

Федеральное государственное автономное образовательное учреждение  
высшего образования

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**

**«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

**Факультет экономических наук**

**Образовательная программа «Экономика»**

## **ПРОЕКТНАЯ РАБОТА**

Влияние технических характеристик и ненаблюдаемых признаков качества  
на формирование цены на рынке мобильных телефонов, пример Кении

Выполнили:

Студенты 3 курса бакалавриата

**Аксёнов Антон Витальевич БЭК201**

**Зайцев Александр Алексеевич БЭК201**

**Трофименко Илья Александрович БЭК202**

	<i>Аксёнов Антон</i>	<i>Зайцев Александр</i>	<i>Трофименко Илья</i>
<i>Предобработка данных</i>	<b>34%</b>	<b>33%</b>	<b>33%</b>
<i>Построение моделей</i>	<b>33%</b>	<b>34%</b>	<b>33%</b>
<i>Проверка гипотез</i>	<b>33%</b>	<b>33%</b>	<b>34%</b>
<i>Прогнозирование</i>	<b>34%</b>	<b>33%</b>	<b>33%</b>
<i>Интерпретация результатов</i>	<b>33%</b>	<b>34%</b>	<b>33%</b>
<i>Составление отчёта</i>	<b>33%</b>	<b>33%</b>	<b>34%</b>

Москва 2023

## 1. Введение

В данном исследовании был поставлен вопрос о зависимости цены от изменения качества продукта. Для товаров, особые свойства и характеристики которых меняются непрерывно, поиск зависимости изменения качества представляет глубокий исследовательский вопрос. В рамках работы предпринят поиск решения данной проблемы с использованием средств эконометрики для изучения взаимосвязи цены и качественных характеристик непрерывно развивающегося и склонного к модификациям товара, как смартфон. Кроме того, отдельно рассматривается часть цены, которая не объясняется качеством товара, и вероятно формируется на основе доверия потребителей к фирме-изготовителю за счёт ненаблюдаемых впечатлений, сформированных рекламной кампанией и предыдущим опытом, что представляет собой некоторый аналог goodwill. Таким образом, с помощью гедонистической ценовой функции проведён анализ формирования цены смартфонов, исследовано влияние технических характеристик, а также проверены выдвинутые гипотезы по нетехническим факторам, отражающим ожидания пользователей. В данном исследовании используются пространственные данные (cross section), для всех моделей и проверки гипотез рассматривается уровень значимости в 5%.

## 2. Гипотезы

Важное наблюдение за поведением человека состоит в его желании к покупке товаров, выпущенных на рынок недавно, то есть существует стремление приобретать новые модели и отказываться от старых без учета их технических характеристик. Каждая новая модель любого мобильного телефона автоматически пользуется повышенным спросом, что и приводит к увеличению цены. Исходя из выдвинутого предположения, в рамках исследования поставлена следующая гипотеза:

***H1:** Год выпуска смартфона положительно влияет на его стоимость. Чем новее устройство, тем оно дороже.*

С другой стороны, наибольшее внимание потребителей привлекает не детальная техническая информация, а бренд рассматриваемого мобильного устройства. Можно сказать, что потребительский спрос растёт на продукцию фирмы-изготовителя, основываясь на прошлой репутации компании, то есть на формирование цены оказывает влияние не только техническая составляющая телефона, но и компания, которая его выпустила. Наиболее известные и авторитетные компании ввиду повышенного спроса устанавливают на свою продукцию наценку. В связи с этим замечанием в рамках данной работы рассматриваются следующая гипотеза:

***H2:** Бренд смартфона значимо влияет на его стоимость. Популярные фирмы имеют возможность повышать цены на свою продукцию, а неизвестные изготовители вынуждены её снижать.*

Также на потребительский спрос, а следовательно, и на цену товара, оказывает влияние опыт предыдущих покупателей из-за наличия неопределенности при выборе товара из множества альтернатив, асимметрии информации на рынке товаров. Следовательно, рейтинг товаров должен оказывать положительное влияние на спрос, а через него и на формирование цен. Исходя из этого соображения, выдвинута гипотеза, что на ценообразование смартфонов влияют отзывы покупателей:

***H3:** Рейтинг покупателей оказывает значимое влияние на ценообразование на рынке мобильных телефонов. Низкий или высокий рейтинг приводит к повышению или снижению спроса на соответствующую модель, что выливается в изменения цены.*

### **3. Выборка**

В данном исследовании рассматривается взаимосвязь между ценой и качеством продукции рынка мобильных телефонов на примере данных одной из торговых площадок Кении, содержащей цены на наиболее продаваемые модели смартфонов за 2021 год. Основная часть данных была взята с сайта Kaggle [\[1\]](#). В качестве наблюдения выступает отдельная модель мобильного телефона. В рамках предобработки выборки были выделены переменные, представленные ниже в Таблице 1. Дополнительно для проверки гипотез с помощью методов парсинга на языке python была добавлена переменная `YEAR_OF_RELEASE`.

Наблюдения, содержащие пропущенные значения, а также выбросы, выявленные методом студентизированных остатков, были удалены из выборки. Кроме того, были удалены наблюдения, которые были получены некорректно ввиду ошибок парсера, что наблюдается на распределении цен и при детальном рассмотрении каждого смартфона. Сет контрольных переменных, которые были включены в модель для построения гедонистической ценовой функции и проверки гипотез, были описаны как значимые в более ранних исследованиях ([Grzybowski & Nicolle, 2021](#)). В качестве зависимой переменной используется натуральный логарифм от цены из-за логнормального распределения цен. Финальная выборка составила 559 наблюдений.

В таблице 2 представлены описательные статистики. Можно заметить, что бренды *XIAOMI*, *HUAWEI* и *SAMSUNG* встречаются гораздо чаще других и в совокупности составляют почти половину всей выборки. Кроме того, медианный год выпуска 2019, то

есть половину выборки составляют более новые модели телефонов, выпущенные в 2019, 2020 и 2021 годах.

Таблица 1  
Определение переменных

Переменная	Определение
<i>Зависимая переменная</i>	
<i>LOG_PRICE</i>	<i>Натуральный логарифм от стоимости смартфона в кенийских шиллингах</i>
<i>Объясняющие переменные</i>	
<i>RATING</i>	<i>Средняя арифметическая оценок, оставленных потребителями на странице модели смартфона по пятибалльной шкале</i>
<i>BRAND_*</i>	<i>1, если смартфон, выпущен под брендом *, 0 иначе</i>
<i>YEAR_OF_RELEASE</i>	<i>Год выпуска смартфона</i>
<i>Контрольные переменные</i>	
<i>DISP_DIAG</i>	<i>Длина диагонали дисплея смартфона в дюймах</i>
<i>CAMERA</i>	<i>Среднее арифметическое количества мегапикселей фронтальной и задней камер мобильного устройства</i>
<i>STORAGE</i>	<i>Объем внутренней памяти устройства в гигабайтах</i>
<i>RAM</i>	<i>Объем оперативной памяти смартфона в гигабайтах</i>
<i>BATTERY_CAPACITY</i>	<i>Объем батареи смартфона в миллиамперах в час</i>
<i>BATTERY_TYPE_LIPO</i>	<i>Тип батареи: 1, если в смартфоне используется батарея типа LiPO, 0 если батарея LiON</i>
<i>IOS_OPER_SYST</i>	<i>Тип операционной системы: 1, если смартфон работает на базе IOS, 0 если на базе Android</i>
<i>OTHER_OPER_SYST</i>	<i>Тип операционной системы: 1, если смартфон работает на базе ОС, отличной от Android или IOS, 0 если на базе Android</i>
<i>OPER_SYST_VERS</i>	<i>Версия операционной системы</i>
<i>DENSITY</i>	<i>Плотность пикселей на дюйм, которую поддерживает матрица дисплея смартфона в 10тыс. пикселях</i>

