

Федеральное государственное автономное образовательное учреждение высшего  
образования  
«Национальный исследовательский университет  
«Высшая школа экономики»

**Отчет**  
**по дисциплине “Эконометрика-2”**  
**Тема “Влияние факторов на кассовые сборы кинокартин”**

Таблица с вкладом участников:

	Выбор темы	Введение и описание статьи	Сбор данных	Обработ ка данных	Описательн ые статистики	Модель из статьи	Дополнительная модель RDD	Квантильная регрессия
Полина	100%	100%	25%	33%	50%	100%		100%
Лена			25%	33%			75%	
Дима			50%	33%	50%		25%	

Москва, 2025

## Введение

В современном мире трудно представить индивида, который ни разу не ходил бы в кино. Основным показателем рыночного успеха фильмов являются кассовые сборы, которые генерируются за счет количества посещений показа в кинотеатрах. Но какие факторы влияют на выбор потребителями киноленты? После премьеры фильма “Барби” в 2023 году у многих могло сложиться впечатление, что для рекордных кассовых сборов в размере 1,45 млрд долларов необходимо затратить на пиар-кампанию сумму, не меньшую затратам на создание всей картины (на создание фильма было потрачено 145 млн долларов, на маркетинг – 150 млн долларов<sup>1</sup> по данным Variety). Однако щедрые траты на создание и пиар фильма не всегда окупаются и приводят к прибыли для кинопродюсеров. Так, при затратах в 250 млн долларов вышедший в 2025 году фильм “Белоснежка” принес создателям только 181 млн долларов<sup>2</sup>, не окупив даже собственных затрат на производство.

По данным Luminata<sup>3</sup> киноиндустрия США в настоящее время переживает кризис, количество выпущенных фильмов в 2024 году упало на 7% по сравнению с предыдущим годом. По оценкам снижение кинопроизводства было связано с кино забастовками американских сценаристов в 2023 году, однако Luminata связывает это еще и с оттоком зрителей на стриминговые сервисы. По данным сервиса сократился и бюджет на производство кинолент. Однако, кассовые сборы у лидера в 2024 были выше, чем в 2023. Так, лидером проката в 2024 стал сиквел “Головоломка 2”, который собрал 1,69 млрд долларов<sup>4</sup>, в то время как “Барби” в 2023 собрал только 1,45 млрд долларов<sup>5</sup>.

**Целью** данной работы является анализ влияния факторов на кассовые сборы для фильмов, выпущенных в 2024 и 2023 году. **Актуальность** настоящего исследования обусловлена с одной стороны снижением кинопроизводства в мире, а с другой — растущей конкуренцией со стороны стриминговых платформ, которые предоставляют зрителям доступ к фильмам без необходимости посещения кинотеатров. Эти изменения в кинопроизводственной и прокатной среде делают анализ факторов, влияющих на

---

<sup>1</sup> Inside ‘Barbie’ s’ Pink Publicity Machine: How Warner Bros. Pulled Off the Marketing Campaign of the Year // Variety URL: <https://variety.com/2023/film/box-office/barbie-marketing-campaign-explained-warner-bros-1235677922/> (дата обращения: 14.04.2025).

<sup>2</sup> Белоснежка // IMDb URL: <https://www.imdb.com/title/tt6208148/> (дата обращения: 8.05.2025).

<sup>3</sup> Luminate Releases 2024 Year-End Film & TV Report // Luminata URL: <https://luminatedata.com/reports/yearend-filmtv-industry-report-2024/> (дата обращения: 8.05.2025).

<sup>4</sup> Головоломка 2 (2024) // Kinopoisk URL: [https://www.kinopoisk.ru/film/5102255/?ysclid=ma8kfcf2nf175937139&utm\\_referrer=yandex.ru](https://www.kinopoisk.ru/film/5102255/?ysclid=ma8kfcf2nf175937139&utm_referrer=yandex.ru) (дата обращения: 8.05.2025).

<sup>5</sup> Inside ‘Barbie’ s’ Pink Publicity Machine: How Warner Bros. Pulled Off the Marketing Campaign of the Year // Variety URL: <https://variety.com/2023/film/box-office/barbie-marketing-campaign-explained-warner-bros-1235677922/> (дата обращения: 14.04.2025)

кассовые сборы, особенно важным для индустрии. В условиях неопределенности и изменения привычек зрителей важно выявить ключевые драйверы кассовых сборов, чтобы эффективно использовать ресурсы и повысить конкурентоспособность в меняющемся рынке кино.

