

● APPLIED
● ECONOMICS
● RESEARCH
● CENTRE
EVIDENCE BASED SOLUTION

Анализ ценообразования на автомобильном рынке Казахстана

Серия «Экономические исследования» №1-2020

Амина Бауыржан, Рустам Бекишев

Нур-Султан, 2020

Серия публикаций AERC (Центр исследований прикладной экономики) предназначена для распространения результатов научно-исследовательских работ работников AERC в целях стимулирования дискуссий в рамках соответствующих тем. Мнения, высказанные в документе, выражают личную позицию авторов и могут не совпадать с официальной позицией AERC.

Анализ ценообразования на автомобильном рынке Казахстана

Авторы: Бауыржан А.Б., Бекишев Р.А.

Февраль 2020 года

AERC-WP-2020-1

Дисклеймер

Представленный материал не является предложением или побуждением продавать или покупать какие-либо активы, в том числе определенные марки автомобилей. Его использование любыми лицами при принятии инвестиционных решений не влечет за собой какой-либо ответственности AERC за возможные потери или убытки, явившиеся следствием таких инвестиционных решений.

© AERC

Любое воспроизводство представленных материалов допускается только с разрешения авторов

Казахстан, г. Нур-Султан, шоссе Қорғалжын, 23/1, ВП-25
тел.: +7 7172 97 20 36

info@aerc.org.kz

www.aerc.org.kz

Анализ ценообразования на автомобильном рынке Казахстана

Амина Бауыржан¹
Рустам Бекишев²

Аннотация

В настоящем исследовании был проведен эмпирический поиск факторов, влияющих на цену автомобиля на рынке Казахстана. Данные по автомобилям были собраны с интернета при помощи технологии web-scraping. Итоговая модель гедонической регрессии включала 57205 наблюдений по 13 различным характеристикам: цена авто, год выпуска, тип кузова, марка авто, объем двигателя, пробег, вид коробки переключения передач, расположение руля, вид потребляемого топлива, регион, в котором продается машина, цвет кузова, а также информация о приводе и растаможке. В результате исследования выявлено: увеличение объема двигателя на каждые 0.1 литра повышает цену авто на 21%; авто с автоматической коробкой передач – на 28% дороже, чем с механической; передний привод – на 14% дороже, чем задний; черный цвет дает надбавку цене на 2.8%, а белый цвет – на 0.9%. Каждый год эксплуатации понижает цену на 7%, каждые дополнительные 100 000 километров пробега понижают цену на 2%, правостороннее расположение руля сокращает цену на 10%; цена нерастаможенных авто на 33% дешевле, чем растаможенных при одинаковых характеристиках авто.

Ключевые слова: гедоническая регрессия, ценообразование на автомобильном рынке, эконометрический анализ, web scraping, дисперсионный анализ (ANOVA), коэффициент инфляции дисперсии (VIF).

Классификация JEL: D41, C51, L62

¹ старший аналитик Департамента прикладных исследований AERC, a.bauyrzhan@aerc.org.kz

² старший аналитик Департамента прикладных исследований AERC, MSc.,
rustam.bekishev@aerc.org.kz

Содержание

Введение	5
Обзор литературы	6
Описание данных	7
Методология исследования	13
Обсуждение результатов	16
Заключение	19
Список использованных источников	20
Приложения	22

Введение

На сегодняшний день автомобиль является неотъемлемой частью жизни современного человека. Анализируя предпочтения населения в выборе авто, можно оценить не только уровень материального благосостояния отдельных физических лиц, но и всего общества. В связи с этим анализ факторов ценообразования на автомобильном рынке является актуальной темой для исследования.

В последние годы в Казахстане, особенно в крупных городах, остро стояла проблема огромного количества старых автомобилей на дорогах. Несмотря на рост отечественного производства, внедрение государством мероприятий по выкупу старых авто и облегчения условий для покупки новых автомобилей при помощи льготных программ, в Казахстане все еще сохраняется повышенный спрос на подержанные авто. Согласно данным Комитета по статистике Министерства национальной экономики РК по состоянию на 1 декабря 2019 г. количество зарегистрированных легковых автомобилей в РК составило 3 768.7 тыс. единиц, в то время как за период январь-ноябрь 2019 г. было зарегистрировано 806.8 тыс. легковых автомобилей. По информации Союза предприятий автомобильной отрасли Казахстана «КазАвтоПром» за аналогичный период времени официальными дилерами было реализовано 63 404 новых автомобилей, на приобретение которых казахстанцы потратили 588.8 млрд. тенге. Примерно 90% зарегистрированных автомобилей приходятся на авто вторичного рынка, также у около 20% автомобилей владельцы сменяются в течение года. Также стоит отметить, что 66% всех зарегистрированных автомобилей старше 10 лет. Так как вторичный автомобильный рынок занимает основную часть общего рынка, в исследовании будут рассмотрены факторы, определяющие рыночную цену на подержанные авто.

На самом деле факторы, влияющие на цену авто на первичном и вторичном рынках, заметно различны. В то время как цены на новые автомобили определенной марки не сильно отличаются у разных дилеров, цена на подержанное авто может сильно варьироваться в зависимости от его состояния, года выпуска, пробега, вида коробки, потребляемого топлива и даже цвета кузова. Из-за сильной дифференциации в ценах на авто на вторичном рынке процесс оценки подержанного авто является трудоемким процессом. Настоящее исследование нацелено на определение факторов ценообразования на вторичном автомобильном рынке Казахстана. Данный анализ поможет определить, какие факторы преобладают в формировании цены вторичного авто для казахстанцев.

Тема исследования может стать точкой преломления для анализа материального благосостояния населения, а также имеет прикладную научную значимость определения состояния казахстанского автомобильного рынка, результаты исследования, несомненно, станут отправной точкой для проведения аналогичных исследований. Более того, данный подход может быть применен в качестве инструмента для оценки подержанных автомобилей профессиональными оценщиками и банками второго уровня.

Структуру данного научного исследования можно представить в следующем виде: обзор литературы, описание данных, методология исследования, обсуждение результатов и заключение.