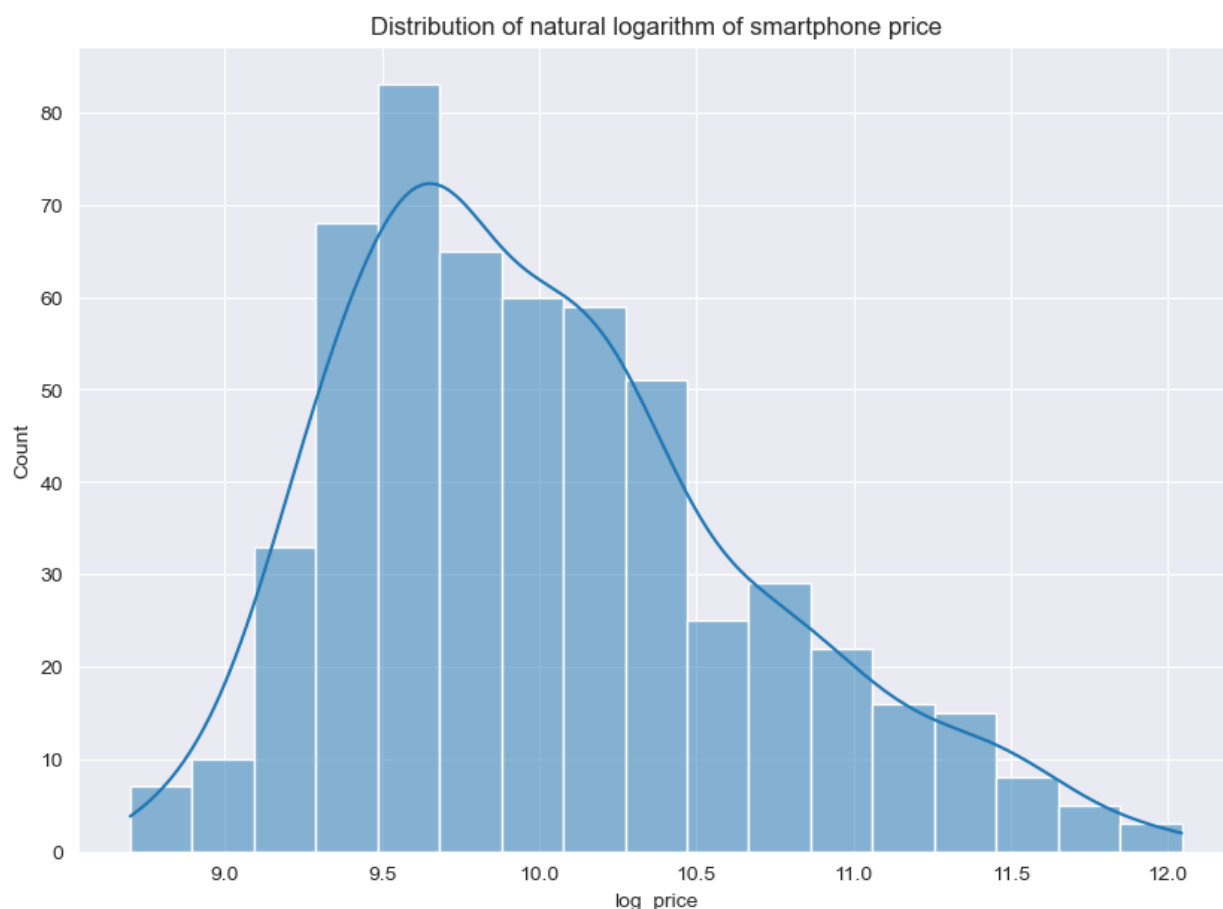
Таблица 2  
Описательные статистики данных

	Obs	Mode	Mean	Median	SD
<i>LOG_PRICE</i>	559	9.31	10.027	9.928	0.647
<i>RATING</i>	559	5	3.845	3.800	0.753
<i>YEAR_OF_RELEASE</i>	559	2019	2018.265	2019	1.536
<i>BRAND_APPLE</i>	559	0	0.047	0	0.211
<i>BRAND_XIAOMI</i>	559	0	0.152	0	0.359
<i>BRAND_VIVO</i>	559	0	0.009	0	0.094
<i>BRAND_TECNO</i>	559	0	0.077	0	0.267
<i>BRAND_SONY</i>	559	0	0.021	0	0.145
<i>BRAND_SAMSUNG</i>	559	0	0.181	0	0.385
<i>BRAND_REALME</i>	559	0	0.032	0	0.177
<i>BRAND_OPPO</i>	559	0	0.054	0	0.226
<i>BRAND_ONEPLUS</i>	559	0	0.029	0	0.167
<i>BRAND_NOKIA</i>	559	0	0.039	0	0.195
<i>BRAND_MOTOROLA</i>	559	0	0.011	0	0.103
<i>BRAND_LG</i>	559	0	0.023	0	0.151
<i>BRAND_LENOVO</i>	559	0	0.020	0	0.139
<i>BRAND_LAVA</i>	559	0	0.004	0	0.060
<i>BRAND_INFINIX</i>	559	0	0.063	0	0.242
<i>BRAND_HUAWEI</i>	559	0	0.123	0	0.329
<i>BRAND_HTC</i>	559	0	0.055	0	0.229
<i>BRAND_HONOR</i>	559	0	0.009	0	0.094
<i>BRAND_GOOGLE</i>	559	0	0.020	0	0.139
<i>BRAND_GIONEE</i>	559	0	0.004	0	0.060
<i>BRAND_CUBOT</i>	559	0	0.007	0	0.084
<i>BRAND_BLACKBERRY</i>	559	0	0.021	0	0.145
<i>BATTERY_TYPE_LIPO</i>	559	1	0.544	1	0.499
<i>DISP_DIAG</i>	559	5.50	6.006	6.200	0.586
<i>CAMERA</i>	559	8	13.050	10	8.653
<i>STORAGE</i>	559	128	83.177	64	77.043
<i>RAM</i>	559	4	4.460	4	2.564
<i>BATTERY_CAPACITY</i>	559	4000	3756.696	4000	864.520
<i>IOS_OPER_SYST</i>	559	0	0.047	0	0.211
<i>OTHER_OPER_SYST</i>	559	0	0.004	0	0.060
<i>OPER_SYST_VERS</i>	559	9	8.438	9	1.615
<i>DENSITY</i>	559	68548.76	56945.095	60048.789	30035.269

#### 4. Построение эконометрической модели для цены товара

В качестве зависимой переменной в исследовании используется логарифм от цены, в то время как регрессоры не логарифмируются, то есть для анализа применяется полулогарифмическая модель. Распределение натурального логарифма цены смартфона представлено на Рис. 1. Как видно на графике ниже визуальное распределение логарифма цены довольно похоже на нормальное, но имеет правостороннюю асимметрию.

