAV TRMK, PIKK, NMTP
		Aggressive (SA)	POLY, VSMO, LSRG, MVID, MSTT, PHST, TATNP, PHOR, AFLT, DIXY, MSRS, GCHE
Big	Book-to-Market (BM)	High (BH)	GAZP, SNGS, SNGSP, TRNFP, HYDR, RSTI, LKOH, ROSN, TATN, AFKS, CHMF, NLMK
		Low (BL)	MGNT, NVTK, MTSS, GMKN, URKA, RTKM, MFON, MOEX, BANE, ALRS, UPRO
	Operating Profitability	Robust (BR)	LKOH, ROSN, TATN, MGNT, MTSS, GMKN, MFON, BANE, GAZP, NVTK, AFKS, TRNFP, ALRS
		Weak (BW)	SNGS, SNGSP, HYDR, CHMF, NLMK, RSTI, URKA, RTKM, UPRO
	Investment	Conservative (BC)	GMKN, URKA, RTKM, BANE, NLMK, UPRO, RSTI, MTSS, AFKS, TRNFP, HYDR, CHMF
		Aggressive (BA)	LKOH, MGNT, SNGS, SNGSP, NVTK, ROSN, MFON, ALRS, MOEX, GAZP, TATN

Таблица 3, Конструирование факторов с разбиением 2x2 на 2014 год

Рисунок 2. Корреляционная матрица для декомпозиции 2x2

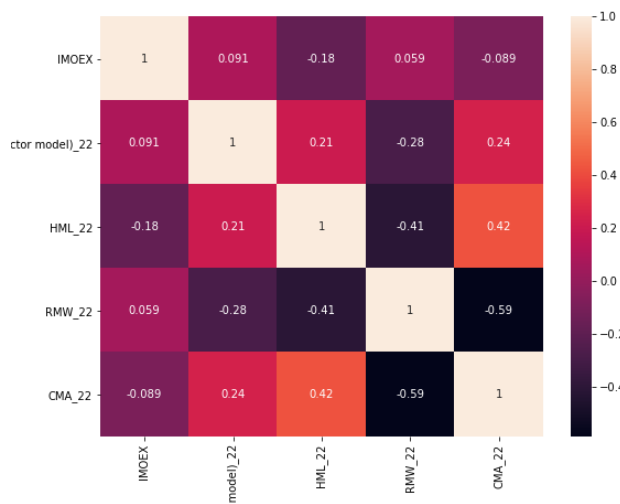
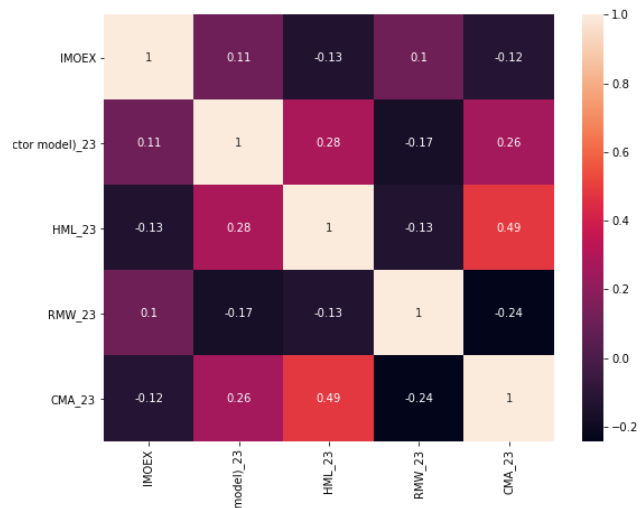