Обзор литературы

Целью данного исследования является оценка степени влияния разных характеристик авто на существующие цены подержанных автомобилей, используя методы гедонического ценообразования. В исследовании «Гедонический анализ цен продажи жилья с помощью полупараметрических методов» авторы прогнозируют цену жилья, используя гедоническую регрессию, основанную на сглаженном сплайн-сглаживании, и демонстрирует ее преимущество по сравнению с прогнозированием цены с помощью традиционных параметрических моделей. (Giudice et al., 2017)

Модель гедонического ценообразования используется во многих странах мира для анализа ценообразования на автомобильном рынке. Среди работ необходимо отметить следующие:

Erdem и Şentürk в своей работе «Гедонический анализ цен на подержанные автомобили в Турции» показали, что такие характеристики автомобилей, как наличие дизельного двигателя, цвет автомобиля, автоматическая коробка передач, люк на крыше, страна-производитель, а также год выпуска автомобиля, имеют значимое влияние на ценообразование (Erdem et al, 2009).

Примечательно, что в странах с высокими ценами на топливо, объем двигателя и вид потребляемого топлива влияет в значительной степени на цену вторичного авто. В исследовании «Does the Swiss Car Market Reward Fuel Efficient Cars? Evidence from Hedonic Pricing Regressions, a Regression Discontinuity Design, and Matching» было доказано, что экономия топлива капитализируется в цену автомобилей (Alberini et al, 2014).

Исследователи из Китая статистически проанализировали данные для прогнозирования цены автомобилей. С помощью 100 000 наблюдений о продаже автомобилей по всему Китаю был проведен эмпирический анализ с помощью сравнения двух алгоритмов: линейной регрессии и случайного леса. Данные двух алгоритмов были использованы для прогнозирования цены автомобилей по различным типам моделей (Chuanan et al., 2017; Kuiper, 2008).

Авторы из Индии, в свою очередь, также подчеркивают необходимость в эконометрической модели, которая может с высокой точностью определить цену для транспортного средства путем оценки его характеристик с учетом различных типов факторов. В своей статье исследователи используют статистическую модель случайного леса, чтобы предсказать цены на автомобили. В ходе проделанной работы создан случайный лес с 500 деревьями решений. По результатам эксперимента точность модели составила примерно 84% (Pal et al., 2018).

Несмотря на то, что, модель гедонического ценообразования используется во многих сферах, существует ограниченное количество исследований, нацеленных на анализ ценообразования на рынке подержанных авто.

После детального изучения трудов, основанных на гедоническом методе, было выявлено, что факторы, влияющие на цену, отличаются в разных исследованиях в зависимости от страны, в которой были собраны данные. В связи с этим особенно важно провести данный анализ для Казахстана с целью выявления факторов, определяющих цену на отечественном автомобильном рынке.

Описание данных

В рамках данного исследования были использованы данные, размещенные в интернете. Веб-сайты являются самой популярной платформой для размещения объявлений о продаже авто в Казахстане. На просторах сети покупатели могут сравнить более сотен тысяч объявлений, чтобы выбрать подходящее авто по таким параметрам, как марка, модель автомобиля, тип двигателя, привод, кузов, цвет, расположение руля, пробег и стоимость.

Для автоматизации сбора данных из интернета была использована технология web scraping. Данный инструмент позволяет за короткое время собрать всю необходимую информацию, размещенную на определенном веб-сайте, в единую базу данных. Эта технология позволяют автоматически извлекать данные и сохранять их для последующего использования посредством написания специального кода. В рамках данного исследования сбор данных производился на платформе R при помощи таких специализированных пакетов, как rvest и RCrawler.

При поиске объявлений в интернете сначала поисковик выдает ссылки на множество автомобилей, только пройдя по определенной ссылке, можно подробно прочитать информацию о каждом авто. Таким образом, сбор данных включал в себя сбор всех ссылок с объявлениями о продаже авто и скачивание информации с каждой ссылки. Стоит отметить, что структуры объявлений заметно отличаются друг от друга, так как не все продавцы указывают характеристики авто в едином порядке. Для структурирования собранных данных была использована технология Power Query.

База данных, используемая в данном исследовании, включает в себя объявления, актуальные на январь 2020 г. Первичная необработанная база содержала 128 574 объявления. Однако, не все наблюдения имеют полную информацию по всем показателям, так как указание всех характеристик автомобиля при размещении объявления является необязательным процессом. Таким образом, после обработки базы данных около половины наблюдений были отфильтрованы и в итоговый анализ были включены 66 130 наблюдений.

Следующим этапом было приведение цен на авто в логарифмическую форму. Далее из набора данных были исключены заметно выделяющиеся значения (выбросы) по таким параметрам, как год выпуска, пробег и стоимость авто. Чтобы выявить выбросы используется метод межквартильного расстояния.

Используемая выборка данных для настоящего исследования содержит 57 205 наблюдений, по которым имеется информация о 13 характеристиках: стоимость авто, год выпуска, тип кузова, марка авто, объем двигателя, пробег, вид коробки двигателя, расположение руля, вид потребляемого топлива, регион, в котором продается машина, цвет кузова, а также информация о приводе и растаможке.

Среди всех размещенных объявлений большая часть авто имеет следующие характеристики: седан белого цвета с объемом двигателя 1,6 марки Lada или Toyota 2005-го года выпуска со средней стоимостью 3,7 млн. тенге.