## Основная работа

Одной из основных работ в изучении влияния факторов на кассовые сборы является «Examining Success in the Motion Picture Industry<sup>6</sup>», написанная Патриксом Топфом. В работе проанализировано 189 кинокартин широкого проката (фильмы, которые показывали в кинотеатрах США и Канады более 600 раз). Автор использовал модели линейной регрессии для анализа влияния на общие кассовые сборы (TR – total revenue) восьми независимых переменных: Production costs (PC), Professional review scores (PR), Star power (SP), Rating (R), Genre (G), Sequel (S), Season (SW), Holiday release (H). Автор приходит к выводу, что значимыми в модели являются 4 переменные: производственные затраты, отзывы критиков и жанры экшн и комедия. При этом коэффициент при переменной PC (прокси-переменная на маркетинговую кампанию) составляет 0,94, то есть каждый вложенный доллар в рекламу увеличивает выручку на сумму менее доллара. Это свидетельствует о том, что в некоторых случаях чересчур сильная пиар-кампания может негативно сказаться на итоговой выручке. Помимо основной модели автор построил дополнительную с включенной перекрестной переменной PC\*PR, модель стала лучше предсказывать кассовые сборы, но появилась сильная мультиколлинеарность между признаками.

## Система гипотез исследования:

На основании проведенного анализа рынка и соответствующей литературы, задач исследования и его предпосылок были сформулированы следующие гипотезы (см. Таблица 1):

Таблица 1. Гипотезы исследования

Нулевая гипотеза		Пояснение
$H_0^{1.1}$	Профессиональные оценки положительно коррелируют с кассовыми сборами	Один из основных выводов статьи, который необходимо проверить

<sup>6</sup> Topf P. Examining success in the motion picture industry //The Park Place Economist. – 2010. – Т. 18. – No. 15. – С. 1.

$H_0^{1.2}$	<b>Бюджет фильма оказывает положительное влияние на кассовые сборы</b>	Еще один из основных выводов статьи, который нужно проверить
$H_0^{1.3}$	<b>Жанр “драма” положительно влияет на кассовые сборы</b>	По данным Luminate <sup>7</sup> драмы являются самым популярным жанром на всех видеосервисах в 2024 году, из-за чего кассовые сборы на них могли быть выше в обоих годах
$H_0^{1.4}$	<b>Одиночные фильмы или приквелы приносят больше кассовых сборов</b>	Согласно теории упадка <sup>8</sup> фильмы-сиквелы собирают в прокате меньше, чем их предшественники
$H_0^{1.5}$	<b>Наличие SJW повестки отрицательно влияет на кассовые сборы для фильмов</b>	Современные стандарты толерантности могут отпугивать зрителей, которые не относятся к меньшинствам
$H_0^{2.1}$	<b>Факторы, влияющие на кассовые сборы в 2023 и 2024 годах не различаются</b>	Так как за год основные потребители кино не изменились, то и их потребительские паттерны не должны были сильно измениться.
$H_0^{2.2}$	<b>После 14 июля 2023 года кассовые сборы для выпущенных фильмов были не выше чем раньше</b>	Из-за забастовки сценаристов и актеров в 2023 году премьеры некоторых картин были отложены из-за чего крупных фильмов с огромными сборами было меньше.

### Сбор и описание данных

Выборка была собрана с помощью методов анализа данных с сайта “Кинопоиск”. Сначала был сформирован начальный список фильмов за 2023 и 2024 годы, при этом отобраны только те ленты, среди стран производства которых было США. Для ранжирования использовано количество зрительских оценок на Кинопоиске по убыванию этого показателя.

Чтобы получить уникальные идентификаторы (ID) каждого фильма, был написан небольшой Python-скрипт, который на основе кода элемента со страницы парсил id всех фильмов (найлены ссылки типа film/sample\_id/ в HTML). Затем был использован неофициальный API Кинопоиска: подготовлен другой скрипт на Python, прописан API-ключ и загружен список уже найденных ID. С помощью этого запроса были получены данные по каждому фильму (бюджет, кассовые сборы, рейтинги и т. д.).

При этом, не у всех фильмов была информация по общемировым сборам и бюджетам на производство, поэтому было решено исключать такие киноленты из датасета. Стоит отметить, что точную структуру бюджетных трат редко разглашают (непосредственное создание фильма, затраты на пиар и тд), поэтому как в

<sup>7</sup> Luminate Releases 2024 Year-End Film & TV Report // Luminata URL: <https://luminatedata.com/reports/yearend-filmtv-industry-report-2024/> (дата обращения: 8.05.2025).