Рисунок 3. Корреляционная матрица для декомпозиции 2x3



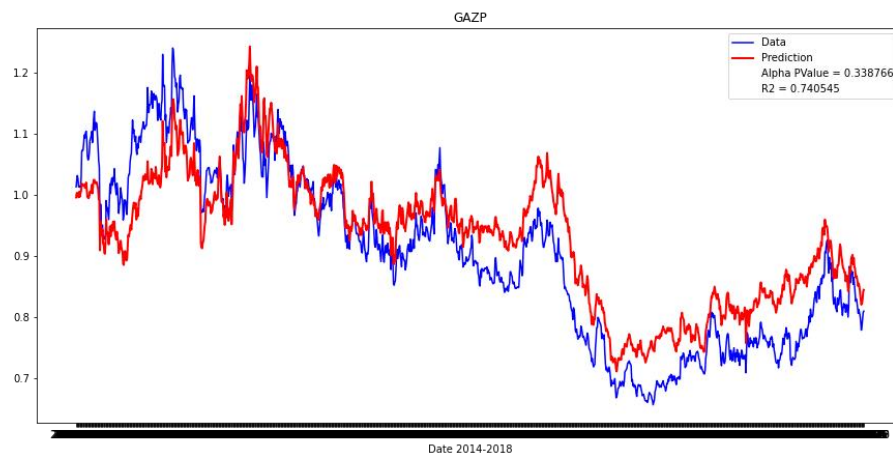


Рисунок 4. Газпром пятифакторная

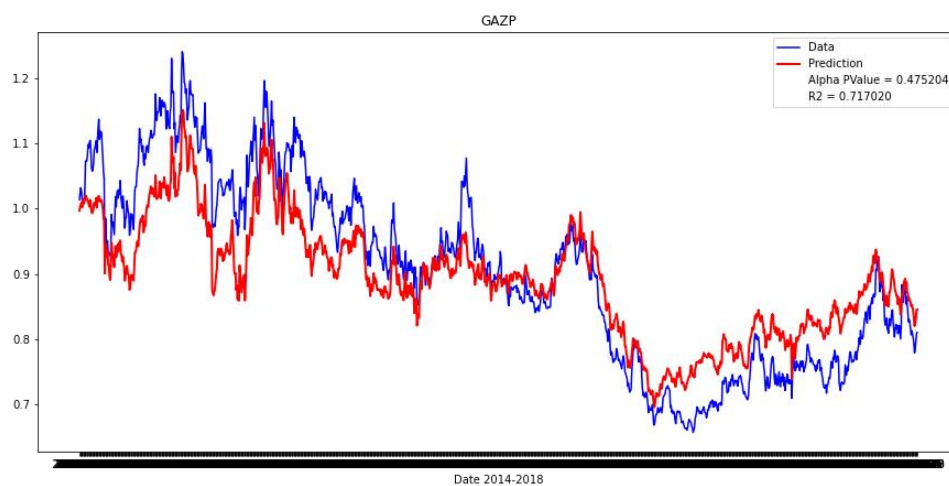


Рисунок 5. Газпром Трёхфакторная

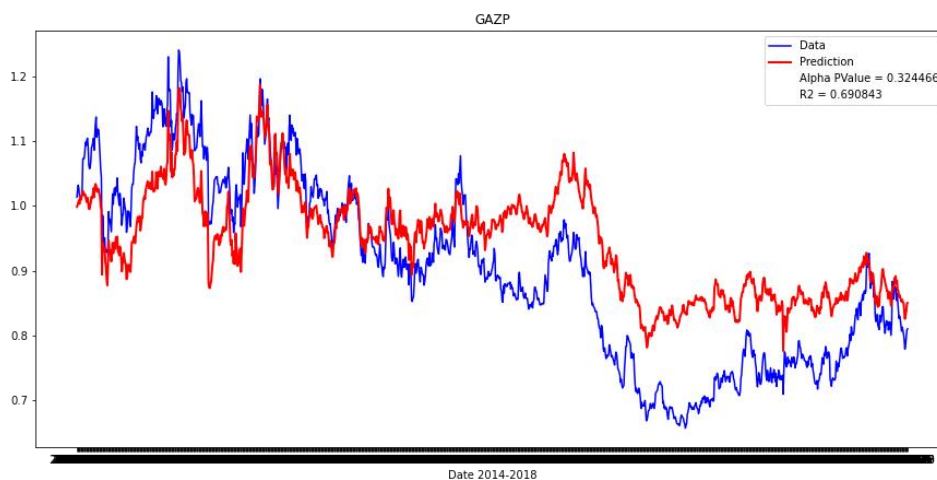


Рисунок 6. Газпром Однофакторная

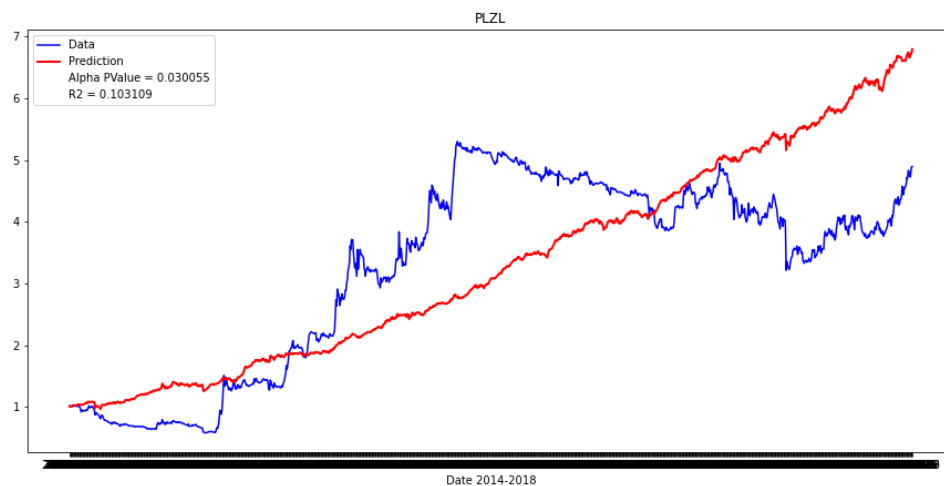


Рисунок 7. Полюс Пятифакторная

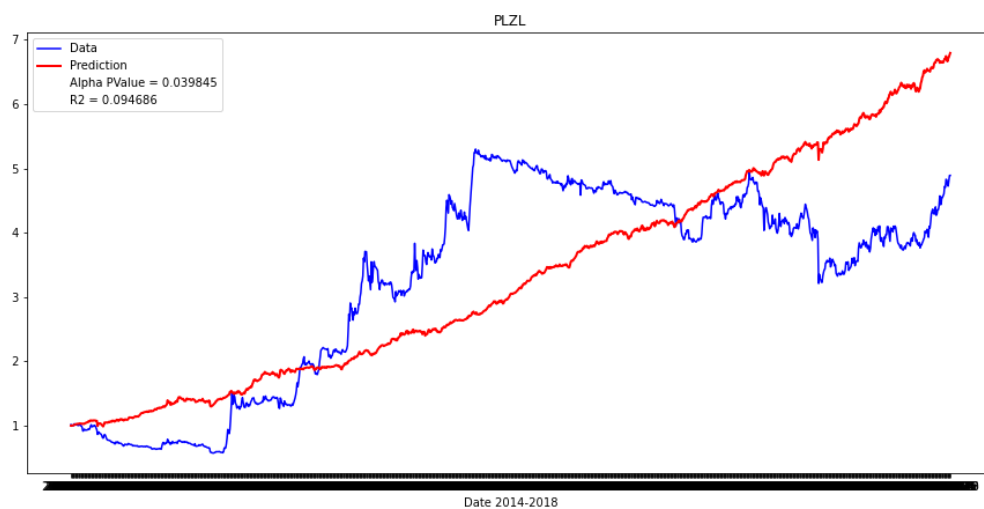


Рисунок 8. Полюс Трёхфакторная

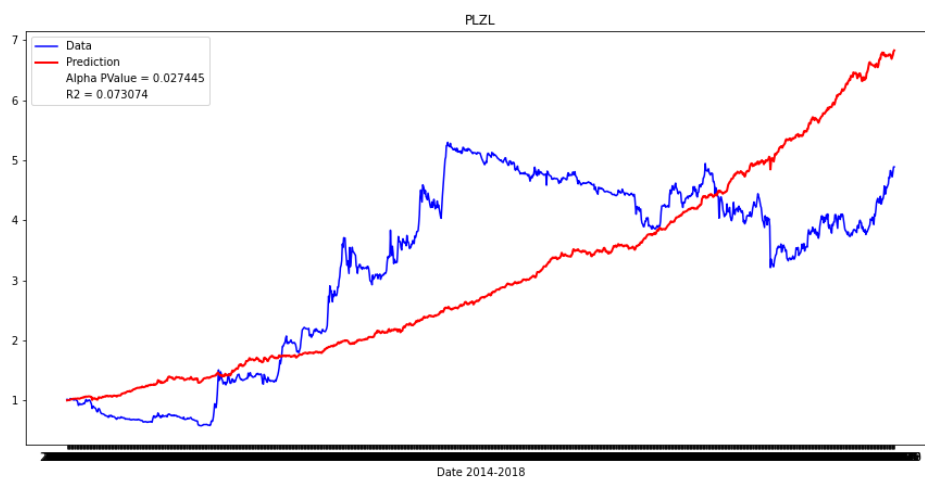


Рисунок 9. Полюс однофакторная

Таблица 6 (Слева бумаги, по которым p -значение коэффициента пересечения меньше 0.05, справа – больше 0.8)

	Дневная	Скольльзящее окно	Месячная		Дневная	Скольльзящее окно	Месячная
CAPM	TATNP, PIKK, MAGN, AKRN, PLZL	-	TATN, TATNP, PIKK, MAGN, AKRN, LKOH	CAPM	MOEX, AFLT	-	GAZP, LSRG
3 Factor	TATNP, PIKK, MAGN, PLZL, MGNT, AKRN	-	TATN, TATNP, PIKK, MAGN, NVTK, LKOH	3 Factor	MOEX, AFLT, MVID	-	GAZP, HYDR
5 Factor	TATNP, PIKK, MAGN, PLZL	-	LKOH, NVTK, TATN, TATNP, PIKK, MAGN, AKRN	5 Factor	ROSN, MVID	-	GAZP, HYDR, MOEX, LSRG