Таблица 1. Показатели стоимости авто в разрезе регионов

Регион	Медиана	Максимум	Минимум	тенге
				Среднее значение
Акмолинская	3 450 000	33 100 000	200 000	4 560 989
Актюбинская	2 000 000	30 000 000	180 000	3 037 407
Алматинская	3 600 000	33 000 000	200 000	4 941 864
Атырауская	2 600 000	30 000 000	250 000	3 348 995
ВКО	2 300 000	30 000 000	185 000	3 326 921
Жамбылская	1 800 000	32 700 000	200 000	2 792 591
ЗКО	2 100 000	33 000 000	200 000	3 059 770
Карагандинская	2 700 000	32 500 000	200 000	3 621 717
Костанайская	2 000 000	25 000 000	180 000	3 152 814
Кызылординская	1 600 000	25 000 000	250 000	2 633 635
Мангистауская	3 200 000	28 000 000	250 000	4 080 484
Павлодарская	2 100 000	29 990 000	220 000	3 251 524
СКО	2 235 000	28 500 000	180 000	3 183 335
ЮКО	1 900 000	33 300 000	200 000	3 003 184

В географическом разрезе база данных охватывает 14 регионов. 23% всех объявлений размещены в Алматинской области, далее располагается Акмолинская область (15%) и ЮКО (11%). Стоит отметить, что общее описание объявлений, цена и другие характеристики заметно различаются от региона к региону. Например, в ЮКО, Жамбылской и Кызылординской областях, где отмечаются самые низкие номинальные денежные доходы населения, средняя цена продаваемых автомобилей является самой низкой по всей республике.

Таблица 2. Описательная статистика показателей в разрезе регионов

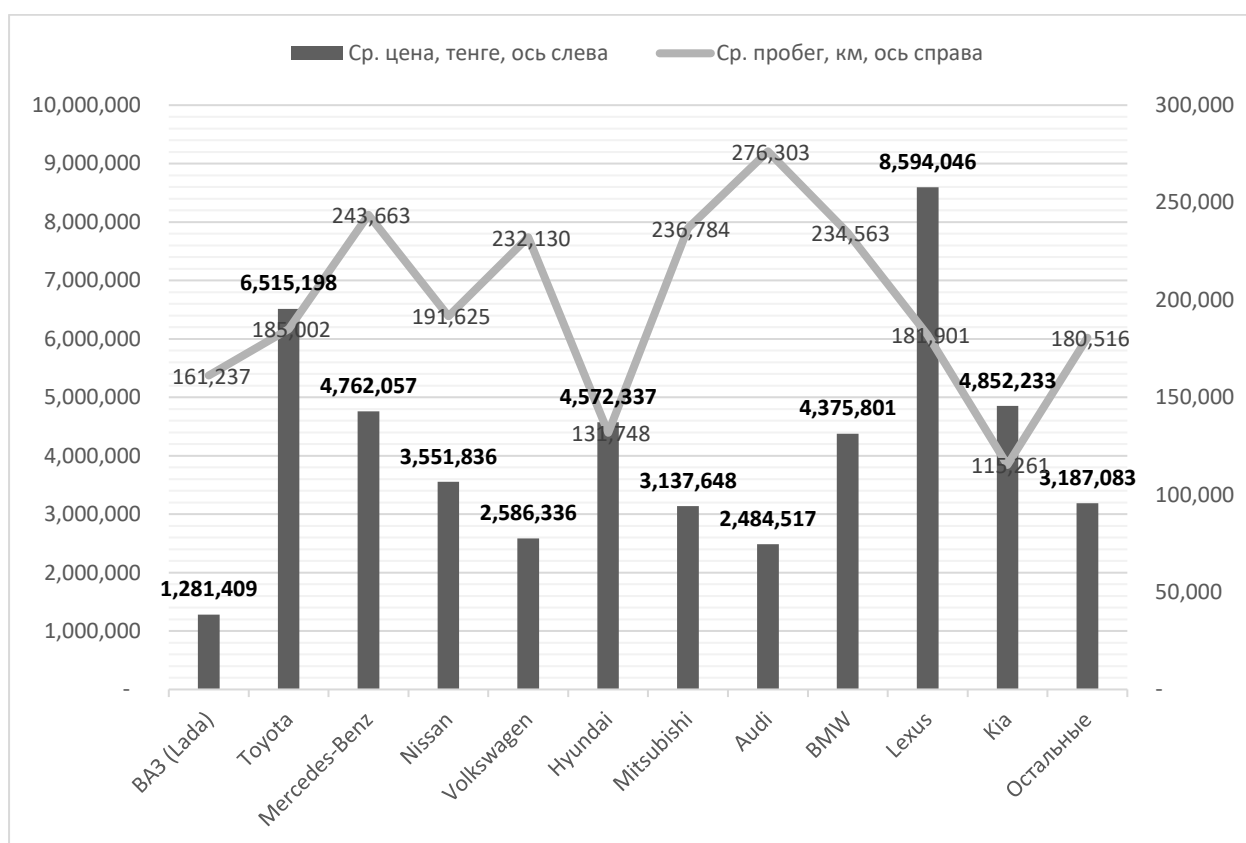
Регион	Кол-во машин, штук	Доля, %	Ср. возраст, лет	Ср. пробег, км	Нерастаможенные, %	Популярные марки
Алматинская	13 255	23%	15	196 877	5%	Toyota – 20,9%; Mercedes – 11%; Nissan – 7%; Mitsubishi – 6,3%
Акмолинская	8 739	15%	13	181 803	17%	Toyota – 18,94%; Lada – 14,5%; Mercedes – 7,3%
ЮКО	5 990	10%	16	203 786	14%	Lada – 23%; Toyota – 9,4%; Daewoo – 9%
Актюбинская	3 871	7%	13	175 463	42%	Lada – 31,5%; Toyota – 14%; Nissan – 6%
Атырауская	3 640	6%	12	160 124	45%	Lada – 16,4%; Toyota – 12,6%; Mercedes – 10,7%
Карагандинская	3 574	6%	15	199 032	13%	Lada – 28,3%; Toyota – 18,9%; Nissan – 6%
ВКО	3 113	5%	15	186 766	17%	Lada – 25%; Toyota – 19,2%; Nissan – 5,9%
Мангистауская	2 960	5%	13	183 839	24%	Toyota – 23,3%; Lada – 14,4%; Mercedes – 7,3%