Рис. 1.  
Гистограмма распределения натурального логарифма цены смартфона



Для того, чтобы убедиться в правильности выбора функциональной формы проведена визуальная оценка связи между зависимой переменной и регрессорами. Графики облака рассеяния представлены на Рис. 2. и Рис. 3. Явной линейной и логлинейной зависимости на данных не прослеживается. Графики облака рассеяния на логарифмированных признаках представлены на Рис. 4. и Рис. 5. Анализируя результаты, можно сказать, что явных зависимостей также не видно. Однако приведённые построения - это проекция данных на плоскость, следует проверить наличие более сложных взаимосвязей.

Рис. 2.  
Облака рассеяния регрессоров в зависимости от переменной Log\_Price

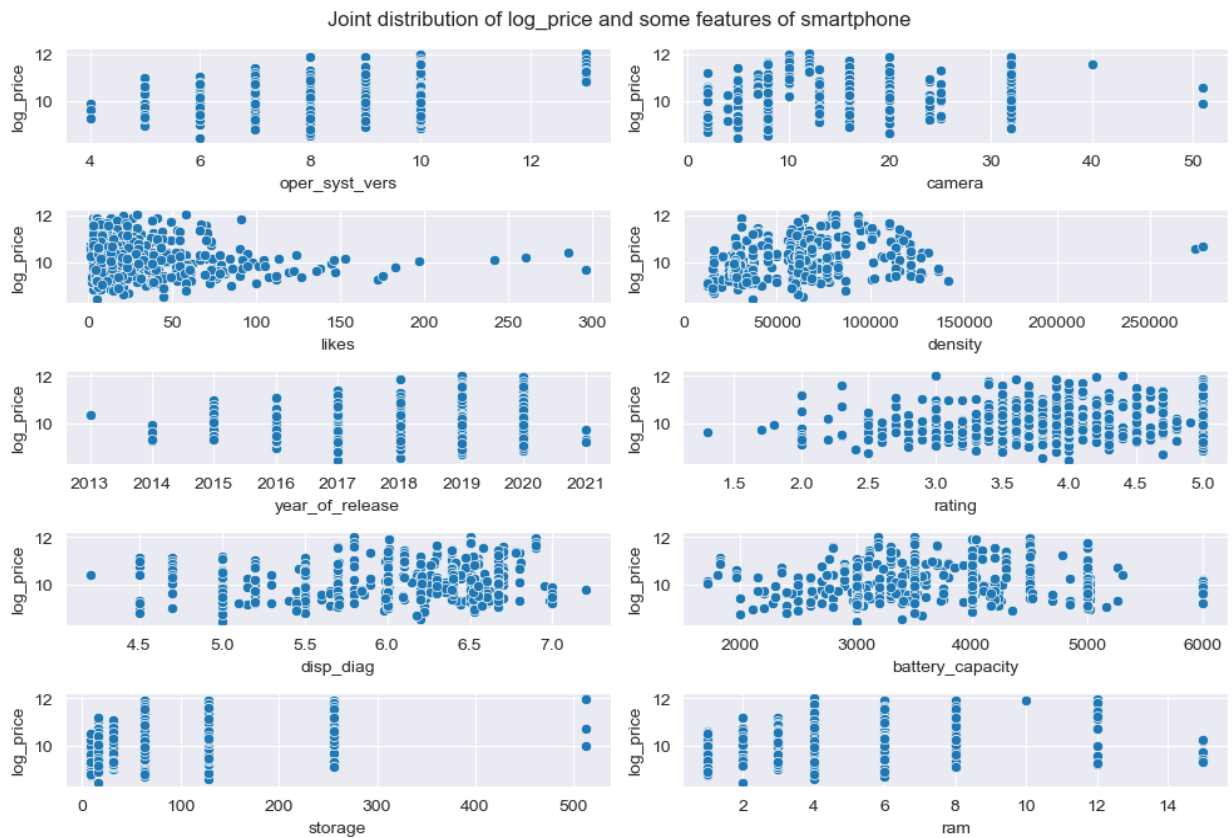


Рис. 3.  
Облака рассеяния регрессоров в зависимости от переменной Price

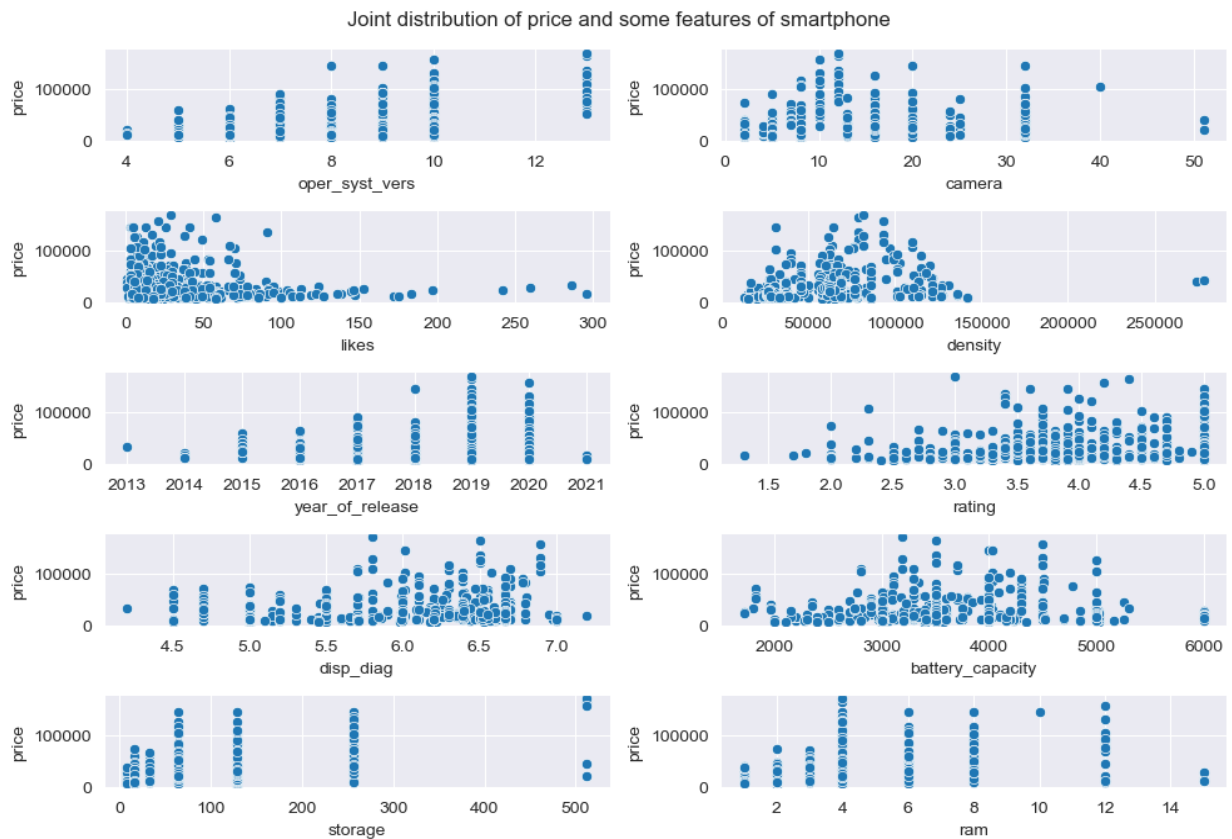


Рис. 4.