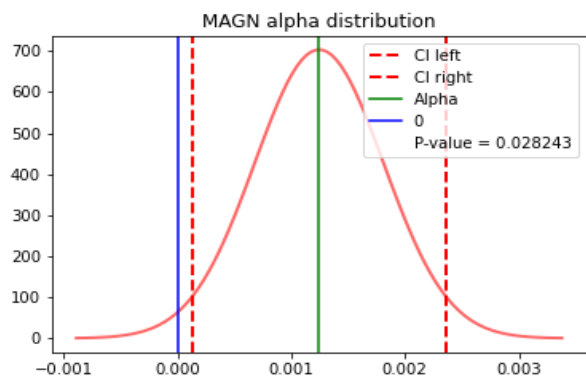


Рисунок 10. Распределение оценок α модели CAPM по дневным данным для MAGN

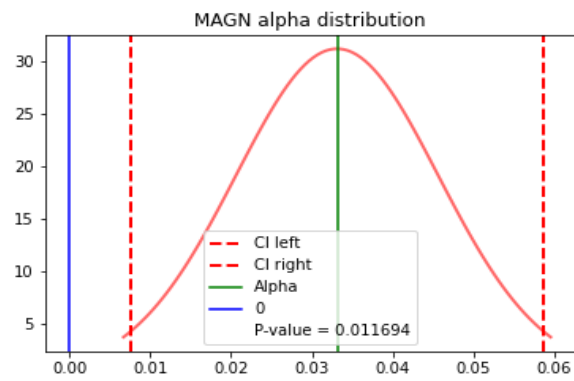


Рисунок 11. Распределение оценок α модели CAPM по месячным данным для MAGN

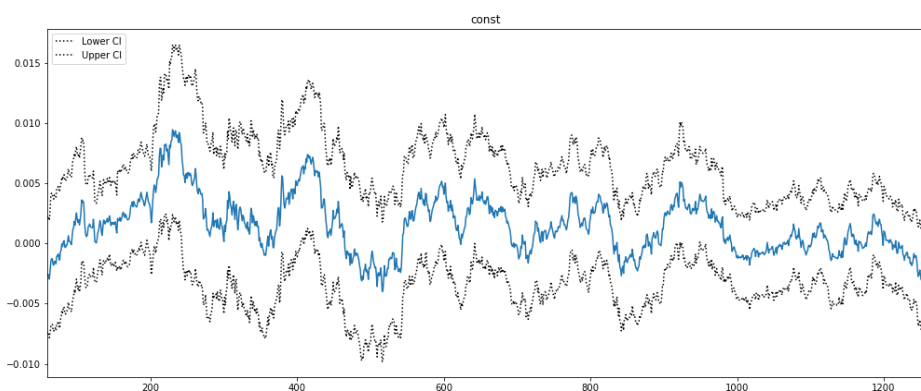


Рисунок 12. Доверительные интервалы оценок α для модели CAPM со скользящим окном для каждого из окон для MAGN

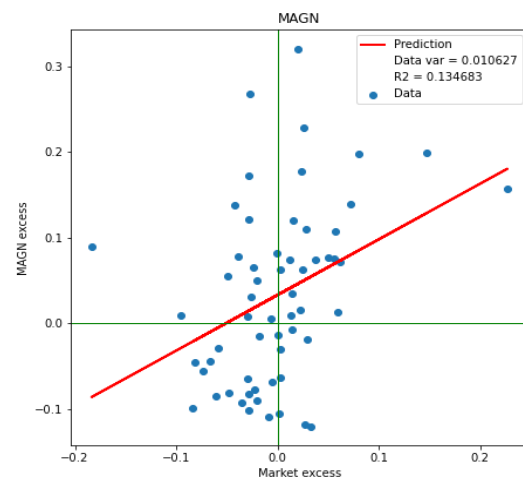


Рисунок 13. Предсказываемая и реальная зависимость доходности от рыночной для модели CAPM по месячным данным для MAGN

Таблица 8. Акции, для которых оценивалась модель

Ticker	Name
AFKS	ПАО АФК "Система", ао
AFLT	ПАО "Аэрофлот", ао
AKRN	ОАО "Акрон", ао
ALRS	АК "АЛРОСА" (ПАО), ао
CHMF	ПАО "Северсталь", ао
FEES	ПАО "ФСК ЕЭС", ао
GAZP	ПАО "Газпром", ао
GMKN	ПАО "ГМК "Норильский никель", ао
HYDR	ПАО "РусГидро", ао
LKOH	ПАО "ЛУКОЙЛ", ао
LSRG	ОАО "Группа ЛСР", ао
MAGN	ПАО "ММК", ао
MGNT	ПАО "Магнит", ао
MOEX	ПАО Московская Биржа, ао
MSNG	ПАО "Мосэнерго", ао
MSTT	ОАО "МОСТОТРЕСТ", ао
MVID	ПАО "М.видео", ао
NLMK	ПАО "НЛМК", ао
NMTP	ОАО "НМТП", ао
NVTK	ПАО "НОВАТЭК", ао
PHOR	ПАО "ФосАгро", ао
PIKK	ПАО "Группа Компаний ПИК", ао
PLZL	ПАО "Полус", ао
ROSN	ПАО "НК "Роснефть", ао
RSTI	ОАО "Россети", ао
SNGS	ОАО "Сургутнефтегаз", ао
SNGSP	ОАО "Сургутнефтегаз", ап
SVAV	ОАО "СОЛЛЕРС", ао
TATN	ПАО "Татнефть" им. В.Д. Шашина, ао
TATNP	ПАО "Татнефть" им. В.Д. Шашина, ап
TRMK	ПАО "ТМК", ао
TRNFP	ПАО "Транснефть", ап
VSMO	ОАО "Корпорация ВСМПО-АВИСМА", ао
VTBR	Банк ВТБ (ПАО), ао
YNDX	Яндекс Н.В., акции иностранного эмитента