ЗКО	2 935	5%	12	169 319	44%	Lada – 34,1%; Toyota – 10,3%; Volkswagen – 7%
Костанайская	2 685	5%	14	174 268	37%	Lada - 29,6%; Volkswagen – 9,1%; Toyota – 7,6%
Кызылординская	1 925	3%	16	199 418	30%	Lada – 25%; Mercedes – 11,4%; Toyota – 11,1%
Павлодарская	1 763	3%	16	205 092	24%	Audi – 13%; Mercedes – 12,5%; Volkswagen – 12%
Жамбылская	1 664	3%	20	252 738	12%	Lada – 22%; Toyota – 16,5%; Nissan – 7,1%
СКО	1 090	2%	15	182 021	36%	Lada – 21,4%; Toyota – 16%; Nissan – 7,4%

Самые старые автомобили продаются в Жамбылской области, где средний пробег авто составляет 252 738 км, что также является максимальным показателем по всей республике. Примечательно то, что больше всего нерастаможенных авто продаются в западных регионах страны: в Атырауской (45%), Западно-Казахстанской (44%) и Актюбинской (42%) областях. Также важно отметить факт того, что предпочтения по маркам авто заметно различаются в разрезе регионов, в то время как большинство все же отдает предпочтение автомобилям марки Lada и Toyota (соответственно, 17,6% и 16,8% по всей республике).

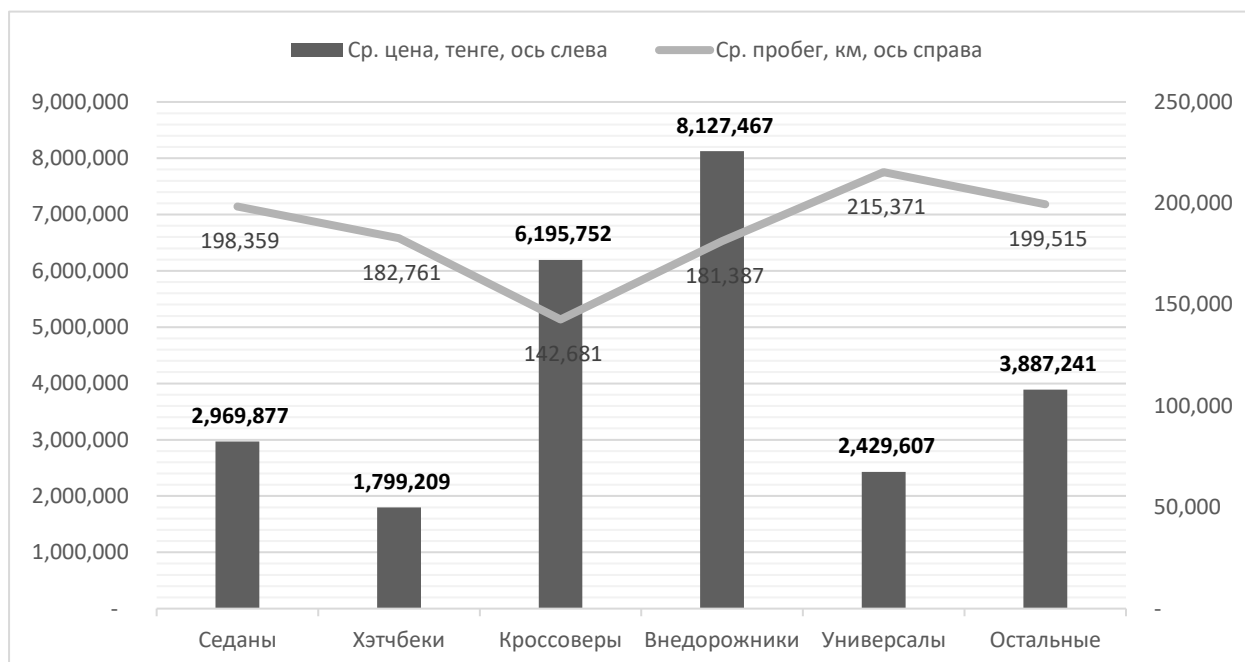
Таким образом, самыми популярными марками авто среди казахстанцев оказались Lada и Toyota, которые представляют более одной трети всего рынка. В тройке самых популярных авто также оказались автомобили марки Mercedes Benz (8% от общего количества размещенных объявлений).

Рисунок 1. Показатели цены авто в разрезе видов марок



Всего на сайте представлены автомобили 92 марок. В первой десятке самых популярных авто самыми дорогими являются представители марки Lexus, а самыми бюджетными оказались автомобили марки Lada.

Рисунок 2. Показатели цены авто в разрезе видов кузова



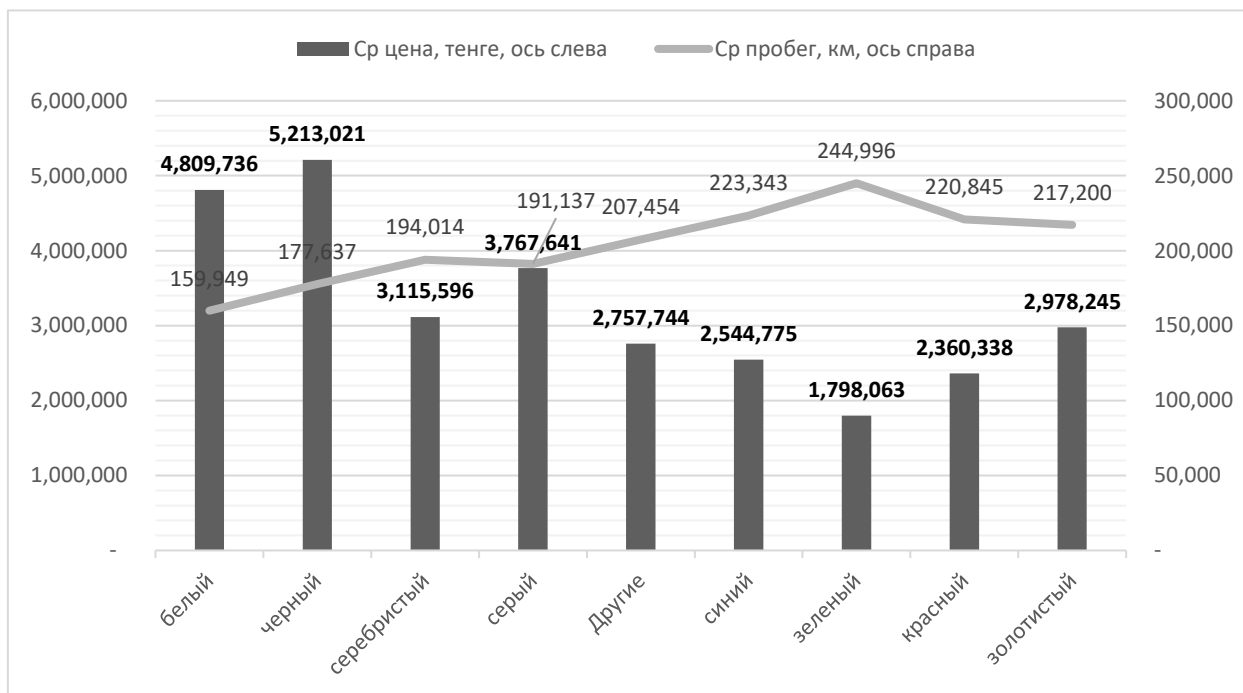
Почти половина (47%) хозяев объявлений нацелена на продажу авто с кузовом седан, средняя цена которых составляет 2,9 млн. тенге. Хэтчбеки (12%), внедорожники (11%) и кроссоверы (11%) оказались почти одинаково популярными видами авто на продажу. Самым дорогим типом кузова являются внедорожники, в среднем их хозяева оценивают свое авто в 8,1 млн. тенге. Примечательно, что у микроавтобусов самый высокий показатель по пробегу: в среднем каждый микроавтобус, выставленный на продажу, проехал 563 637 км и приблизительно 2004 г. выпуска.

