

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

Гордеев Д.С., Косухина Е.А.

**Исследование влияния цен на нефтепродукты на их
потребление в автомобильном транспорте**

Москва 2020

Аннотация. Цель работы заключается в разработке методологии, необходимой для проведения количественной оценки эластичности потребления нефтепродуктов в легковом автомобильном транспорте в зависимости от цен и дохода индивидуумов. На основании полученных результатов предлагаются рекомендации, направленные на ускорение экономического роста в РФ.

Abstract. The aim of the study is to develop methodology to evaluate influence of prices of gasoline fuel and income on fuel consumption of passenger cars. Estimated results will be used for development of recommendations needed for economic growth of Russian Federation.

Гордеев Д.С., старший научный сотрудник лаборатории системного анализа отраслевых рынков ИОРИ Российской академии народного хозяйства и государственной службы при Президенте РФ;

Косухина Е.А., младший научный сотрудник лаборатории системного анализа отраслевых рынков ИОРИ Российской академии народного хозяйства и государственной службы при Президенте РФ.

Данная работа подготовлена на основе материалов научно-исследовательской работы, выполненной в соответствии с Государственным заданием РАНХиГС при Президенте Российской Федерации на 2019 год.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	4
1 РАЗРАБОТКА МОДЕЛИ СПРОСА НА НЕФТЕПРОДУКТЫ АВТОМОБИЛЬНЫМ ВИДОМ ТРАНСПОРТА.....	6
2 ПРОВЕДЕНИЕ КОЛИЧЕСТВЕННОЙ ОЦЕНКИ ЭЛАСТИЧНОСТИ СПРОСА НА НЕФТЕПРОДУКТЫ АВТОМОБИЛЬНЫМ ВИДОМ ТРАНСПОРТА В ЗАВИСИМОСТИ ОТ ЦЕН НА НЕФТЕПРОДУКТЫ.....	16
3 РАЗРАБОТКА РЕКОМЕНДАЦИЙ ПО ЭКОНОМИЧЕСКОЙ ПОЛИТИКЕ НА ОСНОВАНИИ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ.....	56
ЗАКЛЮЧЕНИЕ.....	59
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	61

ВВЕДЕНИЕ

Транспортная отрасль является одной из системообразующих отраслей российской экономики. Ее состояние оказывает прямое влияние на мобильность населения и грузовые перевозки, что особенно важно для крупнейшего государства в мире. Наибольшее количество пассажирских перевозок внутри Российской Федерации осуществляется за счет автомобильного транспорта, как общественного, так и личного. Для обеспечения бесперебойного функционирования данных видов перевозок необходимо обеспечивать достаточные объемы нефтепродуктов, для чего требуется механизм, позволяющий прогнозировать их потребление. Помимо энергетической безопасности, с экономической точки зрения, потребление нефтепродуктов обеспечивает поступления в региональные и федеральный бюджеты за счет акцизных платежей. Увеличение ставки акцизов на нефтепродукты с целью компенсации сокращения поступлений в бюджет, которое проводилось Правительством в последние годы, позволяет в краткосрочном периоде увеличить поступления, так как нефтепродукты являются товаром первой необходимости с неэластичным спросом, однако для долгосрочного периода суммарный эффект для поступлений не очевиден. Потенциальное сокращение потребления нефтепродуктов в связи с ростом цен на них в долгосрочном периоде может быть эластичным, что приведет к сокращению налоговых поступлений в бюджет. Согласно исследованиям, сокращение потребления нефтепродуктов обусловлено несколькими эффектами:

- Эффектом увеличения топливной эффективности транспортных средств;
- Эффектом снижения мобильности населения;
- Эффектом сокращения количества личных транспортных средств и ростом поездок на общественном транспорте.

Эффект увеличения топливной эффективности может заключаться в нескольких аспектах. Первый состоит в том, что с научно-техническим прогрессом происходит улучшение характеристик традиционных двигателей внутреннего сгорания. Учитывая тот факт, что автомобили отличаются длительным сроком службы, в большинстве своем новые автомобили энергоэффективнее тех, чей возраст составляет несколько лет и более. С увеличением цены часть потребителей может предпочесть приобрести новый вид автомобиля, что сократит их затраты на топливо при том же количестве пройденных километров. С другой стороны, потребители могут предпочесть пересесть с больших легковых автомобилей на более компактные, в которых топливная эффективность выше. Еще одна опция заключается в применении альтернативного вида топлива, например, переходе с бензинового автомобиля на дизельный, который считается более энергоэффективным.

Эффект снижения мобильности населения заключается в сокращении количества поездок на личном транспорте. При подорожании топлива расходы на поездки будут увеличиваться, что окажет влияние как на дальность поездок, так и на их частоту.

Отдельная группа населения переориентируется с использования личного транспорта в пользу общественного. В данном случае для той группы лиц, у которой было транспортное средство, количество поездок на нем сократится. Часть тех индивидов, которые рассматривали приобретение транспортного средства для личного пользования, откажется от покупки и будет использовать для перемещения общественный транспорт. Это приведет к сокращению приобретения новых транспортных средств.

Понимание степени влияния изменения цен на данные факторы позволит более качественно подойти к вопросу развития транспортного сектора экономики. Благодаря оценкам эластичностей данных эффектов можно будет прогнозировать поступления в региональный и федеральный бюджеты. Полученные оценки замещения пассажирских перевозок с личного транспорта в пользу общественного позволят прогнозировать пассажиропотоки, как в общественном городском транспорте, так и в междугороднем, что позволит заблаговременно реагировать на потенциальные «узкие места» и увеличивать количество транспортных средств на маршрутах. Более точный прогноз объемов приобретения новых транспортных средств позволит сократить затраты производителей, связанные с хранением невостребованных транспортных средств в складских помещениях.

1 Разработка модели спроса на нефтепродукты автомобильным видом транспорта

Принимая во внимание анализ эмпирических работ, посвященных исследованию спроса на нефтепродукты автомобильным видом транспорта, в данной главе будет представлена логика эконометрических моделей, необходимых для проведения количественной оценки эластичности спроса на бензин в Российской Федерации, а также будут описаны особенности спецификаций этих моделей.

Для целей анализа мы будем в качестве базовой методологии использовать модель, которая была впервые представлена в работах Sweeney (1978) [CITATION Swe78 \l 1049] и Baltagi и Griffin (1983, 1997) [CITATION Заполнитель2 \l 1049], [CITATION Bal97 \l 1049] и широко применяется в данный момент для исследований, связанных с потреблением нефтепродуктов в транспортном секторе экономики. Согласно данной модели совокупное потребление топлива в стране или регионе (зависит от уровня агрегированности данных) рассчитывается на основе средней интенсивности использования автомобиля (среднего пробега), его средней топливной эффективности (в нашем случае расхода топлива на единицу пройденного расстояния) и общего количества используемых автомобилей (формула (1)):

$$Потребление_{бензина} = \frac{Кол - во_{км}}{Кол - во_{авто}} \cdot \frac{Потребление_{бензина}}{Кол - во_{км}} \cdot Кол - во_{авто} \quad (1)$$

Поскольку из упомянутых переменных наблюдаемой является только количество автомобилей, то для моделирования остальных переменных (ненаблюдаемых) необходимы прокси-переменные, которые можно было бы использовать в модели. Как правило, в качестве прокси-переменной для замены переменной среднего пробега в расчете на один автомобиль используется располагаемый доход на душу населения или ВВП на душу населения, а для замены переменной средней топливной эффективности – цена на бензин. Ввиду того, что во многих исследованиях было продемонстрировано, что при увеличении количества транспортных средств в расчете на одно домохозяйство происходит сокращение удельного потребления топлива в расчете на один автомобиль, количество автомобилей как отдельная переменная не является оптимальной. Следовательно, этот факт необходимо учитывать и вместо общего количества автомобилей в экономике использовать переменную количества автомобилей на одного человека. Распространенной практикой является применение авторами в своих исследованиях такой переменной, как

количество транспортных средств на количество водителей, или на количество жителей в возрасте от 18 до 70 лет, или на одно домохозяйство.

После выбора прокси-переменных, отражающих интенсивность использования автомобиля и его топливную эффективность, получаем следующее уравнение спроса на бензин (формула (2)):

$$(GAS)^i = \alpha (Y)^\beta (PG)^\gamma (CAR)^\delta, \quad (2)$$

где

GAS – среднее потребление бензина в расчете на один легковой автомобиль;

Y – реальный валовый региональный продукт в расчете на душу населения;

PG – реальная цена на бензин;

CAR – количество легковых автомобилей на душу населения.

Таким образом, оценка спроса на бензин в расчете на один автомобиль основывается на реальном валовом продукте на душу населения (Y), реальной цене бензина (PG) и количестве легковых автомобилей на душу населения (CAR).

Поскольку для обозначения одних и тех же переменных в разных целях используются разные сокращения, для обеспечения дальнейшей согласованности в рамках проводимого анализа, в таблице представлено соответствие этих обозначений и их смысловой нагрузки.

Таблица 1 – Сокращенные обозначения переменных модели спроса на бензин с непосредственным учетом количества транспортных средств и их расшифровка

Смысл переменной	Обозначение в формулах*	Обозначение в выводе результатов статистических операций*	Единицы измерения
среднее потребление бензина в расчете на один легковой автомобиль	GAS	gaspercar	Тонны
реальный валовый региональный продукт в расчете на душу населения	Y	vgr	рубли
реальная цена на бензин	PG	gasprice	рубли
количество легковых автомобилей на душу населения	CAR	carperdriver	штуки
лаг логарифма среднего потребления бензина в расчете на один легковой автомобиль	GAS_{t-1}	lag1	
* - при взятии логарифма перед названием соответствующей переменной появляется "ln"			

В таблице представлен предполагаемый характер влияния на среднее потребление бензина в расчете на один легковой автомобиль выбранных переменных. Кроме того, в данной таблице выдвинуты гипотезы относительно механизмов такого влияния.

Таблица 2 – Гипотезы о знаках при оценках коэффициентов уравнения регрессии и их обоснование

Объясняющая переменная	Характер влияния на зависимую переменную	Механизм влияния
Y	+	Более высокий реальный региональный продукт на душу населения повышает интенсивность использования автомобилей в результате большей экономической активности
Y	-	на индивидуальном уровне увеличение располагаемого дохода ведет к увеличению количества автомобилей на душу населения
CAR	-	увеличение количества автомобилей на душу населения приводит к снижению интенсивности их использования
PG	-	с ростом цены объем потребления блага сокращается
GAS _{t-1}	+	изменение потребления топлива в зависимости от изменения цены на нефтепродукты является высоко инерционным процессом, поскольку в краткосрочном периоде у потребителя нет возможности сменить транспортное средство

Абсолютно логично предположить, что более высокий валовый региональный продукт повышает интенсивность использования автомобилей в результате большей экономической активности. Однако на индивидуальном уровне увеличение располагаемого дохода ведет к увеличению количества автомобилей на душу населения. А переменная количества легковых автомобилей на душу населения (CAR) должна учитывать снижение интенсивности использования транспортного средства, вызванное ростом количества автомобилей на домохозяйство (статистические данные свидетельствуют о том, что семья, владеющая двумя автомобилями, преодолевает на них меньше, чем вдвое большее расстояние, чем семья с одним автомобилем).