<sup>8</sup> Edward P. Lazear, The Peter Principle: A Theory of Decline

оригинальной статье общий бюджет является прокси переменной на рекламную кампанию. Также неудавшиеся проекты в принципе имеют стимулы преуменьшать или вообще не разглашать бюджет, что ведет к предположению, что выборка изначально может оказаться усеченной. В итоге была собрана информация по 123 фильмам. В качестве основных характеристик выделяются следующие (см. Таблицу 2):

Таблица 2. Перечень анализируемых характеристик

Характеристика	Тип шкалы	Примечания
<b>budget (B)</b>	Количественная	Бюджет фильма в долларах США, в статье является прокси-переменной на маркетинговую кампанию
<b>ratingKinopoisk (V)</b>	Количественная	10-бальная шкала зрительских оценок на платформе Кинопоиск
<b>FilmCritics (C)</b>	Количественная	10-бальная шкала профессиональных критиков
<b>MPAA (M)</b>	Категориальная	Набор возрастных рейтингов, принятых ассоциацией МРАА
<b>genres</b>	Номинальная	Набор дамми-переменных, указывающих жанры фильма
<b>box_office (BO)</b>	Количественная	Общемировые кассовые сборы фильма в долларах США
<b>weekend (W)</b>	Номинальная	Дамми-переменная, равная 1, если премьера фильма в США состоялась в праздничный день
<b>sequel (S)</b>	Номинальная	Дамми-переменная, равная 1, если фильм является продолжением других фильмов
<b>sjw (SJW)</b>	Номинальная	Дамми-переменная, равная 1, если фильм связан с крупными скандалами по теме инклюзивности
<b>star power (SP)</b>	Количественная	Количество выигранных оскаров у актеров и режиссера фильма до его выхода
<b>season (SE)</b>	Номинальная	Дамми-переменная, равная 1, если фильм был выпущен летом или в три последние недели декабря
<b>lenght (L)</b>	Количественная	Длина фильма в минутах

## Описательные статистики

После удаления данных с пропущенными значениями в таблице осталось 120 фильмов, из которых 63 были выпущены в 2023, а 57 – в 2024. Количество фильмов, которые были замечены в крупных скандалах по теме инклюзивности, оказалось равным 30. Всего 46 фильмов были выпущены летом или в последние 3 недели декабря. Количество фильмов, премьера в США которых была в праздничный день, оказалось равным 2, количество фильмов, которые являются сиквелами составляет 45. Большинство кинолент длится от 1,5 до 2 часов.

В среднем фильмы из выборки зарабатывали 261 млн долларов, при средних затратах в 88 млн долларов, при этом как среди кассовых сборов, так и среди бюджетных затрат присутствуют выбросы (см. Приложение 1. Рис 1-4). Сравнивая профессиональные и зрительские оценки (см. Приложение 1. Рис 5), можно сказать, что зрители реже ставят как очень низкие, так и очень высокие оценки, однако их оценки довольно скошены влево. Профессиональные оценки более симметричны. Наиболее распространенный жанр в выборке это экшн (35,7%). Только 24% фильмов пропагандировали инклюзивную повестку. Для дальнейшего построения моделей было принято решение удалить выбросы.

## Оценка моделей из статьи

Для начала была оценена модель, идентичная модели из статьи за исключением дамми-переменной на возрастной рейтинг, так как в выборке нет фильмов с рейтингом G, также в связи с тем, что всего 2 фильма были выпущены в праздничные дни было принято решение исключить переменную из выборки, так как она бы являлась незначимой и не давала бы верно оценивать оставшиеся факторы. Модель имеет следующий функциональный вид:

$$BO = \alpha_0 + \alpha_1(B) + \alpha_2(C) + \alpha_3(SP) + \alpha_4(action) + \alpha_5(comedy) + \alpha_6(S) + \alpha_7(SE)$$

Использование робастных ошибок в форме Уайта решило проблему гетероскедастичности, также в модели отсутствует мультиколлинеарность, высокие значения Condition Number вызвано разным масштабом переменных, при этом изменение масштаба не влияет на коэффициенты модели и их значимость, поэтому данные остались в первоначальном виде. Ошибки в модели распределены ненормально. Так как остатки в модели близки к нормальным и количество данных является довольно большим, то по ЦПТ в пределе остатки стремятся к нормальным и

можно не использовать преобразование Бокса-Кокса, то есть коэффициенты становятся устойчивыми.