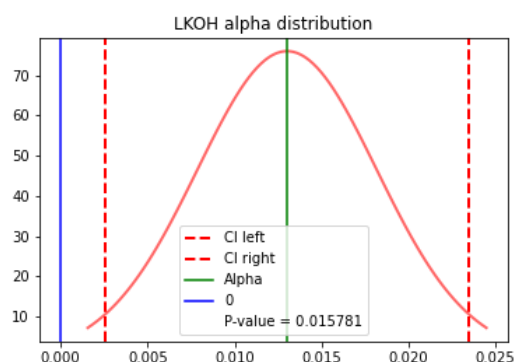


Рисунок 14. Распределение альфа-коэффициента для «Лукойла»

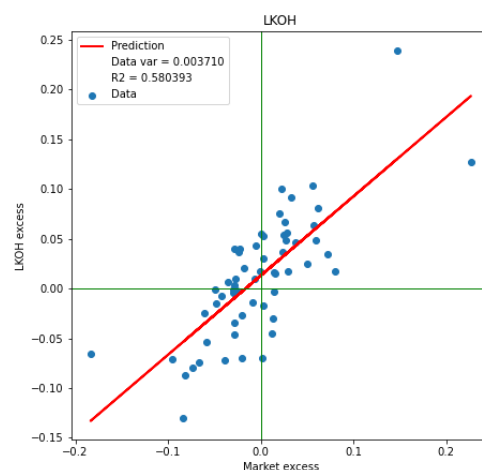


Рисунок 17. Предсказываемая и реальная рыночные доходности для «Лукойла»

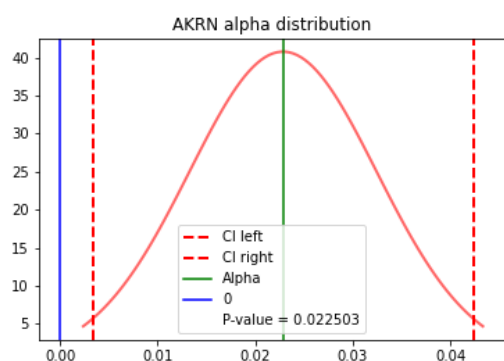


Рисунок 15. Распределение альфа-коэффициента для «Акрона»

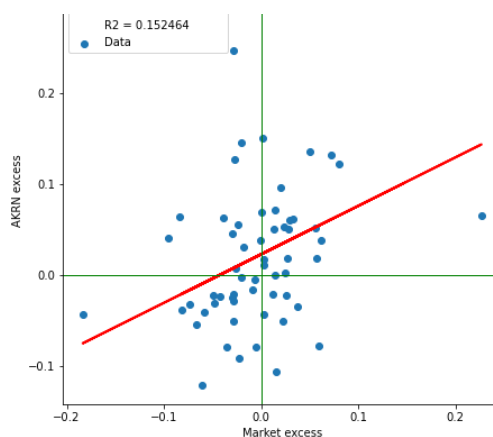


Рисунок 18. Предсказываемая и реальная рыночные доходности для «Акрона»

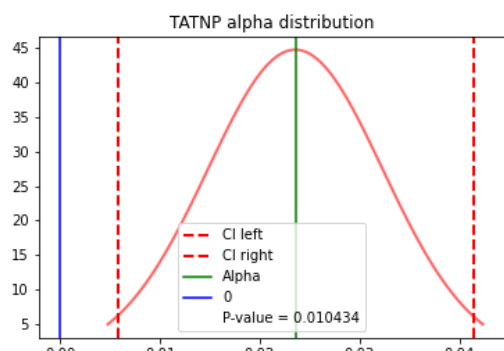


Рисунок 16. Распределение альфа-коэффициента для «Татнефти»

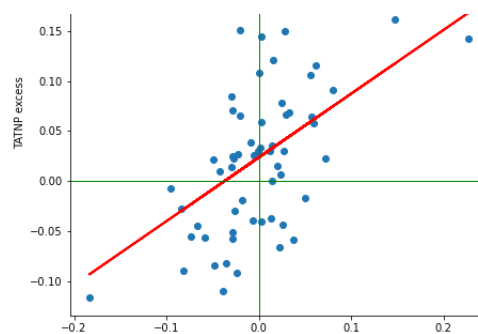


Рисунок 19. Предсказываемая и реальная рыночные доходности для «Татнефти»