Вдобавок, вследствие того факта, что количество автомобилей учитывается в нашей спецификации в расчете на душу населения (в качестве знаменателя переменной CAR используется общая численность населения), а не на количество водителей, не представляется возможным учесть демографические эффекты. Так, например, в работе Schmalensee и Stoker (1999) [CITATION Sch99 \l 1049] авторы на данных США показали, что в результате использования в такой ситуации количества водителей значительно смещает расчетное значение эластичности по доходу в сторону снижения. Таким образом, можем предположить, что найденные нами оценки эластичности спроса на бензин окажутся завышенными относительно истинных значений. Сказанное означает, что при получении дополнительных статистических данных будет возможным устранить данный недостаток для получения более достоверных результатов.

В соответствии с экономической теорией, между ценой на благо и величиной спроса на него существует обратная зависимость, то есть с ростом реальной цены на бензин его потребление будет снижаться.

Что касается эффективности использования топлива, она не наблюдается на уровне агрегированных данных и в течение выбранного периода времени для всех регионов. Она определяется техническими характеристиками двигателя автомобиля, манерой вождения, интенсивностью движения, географическими условиями, весом транспортного средства, использованием кондиционера и прочими факторами. Поскольку автомобильный парк имеет свой срок эксплуатации, который, как ожидается, зависит от экономического роста, топливная эффективность может быть выражена распределенными лагами экономических переменных (Baltagi и Griffin (1983) [CITATION Заполнитель2 \l 1049]). Включение вместо этого в качестве регрессора лага зависимой переменной отражает предложенную спецификацию с распределенными лагами (Baltagi et al. (2003) [CITATION Bal03 \l 1049]).

С точки зрения модели, включение лага потребления бензина в расчете на один автомобиль оправдано принятием в научном сообществе модели регулирования потока Houthakker et al. (1974) [CITATION Hou74 \l 1049]. Согласно этой модели, интенсивность использования автомобиля приближается к желаемому уровню посредством механизма сохранения постоянства привычки. Применительно к спросу на топливо это означает, что адаптация фактического потребления бензина GAS к желаемому уровню $(GAS)^i$ с течением времени считается процессом первого порядка (формула (3)):

$$\frac{GAS_t}{GAS_{t-1}} = \left(\frac{GAS_t^i}{GAS_{t-1}} \right)^\theta, \quad 0 < \theta < 1 \quad (3)$$

где θ – коэффициент адаптации текущего уровня потребления топлива к желаемому.

После подстановки значения $(GAS)^i$ из уравнения (1) в (3), логарифмирования и классификации данных по региональному (i) и временному (t) признаку, можно получить часто используемое динамическое уравнение спроса на бензин (формула (4)):

$$\ln GAS_{i,t} = \theta \ln \alpha + (1 - \theta) \ln GAS_{i,t-1} + \theta \beta \ln Y_{i,t} + \theta \gamma \ln PG_{i,t} + \theta \delta \ln CAR_{i,t} + u_{i,t}, \quad (4)$$

где $u_{i,t}$ – вектор возмущений.

$u_{i,t}$ определяется в соответствии с двунаправленной моделью ошибок (формула (1)):

$$u_{i,t} = \mu_i + \lambda_t + \varepsilon_{i,t}, \quad i=1, \dots, I, \quad t=1, \dots, T \quad (5)$$

где

μ_i – региональный эффект,

λ_t – временной эффект,

$\varepsilon_{i,t}$ – белый шум.

Региональный и временной эффекты в модели могут быть как фиксированными, так и случайными. Уравнение (4) является нашей базовой моделью спроса на бензин в краткосрочном периоде с непосредственным учетом количества транспортных средств. Согласно уравнению (4), краткосрочные значения эластичностей спроса на бензин в расчете на один автомобиль по доходу на душу населения, реальной цене и количеству автомобилей на душу населения составляют $\theta\beta$, $\theta\gamma$ и $\theta\delta$ соответственно. Аналогичные долгосрочные эластичности определяются значениями β , γ и δ соответственно, скорость адаптации к долгосрочному равновесию θ находим из коэффициента при лаге зависимой переменной. Кроме того, так как количество транспортных средств входит в уравнение и в качестве независимой переменной, и в качестве зависимой, эластичность совокупного спроса на бензин в зависимости от изменения общей численности автомобильного парка для краткосрочного и долгосрочного периодов представляется в виде $(1+\theta\delta)$ и $(1+\delta)$ соответственно.

Важно отметить, что механизмы адаптации потребителей к изменению цен в краткосрочном и долгосрочном периодах отличаются. В краткосрочном периоде изменение цен на бензин приводит главным образом к изменению интенсивности использования автомобилей, в то время как в долгосрочном периоде потребители будут

стремиться не только покупать более энергоэффективные автомобили с бензиновыми двигателями, но и переходить на дизельные автомобили, поскольку операционные затраты собственников дизельных автомобилей при больших пробегах существенно ниже. Автомобили с бензиновым двигателем, которые используются более интенсивно (то есть чей годовой пробег выше средних значений), будут заменены в первую очередь. Постепенно автомобили с бензиновым двигателем останутся у тех, кто пользуется ними менее интенсивно, что, безусловно, снизит показатель потребления бензина в расчете на одну машину с бензиновым двигателем.

Представим ниже некоторые критические замечания применительно к представленной модели, а также покажем возможное ее расширение, которое позволит эти недостатки исправить, и которое на данном этапе не было возможным реализовать из-за отсутствия статистических данных по дизельным автомобилям.

Итак, несмотря на логичность и обоснованность уравнения спроса (), в нем существует ряд недостатков. Во-первых, оно не учитывает упомянутое выше косвенное влияние постепенно растущего числа дизельных автомобилей на среднее потребление бензина в расчете на один автомобиль. Во-вторых, в уравнении () общее количество легковых автомобилей, используемое для вычисления и зависимой, и объясняющей переменных, включает в себя в том числе и автомобили, работающие на дизельном топливе. То есть в данной спецификации уравнения спроса на бензин имеются пропущенные переменные, вследствие чего оценки эластичности соответственно смещены.

С целью устранения представленных недостатков в уравнении спроса на бензин, возможно видоизменить уравнение () путем использования в качестве зависимой переменной потребления бензина в расчете на один автомобиль с бензиновым двигателем (GASG), а в качестве объясняющей переменной – количества легковых автомобилей с бензиновым двигателем на душу населения (или, если такие данные доступны, на одного водителя) (CARG). Кроме того, к регрессорам добавляется количество дизельных легковых автомобилей в расчете на душу населения (или на одного водителя) (CARD). В результате получаем уравнение ():

$$\ln GASG_{i,t} = \theta \ln \alpha + (1 - \theta) \ln GASG_{i,t-1} + \theta \beta \ln Y_{i,t} + \theta \gamma \ln PG_{i,t} + \theta \delta \ln CARG_{i,t} + \theta \phi \ln CARD_{i,t} + u_{i,t}, \quad (6)$$

где ϕ – коэффициент долгосрочной эластичности спроса на бензин в расчете на один бензиновый автомобиль в зависимости от числа дизельных автомобилей на душу населения (или на одного водителя). Эта величина отражает эффект перехода тех, кто более интенсивно использует автомобили с бензиновыми двигателями, на автомобили с дизельными двигателями, сокращая среднее потребление бензина на один автомобиль с бензиновым двигателем. Аналогичным образом, коэффициент δ показывает долгосрочную эластичность спроса на бензин в расчете на один бензиновый автомобиль в зависимости от количества бензиновых автомобилей на душу населения (или на одного водителя), при условии, что распределение интенсивности использования автомобилей индивидуумами остается неизменным. На первый взгляд не очевидно, какой эффект будет доминировать, и является ли одна из переменных лишней. В работе Рок (2010) [CITATION Roc10 \l 1049], например, F-тест на совместную значимость включенных переменных отклоняет нулевую гипотезу о том, переменная CARD является лишней в моделях как с однонаправленными, так и с двунаправленными фиксированными или случайными эффектами. Мы же в настоящий момент не можем реализовать подобное расширение модели ввиду отсутствия статистических данных по дизельным автомобилям.

Также следует объяснить, почему проблема эндогенности, которая влечет за собой несостоятельность и смещенность оценок МНК, данной спецификации не свойственна. Известно, что переменная количества легковых автомобилей на душу населения зависит от уровня дохода на душу населения (Dahl и Sterner (1991) [CITATION Dah91 \l 1049], а также Dargay et al. (2007) [CITATION Dar07 \l 1049]), что вызывает озабоченность по поводу возможной эндогенности в спецификациях () и (). Безусловно, верно, что более высокий совокупный доход, как показатель большей экономической активности и благосостояния населения, приводит к увеличению автомобильного парка в долгосрочной перспективе. Однако в краткосрочной перспективе более высокий доход в первую очередь влияет на интенсивность использования уже имеющегося парка автомобилей, особенно автомобилей, которые используются для работы и деловых нужд. На уровне домохозяйств рост располагаемого дохода ведет к увеличению количества автомобилей на душу населения, что снижает средний показатель интенсивности использования автомобилей. Помимо этого, более высокий уровень продаж автомобилей повышает ВВП (особенно этот эффект заметен в странах с крупными автомобильными производствами). Таким образом, причинно-следственная связь между двумя обозначенными переменными не обязательно является однонаправленной и непосредственно прямой. Исключение переменных, отвечающих за количество транспортных средств, или замена этих переменных слабыми инструментами может привести к потере объясняющей силы моделей. В соответствии с

хорошо известными спецификациями моделей в работах (Baltagi and Griffin (1983, 1997) [CITATION Заполнитель2 \l 1049], [CITATION Bal97 \l 1049], Baltagi et al. (2003) [CITATION Bal03 \l 1049], и Johansson и Schipper (1997) [CITATION Joh97 \l 1049]), мы предполагаем, что переменные цены, дохода и количества автомобилей на душу населения (или на одного водителя) являются экзогенными для потребления топлива. Следовательно, корреляция между доходом и количеством автомобилей влечет за собой скорее коллинеарность, чем проблему эндогенности, что приводит к увеличению стандартных ошибок, но не к несостоятельности оценок коэффициентов.

Ниже представлены методы оценивания, с помощью которых будет проанализировано уравнение спроса на бензин. Предполагая экзогенность регрессоров, к базе данных были применены следующие стандартные методы оценивания: сквозная регрессия, оцененная обычным методом наименьших квадратов, которая не учитывает какие-либо эффекты; регрессия с фиксированными эффектами, оцененная обычным методом наименьших квадратов; регрессия с фиксированными эффектами и поправкой на смещение; а также обобщенный метод наименьших квадратов, при котором предполагается, что специфичные для объектов эффекты случайны. Кроме того, мы использовали доступный (реализуемый) обобщенный метод наименьших квадратов, учитывающий гетероскедастичность и кросс-секционную (пространственную) корреляцию, а также характерную для панельных данных структуру ошибок AR (1), при этом регрессия оценивается с индивидуальными фиксированными эффектами и имеет более расширенную структуру ошибок, чем при упомянутом ранее обобщенном методе наименьших квадратов.

Поскольку построенная модель является динамической и учитывает индивидуальные региональные эффекты, лаг зависимой переменной коррелирует с вектором ошибок и, следовательно, приводит к несостоятельности оценок сквозной регрессии, оцененной методом наименьших квадратов; регрессии с фиксированными эффектами, оцененной методом наименьших квадратов и регрессии со случайными эффектами, оцененной обобщенным методом наименьших квадратов. С целью учета данной особенности модели в научных работах используют метод инструментальных переменных, который обеспечивает состоятельность, но не обязательно эффективность оценок параметров модели. Инструментами в данном случае, как правило, являются экзогенные переменные и их запаздывающие значения (лаги).