Таблица 3. Описательная статистика показателей в разрезе видов кузова

Кузов	Кол-во машин, штук	Доля от общего, %	Средний возраст, лет	Доля нерастаможенных, %	Доля авто, потребляющих газ-бензин, %
Седаны	27 096	47%	15	23%	15%
Хэтчбеки	6 802	12%	14	33%	10%
Кроссоверы	6 499	11%	10	9%	17%
Внедорожник и	6 231	11%	13	6%	26%
Универсалы	3 901	7%	17	20%	19%
Остальные	6 675	12%	15	25%	25%

Каждая четвертая продаваемая машина имеет белый цвет кузова, и их средняя цена составляет 4,8 млн тенге. Дороже всего стоят авто черного цвета (18% от общего числа объявлений) и их средняя цена составляет примерно 5,21 млн. тенге.

Рисунок 3. Показатели цены авто в разрезе цвета кузова



Цена авто заметно дифференцируется в зависимости от разных характеристик. Однако, для подтверждения статистической значимости этих взаимосвязей необходимо проведение статистических тестов.

Методология исследования

Наиболее распространенным эконометрическим методом, который используется для изучения ценообразования, является гедонический анализ. Суть гедонического анализа заключается в исследовании взаимосвязи цен на определенные товары. Гедоническая гипотеза была сформулирована еще в конце XX века и состоит в предположении о том, что рыночная цена продукта зависит от ряда его характеристик (Rosen, 1974). Одной из основных сфер применения гедонического подхода с помощью гедонических регрессий является измерение точности ценовых индексов, в том числе инфляции [Griliches 1971; Silver 1995; Feenstra 1995; Diewert 2001; Silver, Heravi, 2001]. Многие исследования в данной сфере связаны с математическим построением гедонических функций с использованием регрессионных уравнений. Наиболее простой и удобной формой является линейная в связи с тем, что нелинейную зависимость очень сложно определить (Day et al., 2004). Эконометрический анализ позволяет исследователям не только подтвердить или отклонить гедоническую гипотезу, но и в случае ее подтверждения, выявить силу влияния той или иной характеристики на цены исследуемого продукта.

Данное исследование позволяет посмотреть на состояние ценообразования на вторичном автомобильном рынке Казахстана, оценить вклад других различных ценообразующих характеристик транспортного средства в формирование его цены. Для оценки степени влияния нескольких факторов на цену исследуемого продукта в качестве модели гедонической ценовой регрессии был использован стандартный OLS (МНК).

Предварительно был проведен тест на коллинеарность среди объясняющих факторов. Цена определяется как зависимая переменная и регрессирует на набор независимых переменных, которые, как считается, влияют на цену, основываясь на экономической теории, интуиции исследователя или исследованиях потребителей. Стоит заметить, что выборки наблюдений, применяемые в настоящей работе, являются перекрестными данными. К тому же часть объясняющих переменных, используемых в регрессии, является непрерывными: пробег, объем двигателя, возраст авто и часть являются категориальными: вид марки, вид кузова, цвет кузова, регион, тип потребляемого топлива, коробка переключения передач, расположение руля, привод и информация касательно растаможки. Таким образом, ценообразование можно представить в виде лог-линейной функции от наблюдаемых характеристик авто и ошибки.

$$\ln(P_i) = \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i$$

где P_i – это цена автомобиля, x_j – характеристика транспортного средства, β_j – статистическая оценка коэффициента, соответствующего характеристике транспортного средства, количество наблюдений $i \in [1; 57205]$, ε_i – отклонение наблюдаемого значения от истинного значения переменной.

Мультиколлинеарность является довольно распространенной проблемой, и она возникает, когда при моделировании множественной регрессии две или более объясняющих переменных приблизительно линейно связаны, что может привести

к серьезным проблемам оценок регрессии, оценки регрессии могут быть статистически незначимыми, в то время как модель будет иметь высокие значения R-квадрата (R-squared) и F-статистики.

Таким образом, необходимо выявить коллинеарность при помощи специальных методов. Коллинеарность может быть обнаружена следующими способами:

1) Выявление корреляции между каждой парой объясняющих переменных (попарная корреляция). Если две переменные сильно коррелированы, то это может быть возможным источником мультиколлинеарности.

Формула парной корреляции Пирсона выражается в следующем виде:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

где \bar{x} и \bar{y} являются средними значениями соответствующих переменных x и y

Высокое значение корреляции показывает, что переменные имеют сильную зависимость (Benesty et al., 2009).