Так как модель в статье была линейной, то для сравнения результатов никаких изменений проводиться не будет. Так, модель способна предсказать 61,2% дисперсии кассовых сборов, она статистически значима. Значимыми в модели являются константа, бюджет, оценка критиков и сиквел (см. Приложение 2). Примечательно, что увеличение бюджета на 1 доллар приводит к росту кассовых сборов на 1,96 доллара, то есть окупаемость вложений для продюсеров практически в два раза, что отчасти противоречит статье. Также в отличие от статьи факторы, отвечающие за жанр не являются значимыми. Увеличение оценки критиков на 1 приводит к увеличению кассовых сборов на 38 млн долларов.

Однако, для попытки нормализовать остатки были оценены еще две модели: полулогарифмическая с логарифмированными кассовыми сборами и логарифмическая с логарифмированными кассовыми сборами и бюджетом. Обе модели избавлены от гетероскедастичности и мультиколлинеарности. Сравнивая логарифмическую с полулогарифмической моделью, можно сказать, что логарифмическая имеет большую предсказательную силу, поэтому в дальнейшем будем сравнивать ее с линейной.

По результатам РЕ теста было получено, что гипотеза о  $\delta_{lin} = 0$  отвергается, а  $\delta_{log} = 0$  не отвергается, поэтому наилучшей моделью будет линейная в логарифмах, с функциональной формой:

$$\ln(BO) = \alpha_0 + \alpha_1 \ln(B) + \alpha_2(C) + \alpha_3(SP) + \alpha_4(action) + \alpha_5(comedy) + \alpha_6(S) + \alpha_7(SE)$$

В ней значимыми являются логарифмированный бюджет, профессиональные оценки, жанр экшн и сиквел (см. Приложение 3). Так, увеличение бюджета на 1% приводит к увеличению кассовых сборов на 0,81%, что говорит об убывающей отдаче от масштаба. При увеличении профессиональной оценки на 1 кассовые сборы увеличиваются на 33,1%. Фильмы сиквелы собирают на 179% больше, чем самостоятельные фильмы и приквелы. Если в жанрах имеется экшн, то кассовые сборы будут ниже на 37,9%.

При добавлении перекрестной переменной  $B \cdot C$  и устранении мультиколлинеарности посредством центрирования факторов в линейной МНК модели значимые коэффициенты не изменились даже при ненормальных остатках, что говорит о том, что объема выборки достаточно для устойчивости коэффициентов. При этом, дополнительно была оценена модель с преобразованием Бокса-Кокса для

устранения ненормальности остатков модели, по ее результатам значимые коэффициенты остались прежними, то есть константа, бюджет, оценки критиков и сиквел. Линейная модель после добавления перекрестной переменной стала иметь лучшую предсказательную способность (66,9%), как и в оригинальной статье. Перекрестный коэффициент оказался значимым (см. Приложение 4), как и бюджет, профессиональные оценки и сиквел.

Помимо модели из статьи была проанализирована еще одна модель, включающая дополнительные факторы: зрительская оценка, дамми-переменная на инклюзивность и некоторые жанры, которые упоминались в статье. После оценки разных моделей и проведения РЕ теста была выбрана логарифмическая модель, она имеет следующий функциональный вид:

$$\ln(BO) = \alpha_0 + \alpha_1 \ln(B) + \alpha_2(C) + \alpha_3(V) + \alpha_4(L) + \alpha_5(S) + \alpha_6(SE) + \alpha_7(SJW) + \alpha_8(SP) + \alpha_9(action) + \alpha_{10}(comedy) + \alpha_{11}(drama)$$

Модель способна объяснить 58,9% дисперсии логарифма кассовых сборов, что является хорошей предсказательной силой, однако переменные SJW и drama не являются значимыми (см. Приложение 5). Статистически значимыми на 5% уровне являются логарифм бюджета, сиквел и жанр экшн. Так, увеличение бюджета производства на 1% увеличивает кассовые сборы на 0,72% при прочих равных. Сиквелы приносят киностудиям на 1,96% больше, чем самостоятельные фильмы и приквелы.

## Выводы после оценки МНК моделей

Относительно гипотез, связанных с проведенным ранее исследованием, можно сказать, что гипотеза  $H_0^{1.1}$  не может быть отвергнута в силу значимости показателя во всех моделях. Гипотеза  $H_0^{1.2}$  не отвергается только в некоторых моделях, поэтому требуется дополнительная оценка других функциональных форм. Гипотеза  $H_0^{1.3}$  не может быть ни отвергнута, ни принята, так как влияние жанра драма на кассовые сборы не обнаружено, как и гипотеза  $H_0^{1.5}$ . Гипотеза  $H_0^{1.4}$  отвергается, так как коэффициент при переменной имеет положительный знак.