Таким образом, далее также используется регрессия с фиксированными эффектами, оцененная двухшаговым методом наименьших квадратов. Использование данного метода подразумевает оценивание уравнения регрессии в отклонениях по времени от среднего

для каждого региона с помощью двухшагового метода наименьших квадратов. Таким образом представляется возможным корректно учесть индивидуальные эффекты. Инструментами в данном случае являются экзогенные переменные и их лаги в уровнях.

Следующим был использован метод FD-2SLS, который был разработан Anderson и Hsiao (1981) [CITATION And82 \l 1049] и заключается в оценивании записанного в первых разностях уравнения с помощью двухшагового МНК. Это позволяет устранить ненаблюдаемую региональную неоднородность. Инструментами здесь выступают экзогенные переменные, их лаги в уровнях и второй лаг уровня объясняемой переменной.

Из ограничений данного метода следует отметить, что если случайные ошибки в модели автокоррелированы, то для получения состоятельных оценок необходимо использовать более глубокие лаги зависимой переменной для инструментирования лага ее первой разности.

Затем был применён метод FD-GMM, предложенный в работе Arellano и Bond (1991) [CITATION Are91 \l 1049]. Данный метод заключается в оценивании регрессии в первых разностях с помощью обобщенного метода моментов (мы использовали одношаговый). Инструментами для эндогенной переменной в первых разностях в данном случае являются экзогенные переменные, их лаги в уровнях, второй и другие лаги уровня зависимой переменной. Такое обилие инструментов дает возможность получить более эффективные оценки, чем, например, метод Андерсона – Хсиао.

Дальше мы обратились к системному обобщенному методу моментов. Идея предложенного в работе Blundell и Bond (1998) [CITATION Blu98 \l 1049] метода sys-GMM состоит в использовании в качестве инструментов для уравнения в уровнях первых разностей, а для уравнения в первых разностях – переменных в уровнях. Преимущество данного метода заключается в нивелировании смещенности оценок при работе с малыми выборками. Такая смещенность имеет место при применении методов Андерсона – Хсиао и Ареллано – Бонда вследствие недостаточного количества инструментов. В то же время при оценивании методом Бланделла – Бонда с увеличением числа регрессоров условия на моменты приближаются к числу наблюдений в малых выборках. Слишком большое количество инструментов приводит к сверхподгонке инструментлируемой переменной, и полученные оценки оказываются смещенными в сторону оценок МНК (как показано, например, в работе Baltagi (2005) [CITATION Bal05 \l 1049]).

В работе Bun, Kiviet (2006) [CITATION Bun06 \l 1049] авторы показали, что желаемые свойства оценок, полученных методами Ареллано – Бонда, Бланделла – Бонда, Андерсона – Хсиао и в регрессии с фиксированными эффектами, оцененной двухшаговым

МНК, достигаются асимптотически при большом количестве объектов, а иначе могут оказаться смещенными и неэффективными.

Не так давно в эконометрической литературе появился и стал активно использоваться альтернативный подход, основанный на оценивании регрессии с фиксированными эффектами и поправкой на смещение (LSDVc). Метод LSDV подразумевает оценивание коэффициентов авторегрессионной модели панельных данных с фиктивными переменными для индивидуальных детерминированных эффектов методом наименьших квадратов. В исследовании Nickell (1981) [CITATION Nic81 \l 1033] было продемонстрировано, что такие оценки несостоятельны для панелей с конечным количеством временных тактов. Однако, в работах Arellano и Bond (1991) [CITATION Are91 \l 1033]; Kiviet (1995) [CITATION Kiv95 \l 1033] показано, что хоть оценки, полученные с помощью LSDV, и не состоятельны, но имеют относительно небольшую дисперсию по сравнению с оценками методов инструментальных переменных (Андерсона – Хсiao и в регрессии с фиксированными эффектами, оцененной двухшаговым МНК) и методов моментов (Ареллано – Бонда, Бланделла – Бонда).

Учитывая этот факт, широкое распространение получил предложенный в работе Bruno (2005) [CITATION Bru05 \l 1033] метод LSDV corrected (LSDVc), при котором проводится диагональная коррекция оценок LSDV и используется процедура бутстрапирования при нахождении оценок стандартных ошибок. Поправки можно получить, основываясь на оценках, полученных методом Андерсона – Хсiao, или Ареллано – Бонда, или Бланделла – Бонда.

В данном случае был применен метод оценивания регрессии с фиксированными эффектами и поправкой на смещение, используя условия аппроксимации до порядка T^{-1} (например, Bun и Kiviet (2003) [CITATION Bun03 \l 1049]). Начальные значения истинных коэффициентов были получены с помощью метода Ареллано – Бонда, обеспечивающего состоятельные оценки. Также этот метод хорошо себя зарекомендовал при работе с несбалансированными панелями.

2 Проведение количественной оценки эластичности спроса на нефтепродукты автомобильным видом транспорта в зависимости от цен на нефтепродукты

В данной главе представлено описание статистической базы данных, на основе которой будет проведена количественная оценка эластичности спроса на бензин в Российской Федерации. Кроме того, в данной главе будут представлены процедуры проведения оценивания и интерпретацию полученных результатов. Итоговые оценки представлены для методологий спроса на бензин с непосредственным учетом количества транспортных средств, модели коррекции ошибок для спроса на бензин с использованием свойства коинтеграции временных рядов, а также системы одновременных уравнений спроса и предложения, логика и особенности спецификации которых изложены в главе 4.

Для проведения анализа были использованы статистические базы данных Федеральной службы государственной статистики, Софист ВШЭ и Центрального банка Российской Федерации. Статистическая информация включает в себя такие показатели, как объемы поставок бензина, средние потребительские цены на бензин, средние цены на дизельное топливо; цены на сырую нефть; валовый региональный продукт; численность населения; количество легковых автомобилей; значение кривой бескупонной доходности государственных облигаций сроком до погашения в один год и индекс потребительских цен. Необходимо отметить, что база данных представляет из себя панель, содержит ежемесячные данные по 68 регионам Российской Федерации с 2010 по 2016 год и является сбалансированной (имеются наблюдения для всех объектов в каждый рассматриваемый момент времени). Для тех переменных, статистика которых публикуется с периодичностью в один год, была использована линейная интерполяция для приведения их к периодичности в один месяц. К показателям с явно выраженной сезонностью (например, к объемам поставок бензина) применялось сезонное сглаживание.

Зависимая переменная нашего анализа – удельный расход бензина на одно легковое транспортное средство. Объясняющие переменные включают в себя: реальный валовый региональный продукт в расчете на душу населения; реальная цена на бензин и количество легковых автомобилей на душу населения. В таблице представлены описательные статистики используемых переменных.

Таблица 3 – Описательные статистики переменных модели спроса на бензин с непосредственным учетом количества транспортных средств (68 регионов)

Переменная		Среднее значение	Стандартное отклонение	Минимальное значение	Максимальное значение	Медиана	Наблюдения
Удельное потребление бензина	overall	154.2799	166.6924	4.879865	1351.625	92.3216	N = 5712
	between		161.935	7.808304	1051.543		n = 68
	within		44.0968	-218.4145	693.9488		T = 84
ВРП на душу населения	overall	19414.61	14081.03	6520.462	96134.14	16319.16	N = 5712
	between		14017.49	8292.041	86407.31		n = 68
	within		2154.299	-3057.958	30500.86		T = 84
Цены на бензин	overall	21.82024	1.735587	19.38	32.29128	21.34201	N = 5712
	between		1.711444	19.9251	29.22317		n = 68
	within		0.3546687	20.07969	25.9673		T = 84
Количество авто на душу населения	overall	204.2139	136.1974	23.03079	696.4721	182.3058	N = 5712
	between		133.4561	24.08193	628.9122		n = 68
	within		31.59196	39.93445	332.4967		T = 84

Как видно из таблицы, для некоторых переменных характерен довольно сильный разброс и высокий показатель вариации, что может вызывать гетероскедастичность в будущей модели. Так, например, переменные ВРП на душу населения и количество автомобилей на душу населения характеризуются высоким стандартным отклонением.

Используя медианные значения, представляется возможным описать типичное наблюдение выборки 2010-2016 годов: один автомобиль в месяц потребляет 92.3216 тонны бензина по цене 21.34201 рубль за литр; реальный ВРП за месяц на душу населения составляет 16319.16 рублей; количество авто на душу населения – 0.182.

В целях дальнейшего анализа и интерпретации результатов оценивания уравнения спроса на бензин в Российской Федерации в терминах эластичности проведем логарифмирование всех используемых переменных. Описательные статистики для логарифмов статистических данных представлены в таблице.

Таблица 4 – Описательные статистики переменных модели спроса на бензин с непосредственным учетом количества транспортных средств (68 регионов)

Переменная		Среднее значение	Стандартное отклонение	Минимальное значение	Максимальное значение	Медиана	Наблюдения
Логарифм удельного потребления бензина	overall	4.600168	0.9534307	1.585118	7.209063	4.525278	N = 5712
	between		0.926117	2.033193	6.946075		n = 68
	within		0.2525916	3.519671	5.701073		T = 84
Логарифм ВРП на душу населения	overall	4.60059	0.9531444	1.585118	7.209063	4.525547	N = 5644
	between		0.9257923	2.033149	6.946874		n = 68
	within		0.2526806	3.523918	5.70532		T = 83
Логарифм ВРП на душу населения	overall	9.733506	0.4712916	8.782701	11.4735	9.700096	N = 5712
	between		0.4643285	9.021053	11.36212		n = 68

Продолжение таблицы

	within		0.0982249	9.103589	9.971735		T = 84
Логарифм цен на бензин	overall	3.079993	0.0735435	2.964242	3.474797	3.060678	N = 5712
	between		0.072346	2.991895	3.37441		n = 68
	within		0.0158359	3.001393	3.230009		T = 84
Логарифм количества авто на душу населения	overall	5.095561	0.6957548	3.136832	6.546028	5.205686	N = 5712
	between		0.6881601	3.181324	6.442392		n = 68
	within		0.1318819	4.338507	5.53757		T = 84

Проверка объясняемой переменной на нормальность является важным условием при проведении эконометрических оценок. Часто из-за отсутствия нормальности зависимой переменной распределение ошибок также получается не нормальным, что мешает проведению классических t- и F-тестов. Разумеется, не все статистики критериев распределены согласно нормальному закону, и тем не менее большая часть из них или имеют гауссовское распределение, или иное распределение, но связанное с гауссовским и рассчитываемое на его основе. К таким относятся, например, распределение Стьюдента, F и хи-квадрат - распределения. Как правило, необходимым условием для этих статистик является нормальное в совокупности распределение анализируемых переменных. В связи с вышесказанным необходимо рассмотреть, насколько близко распределение зависимой переменной к нормальному. Кроме того, важно заметить, что по мере увеличения размера выборки параметры нормальности становятся более ограничительными, и становится все труднее сделать вывод, что данные распределены нормально, поскольку обнаруживаются даже небольшие отклонения от нормальности. Поэтому для выборок очень большого размера проверка нормальности становится менее критичной.

Существует довольно много методов для оценки того, являются данные нормально распределенными или нет. Они разделяются на две большие категории: графические и статистические. Вот примеры некоторых общепринятых методов: графические: нормальный график квантиль-квантиль (диаграммы вероятности Q-Q); совокупные частотные (P-P) графики; статистические: W / S тест; тест Харке-Бера; тест Шапиро-Вилкса, тест Колмогорова-Смирнова, тест д'Агостино.

Графические методы обычно не очень полезны, когда размер выборки невелик. Они могут давать неправильное представление о нормальности, например, судя по графику, данные могут не выглядеть нормально распределенными, но статистически их распределение не будет отличаться от нормального.