Облака рассеяния логарифмов регрессоров в зависимости от переменной Log\_Price

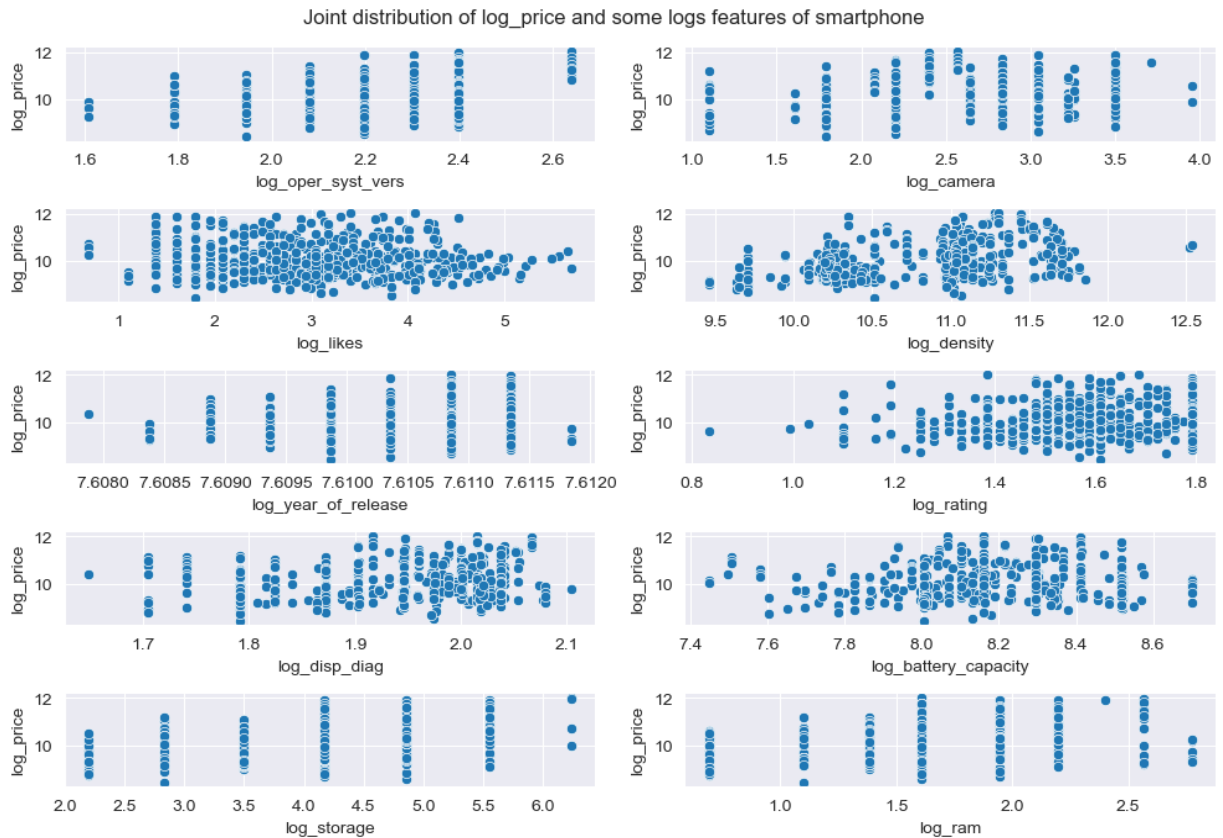
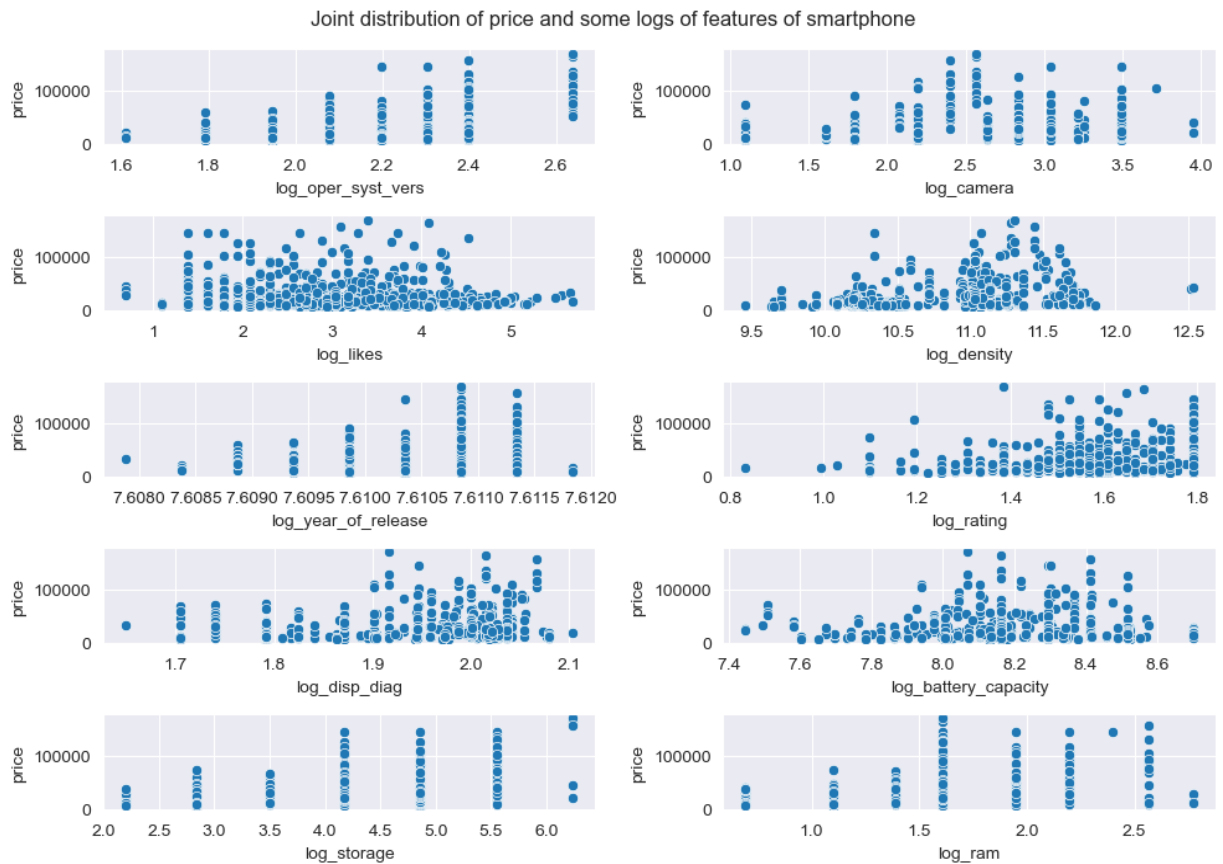


Рис. 5.