## Построение дополнительных моделей и проверка гипотез:

Для проверки гипотезы  $H_0^{2.1}$  был проведен тест Чоу и оценена логарифмическая модель с включением индикаторов, для оценки влияния 2024 года. По результатам



обоих тестов факторы влияния для логарифмической модели в 2023 и 2024 году не различаются. Гипотеза **не отвергается**. **F-статистика: 1.238, p-value: 0.28.**

## Построение модели RDD

Для проверки следующей гипотезы была построена модель разрывного регрессионного дизайна, где в качестве даты разрыва взято 14.07.2023 — начало актерской забастовки. Модель позволяет оценить эффект воздействия конкретного события на кассовые сборы, поэтому является оптимальной для данной задачи.

**Результат оценивания:** t-статистика: 0.818, p-value: 0.792. Гипотезу  $H_0^{2,2}$  **не отвергли** статистически — p-value выше 0.05. После 14.07.2023 и на графике, и через проверку тестом показывается снижение сборов после забастовки. **Возможные причины:** задержки и срывы маркетинга (забастовка ограничила продвижение фильмов, актеры не могли участвовать в съемке контента), смещение крупных релизов — удар по крупным ожидаемым фильмам (студии могли отложить релизы, опасаясь слабого старта и поэтому перенесли их на 2025 год и позднее), ухудшение качества (некоторые сценарии не были доработаны из-за прекращения работы сценаристов, сделаны быстро), общее снижение интереса к кинотеатрам в 2023—2024 годах (постковидное падение, конкуренция с онлайн-платформами).

На Рис.6 (см. Приложение 1) до 14.07.2023 синяя линия показывает сильный положительный тренд — кассовые сборы увеличиваются с течением времени. После забастовки (красная линия): тренд ослабевает, линия почти плоская, кассовые сборы стабилизируются или даже слегка снижаются. Разрыв показывает визуальное снижение — уровень сборов ниже, чем можно было бы ожидать по продолжению предыдущего тренда.

Возможные минусы использования модели RDD: большая дисперсия сборов — данные сильно "шумят", поэтому сложно зафиксировать чёткий эффект и эндогенность — возможно, сами студии отбирали даты релизов в зависимости от ожиданий, что искажает чистый эффект даты и изначальный датасет.

Рис.7 (см. Приложение 1) плацебо-тест с использованием бутстрапа проверяет: не является ли наблюдаемый скачок случайным шумом, сравнивая с фальшивыми данными (смещаем фиктивных дат относительно забастовки). Синяя область: 95% доверительный интервал (надежность оценки). Почти все оценки скачков колеблются около 0 или отрицательны — нет значительных положительных "фейковых" скачков.

Разрыв в настоящей дате (0 дней смещения) не больше других случайных скачков и лежит внутри доверительного интервала. **Микровывод:** наблюдаемое значение кассовых сборов может быть вызвано другими факторами кроме забастовки, однако снижение очевидно и статистически значимо.

### **Построение модели квантильной регрессии**

Так как выборка сильно распределена по кассовым сборам, и величина максимальных кассовых сборов в несколько сотен раз превышает величину минимальных, то для лучшей оценки было решено построить квантильную регрессию. Для определения оптимального числа квантилей было проведено несколько тестов: через подбор по оптимальным значениям AIC, BIC,  $R^2$  и с помощью построения дерева для определения оптимальных точек различия. При выборе числа квантилей и границ разрыва необходимо учитывать количество наблюдений, которое в них попадет. Так, было принято решение остановиться на 3 разбиениях. Примечательно, что разбиения проходили по 0.15, 0.5 и 0.85 квантилям для большей точности результатов. Так, выборка разделена на неудачные фильмы, средние и блокбастеры (см. Приложение 6).

Для нижнего квантиля коэффициент при бюджете составляет 0.94, что свидетельствует о том, что каждый вложенный доллар приносит только 94 цента. Подобные результаты получил автор статьи, которую мы анализировали, из чего можно сделать вывод, что его выборка была наполнена не сильно прибыльными фильмами. Все остальные переменные не оказывают особого статистического влияния. Так, для фильмов с низкими сборами только бюджет определяет кассовые сборы, это может свидетельствовать о низкой окупаемости инвестиций в производство в данной категории. На 10% уровне значимым является наличие SJW повестки.

Для средних и наиболее кассовых фильмов влияние бюджета становится сильнее. Так, вложение доллара в средний фильм вырастает в 2 доллара кассовых сборов, а для блокбастеров – в 4,42. Комедии влияют на кассовые сборы только для блокбастеров, влияние жанров для других категорий не обнаружено. Примечательно, что зрительские оценки не являются значимыми ни в одной модели. Также, что практически все коэффициенты модели имеют нечеткую структуру (см. Рис 8) и только бюджет имеет восходящий тренд.