Бумаги с высоким коэффициентом детерминации

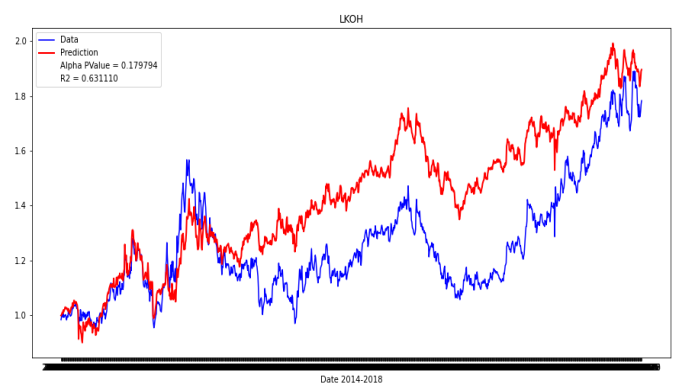
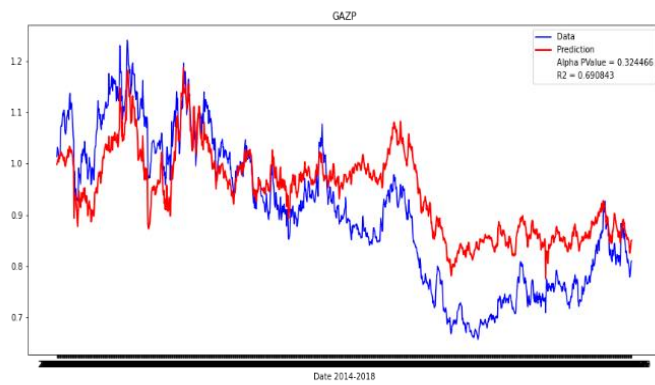


Рисунок 20. CAPM

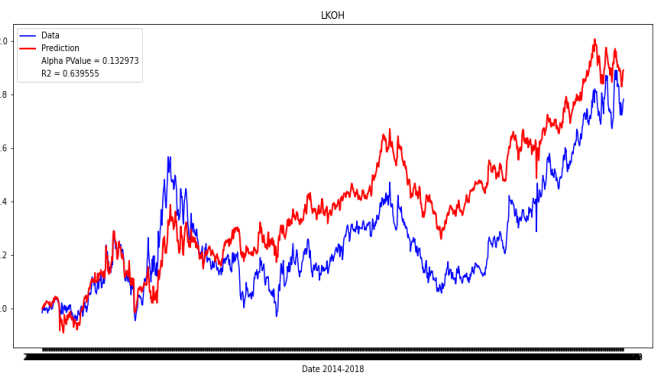
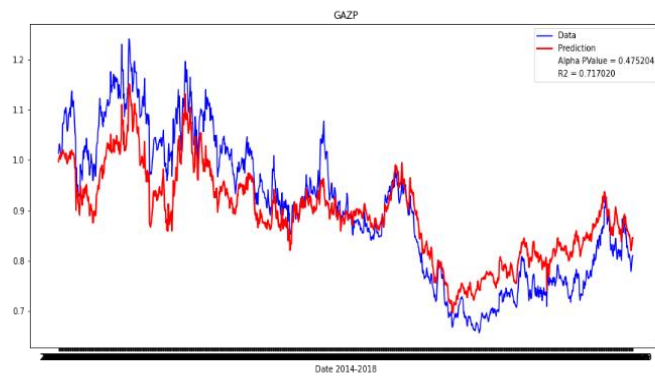


Рисунок 21. 3 Factor Daily

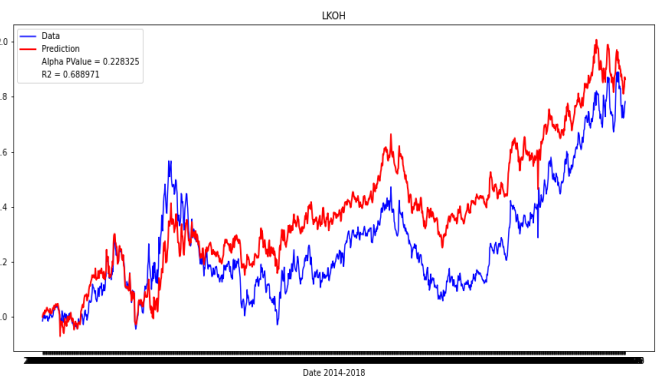
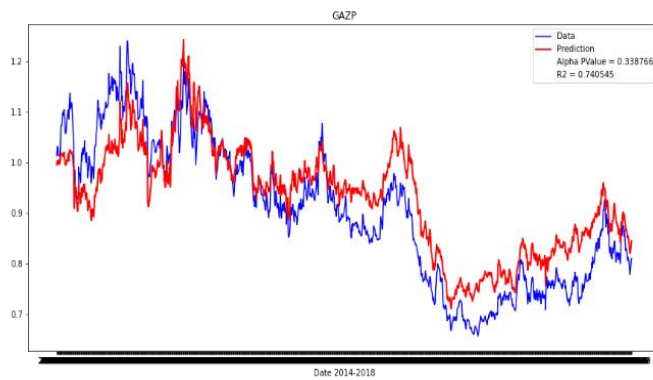