Облака рассеяния регрессоров в зависимости от зависимой переменной Price

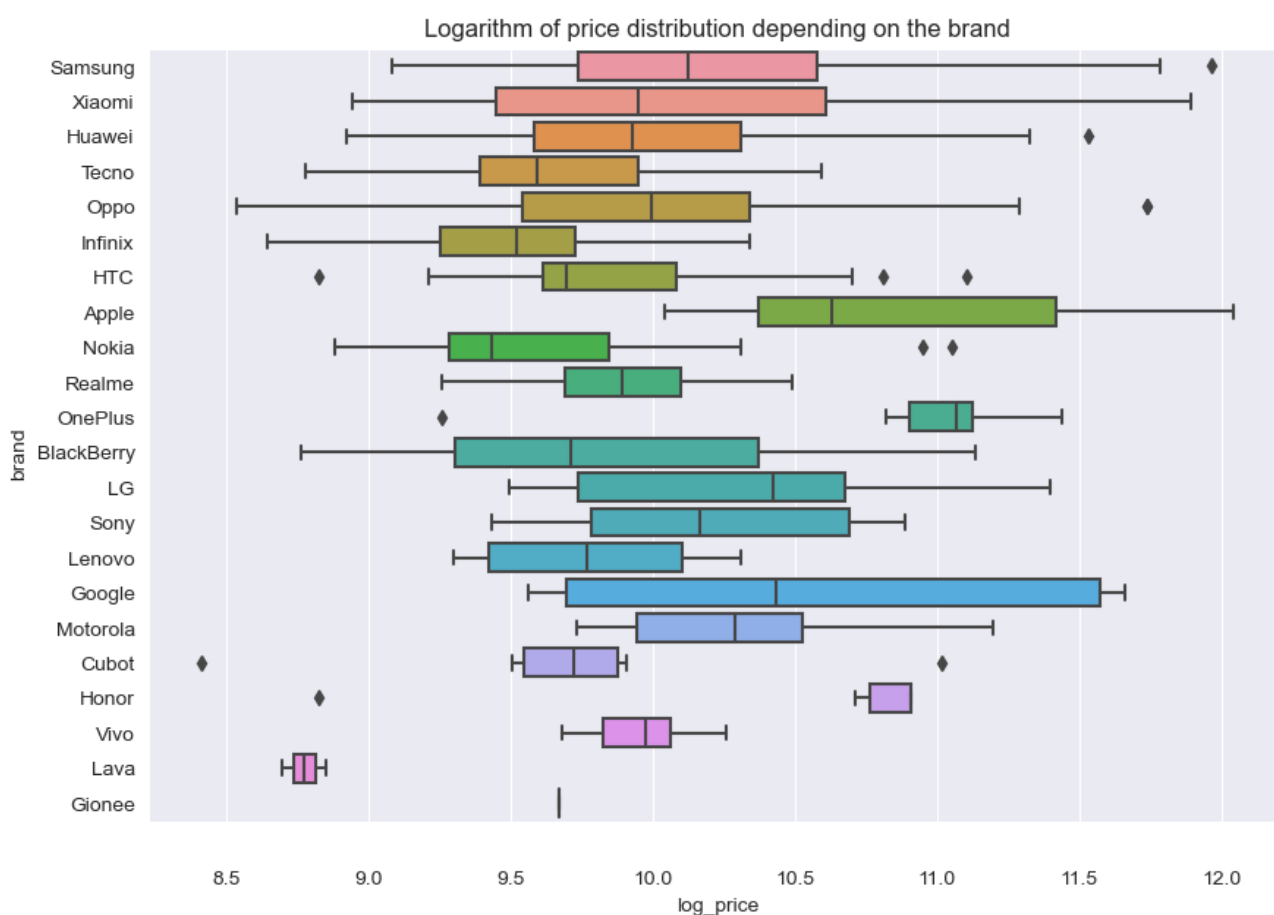




Для того, чтобы определить верную функциональную форму зависимой переменной и регрессоров, применяется РЕ-тест. Согласно его результатам на любом разумном уровне значимости нулевая гипотеза о незначимости логарифмической части в линейной модели отвергается ( $p\text{-value} \approx 0$ ), в то время как незначимость линейной части в полулогарифмической модели не отвергается ( $p\text{-value} \approx 0,52$ ). Таким образом, линейная часть в полулогарифмической модели не позволила лучше объяснить данные, в то время как логарифмическая часть в линейной показала значимый результат. Следовательно, по РЕ-тесту необходимо выбрать полулогарифмическую модель, что соответствует первоначальному выбору. Далее все исследования проводятся с применением полулогарифмических моделей.

После выбора функциональной формой зависимой переменной, необходимо обратить более пристальное внимание на объясняющие переменные. На Рис. 6. представлено распределение логарифма цены смартфонов в зависимости от их бренда. Отметим, что смартфоны фирм Apple и Google в среднем дороже, чем остальные. Кроме того, другие известные бренды: Samsung и Xiaomi выставляют на рынок гораздо более широкую линейку смартфонов, поэтому их стоимость сильно различается.

Рис. 6.  
Распределение переменной Log\_Price в зависимости от бренда в виде ящиков с усами





Одной из предпосылкой теоремы Гаусса-Маркова является гомоскедастичность дисперсии ошибок. В рамках проверки данной предпосылки для исследуемых данных был проведен графический и статистический анализы. В рамках графического анализа можно заметить снижение разброса остатков, график представлен на Рис. 8. Кроме того, на облаках рассеяния многих признаков заметно увеличение или уменьшение разброса (Рис.2-3). К примеру, для признака RATING увеличивается разброс данных, а для BATTERY\_CAPACITY уменьшается, что подтверждает гипотезу о наличие гетероскедастичности. Также были проведен тесты Уайта ( $p\text{-value} \approx 0$ ), Бройша-Пагана ( $p\text{-value} \approx 0$ ) и Голдфельда-Куандта ( $p\text{-value} \approx 0.108$ ). Относительно тестов Уайта и Бройша-Пагана гипотеза о гомоскедастичности остатков отвергается при любом разумном уровне значимости.

Общая формула модели:

$$\begin{aligned} \log\_Price = & \beta_0 + \beta_1 * RATING + \beta_2 * YEAR\_OF\_RELEASE + \sum_i \alpha_i * Brand_i \\ & + \beta_3 * DISP\_DIAG + \beta_4 * CAMERA + \beta_5 * STORAGE + \beta_6 * RAM \\ & + \beta_7 * BATTERY\_CAPACITY + \beta_8 * BATTERY\_TYPE\_LIPO \\ & + \beta_9 * IOS\_OPER\_SYST + \beta_{10} * OPER\_SYST\_VERS + \beta_{11} * DENSITY + \epsilon \end{aligned}$$

Оценим четыре спецификации модели, в трёх из которых включена одна из объясняющих переменных, а в четвёртой модели оцениваются все три объясняющие переменные вместе. Результаты МНК оценок спецификаций модели представлены в Таблице 3. Для всех моделей стандартные ошибки были оценены в робастной форме Уайта, чтобы избежать проблемы гетероскедастичности. Все модели оказались значимыми в целом согласно F-статистике, верно специфицированными в соответствии с тестом Рамсея на 5% уровне значимости ( $p\text{-value} > 0.098$ ), однако согласно критерию Акаике и скорректированному  $R^2$  модель, которая включает все объясняющие переменные, лучше объясняет данные. Кроме того, остатки моделей, в которые включена одна из объясняющих переменных, на 5% уровне значимости не имеют нормального распределения согласно тесту Харке-Бера ( $p\text{-value} < 0.04$ ), а также наблюдается некоторая автокорреляция на остатках модели в соответствии с тестом Дарбина-Уотсона ( $DW < 1.5$ ), что нарушает предпосылки теоремы Гаусса-Маркова о нормальности распределения ошибок и об отсутствии автокорреляции остатков, в то время как на полной спецификации не отвергается нормальность распределения остатков ( $p\text{-value} < 0.075$ ) и значение статистики Дарбина-Уотсона на граничных значениях для отсутствия автокорреляции ( $DW \approx 1.5$ ). Поскольку остатки является оценкой ошибок, для модели (4) не нарушаются предпосылки Гаусса-Маркова о нормальности распределения ошибок и отсутствии автокорреляции.