## Приложения

### Приложение 1.

Рис. 1-4 Voxplot и гистограммы для кассовых сборов и бюджета (млн долларов)

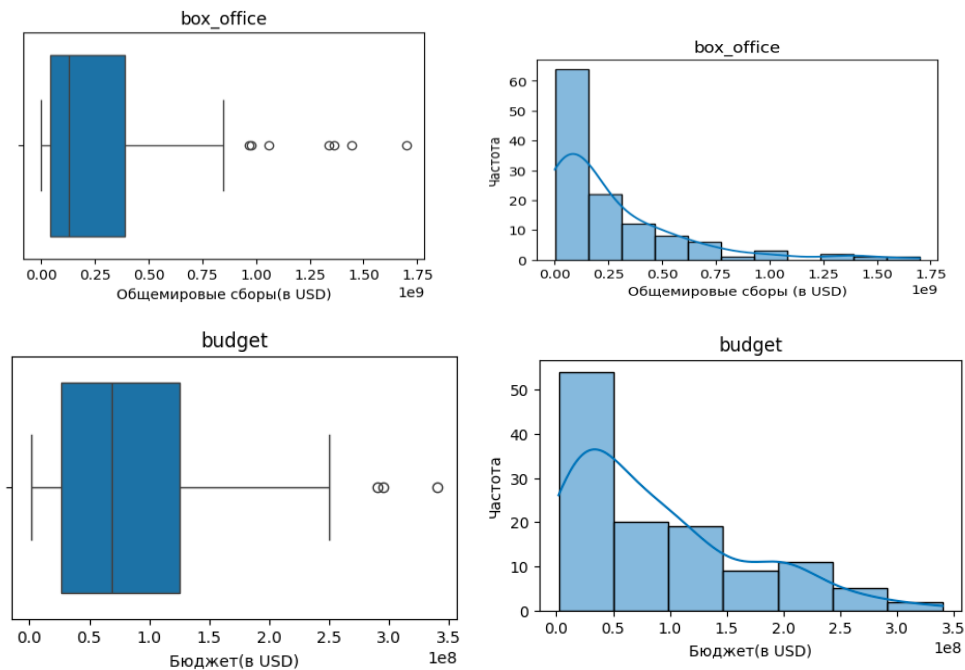


Рис. 5 Сравнительная гистограмма для зрительских и профессиональных оценок

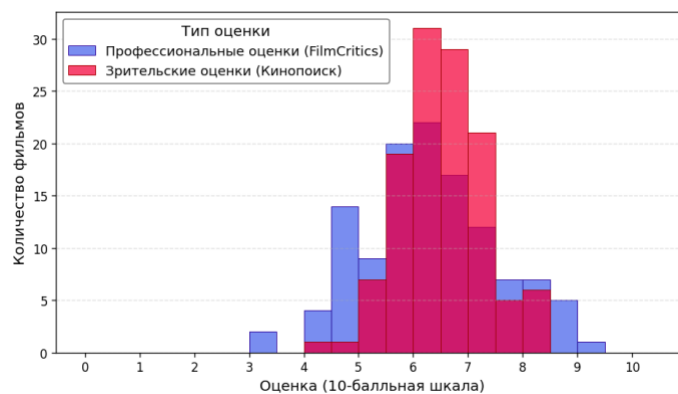
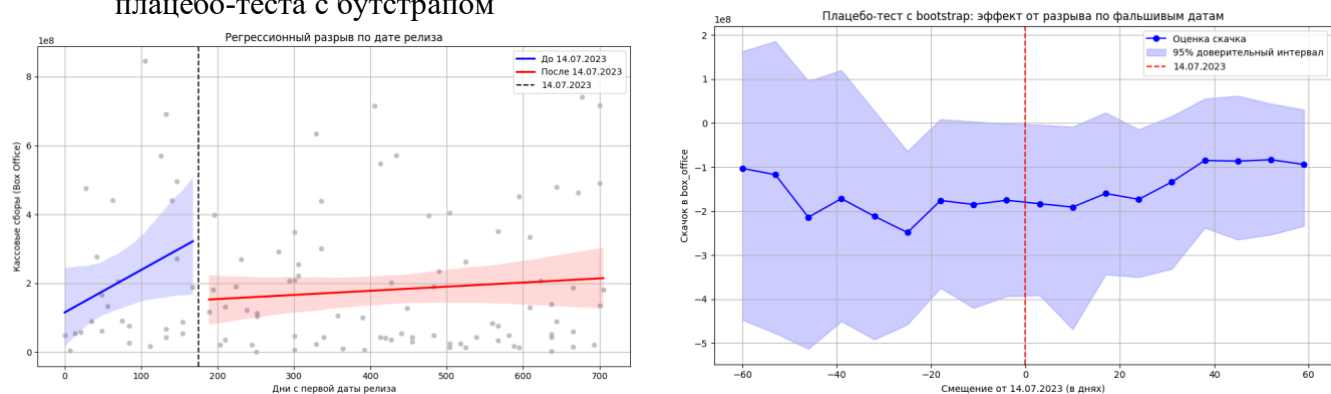