Рисунок 22. 5 Factor Daily

Бумаги со средним коэффициентом детерминации

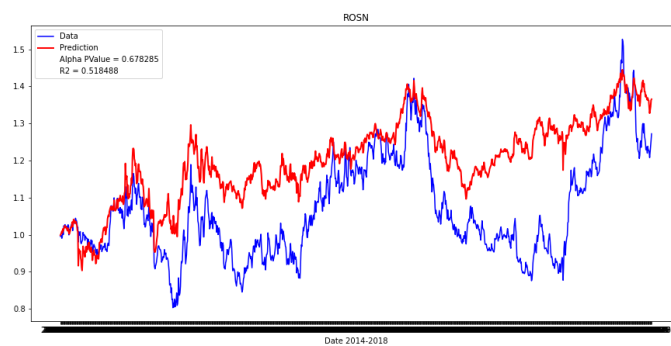
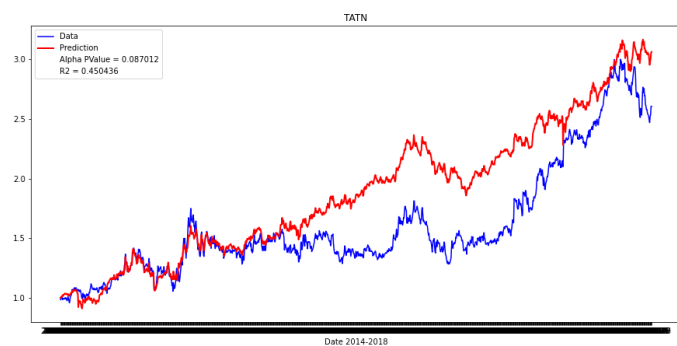


Рисунок 23. CAPM

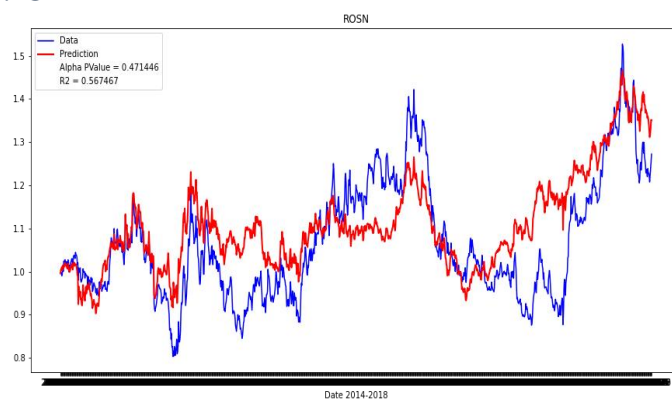
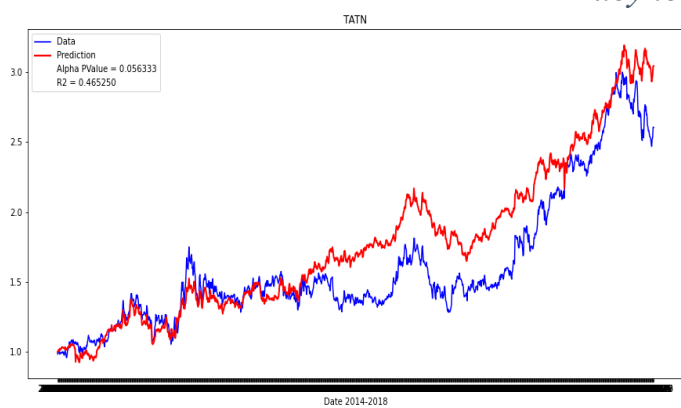


Рисунок 34. 3 Factor Daily

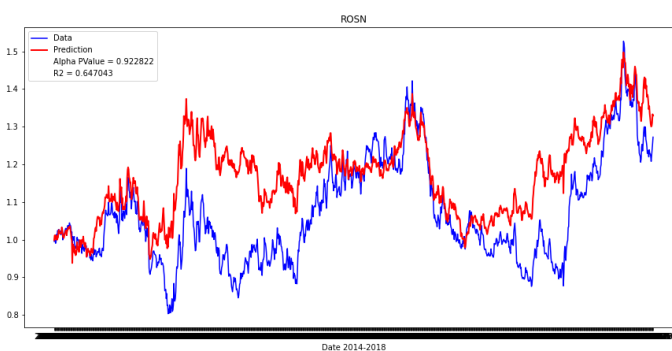
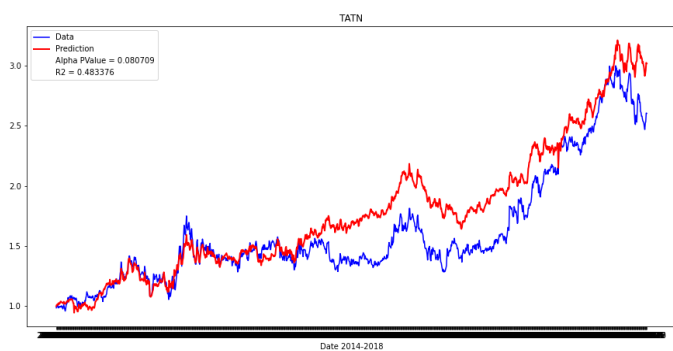