Таким образом, оптимальная модель включает все три объясняющие переменные. На Рис. 8. представлено облако рассеяния прогнозных значений зависимой переменной. В теории должна наблюдаться горизонтальная прямая остатков, но на исследуемых данных вид прямой указывает на наличие некоторой сложной нелинейной связи в остатках.

Таблица 3  
Результаты МНК оценки

<i>Dependent variable: LOG_PRICE</i>				
	(1)	(2)	(3)	(4)
RATING	0.074** (0.03)			0.064** (0.029)
YEAR_OF_RELEASE		0.032 (0.022)		0.035* (0.020)
BRAND_APPLE			0.949*** (0.128)	0.974*** (0.122)
BRAND_BLACKBERRY			0.217 (0.227)	0.180 (0.216)
BRAND_CUBOT			0.224** (0.095)	0.315*** (0.092)
BRAND_GIONEE			-0.175* (0.094)	-0.141 (0.113)
BRAND_GOOGLE			0.604** (0.259)	0.576** (0.260)
BRAND_HTC			-0.032 (0.106)	0.001 (0.104)
BRAND_HONOR			0.767*** (0.168)	0.734*** (0.171)
BRAND_HUAWEI			-0.060 (0.081)	-0.038 (0.082)
BRAND_INFINIX			-0.172** (0.076)	-0.152** (0.077)
BRAND_LG			0.420***	0.452***

			(0.152)	(0.154)
BRAND_LAVA			-0.592***	-0.639***
			(0.101)	(0.096)
BRAND_LENOVO			-0.079	-0.023
			(0.085)	(0.088)
BRAND_MOTOROLA			0.580***	0.647***
			(0.188)	(0.194)
BRAND_NOKIA			-0.088	-0.113
			(0.086)	(0.088)
BRAND_ONEPLUS			0.630***	0.621***
			(0.101)	(0.100)
BRAND_OPPO			-0.064	-0.034
			(0.093)	(0.095)
BRAND_REALME			-0.001	-0.047
			(0.104)	(0.107)
BRAND_SAMSUNG			0.165**	0.180**
			(0.072)	(0.072)
BRAND_SONY			0.340**	0.368**
			(0.160)	(0.154)
BRAND_TECNO			0.009	0.039
			(0.087)	(0.088)
BRAND_VIVO			0.097	0.143**
			(0.067)	(0.067)
CONST	9.425***	-54.598	9.006***	-61.279
	(0.356)	(43.650)	(0.344)	(40.402)
CAMERA	0.005*	0.005*	0.009***	0.008***
	(0.003)	(0.003)	(0.003)	(0.003)
DENSITY	0.035***	0.041***	0.022***	0.025***
	(0.008)	(0.008)	(0.008)	(0.008)
DISP_DIAG	-0.011	-0.032	0.067	0.013

	(0.075)	(0.077)	(0.071)	(0.069)
RAM	0.045***	0.045***	0.055***	0.056***
	(0.013)	(0.013)	(0.015)	(0.014)
STORAGE	0.003***	0.003***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
OTHER_OPER_SYST	0.032	0.062	0.073	0.218
	(0.355)	(0.438)	(0.488)	(0.451)
BATTERY_CAPACITY	-0.103***	-0.111***	-0.050	-0.054
	(0.039)	(0.041)	(0.039)	(0.040)
BATTERY_TYPE_LIPO	0.039	0.037	0.049	0.035
	(0.059)	(0.059)	(0.054)	(0.054)
Observations	559	559	559	559
R <sup>2</sup>	0.434	0.568	0.430	0.576
Adjusted R <sup>2</sup>	0.424	0.544	0.420	0.551
F Statistic	46.719***	23.958***	45.936***	23.057***
AIC	801.4	805.5	690.5	684.2
Durbin-Watson	1.272	1.280	1.465	1.497
Jarque-Bera	10.93***	9.9***	6.748**	5.18*
Ramsey's RESET	1.502	1.542	0.575	2.743*

Note: \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

Рис. 8.  
Облако рассеяния прогнозных значений в зависимости от остатков модели (4)



## 5. Эндогенность

Относительно признаков в выборке можно отметить, что помимо характеристик, выбранных в рамках эконометрической модели, существует также множество других регрессоров, отражающих производительность смартфона. К примеру: графический и центральный процессоры. В то же время такая характеристика емкость батареи (BATTERY\_CAPACITY) может иметь корреляцию с производительностью процессоров. Например, для более производительных процессоров может быть необходимо большее энергопотребление, следовательно, необходима повышенная емкость батареи. Следовательно, так как остатки могут объясняться признаками, не входящими в эконометрическую модель, а также быть коррелированы с признаками модели, то может возникать проблема эндогенности.

Для решения данной проблемы использовался метод IV, инструментальной переменной в которой является оценка производительности смартфона относительно других моделей, выраженная в процентах (SPECS\_SCORE). В рамках данного метода необходимо проверить, действительно ли наблюдается проблема эндогенности, поэтому была оценена модель, в которой эндогенной переменной являлась емкость батареи. После первого шага значение F-статистики значительно превышает 105 ( $F \approx 3655$ ) для исследуемого эндогенного признака, следовательно, инструмент является релевантным.

В качестве проверки наличия эндогенности был использован тест Хаусмана, который проверяет наличие эндогенности с помощью релевантных инструментов и эндогенных признаков. После проверки предполагаемого эндогенного регрессора, значение  $p\text{-value} > 0.8$ , следовательно, для любого разумного уровня значимости не отвергается нулевая гипотеза, что отсутствует эндогенность на данных. Тест Хаусмана в случае отвержения нулевой гипотезы не описывают точную причину отвержения, поэтому объяснением данного результата может являться либо отсутствие эндогенности регрессоров, либо невалидность инструмента. Для проверки валидности инструмента необходимо провести тест Саргана-Хансена, однако ввиду наличия только одной инструментальной переменной, тест неприменим на наших данных, так как он требует строго больше инструментов, чем эндогенных переменных. Следовательно, валидность SPECS\_SCORE проверить не удалось, то есть невозможно сделать однозначный вывод о результате теста Хаусмана.

## **6. Результаты оценивания множественной линейной регрессии**

Проинтерпретируем результаты оптимальной линейной модели, полученные выше. На 5% уровне значимости оказались статистически значимыми следующие регрессоры: RATING, BRAND\_APPLE, BRAND\_CUBOT, BRAND\_GOOGLE, BRAND\_HONOR, BRAND\_INFINIX, BRAND\_LG, BRAND\_LAVA, BRAND\_MOTOROLA, BRAND\_ONEPLUS, BRAND\_SAMSUNG, BRAND\_SONY, BRAND\_VIVO, CAMERA, DENSITY, RAM, STORAGE. Учитывая, что исследуемая зависимая переменная выражена в логарифмах, можно сказать, что в среднем при прочих равных на 5% уровне значимости повышение рейтинга смартфона на 1 балл приводит к его удорожанию на 6.4%, увеличение качества камеры на 1 мегапиксель приводит к росту цены на 0.8%, увеличение плотности пикселей на дюйм дисплея на 10000 единиц увеличивает цену на 2.5%, расширение объёма оперативной памяти на 1 гигабайт увеличивает цену смартфона на 5.6%, а рост объёма внутренней памяти на 1 гигабайт приводит к повышению цены смартфона на 0.2%, смартфоны от фирмы Apple стоят на 164.9% дороже смартфонов под брендом Xiaomi, который был выбран базовым, принадлежность к фирме Cubot увеличивает стоимость смартфона на 37.02% относительно Xiaomi, бренд Google повышает цену на 77.9% относительно Xiaomi, бренд Honor на 108.3%, бренд LG на 57.1%, бренд Motorola на 91%, бренд OnePlus на 86.07%, бренд Samsung на 19.7%, бренд Sony на 44.5%, бренд Vivo на 15.4%, бренд Infinix наоборот снижает стоимость смартфона относительно бренда Xiaomi на 16.4%, бренд Lava уменьшает цену на 89.45%.