Необходимость проверки на нормальность ошибок в регрессионных моделях обуславливается как методологическими, так и концептуальными причинами. С чисто методологической точки зрения, отсутствие гауссовского распределения иногда ставит под сомнение надежность простых процедур оценивания и тестирования и требует либо более

совершенных методов, предполагающих альтернативный вид распределения, либо применения робастных методов, желаемые свойства которых не зависят от особенностей распределения. В целом, вопрос о том, что, возможно, следует более адекватно отражать распределение ошибок, используя перекошенное и / или островершинное распределение, является важным в статистической науке.

Предположение о нормальности также играет решающую роль в валидности тестов на правильность спецификации и прогнозировании. В работе Blanchard и Matyas (1996) [CITATION Bla96 \l 1049], посвященной панельным данным, авторы изучают последствия ненормальности распределения компонентов ошибок для выполнения некоторых тестов. В исследовании Montes-Rojas и Sosa-Escudero (2011) [CITATION Mon11 \l 1049] показано, что такая ненормальность сильно влияет на эффективность тестов на гетероскедастичность для панельных данных, описанных в трудах Holly и Gardiol (2000) [CITATION Hol00 \l 1049] и Baltagi et al. (2006) [CITATION Bal06 \l 1049]. Несмотря на эти ограничения, гауссовское распределение (ошибок) широко используется в тестировании спецификаций моделей с однонаправленной компонентой ошибки; например, тесты для пространственных моделей панельных данных Baltagi et al. (2003) [CITATION Bal031 \l 1049] и Baltagi et al. (2007) [CITATION Bal07 \l 1033].

В то время как тестирование асимметрии и эксцесса распределений кросс-секционных данных и временных рядов широко представлено в научной литературе, в частности в работах Erickson и Whited (2010) [CITATION Eri00 \l 1033], Bai и Ng (2005) [CITATION Bai05 \l 1033] и Bera и Premaratne (2001) [CITATION Ber15 \l 1033], в отношении панельных данных подобных работ мало. Естественная сложность заключается в том, что, в отличие от кросс-секционных данных и временных рядов, в простых моделях компонентов ошибок отсутствие гауссовского распределения может возникать в более чем одном компоненте. Таким образом, дополнительной задачей, связанной с обнаружением отклонений от нормального распределения, является определение того, какой компонент это отклонение вызывает. Данная тема освещена в исследованиях Gilbert (2002) [CITATION Gil02 \l 1033] (используются смешанные моменты) и Meintanis (2011) [CITATION Mei11 \l 1033] (предлагается комбинированный тест на нормальность обоих компонентов совместно, основанный на эмпирических характеристических функциях).

В работе Galvao et al. (2013) [CITATION Gal13 \l 1033] представлены тесты на асимметрию (отсутствие симметрии), эксцесс и нормальность для моделей панельных данных с однонаправленной компонентой ошибки, задаваемой формулой ():

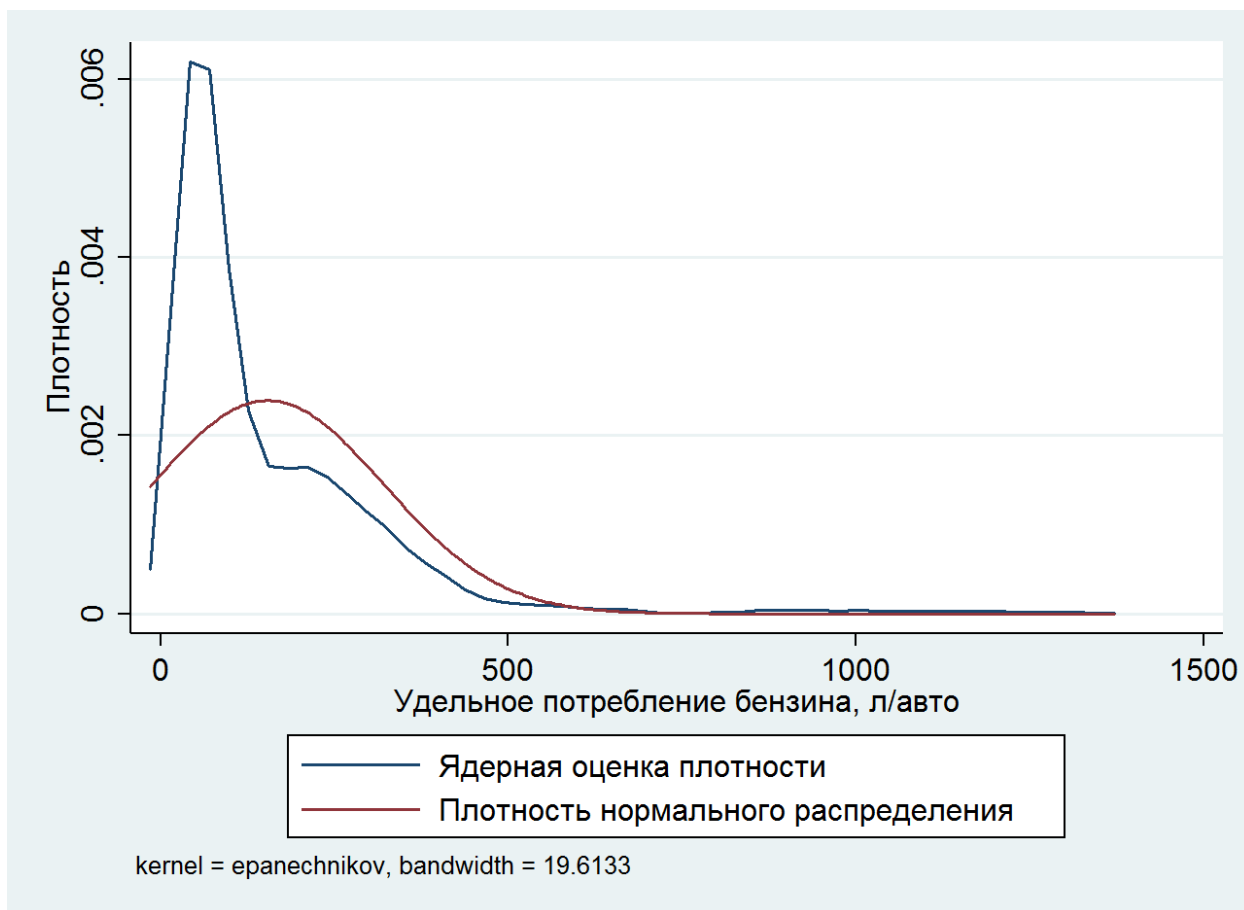
$$y_{it} = x_{it}b + u_i + e_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T, \quad (7)$$

где b – это вектор параметров;

u_i и e_{it} здесь отвечают за индивидуальную и остаточную компоненты ошибки, среднее значение каждой из которых равно нулю.

Эти тесты важны из практических соображений, потому что в случае панельных данных стандартный тест Харке – Бера не позволяет отделить отклонения от нормального распределения индивидуальных компонентов и компонента остатков. Рассматриваемые тесты построены на основе условий на моменты within и between преобразований остатков МНК. Эти условия используются для разработки тестов на асимметрию и эксцесс индивидуальных компонентов и компонента остатков по отдельности и совместно. Данные тесты особенно полезны в случае, когда число объектов наблюдения, N , стремится к бесконечности, а количество временных периодов, T , фиксировано и может быть небольшим.

Чтобы посмотреть, каким образом распределена переменная, отвечающая за среднее потребление бензина в расчете на один легковой автомобиль, сначала был построен график ее ядерной функции плотности. Этот график приведен на рисунке .



Источник: составлено авторами

Рисунок 1 – Сравнение ядерной функции плотности переменной, отвечающей за среднее потребление бензина в расчете на один легковой автомобиль, и нормального распределения (68 регионов)

На рисунке можно заметить, что распределение переменной среднего потребления бензина в расчете на один легковой автомобиль значительно отличается от нормального, однако визуально очень напоминает логнормальное, следовательно, есть основания предположить, что логарифм этой переменной имеет нормальное распределение.

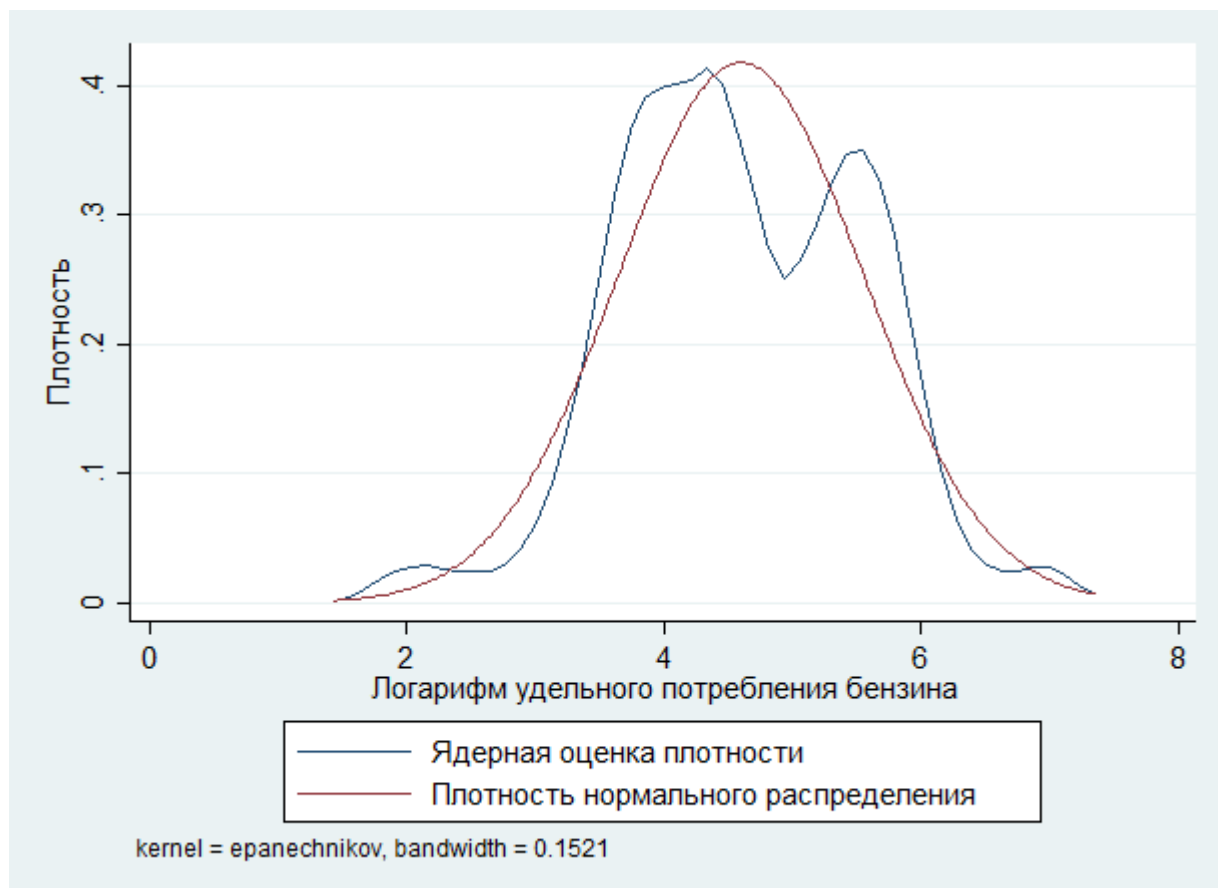
Затем был проведен описанный в работе Galvao et al. (2013) [CITATION Gal13 \l 1049] адаптированный к особенностям данных панельной структуры тест Харке-Бера на нормальность распределения ошибок наблюдений. Нулевая гипотеза состоит в том, что асимметрия и эксцесс равны соответствующим значениям нормального распределения. Результаты тестирования приведены в таблице .