2) Использование VIF (Variance Inflation Factor). Чем больше значение данного критерия, тем более тот или иной регрессор коллинеарен с остальными. Стандартным значением VIF-критерия, который является критическим, примерно должно превышать 10.

Коэффициент инфляции дисперсии (VIF) количественно определяет степень корреляции между одним предиктором и другими предикторами в модели.

Степень, в которой предиктор коррелирует с другими переменными в регрессии, может быть определена количественно при помощи меры R-квадрат (R-squared) статистики регрессии, где переменная предсказывается всеми другими переменными модели (Akinwande et al., 2015).

Формула коэффициента инфляции дисперсии:

$$VIF_j = \frac{1}{1 - R_j^2}$$

где R_j^2 показывается коэффициент детерминации, когда в качестве зависимой переменной выступает регрессор.

На следующем этапе для изучения влияния нескольких факторов, объясняющих зависимую переменную, был проведен дисперсионный анализ ANOVA test. Суть этого теста заключается в определении статистически значимых связей между переменными. В результате дисперсионного анализа отобран ряд переменных, которые являются статистически значимыми при объяснении зависимой переменной.

Согласно теории, уравнение множественной регрессии значимо, если значение F:

$$F = \frac{(RSS(\beta_q) - RSS(\beta_p))/(p - q)}{RSS(\beta_p)/(n - p)}$$

где RSS – сумма квадратов отклонений, n – число наблюдений в модели, p и q количество оцениваемых параметров больше критического значения критерия Фишера $F_{\alpha;p-q,n-p}$. В противном случае принимается нулевая гипотеза о незначимости параметров.

Обсуждение результатов

В результате проверки на коллинеарность между непрерывными переменными не было обнаружено мультиколлинеарных взаимосвязей. Изначально было сделано предположение, что такие показатели, как возраст авто и пробег, будут связаны между собой, однако эта гипотеза не была подтверждена. Соответственно, в итоговую модель были включены все непрерывные переменные.

Таблица 1. Корреляционная таблица для непрерывных переменных

	Объем двигателя	Пробег	Возраст авто
Объем двигателя	1.00	0.03	0.03
Пробег	0.03	1.00	0.52
Возраст авто	0.03	0.52	1.00

Результаты проверки на мультиколлинеарность VIF (Variance Inflation Factor) представлены в нижеследующей таблице. Согласно полученным результатам, никакие переменные не коррелируют между собой, так как показатель в пределах допустимых значений.

Таблица 2. Результаты VIF (Variance Inflation Factor) теста для переменных, используемых в регрессионной модели

Показатели	GVIF	Df	GVIF ^{1/(2*Df)}
Объем двигателя	2.32	1	1.52
Пробег	1.41	1	1.19
Возраст авто	2.21	1	1.49
Коробка двигателя	2.40	5	1.09
Расположение руля	1.40	1	1.18
Привод	4.92	2	1.49
Растаможен	1.31	1	1.14
Потребляемое топливо	1.59	4	1.06
Регион	1.64	13	1.02
Вид кузова	4.79	5	1.17
Вид марки	6.97	11	1.09
Цвет кузова	1.35	8	1.02

В связи с тем, что не были обнаружены мультиколлинеарные взаимосвязи, все показатели были включены в итоговую модель. Для того, чтобы проверить целесообразность включения каждого параметра в итоговую модель, был проведен дисперсионный анализ ANOVA (Analysis of Variance). Данный тест позволяет определить статистическую взаимосвязь между показателями и зависимой переменной. Результаты теста приведены в следующей таблице.

Таблица 3. Результат проведения ANOVA (Analysis of Variance) теста для объясняющих переменных регрессионной модели

Показатели	Df	Sum Sq	Mean Sq	F value	Pr(>F)
------------	----	--------	---------	---------	--------

Объем двигателя	1	14121.30	14121.30	147619.16	< 2.2e-16***
Пробег	1	4896.80	4896.80	51189.79	< 2.2e-16***
Возраст авто	1	11250.90	11250.90	117612.59	< 2.2e-16***
Трансмиссия	5	4305.40	861.10	9001.44	< 2.2e-16***
Расположение руля	1	60.40	60.40	631.05	< 2.2e-16***
Привод	2	476.80	238.40	2492.12	< 2.2e-16***
Раस्ताможен	1	2282.70	2282.70	23863.02	< 2.2e-16***
Потребляемое топливо	4	274.80	68.70	718.12	< 2.2e-16***
Регион	13	96.80	7.40	77.86	< 2.2e-16***
Вид кузова	5	557.50	111.50	1165.66	< 2.2e-16***
Вид марки	11	3775.30	343.20	3587.81	< 2.2e-16***
Цвет кузова	8	14.40	1.80	18.87	< 2.2e-16***
Residuals	57150	5467.00	0.10		
Уровень значимости: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Как видно из таблицы 3, все показатели оказались статистически значимыми на высоком уровне в определении цены авто. Соответственно, в итоговую модель были включены все данные переменные для определения степени влияния уровней каждого показателя в модели ценообразования.

В исследуемой гедонической модели контрольная группа представляет собой растаможенные автомобили с механической коробкой передач и обладает такими характеристиками, как задний привод, двигатель, потребляющий бензин, левостороннее расположение руля и из Карагандинской области. Причиной, по которой данные характеристики были выбраны в качестве контрольной, является то, что данная группа представляет собой среднерыночную цену авто на казахстанском рынке.

Согласно полученным результатам гедонической модели ценообразования (Приложение 1), такие факторы, как пробег и возраст, отрицательно влияют на цену автомобиля. К примеру, в среднем продавцы понижают цену на 7% от среднерыночной за каждый год подержанного авто. Более того, каждые 100 000 км пробега понижают цену на 2%. Напротив, чем больше объем двигателя, тем выше средняя цена автомобиля.