Таким образом, гипотеза H1 отвергается на 5% уровне значимости, то есть год выпуска в среднем при прочих равных не влияет на цену смартфона. Гипотеза H2 напротив не отвергается, хотя не все бренды оказались статистически значимыми, всё же принадлежность к тем или иным фирмам-изготовителям может влиять на цену смартфона как положительно, так и отрицательно. Гипотеза H3 также не отвергается, так как переменная RATING оказалась значимой, более того, её коэффициент положителен, то есть с увеличением рейтинга растёт и цена смартфонов.

## **7. Квантильная регрессия**

Можно заметить, что, согласно ранее проведенному графическому анализу, логарифм цены смартфона имеет правостороннюю асимметрию, поэтому медиана может лучше объяснить данные оценкой логарифма цены смартфона, в сравнении с оценкой математического ожидания, которая получается в результате применения метода наименьших квадратов. Для получения оценки условной медианы логарифма цены



смартфона с заданными характеристиками можно воспользоваться квантильной регрессией, которая кроме всего прочего в сравнении с обычной линейной регрессией, оценённой с помощью МНК, менее чувствительна к выбросам, которые в используемых данных, безусловно, остались даже после тщательной работы авторов по избавлению от них.

В качестве спецификации модели выберем наиболее полную из всех ранее описанных, включающую все переменные интереса и сет контрольных переменных. Результаты оценивания коэффициентов модели (6) приведены в Таблице 4, где также показаны результаты оценивания квантильных регрессий, рассматривающих квантили уровня 0.1 и 0.9 (модели (5) и (7) соответственно).

Таблица 4  
Результаты оценивания квантильной регрессии

	<i>Dependent variable: LOG_PRICE</i>		
	(5)	(6)	(7)
CONST	0.003 (0.001)	0.001 (0.001)	0.000 (0.001)
OTHER_OPER_SYST	0.000 (0.236)	-0.000 (0.304)	-0.003 (0.371)
BATTERY_CAPACITY	-0.030 (0.036)	-0.019 (0.033)	-0.053 (0.061)
BATTERY_TYPE_LIPO	0.004 (0.062)	0.060 (0.045)	0.032 (0.091)
BRAND_APPLE	0.000 (0.116)	0.747*** (0.103)	0.741*** (0.230)
BRAND_BLACKBERRY	0.192 (0.248)	-0.141 (0.141)	-0.149 (0.312)
BRAND_CUBOT	0.000 (0.120)	0.106 (0.203)	0.000 (0.160)
BRAND_GIONEE	0.387	0.000	-0.000

	(0.738)	(0.280)	(0.185)
BRAND_GOOGLE	0.229	-0.000	0.511
	(0.168)	(0.131)	(0.356)
BRAND_HTC	0.015	-0.169*	-0.341*
	(0.137)	(0.087)	(0.179)
BRAND_HONOR	0.000	0.002	0.412
	(0.104)	(0.181)	(0.443)
BRAND_HUAWEI	-0.054	-0.191***	-0.188
	(0.091)	(0.066)	(0.119)
BRAND_INFINIX	0.006	-0.314***	-0.534***
	(0.122)	(0.081)	(0.146)
BRAND_LG	0.258	0.269**	-0.002
	(0.187)	(0.123)	(0.247)
BRAND_LAVA	-0.391	-0.000	-0.000
	(0.738)	(0.285)	(0.218)
BRAND_LENOVO	0.064	-0.208	-0.000
	(0.154)	(0.128)	(0.111)
BRAND_MOTOROLA	0.000	0.263	0.000
	(0.112)	(0.170)	(0.512)
BRAND_NOKIA	-0.014	-0.226**	-0.537***
	(0.137)	(0.098)	(0.200)
BRAND_ONEPLUS	-0.000	0.370***	0.078
	(0.092)	(0.112)	(0.226)
BRAND_OPPO	-0.120	-0.134	-0.314**
	(0.116)	(0.084)	(0.151)
BRAND_REALME	0.019	-0.309***	0.035
	(0.151)	(0.104)	(0.188)
BRAND_SAMSUNG	0.182**	-0.013	-0.066

	(0.087)	(0.059)	(0.104)
BRAND_SONY	0.000	0.198	-0.013
	(0.109)	(0.131)	(0.307)
BRAND_TECNO	0.039	-0.184**	0.088
	(0.112)	(0.079)	(0.143)
BRAND_VIVO	0.552*	-0.068	0.000
	(0.311)	(0.182)	(0.135)
CAMERA	0.012***	0.004*	0.007
	(0.004)	(0.002)	(0.005)
DENSITY	0.068	0.025***	0.044***
	(0.008)	(0.007)	(0.011)
DISP_DIAG	0.191***	0.052	-0.101
	(0.072)	(0.057)	(0.129)
RAM	0.003	0.058***	0.067***
	(0.015)	(0.010)	(0.015)
RATING	-0.001	0.060**	0.125**
	(0.035)	(0.025)	(0.052)
STORAGE	0.002***	0.003***	0.003***
	(0.000)	(0.000)	(0.001)
YEAR_OF_RELEASE	0.004***	0.004***	0.005***
	(0.000)	(0.000)	(0.000)
Observations	559	559	559
Note: *p<0.1; **p<0.05; ***p<0.01			

Полученные результаты медианной регрессии можно сравнить с результатами модели с полной спецификацией, оцененной с помощью МНК: в медианной регрессии коэффициент при годе выпуска смартфона является значимым на любом разумном уровне значимости, в отличие от МНК оценки, где он значим только на 10% уровне значимости. Также отметим, что количество значимых коэффициентов перед брендами существенно

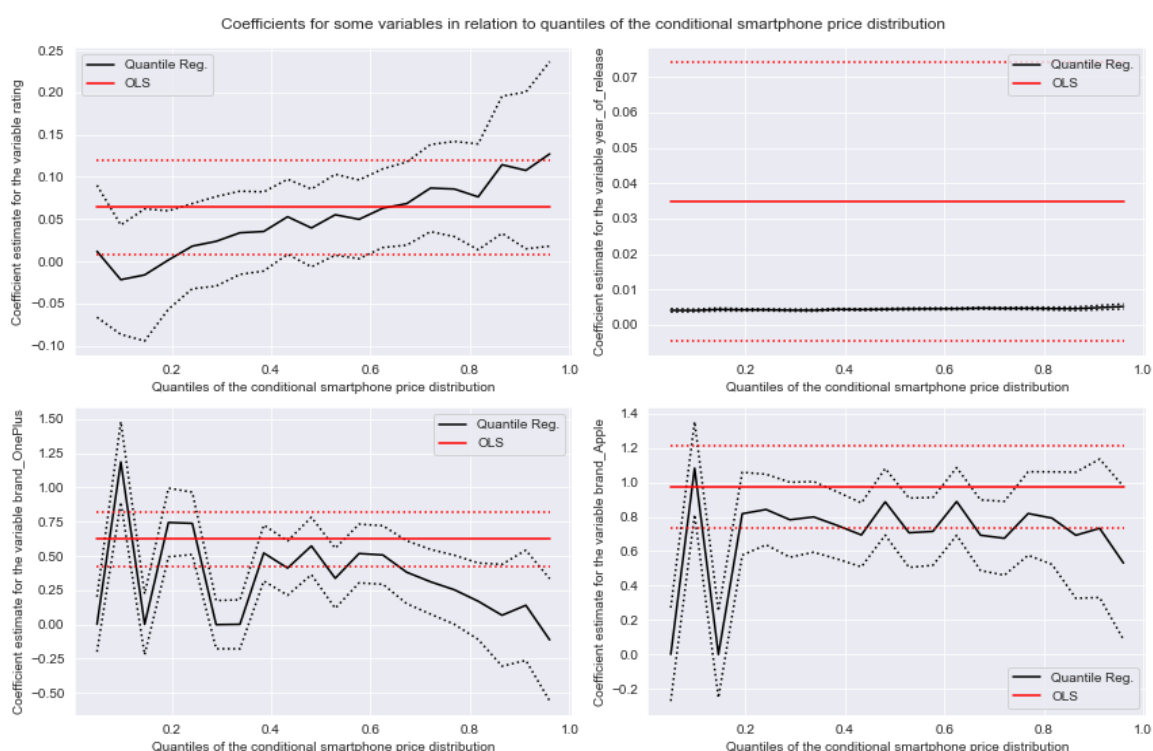
снизилось. Рейтинг смартфона также, как и в основной модели, является значимым на 5% уровне значимости.