Таблица 5 – Результаты теста Харке-Бера на нормальность панельных данных для переменной среднего потребления бензина в расчете на один легковой автомобиль (68 регионов)

Количество наблюдений = 5,712						
Количество репликаций процедуры бутстрап = 500						
(Реплицирование на основе 68 групп по объектам наблюдения)						
	Коэффициент (наблюдаемые)	Стандартная ошибка (бутстрап)	z	P> z	[95% Доверительный интервал]	
Асимметрия_e	1.155468	0.7229441	1.6	0.110	-0.2614769	2.572412
Экссесс_e	26.98853	10.47966	2.58	0.010	6.448768	47.52829
Асимметрия_u	2.940921	0.9087778	3.24	0.001	1.159749	4.722093
Экссесс_u	12.3659	6.071112	2.04	0.042	0.4667375	24.26506
Совместный тест на нормальность e:			chi2(2) = 9.19	Prob > chi2 = 0.0101		
Совместный тест на нормальность u:			chi2(2) = 14.62	Prob > chi2 = 0.0007		

Как видно из таблицы, ни индивидуальную, ни остаточную компоненты ошибки нельзя признать нормальными – совместные тесты на нормальность эту гипотезу отклоняют. Только тест на равенство асимметрии остаточной компоненты ошибок нормальной (то есть нулю) позволяет принять нулевую гипотезу.

Далее для проверки на нормальность был использован логарифм переменной среднего потребления бензина в расчете на один легковой автомобиль. График его ядерной функции плотности представлен на рисунке.



Источник: составлено авторами

Рисунок 2 – Сравнение ядерной функции плотности переменной, отвечающей за среднее потребление бензина в расчете на один легковой автомобиль, и нормального распределения

Рисунок свидетельствует о том, что проделанная ранее процедура позволила получить асимметрию и эксцесс распределения логарифма переменной, отвечающей за среднее потребление бензина в расчете на один легковой автомобиль, гораздо более близкими к асимметрии и эксцессу нормального распределения. Для того, чтобы сделать более точный вывод о соответствии распределения данной переменной нормальному, необходимо провести формальное тестирование (результаты представлены в таблице).

Таблица 6 – Результаты теста Харке-Бера на нормальность панельных данных для переменной логарифма среднего потребления бензина в расчете на один легковой автомобиль (68 регионов)

Количество наблюдений = 5,712						
Количество репликаций процедуры бутстрап = 500						
(Реплицирование на основе 68 групп по объектам наблюдения)						
	Коэффициент (наблюдаемые)	Стандартная ошибка (бутстрап)	z	P> z	[95% Доверительный интервал]	
Асимметрия_e	-0.0874208	0.0732707	-1.19	0.233	-0.2310288	0.0561871
Эксцесс_e	1.543336	0.3043958	5.07	0.000	0.9467308	2.13994
Асимметрия_u	0.0040113	0.2962426	0.01	0.989	-0.5766135	0.584636

Продолжение таблицы

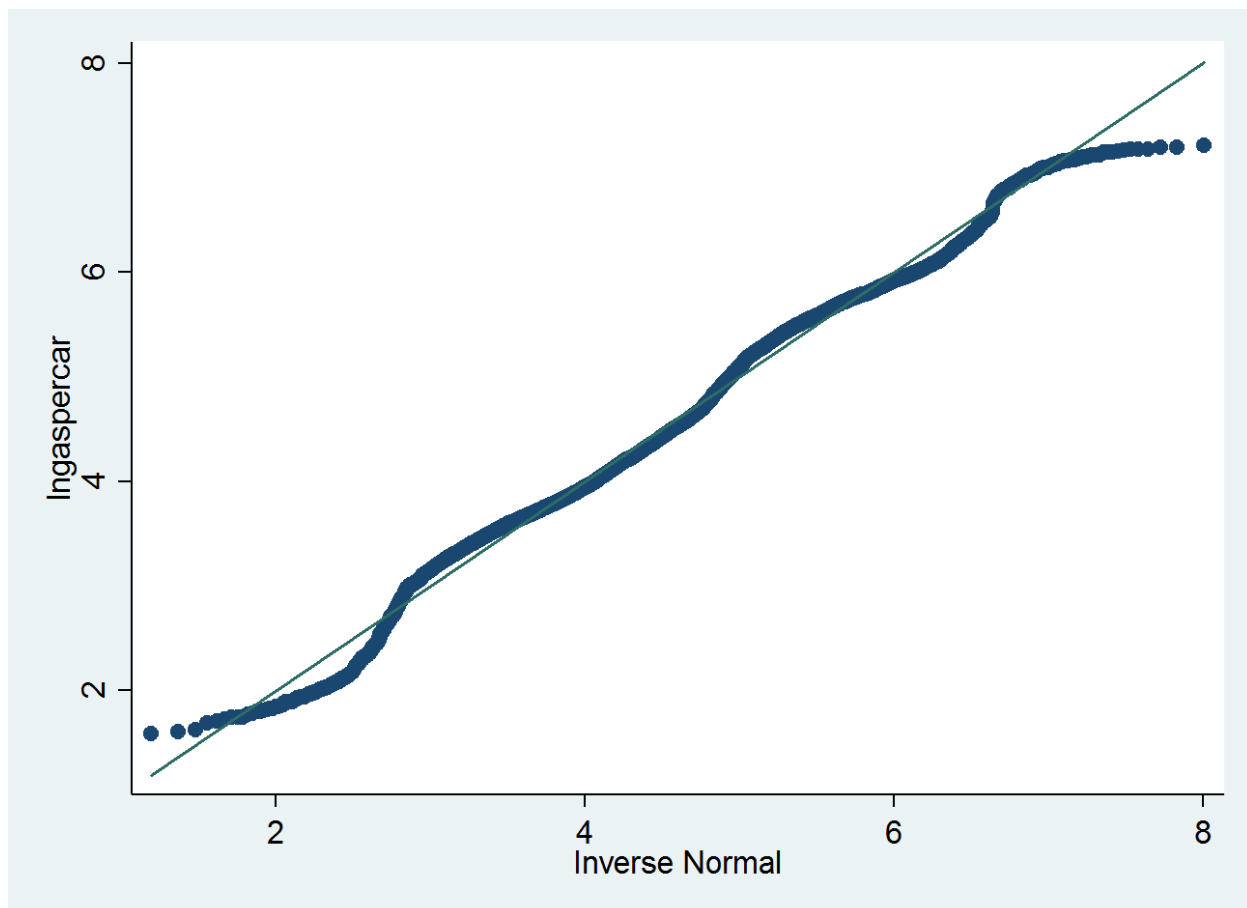
Эксцесс_u	-0.172958	0.4442634	-0.39	0.697	-1.043698	0.6977823
Совместный тест на нормальность e:			chi2(2) = 27.13		Prob > chi2 = 0.0000	
Совместный тест на нормальность u:			chi2(2) = 0.15		Prob > chi2 = 0.9269	

В соответствии с результатами тестов и асимметрия, и эксцесс компонента u равны нормальным. Это ожидаемо приводит к тому, что совместный тест на нормальность u дает все основания принять нулевую гипотезу. Компонент e же имеет соответствующий нормальному распределению 3 момент и несоответствующий нормальному распределению 4 момент, и совместный тест на нормальность показывает, что нулевую гипотезу следует отвергнуть. Таким образом можно отметить, что после логарифмирования зависимая переменная приобрела более близкое к нормальному распределение.

Далее был рассмотрен еще один графический метод тестирования нормальности распределения данных – нормальный график квантиль-квантиль. В общем случае график квантиль-квантиль (Q-Q) демонстрирует взаимное расположение друг относительно друга квантилей из двух распределений. Нормальный же график квантиль-квантиль отображает эмпирические значения по отношению к нормально распределенным данным, изображенных линией с углом наклона 45 градусов. В случае, когда распределение

наблюдаемых значений соответствует нормальному, точки располагаются на биссектрисе первого координатного угла. Если распределение данных отличается от нормального, точки будут отклоняться от этой биссектрисы.

На рисунке можно увидеть, что основные отклонения от «нормальной» линии появляются на «хвостах» распределения, в большей степени на левом «хвосте», где среднее потребление бензина в расчете на один автомобиль низкое.



Источник: составлено авторами

Рисунок 3 – Нормальный график квантиль-квантиль для сравнения распределения логарифма удельного потребления бензина с нормальным распределением

Кроме того, из сравнения ядерной функции плотности переменной, отвечающей за среднее потребление бензина в расчете на один легковой автомобиль, и ее логарифма с нормальным распределением, а также нормального графика квантиль-квантиль можно отметить, что наиболее значительно наши данные отклоняются от нормальности на левом «хвосте» распределения, правый «хвост» отклоняется в меньшей степени.

На следующем этапе анализа был рассмотрен вопрос существования наблюдений (а точнее, таких регионов), для которых заметно значительное отклонение от общей картины – перейдем к обнаружению выбросов. Выбросы – это точки данных (наблюдения), которые находятся за пределами общего линейного паттерна, срединная линия которого является линией регрессии. Эмпирическое правило заключается в том, что выбросы – это

точки, у которых стандартизированный остаток превышает 3,3 (соответствует альфа-уровню 0,001). Удаление выбросов из анализируемого массива данных иногда может существенно повлиять на объясняющую силу и качество подгонки регрессионной модели. Выбросы следует удалять, если есть основания полагать, что другие переменные, не входящие в модель, объясняют, почему случаи выбросов являются необычными, то есть почему для этих случаев нужна отдельная модель. Или напротив, наличие выбросов может свидетельствовать о том, что в модель следует добавить дополнительные объясняющие переменные (то есть спецификация модели нуждается в уточнении). Альтернативой удалению выбросов из анализа является использование устойчивой (робастной к выбросам) регрессии, алгоритм которой учитывает выбросы с меньшим весом, но все же не исключает их из анализа.

Далее рассмотрены некоторые методы обнаружения нетипичных наблюдений и применим их к применяемой статистической панели данных.