Рисунок 25. 5 Factor Daily

Бумаги с низким коэффициентом детерминации

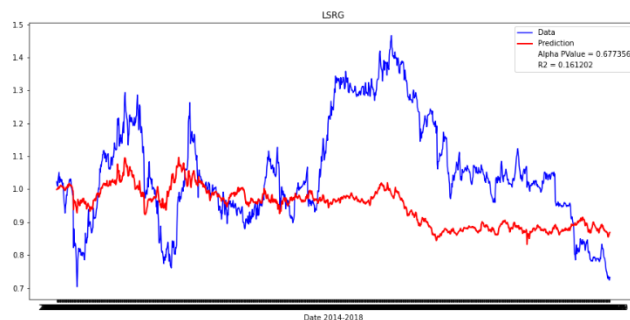
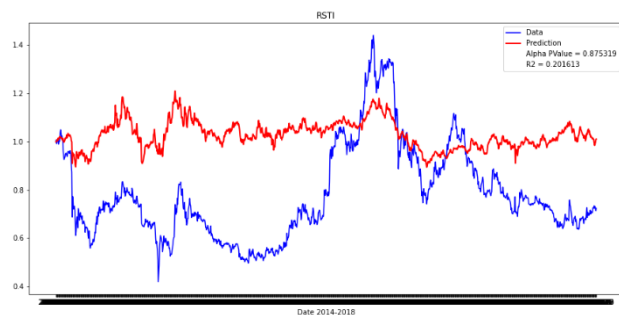


Рисунок 26. CAPM

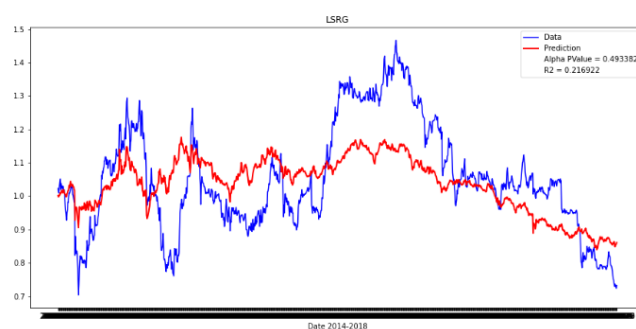
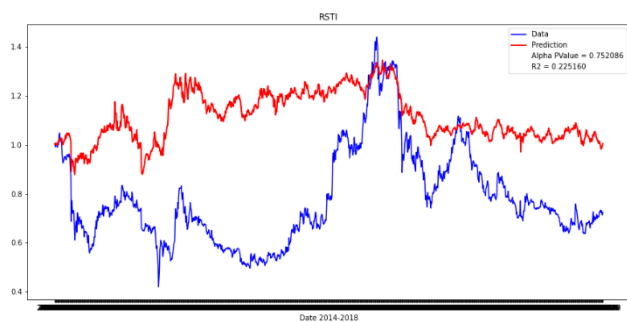


Рисунок 27. 3 Factor Daily

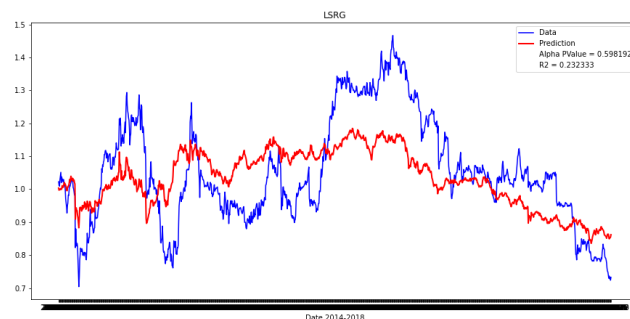
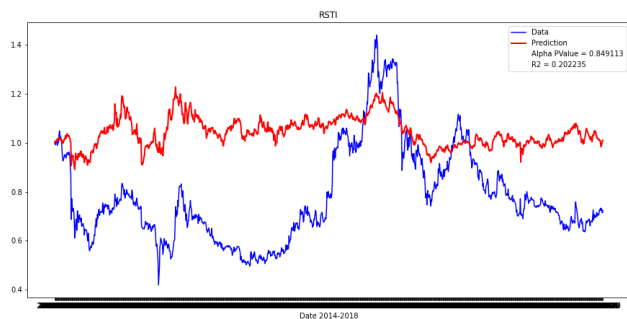


Рисунок 28. 5 Factor Daily