Рис. 6-7 График RDD модели для данных до и после забастовки и график реализации плацебо-теста с бутстрапом



## Приложение 2. Оценка обычной МНК модели из статьи

	coef	std err	z	P> z	[0.025	0.975]
const	-2.297e+08	7.89e+07	-2.911	0.004	-3.84e+08	-7.51e+07
budget	1.9649	0.374	5.247	0.000	1.231	2.699
FilmCritics	3.811e+07	1.15e+07	3.304	0.001	1.55e+07	6.07e+07
star_power	-5.616e+07	3.77e+07	-1.491	0.136	-1.3e+08	1.77e+07
action movie	-2.311e+07	3.13e+07	-0.738	0.461	-8.45e+07	3.83e+07
comedy	-7.624e+06	3.42e+07	-0.223	0.824	-7.47e+07	5.94e+07
sequel	1.27e+08	2.9e+07	4.377	0.000	7.01e+07	1.84e+08
season	2.47e+07	2.8e+07	0.881	0.379	-3.03e+07	7.97e+07

## Приложение 3. Оценка обычной логарифмической модели

	coef	std err	z	P> z	[0.025	0.975]
const	2.0189	2.547	0.793	0.428	-2.972	7.010
log_budget	0.8160	0.142	5.736	0.000	0.537	1.095
FilmCritics	0.2866	0.087	3.303	0.001	0.117	0.457
star_power	-0.1974	0.195	-1.015	0.310	-0.579	0.184
action movie	-0.4774	0.230	-2.079	0.038	-0.927	-0.027
comedy	-0.1799	0.215	-0.837	0.403	-0.601	0.241
sequel	1.0280	0.220	4.663	0.000	0.596	1.460
season	0.1051	0.186	0.567	0.571	-0.259	0.469

## Приложение 4. Оценка МНК модели с перекрестной переменной

	coef	std err	z	P> z	[0.025	0.975]
const	-2.63e+08	6.52e+07	-4.035	0.000	-3.91e+08	-1.35e+08
budget	1.9945	0.349	5.708	0.000	1.310	2.679
FilmCritics	4.282e+07	9.24e+06	4.636	0.000	2.47e+07	6.09e+07
star_power	-4.736e+07	3.25e+07	-1.458	0.145	-1.11e+08	1.63e+07
action movie	-2.71e+07	2.97e+07	-0.914	0.361	-8.52e+07	3.1e+07
comedy	-2.726e+06	3.2e+07	-0.085	0.932	-6.54e+07	5.99e+07
sequel	1.268e+08	2.77e+07	4.575	0.000	7.25e+07	1.81e+08
season	3.259e+07	2.52e+07	1.294	0.196	-1.68e+07	8.2e+07
B*C_2	0.7238	0.183	3.959	0.000	0.365	1.082

## Приложение 5. Оценка дополнительной логарифмической модели

	coef	std err	z	P> z	[0.025	0.975]
const	2.7206	2.728	0.997	0.319	-2.626	8.067
log_budget	0.7235	0.163	4.429	0.000	0.403	1.044
FilmCritics	0.1874	0.140	1.336	0.181	-0.087	0.462
ratingKinopoisk	0.2489	0.182	1.366	0.172	-0.108	0.606
Length	-0.0009	0.007	-0.138	0.890	-0.014	0.012
sequel	1.0878	0.234	4.650	0.000	0.629	1.546
season	0.0691	0.195	0.354	0.723	-0.313	0.452
SJW	0.3924	0.290	1.353	0.176	-0.176	0.961
star_power	-0.1628	0.234	-0.696	0.486	-0.621	0.295
action movie	-0.4918	0.246	-1.997	0.046	-0.974	-0.009
comedy	-0.3197	0.237	-1.351	0.177	-0.783	0.144
drama	-0.0402	0.257	-0.157	0.875	-0.544	0.463