Показатель воздействия наблюдения или «разбалансировка» (leverage statistic, h или $\hat{h}$ -value), позволяет выявить наблюдения, которые влияют на регрессионную модель больше, чем другие. В работе Belsley, Kuh и Welsch (1980) [CITATION Bel80 \l 1033]

авторы определяют показатель воздействия (h_i) i -го наблюдения по формуле (8):

$$h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{(n-1)S_x^2}, \quad (8)$$

где

n – количество наблюдений;

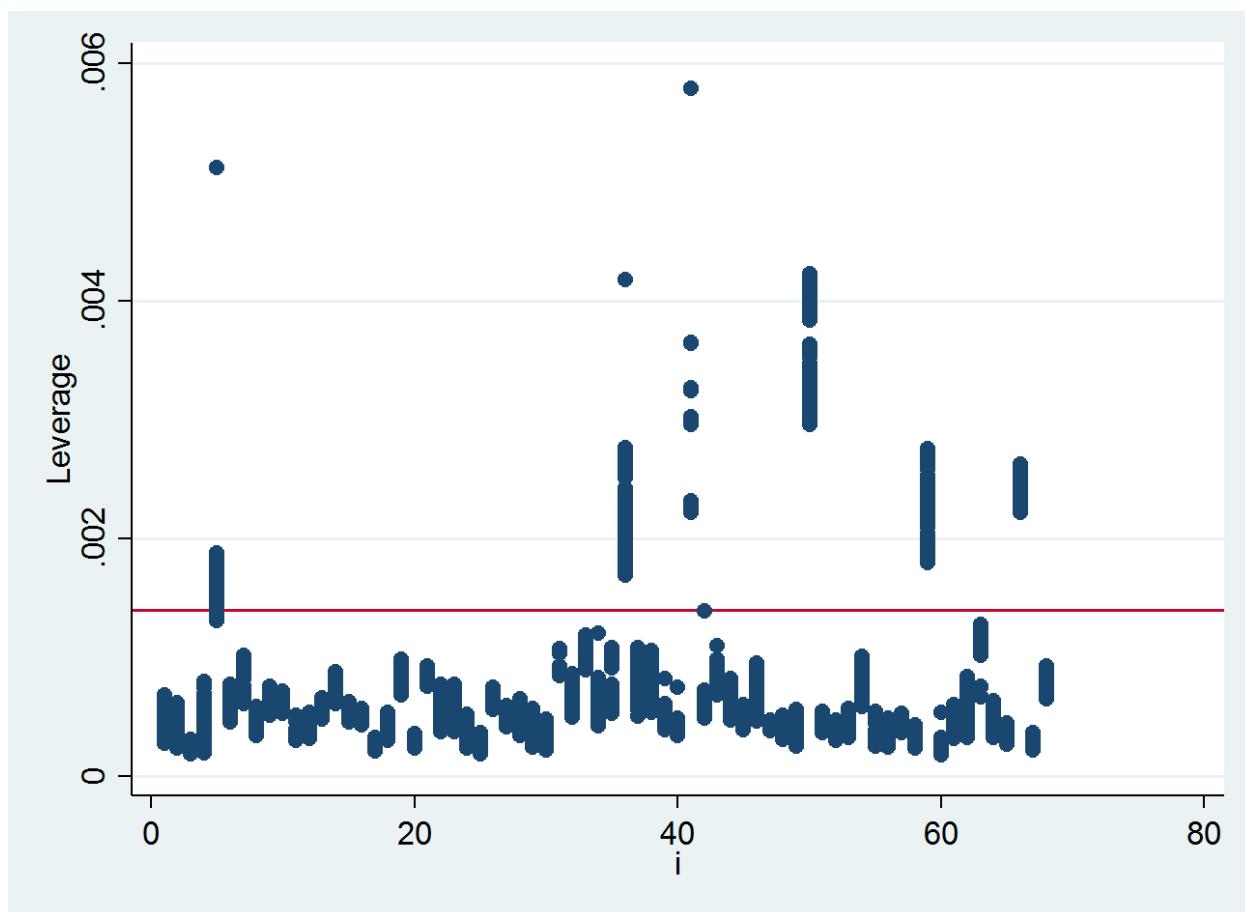
S_x^2 – дисперсия регрессора x .

Разбалансировка позволяет оценить, насколько далеко значение независимой переменной от среднего значения: чем дальше наблюдение, тем больший у него показатель воздействия. Из формулы (8) можно увидеть, что показатель воздействия снижается с увеличением размера выборки (любая отдельно взятая точка оказывает меньшее влияние) и дисперсии независимой переменной (опять же, любая отдельно взятая точка имеет меньшее влияние).

Показатель воздействия наблюдения варьируется от 0 (не влияет на модель) до 1 (полностью определяет модель) ($0 < h < 1$). Наблюдение можно считать проблемным, если $h_i > (2k/n)$, где k – количество параметров, включая константу. Эмпирическое правило

заключается в том, что если показатель воздействия наблюдения ниже 0.2, такое наблюдение не является проблемным, но, если показатель воздействия наблюдения выше 0.5, данное наблюдение следует изучить на предмет возможности ошибки измерения или необходимости моделировать такие случаи отдельно.

На рисунке приведены расчетные значения показателя воздействия наблюдения для логарифма среднего потребления бензина в расчете на один автомобиль по регионам (i) Российской Федерации.



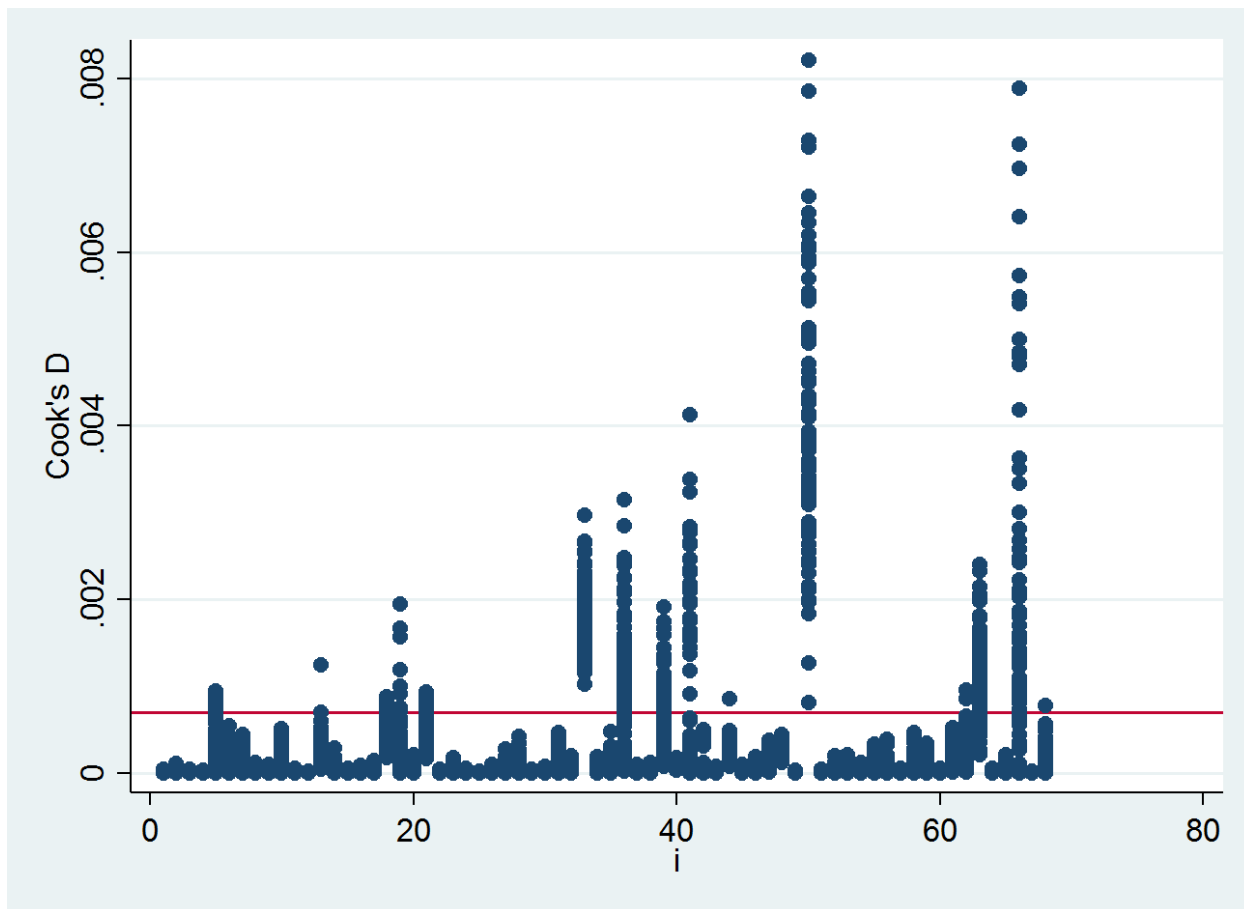
Источник: составлено авторами

Рисунок 4 – Показатели воздействия наблюдения для логарифма среднего потребления бензина в расчете на один автомобиль по регионам Российской Федерации

Исходя из расчета $h_i: 2 \cdot 4 / 5712 = 0.0014$. На основании рисунка выше можно отметить, что точек со значением показателя воздействия, превышающим 0.0014, несколько. Они относятся к следующим регионам: Забайкальский край, Республика Дагестан, Республика Саха (Якутия), Сахалинская область, Тюменская область, город Москва.

Расстояние Кука (Cook's distance, D) является еще одной мерой влияния наблюдения. Расстояние Кука измеряет последствие удаления данного наблюдения из анализа. Наблюдения с большими значениями D , чем у остальных, – это наблюдения с необычным показателем воздействия наблюдения. $D > 4 / n$ – критерий, согласно которому

наблюдение можно признать проблемным, то есть выбросом. Расстояние Кука для логарифма среднего потребления бензина в расчете на один автомобиль по регионам Российской Федерации продемонстрировано на рисунке .



Источник: составлено авторами

Рисунок 5 – Расстояние Кука для логарифма среднего потребления бензина в расчете на один автомобиль по регионам Российской Федерации

Для применяемой в расчетах статистической базы критическое значение составляет $4/5712=0.0007$. На представленном рисунке точек со значением расстояния Кука, превышающим 0.0007, несколько. Они относятся к следующим регионам: Забайкальский край, Костромская область, Ленинградская область, Амурская область, Московская область, Республика Адыгея, Республика Дагестан, Республика Марий Эл, Республика Саха (Якутия), Республика Хакасия, Сахалинская область, Брянская область, Хабаровский край, город Москва, город Санкт-Петербург.

DFBETAS, представляет из себя еще одну методологию, которая позволяет провести оценку степени влияния наблюдения, которая рассчитывается по формуле ():

$$DFBETAS_{ij} = DFBETA_{ij} / (se(i)\sqrt{C_{jj}}) \quad (9)$$

где $se_{(i)}$ – стандартное отклонение ошибки регрессии, не учитывающей i -тое наблюдение;

C_{jj} – диагональный элемент матрицы $(X'X)^{-1}$;

$$DFBETA_{ij} = \hat{B}_j - \hat{B}_{j(i)},$$

а $\hat{B}_j$ – ожидаемое значение j -того параметра.

Наблюдение может считаться значительным выбросом, если на малых и средних выборках $DFBETAS$ больше 1, а на больших выборках $|DFBETAS| > 2 / \sqrt{k/n}$, где k – количество параметров.