Далее оценивалась квантильная регрессия для 20 квантилей уровней от 0.05 до 0.96 и рассматривалось, как будут меняться оценки коэффициентов перед некоторыми переменными, которые представляют наибольший интерес, в частности, RATING, DISP\_DIAG, BRAND\_SAMSUNG и Brand\_Apple. Результат представлен на графике зависимости оценки коэффициентов квантильной регрессии от оцениваемого квантиля на Рис. 9.

Можно заметить, что с ростом уровня оцениваемого квантиля растет и коэффициент при рейтинге смартфона, то есть рейтинг в большей степени определяет верхнюю границу цены, которая может быть на него установлена, нежели ограничивает цену снизу. Оценка коэффициента при годе выпуска практически не меняется с изменением уровня квантиля, причем МНК-оценка данного коэффициента существенно превышает оценку, полученную с помощью квантильной регрессии.

Оценка коэффициента при бренде OnePlus почти для всех квантилей ниже, чем соответствующая МНК-оценка, то же верно и для бренда Apple. Это можно объяснить тем, что квантильная регрессия менее чувствительна к выбросам, а в датасете присутствовали телефоны брендов Apple и OnePlus с крайне высокой ценой.

Рис. 9.  
График зависимости оценки коэффициентов квантильной регрессии от квантиля



## 8. Прогнозирование

Для прогнозирования было осуществлено предсказание логарифма цены смартфона для тех моделей, которых не было в изначальных данных. Так как данные были собраны с сайта маркетплейса Кении [3], то новые наблюдения были взяты там же. Для того, чтобы цены мобильных устройств находились на уровне 2021 года, в котором собирался исследуемый датасет, был использован веб-инструмент [4], позволяющий рассмотреть версию сайта 2021 года, для получения корректных новых данных. В частности, в качестве “тестовой” выборки были выбраны 3 смартфона: [Xiaomi Poco X3 6GB/64G](#), [OnePlus 9R 12GB/256GB](#), [Apple iPhone SE \(2020\) 256GB](#). Логарифмы цен для 3 представленных мобильных устройств были спрогнозированы 4 моделями: основная модель, квантильные регрессии условного квантиля логарифма цены смартфона уровня 0.1, 0.5 и 0.9 соответственно, результаты представлены в Таблице 5.

Таблица 5  
Результаты прогнозирования моделей

	Линейная регрессия	Квантильная регрессия Q=0.1	Квантильная регрессия Q=0.5	Квантильная регрессия Q=0.9	<i>LOG_PRICE</i>
Poco X3	10.025	9.325	10.123	10.783	10.146
OnePlus 9R	11.343	9.930	11.278	11.576	11.142
iPhone SE	11.294	9.753	11.155	11.998	11.184
MSE	0.022	1.398	0.007	0.041	
MAE	0.144	1.155	0.063	0.063	

Заметим, что медианная регрессия даёт прогнозы существенно лучше оптимальной линейной модели согласно MSE и MAE, это может быть связано с её устойчивостью к выбросам, которые сильно влияют на оценки линейной регрессии, с изменением распределения или небольшим количеством данных на тестовой выборке.

## 9. Заключение

Таким образом, в данном исследовании были выделены статистически значимые технические характеристики, влияющие на формирование цены смартфонов. К ним относится объём оперативной и внутренней памяти, а также разрешение камеры и матрицы дисплея. Количественное увеличение названных параметров ведёт к повышению стоимости мобильного устройства ввиду его качественного улучшения. Кроме того, была проанализирована статистическая значимость некоторых несвязанных с техническими характеристиками девайса регрессоров с целью проверки поставленных ранее гипотез. Применения наиболее оптимальной спецификации линейной регрессии позволило показать

значимое влияние рейтинга устройства и его бренда на его стоимость, не отвергая гипотезы H2 и H3. Однако год выпуска смартфона оказался статистически незначимым на 5% уровне значимости при использовании линейной регрессии, отвергая гипотезу H1. Модель линейной регрессии показала достаточно качественные результаты при прогнозировании на новых данных согласно MSE.

Дополнительно была оценена квантильная регрессия на разных уровнях. В рамках прогнозирования наиболее качественной оказалась медианная квантильная регрессия, её прогнозы оказались значительно лучше, чем у линейной модели, в обеих метриках: MSE и MAE. Помимо этого, по результатам медианной квантильной регрессии год выпуска смартфона оказался положительным и значимым на любом разумном уровне значимости, что подтверждает гипотезу H1.

Можно сделать вывод, что на цену смартфонов влияют не только технические характеристики, являющиеся показателем качества продукта, но и иные необъективные признаки, оказывающие воздействие на потребителей. Данное исследование подвержено ряду ограничений: относительно небольшое количество технических характеристик, отсутствие временного ряда для цен для отображения динамики изменения стоимости конкретной модели, недостаточное количество валидных инструментов для снижения влияния эндогенности, возможное наличие невыявленных источников эндогенности. Устранение этих ограничений – задача будущих исследований.

## 10. Список литературы

1. Kaggle: Data Science Community // URL: [https://www.kaggle.com/datasets/lyraxvinns/mobile-phones-specifications-and-prices-in-kenya?select=PhonesPriceInKenya\\_v2.csv](https://www.kaggle.com/datasets/lyraxvinns/mobile-phones-specifications-and-prices-in-kenya?select=PhonesPriceInKenya_v2.csv) (Дата обращения: 10.05.2023)
2. Lukasz Grzybowski, Ambre Nicolle (2021) Estimating Consumer Inertia In Repeated Choices Of Smartphones // The Journal Of Industrial Economics vol. LXIX
3. Price in Kenya: Electronics Store // URL: [www.priceinkenya.com/](http://www.priceinkenya.com/) (Дата обращения: 10.05.2023)
4. Wayback Machine: Internet Archive // URL: <http://web.archive.org/> (Дата обращения: 10.05.2023)
5. Берндт Э. Р. (2005) Практика эконометрики: классика и современность. // М.: ЮНИТИ-ДАНА - 863 с. Глава 4.