Левостороннее расположения руля и тип двигателя «механика» также снижают цену автомобилей. Автомобили с трансмиссией «автомат» на 28% дороже, чем с трансмиссией «механика». Правостороннее расположение руля понижает цену на 10%. Другими словами, одинаковые авто по всем параметрам, кроме расположения руля, будут в среднем иметь разницу в цене в размере 10%.

Также статистически было доказано, что в среднем нерастаможенные автомобили стоят на 33% дешевле, чем растаможенные при одинаковых других характеристиках авто.

Было выявлено, что регион, в котором размещается объявление, лишь незначительно влияет на цену авто, так как покупатели могут сравнивать предложения из разных регионов, что приводит к общему равновесию на рынке. Тем не менее, дороже всего за автомобили просят продавцы Алматинской области, что может быть обусловлено высоким спросом в данном регионе. Одну и ту же комплектацию автомобиля, которую продают в Алматинской области, можно купить в Актюбинском регионе на 8% дешевле.

Примечательно то, что марки авто значительно влияют на цену автомобилей. Например, марка Lexus добавляет 52% при прочих равных характеристиках авто. Марка автомобиля Lada стоит на 50% дешевле. Цена седана марки Lexus, BMW, Mercedes Benz, Toyota с трансмиссией «автомат», левосторонним расположением руля и задним приводом одинакового года выпуска и имеющая идентичный пробег и объем двигателя, будет в разы дороже аналогичного авто марки Hyundai, Kia, Mitsubishi или Lada.

Значимость цвета кузова в определении цены авто также была статистически доказана. Несмотря на то, что цвет кузова не так сильно изменяет цену авто, этот показатель все же является важным критерием ценообразования. К примеру, в среднем черный цвет авто дает надбавку к цене в размере 2.8%, а белый цвет кузова – в размере 0.9%. Наименее предпочтительным цветом, который дает наибольшую скидку в размере 6.5%, является золотистый цвет кузова.

Заклучение

Проведенный анализ на основе собранных данных помог выявить факторы, влияющие на ценообразование на казахстанском вторичном автомобильном рынке. Были выявлены характеристики авто, за которые казахстанцы готовы переплачивать. Большую роль в ценообразовании играют такие факторы, как вид марки, кузова и коробка двигателя. Это может объясняться тем, что определенные виды кузова и марки стоят дороже и на первичном рынке. К примеру, машины марки Lexus стоят на 52% дороже среднерыночной цены, в то время как автомобили Lada оцениваются на 50% дешевле. Наличие коробки переключения передач «механика» на 27% снижает цену авто по сравнению с трансмиссией «автомат». Самую большую надбавку к цене среди всех видов кузова дают внедорожники, которые в среднем стоят на 4.6% дороже. Примечательным является тот факт, что каждый 0.1 литр в объеме двигателя увеличивает цену авто на 2.1%.

Автомобиль с левым расположением руля в среднем на 10% стоит дороже, чем идентичное авто с правосторонним расположением руля. Более того, автовладельцы оценивают авто с передним приводом на 5.6% дороже, чем с задним приводом.

Стоит отметить, что дешевле всего автовладельцы оценивают свой транспорт в западных регионах страны. К примеру, идентичное авто по всем характеристикам в Актюбинской области продают на 7.5% дешевле по сравнению с продавцами авто в Карагандинской области. Самую большую надбавку к цене среди всех цветов дает черный цвет кузова в размере 2.8%.

Полученные результаты гедонической модели позволяют рассчитать примерную цену автомобиля, который имеет определенные характеристики. Данное исследование позволяет создать калькулятор для вычисления среднерыночной цены авто с любыми заданными характеристиками. Соответственно, данный подход может быть использован при прогнозировании цен на авто и является полезным инструментом при оценке авто разными организациями (профессиональные оценщики, банки).

Впоследствии, сбор данных может производиться систематически, что позволит строить эконометрические модели на основе временных рядов для выявления других объясняющих факторов в ценообразовании на автомобильном рынке Казахстана. В будущем, при систематизации сбора данных, возможно будет прогнозировать состояние автомобильного рынка Республики Казахстан.

Список литературы и использованных источников

1. Giudice V., Manganelli B., Paola P. (2017) // Hedonic Analysis of Housing Sales Prices with Semiparametric Methods // International Journal of Agricultural and Environmental Information Systems. 8. 65-77. 10.4018/IJAEIS.2017040105.
2. Cumhuri E., Şentürk I. (2009). // A Hedonic Analysis of Used Car Prices in Turkey // International Journal of Economic Perspectives. 3. 141-149.
3. Alberini A., Bareit M., Filippini M. (2014) // Does the Swiss Car Market Reward Fuel Efficient Cars? Evidence from Hedonic Pricing Regressions, a Regression Discontinuity Design, and Matching // SSRN Electronic Journal. 10.2139/ssrn.2403696.
4. Shonda K. (2008) // Introduction to Multiple Regression: How Much Is Your Car Worth?. J Stat Educ. 16. 10.1080/10691898.2008.11889579.
5. Chen C., Hao L., Xu C. (2017) // Comparative analysis of used car price evaluation models // AIP Conference Proceedings. 1839. 020165. 10.1063/1.4982530.
6. Pal N., Kohli P., Sundararaman D., Palakurthy S. (2019) // How Much Is My Car Worth? A Methodology for Predicting Used Cars' Prices Using Random Forest: Proceedings of the 2018 // Future of Information and Communication Conference (FICC), Vol. 1. 10.1007/978-3-030-03402-3_28.
7. Rosen S. (1974) // Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition // Journal of Political Economy, 82(1), 34-55.
8. Day B., Ian B., Iain L. (2004) // Nonlinearity in hedonic price equation: An estimation strategy using model-based clustering // No. 04-02. CSERGE Working Paper EDM.
9. Akinwande M., Hussaini G., Agboola S. (2015) // "Variance inflation factor: as a condition for the inclusion of suppressor variable (s) in regression analysis." // Open Journal of Statistics 5, no. 07: 754.
10. Benesty J., Jingdong C., Yiteng H., Israel C. (2009) // "Pearson correlation coefficient." In Noise reduction in speech processing, pp. 1-4 // Springer, Berlin, Heidelberg.
11. Diewert E. (2001) // Hedonic Regressions: A Consumer Theory Approach // Paper presented at the Sixth Meeting of the International Working Group on Price Indices, Canberra, Australia, 2-6 April 2001.
12. Griliches Z. (ed.) (1971) // Price Indexes and Quality Change: Studies in New Methods of Measurement, Cambridge, Harvard University Press
13. Silver M. (1999) // An evaluation of the use of hedonic regressions for basic components of consumer price indices // Review of Income and Wealth Series 45, Number 1, March 1999
14. Silver M., Heravi S. (2000) // The Measurement of Quality-Adjusted Price Changes // paper presented at the NBER Conference on Scanner Data and Price Indexes, September 15-16, 2000 at Arlington, Virginia, forthcoming in an CRIW-NBER Volume edited by R.Feenstra and M. Shapiro.
15. Комитет по статистике, МНЭ РК, Статистика транспорта "О количестве легковых автомобилей в Республике Казахстан", 1 декабря 2019г.