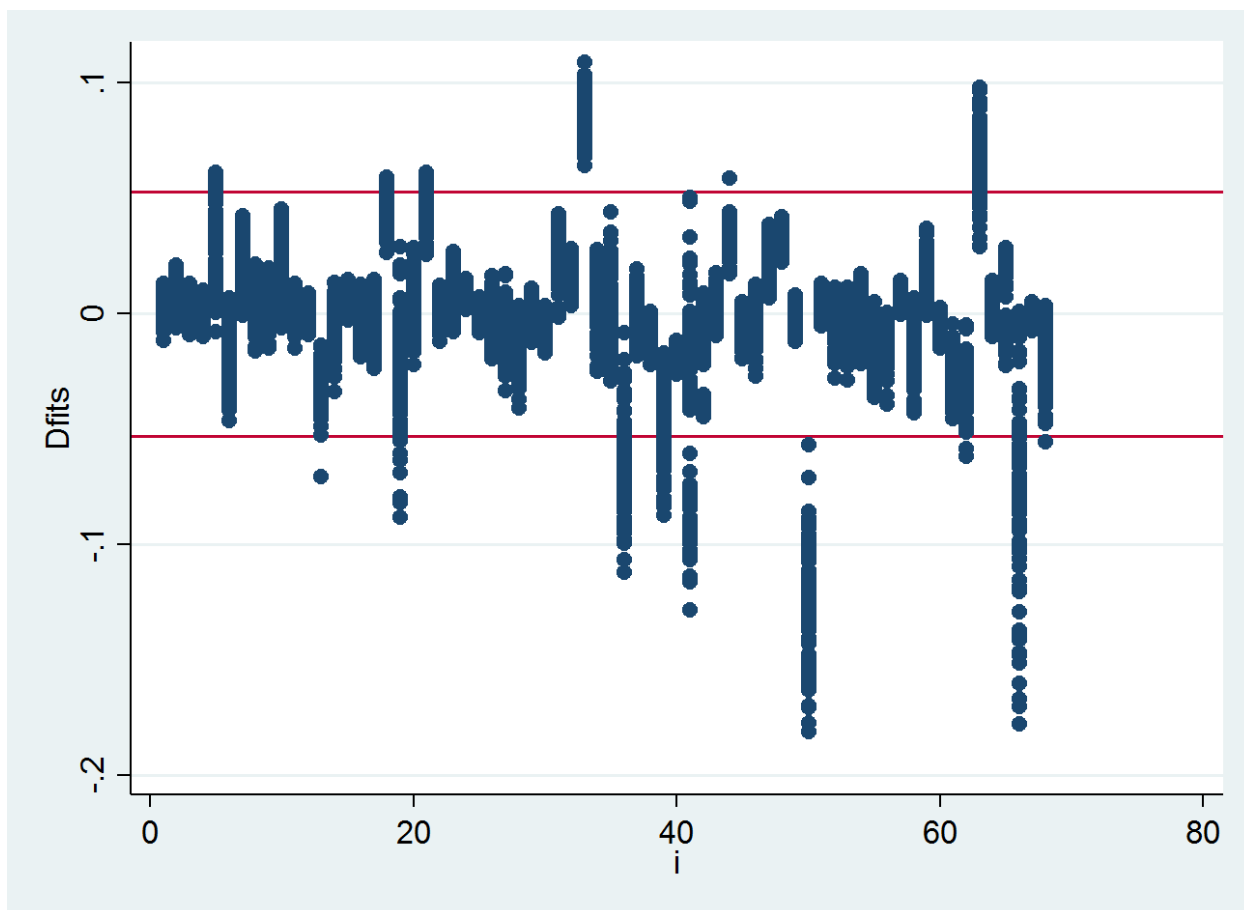
$DFFIT$ измеряет, насколько изменяются результаты оценивания вследствие исключения из анализа конкретного наблюдения. $DFFITS$ определяется по формуле :

$$dfits_i = Rstudent \sqrt{\frac{h_i}{1-h_i}} \quad (10)$$

где $Rstudent$ – студентизированный остаток.

Наблюдение может считаться значительным выбросом, если на малых и средних выборках $DFFITS$ больше 1, а на больших выборках $DFFITS > 2 * \sqrt{k/n}$, где k – количество параметров, включая константу.

$DFFITS$ для логарифма среднего потребления бензина в расчете на один автомобиль по регионам Российской Федерации наглядно представлены на рисунке .



Источник: составлено авторами

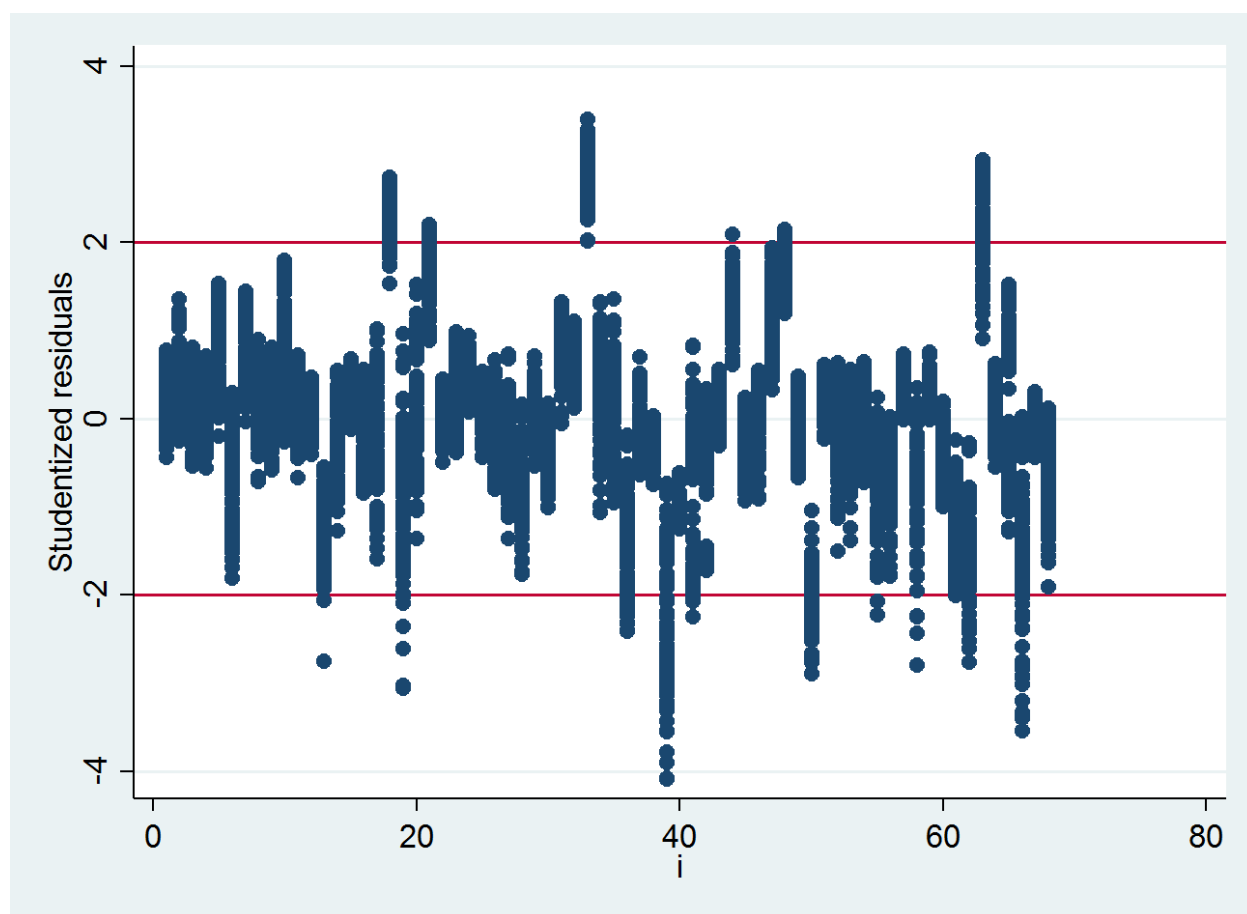
Рисунок 6 – DFITS для логарифма среднего потребления бензина в расчете на один автомобиль по регионам Российской Федерации

Исходя из расчета, пороговое значение составляет: $2 \cdot \sqrt{4/5712} = 0.053$. На данном рисунке точек со значением DFITS, по модулю превышающим 0.053, несколько. Они относятся к следующим регионам: Забайкальский край, Костромская область, Ленинградская область, Амурская область, Московская область, Республика Адыгея, Республика Дагестан, Республика Марий Эл, Республика Саха (Якутия), Республика Хакасия, Сахалинская область, Брянская область, Хабаровский край, город Москва, город Санкт-Петербург.

Стьюдентизированные остатки и удаленные стьюдентизированные остатки также используются для обнаружения выбросов с высоким показателем воздействия наблюдения. «Стьюдентизированный остаток» – это наблюдаемый остаток, деленный на стандартное отклонение. «Удаленный стьюдентизированный остаток», также называемый «jackknife остаток», – это наблюдаемый остаток, деленный на стандартное отклонение, рассчитанное с учетом наблюдения, оставленного вне анализа. Анализ выбросов обычно фокусируется на удаленных остатках. Другими синонимами для их обозначения являются внешний стьюдентизированный остаток или, что может вводить в заблуждение, стандартизованный остаток. Для каждого остатка рассчитывается значение t-статистики с

количеством степеней свободы $df = n - k - 1$, где k – количество независимых переменных. В случае, если расчетное значение t -статистики превышает критическое значение для данного уровня значимости (например, 0.05/ 5%), наблюдение считается выбросом. На графике удаленных студентизированных остатков, в отличие от графика с обычными остатками, можно нарисовать линии на уровне в плюс и минус две единицы, чтобы выделить наблюдения за пределами диапазона, в котором, как правило, лежат 95% наблюдений; точки, существенно отдаленные от прямой линии, являются потенциально проблематичными с точки зрения показателя воздействия наблюдения.

Студентизированные остатки для логарифма среднего потребления бензина в расчете на один автомобиль по регионам Российской Федерации изображены на рисунке .



Источник: составлено авторами

Рисунок 7 – Студентизированные остатки для логарифма среднего потребления бензина в расчете на один автомобиль по регионам Российской Федерации

На данном рисунке точек со значением студентизированных остатков, по модулю превышающим 2, несколько. Они относятся к следующим регионам: Костромская область, Ленинградская область, Амурская область, Московская область, Республика Адыгея, Республика Дагестан, Республика Марий Эл, Республика Саха (Якутия), Республика Хакасия, Самарская область, Сахалинская область, Тамбовская область, Тульская область, Ульяновская область, Брянская область, Хабаровский край, город Москва.

Графики частной регрессии, также называемые кривыми частичной регрессии или графиками добавленных переменных, являются еще одним способом обнаружения нетипичных наблюдений. Графики частной регрессии представляют собой серию двумерных графиков регрессии зависимой переменной на каждую из независимых переменных по отдельности.

По итогам проведенных тестов из анализа было удалено четыре региона (г. Москва, Республика Марий Эл, Республика Дагестан, Забайкальский край) как выбросы.

После удаления выбросов зависимая переменная была повторно протестирована на нормальность и ниже можно узнать, удалось ли улучшить ситуацию. Результаты приведены в таблице .

Таблица 7 – Результаты теста Харке-Бера на нормальность панельных данных для переменной среднего потребления бензина в расчете на один легковой автомобиль (64 региона)

Количество наблюдений = 5,376						
Количество репликаций процедуры бутстрап = 500						
(Реплицирование на основе 64 групп по объектам наблюдения)						
	Коэффициент (наблюдаемые)	Стандартная ошибка (бутстрап)	Z	P> z	[95% Доверительный интервал]	
Асимметрия_e	0.1173685	0.2340742	0.50	0.616	-0.3414085	0.5761456
Эксцесс_e	12.69197	4.496257	2.82	0.005	3.879472	21.50447
Асимметрия_u	3.117897	1.032898	3.02	0.003	1.093455	5.14234
Эксцесс_u	13.85125	6.599564	2.10	0.036	0.9163431	26.78616
Совместный тест на нормальность e:			chi2(2) = 8.22	Prob > chi2 = 0.0164		
Совместный тест на нормальность u:			chi2(2) = 13.52	Prob > chi2 = 0.0012		

Согласно результатам тестирования, изложенным в таблице , нет оснований отвергнуть нулевую гипотезу о том, что асимметрия остаточной компоненты ошибки соответствует асимметрии нормального распределения. Остальные же тесты отвергают нулевую гипотезу: эксцесс остаточной компоненты ошибки, а также асимметрия и эксцесс индивидуальной компоненты ошибки не равны соответствующим параметрам нормального распределения. В таблице представлены результаты тестирования на нормальность переменной логарифма среднего потребления бензина в расчете на один легковой автомобиль с помощью теста Харке-Бера для панельных данных.