## Приложение 6. Оценка квантильной регрессии.

=== Квантильная регрессия для q = 0.15 ===

QuantReg Regression Results

Dep. Variable:	box_office	Pseudo R-squared:	0.1461			
Model:	QuantReg	Bandwidth:	1.617e+08			
Method:	Least Squares	Sparsity:	3.743e+08			
Date:	Thu, 08 May 2025	No. Observations:	120			
Time:	13:34:56	Df Residuals:	108			
		Df Model:	11			
	coef	std err	t	P> t	[0.025	0.975]
const	-684.5462	1479.479	-0.463	0.645	-3617.131	2248.038
budget	0.9741	0.167	5.832	0.000	0.643	1.305
FilmCritics	-4158.9471	9564.556	-0.435	0.665	-2.31e+04	1.48e+04
ratingKinopoisk	-4406.0746	9500.035	-0.464	0.644	-2.32e+04	1.44e+04
Length	-7.734e+04	1.74e+05	-0.444	0.658	-4.23e+05	2.68e+05
sequel	95.5222	182.993	0.522	0.603	-267.202	458.246
season	-434.3918	501.912	-0.865	0.389	-1429.268	560.485
SJW	96.0857	57.281	1.677	0.096	-17.456	209.627
star_power	-62.9512	255.807	-0.246	0.806	-570.006	444.103
action movie	-97.6510	191.690	-0.509	0.611	-477.614	282.312
comedy	-251.9529	302.231	-0.834	0.406	-851.027	347.122
drama	-237.2373	665.551	-0.356	0.722	-1556.475	1082.001

=== Квантильная регрессия для q = 0.5 ===

QuantReg Regression Results

Dep. Variable:	box_office	Pseudo R-squared:	0.2717			
Model:	QuantReg	Bandwidth:	1.200e+08			
Method:	Least Squares	Sparsity:	3.179e+08			
Date:	Thu, 08 May 2025	No. Observations:	120			
Time:	13:34:56	Df Residuals:	108			
		Df Model:	11			
	coef	std err	t	P> t	[0.025	0.975]
const	655.2328	1666.650	0.393	0.695	-2648.357	3958.822
budget	2.0618	0.204	10.102	0.000	1.657	2.466
FilmCritics	5828.2122	1.08e+04	0.541	0.590	-1.55e+04	2.72e+04
ratingKinopoisk	4479.1045	1.07e+04	0.419	0.676	-1.67e+04	2.57e+04
Length	1.343e+05	1.96e+05	0.685	0.495	-2.55e+05	5.23e+05
sequel	-33.6017	206.144	-0.163	0.871	-442.215	375.011
season	-33.6017	565.409	-0.059	0.953	-1154.341	1087.138
SJW	2.52e-06	64.528	3.91e-08	1.000	-127.906	127.906
star_power	-4.914e-07	288.170	-1.71e-09	1.000	-571.203	571.203
action movie	-33.6017	215.941	-0.156	0.877	-461.634	394.431
comedy	-1.064e-06	340.467	-3.13e-09	1.000	-674.864	674.864
drama	688.8345	749.751	0.919	0.360	-797.302	2174.971

=== Квантильная регрессия для  $q = 0.85$  ===

#### QuantReg Regression Results

Dep. Variable:	box_office	Pseudo R-squared:	0.2884
Model:	QuantReg	Bandwidth:	1.650e+08
Method:	Least Squares	Sparsity:	5.381e+08
Date:	Thu, 08 May 2025	No. Observations:	120
Time:	16:31:23	Df Residuals:	108
		Df Model:	11

	coef	std err	t	P> t	[0.025	0.975]
const	1339.1865	2112.010	0.634	0.527	-2847.185	5525.558
budget	4.4288	0.236	18.778	0.000	3.961	4.896
FilmCritics	1.138e+04	1.37e+04	0.833	0.407	-1.57e+04	3.84e+04
ratingKinopoisk	8598.6930	1.36e+04	0.634	0.527	-1.83e+04	3.55e+04
Length	1.861e+05	2.49e+05	0.749	0.456	-3.07e+05	6.79e+05
sequel	-55.7994	261.229	-0.214	0.831	-573.601	462.003
season	-55.7994	716.498	-0.078	0.938	-1476.022	1364.423
SJW	-55.7994	81.771	-0.682	0.496	-217.884	106.285
star_power	-2.654e-08	365.174	-7.27e-11	1.000	-723.839	723.839
action movie	-55.7994	273.645	-0.204	0.839	-598.211	486.612
comedy	1394.9859	431.446	3.233	0.002	539.785	2250.187
drama	1394.9859	950.099	1.468	0.145	-488.274	3278.246

Рис. 8 Изменение коэффициентов в квантильной регрессии

Изменение коэффициентов квантильной регрессии по квантилям