16. КазАвтоПром, Союз предприятий автомобильной отрасли Казахстана,
пресс-релиз от 12 декабря 2019г.

Результат проведенной гедонической регрессионной модели

Категории	Уровни	Оценка	Стандартная ошибка	t - статистика	Pr(> t)
Константа		15.16	1.068e-02	1418.438	< 2e-16***
Объем двигателя		0.2152	2.146e-03	100.286	< 2e-16***
Пробег		-0.0000002	1.106e-08	-21.342	< 2e-16***
Возраст авто		-0.06786	2.459e-04	-275.918	< 2e-16***
Коробка двигателя	автомат	0.2755	3.889e-03	70.857	< 2e-16***
	АКПП	0.1995	2.189e-01	0.912	0.361962
	вариатор	0.2686	9.773e-03	27.480	< 2e-16***
	робот	0.3348	1.602e-02	20.893	< 2e-16***
	типтроник	0.2714	5.197e-03	52.228	< 2e-16***
Расположение руля	справа	-0.1011	6.252e-03	-16.168	< 2e-16***
Привод	передний привод	0.05614	4.942e-03	11.359	< 2e-16***
	полный привод	0.1446	6.040e-03	23.941	< 2e-16***
Растоможен	не растоможен	-0.3295	3.652e-03	-90.248	< 2e-16***
Потребляемое топливо	газ	-0.09232	2.339e-02	-3.946	7.96e-05***
	газ-бензин	-0.006899	3.859e-03	-1.788	0.073828 .
	гибрид	0.006699	2.788e-02	0.240	0.810117
	дизель	0.09618	6.318e-03	15.222	< 2e-16***
Регион	Акмолинская	-0.004258	6.170e-03	-0.690	0.490058
	Актюбинская	-0.07493	7.358e-03	-10.183	< 2e-16***
	Алматинская	0.008355	5.883e-03	1.420	0.155504
	Атырауская	-0.06377	7.416e-03	-8.598	< 2e-16***
	ВКО	0.004227	7.635e-03	0.554	0.579791
	Жамбылская	-0.01874	9.228e-03	-2.031	0.042293**
	ЗКО	-0.0543	7.826e-03	-6.938	4.03e-12***
	Костанайская	-0.02682	7.970e-03	-3.365	0.000766***

Категории	Уровни	Оценка	Стандартная ошибка	t - статистика	Pr(> t)
	Кызылординская	-0.06368	8.812e-03	-7.227	5.01e-13***
	Мангистауская	-0.06504	8.026e-03	-8.104	5.43e-16***
	Павлодарская	0.003787	9.025e-03	0.420	0.674754
	СКО	-0.01466	1.075e-02	-1.364	0.172678
	ЮКО	-0.02436	6.583e-03	-3.701	0.000215***
Вид кузова	Внедорожники	0.004596	6.723e-03	0.684	0.494229
	Кроссоверы	-0.03774	6.461e-03	-5.841	5.22e-09***
	Седаны	-0.2465	4.881e-03	-50.503	< 2e-16***
	Универсалы	-0.03459	6.588e-03	-5.250	1.53e-07***
	Хэтчбеки	-0.2438	6.133e-03	-39.756	< 2e-16***
Вид марки	Audi	0.4034	7.894e-03	51.106	< 2e-16***
	BMW	0.5009	8.252e-03	60.702	< 2e-16***
	Hyundai	0.2042	7.030e-03	29.047	< 2e-16***
	Kia	0.209	7.923e-03	26.374	< 2e-16***
	Lexus	0.5225	8.162e-03	64.019	< 2e-16***
	Mercedes-Benz	0.5337	6.599e-03	80.873	< 2e-16***
	Mitsubishi	0.1792	7.391e-03	24.240	< 2e-16***
	Nissan	0.1914	6.287e-03	30.447	< 2e-16***
	Toyota	0.5274	4.480e-03	117.730	< 2e-16***
	Volkswagen	0.2814	6.331e-03	44.447	< 2e-16***
	BA3 (Lada)	-0.5017	4.656e-03	-107.754	< 2e-16***
Цвет кузова	белый	0.009479	5.063e-03	1.872	0.061196 .
	зеленый	-0.00923	6.969e-03	-1.324	0.185396
	золотистый	-0.06489	1.024e-02	-6.338	2.34e-10***
	красный	-0.005372	9.487e-03	-0.566	0.571236
	серебристый	-0.005124	5.170e-03	-0.991	0.321560
	серый	-0.01124	5.816e-03	-1.933	0.053200 .
	синий	0.01233	6.190e-03	1.992	0.046421*
	черный	0.02816	5.255e-03	5.359	8.39e-08***

Категории	Уровни	Оценка	Стандартная ошибка	t - статистика	Pr(> t)
	Уровень значимости: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
	Стандартная ошибка остатков: 0.3093 на 57150 степеней свободы				
	R-квадрат: 0.8851, Скорректированный R-квадрат: 0.885				
	F-статистика: 8306 на 53 и 57150 DF, p-значение: < 2.2e-16				