Таблица 8 – Результаты теста Харке-Бера на нормальность панельных данных для переменной логарифма среднего потребления бензина в расчете на один легковой автомобиль (64 региона)

Количество наблюдений = 5,376						
Количество репликаций процедуры бутстрап = 500						
(Реплицирование на основе 64 групп по объектам наблюдения)						
	Коэффициент (наблюдаемые)	Стандартная ошибка (бутстрап)	z	P> z	[95% Доверительный интервал]	

Асимметрия_e	-0.1202751	0.0649681	-1.85	0.064	-0.2476102	0.0070601
Эксцесс_e	1.380989	0.298341	4.63	0.000	0.7962516	1.965727
Асимметрия_u	0.0104198	0.3392239	0.03	0.975	-0.6544468	0.6752865
Эксцесс_u	-0.0911167	0.5426598	-0.17	0.867	-1.15471	0.9724769
Совместный тест на нормальность e:			chi2(2) = 24.85		Prob > chi2 = 0.0000	
Совместный тест на нормальность u:			chi2(2) = 0.03		Prob > chi2 = 0.9855	

Из таблицы видно, что совместный тест на нормальность индивидуальной компоненты ошибок свидетельствует в пользу принятия нулевой гипотезы, а совместный тест на нормальность остаточной компоненты ошибок – за ее отклонение в силу несоответствия эксцесса e нормальному. Таким образом, существенных улучшений в терминах нормальности удаление выбросов не дало.

Помимо этого, в регрессионной модели могут присутствовать и другие проблемы, такие как гетероскедастичность, автокорреляция и эндогенность.

Автокорреляция в рамках панельных данных имеет два подвида: автокорреляция по времени и пространственная автокорреляция.

Для того, чтобы проверить наличие временной автокорреляции, наиболее часто используется тест серийных автокорреляций Вулдриджа. В качестве нулевой гипотезы предполагается, что в модели отсутствует автокорреляция первого порядка. Заметим, что, как и в случае с тестом Бройша-Годфри, он проверяет автокорреляцию только первого порядка.

Тест Дарбина-Уотсона имеет ряд ограничений, наиболее значимым из которых в нашем случае является то, что этот тест нельзя применить в отношении авторегрессионных моделей, которые в числе регрессоров содержат первый и/или последующий лаги зависимой переменной. Кроме того, тест Дарбина-Уотсона позволяет обнаружить только автокорреляцию первого порядка. Наличие в модели свободного члена является обязательным условием; периодичность наблюдений должна быть одинаковой (пропуски в наблюдениях не допускаются). Панельный тест Дарбина – Уотсона и тест Бокса – Пирса также не применимы из-за большого количества наблюдений в выборке.

Если возмущения линейной модели автокоррелированы, оценки параметров коэффициентов, полученные с помощью обычного МНК, неэффективны, но несмещены. Однако в динамическом уравнении, где в качестве регрессоров используются запаздывающие значения зависимой переменной, оценки, полученные с помощью обычного МНК, смещены и, как правило, несостоятельны. По этой причине важно иметь рабочие тесты на автокорреляцию, особенно когда это динамическая модель, которую предлагается оценивать с помощью МНК. Так как мы имеем среднее количество как объектов ($N < 300$), так и временных периодов ($T = 84$), тестирование пространственной

автокорреляции не имеет особого смысла, так как тесты перестают работать. Так, например, классический тест на пространственную автокорреляцию LM тест Бройша-Пагана не применим в нашем случае, так как количество домохозяйств в нашей выборке слишком велико.

Далее перейдем к выявлению присутствия в модели гетероскедастичности. Для этого проверили модель с фиксированными эффектами при помощи модифицированного теста Вальда. В качестве нулевой гипотезы будет тестироваться предположение о постоянстве σ_i^2 .

Гетероскедастичность в рамках моделей панельных данных может порождаться обеими компонентами ошибки. Гетероскедастичность может возникнуть из-за непостоянной дисперсии индивидуальных эффектов μ_i а также из-за непостоянной дисперсии случайного шока ε_{it} .

Для того, чтобы протестировать наличие гетероскедастичности, вызванной непостоянной дисперсией индивидуальных эффектов, использовался модифицированный тест Вальда. В отличие от ряда других тестов, данный тест не требует нормальности распределения ошибок. Однако, при небольшом количестве временных периодов, но большом объеме выборки, данный тест имеет довольно низкую мощность (т.е. он склонен не отвергать нулевую гипотезу об отсутствии гетероскедастичности). Статистика теста Вальда имеет следующий вид (формула (11)):

$$W = \frac{N \sum_{i=1}^N (\sigma_i^2 - \sigma^2)^2}{\sigma^4} : \chi_N^2, \quad (11)$$

где $\sigma^2 = \frac{1}{NT} \sum_{i=1}^N u_{it}^2$, u_{it}^2 – квадрат остатков модели within.

Проблема гетероскедастичности может быть нивелирована посредством изменения спецификации модели и оценки модели с робастными (устойчивыми) ошибками. Решение проблемы гетероскедастичности позволяет получить более точное представление о значимости коэффициентов модели без изменения их абсолютных значений. Применение оценки модели с робастными ошибками может привести к тому, что значимость ряда коэффициентов изменится.

Проблема эндогенности — наличие связи между регрессорами и ошибкой модели — крайне важна в контексте исследования регрессионного моделирования по причине

возможной неадекватности результатов модели, возникающей в следствии эндогенности регрессоров. По этой причине авторы исследования проверяют собственную модель на предмет присутствия в ней проблемы эндогенности.

В рамках рассматриваемого исследования эндогенными переменными являются цены на бензин и объемы спроса и предложения, поскольку они определяются на рынке бензина одновременно.

В качестве инструментов для решения проблемы эндогенности нам необходимо выбрать такие переменные, которые на текущий момент не включены в модель, но при этом могут потенциально быть связаны с текущими объясняющими переменными из рассматриваемой модели. Наиболее очевидными претендентами на роль инструментов являются шоки предложения, которые на данный момент действительно не присутствуют в модели, но тем не менее обладают высоким потенциалом влияния в отношении переменных, отвечающих за объем предложения бензина.

В таблице приведены описательные статистики переменных, используемых в модели спроса на бензин.

Таблица 9 – Описательные статистики логарифмов переменных модели спроса на бензин с непосредственным учетом количества транспортных средств (64 региона)

Переменная		Среднее значение	Стандартное отклонение	Минимальное значение	Максимальное значение	Медиана	Наблюдения
Логарифм удельного потребления бензина	overall	4.614868	0.9306509	1.585118	7.209063	4.614868	N = 5376
	between		0.9059327	2.033193	6.946075		n = 64
	within		0.2409778	3.534372	5.715773		T = 84
Логарифм лага удельного потребления бензина	overall	4.615373	0.9303311	1.585118	7.209063	4.615373	N = 5312
	between		0.9056283	2.033149	6.946874		n = 64
	within		0.240866	3.538702	5.720103		T = 83
Логарифм ВРП на душу населения	overall	9.732987	0.4510271	8.782701	11.4735	9.732987	N = 5376
	between		0.4434432	9.021053	11.36212		n = 64
	within		0.0990964	9.103069	9.971216		T = 84

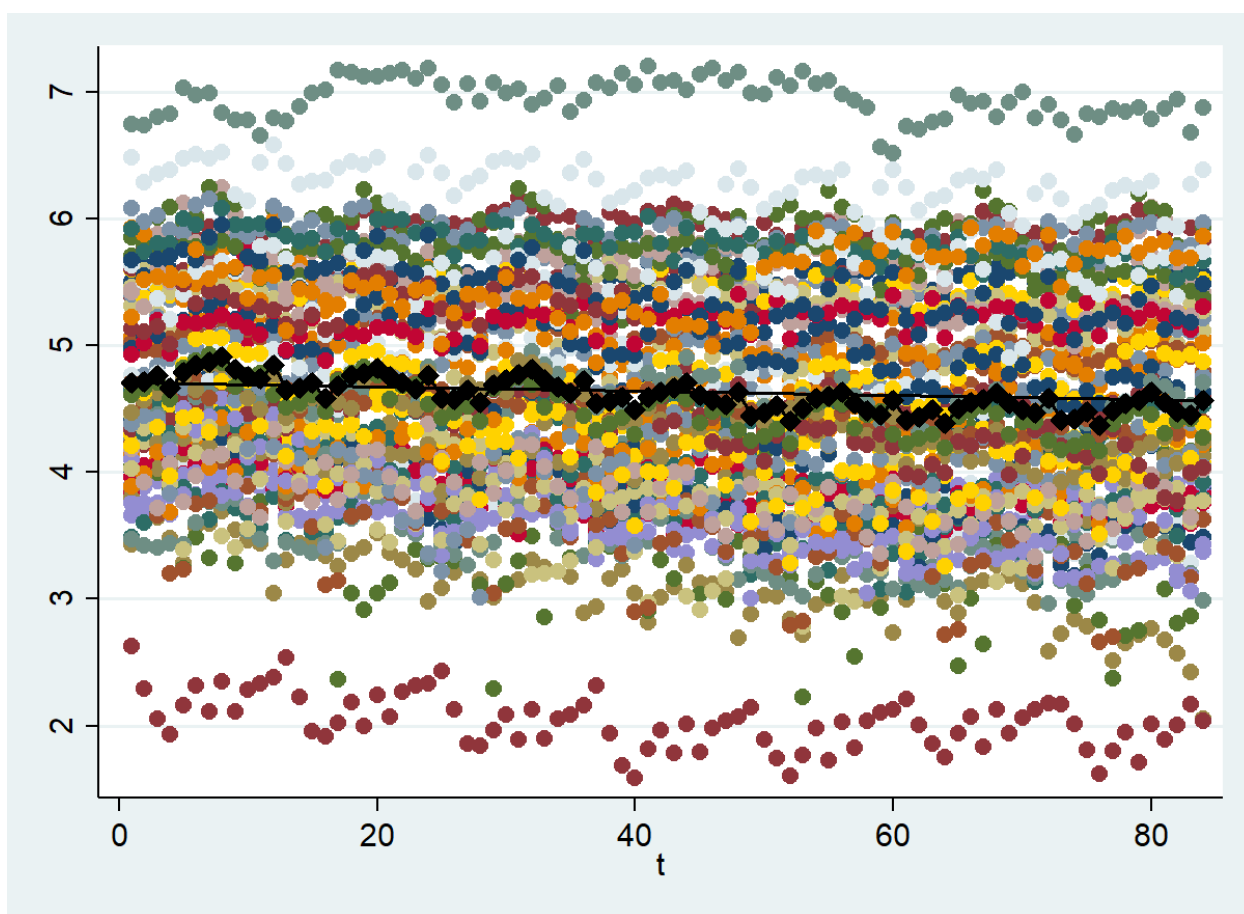
Продолжение таблицы

Логарифм цен на бензин	overall	3.076711	0.071678	2.964242	3.474797	3.076711	N = 5376
	between		0.070498	2.991895	3.37441		n = 64
	within		0.0156369	2.998112	3.223843		T = 84
Логарифм количества авто на душу населения	overall	5.108821	0.6706015	3.47994	6.546028	5.108821	N = 5376
	between		0.662349	3.803948	6.442392		n = 64
	within		0.1333218	4.351767	5.550831		T = 84

Из таблицы описательных статистик видно, что дисперсия логарифма удельного потребления бензина по регионам ("between") превосходит дисперсию по времени ("within"). Такая особенность присуща всем переменным модели. Это означает, что скорее

объекты выборки отличаются друг от друга, чем что-то происходит с ними во времени. Это дает основания предположить, что лучше будет описывать имеющиеся данные модель, учитывающая индивидуальные детерминированные эффекты.

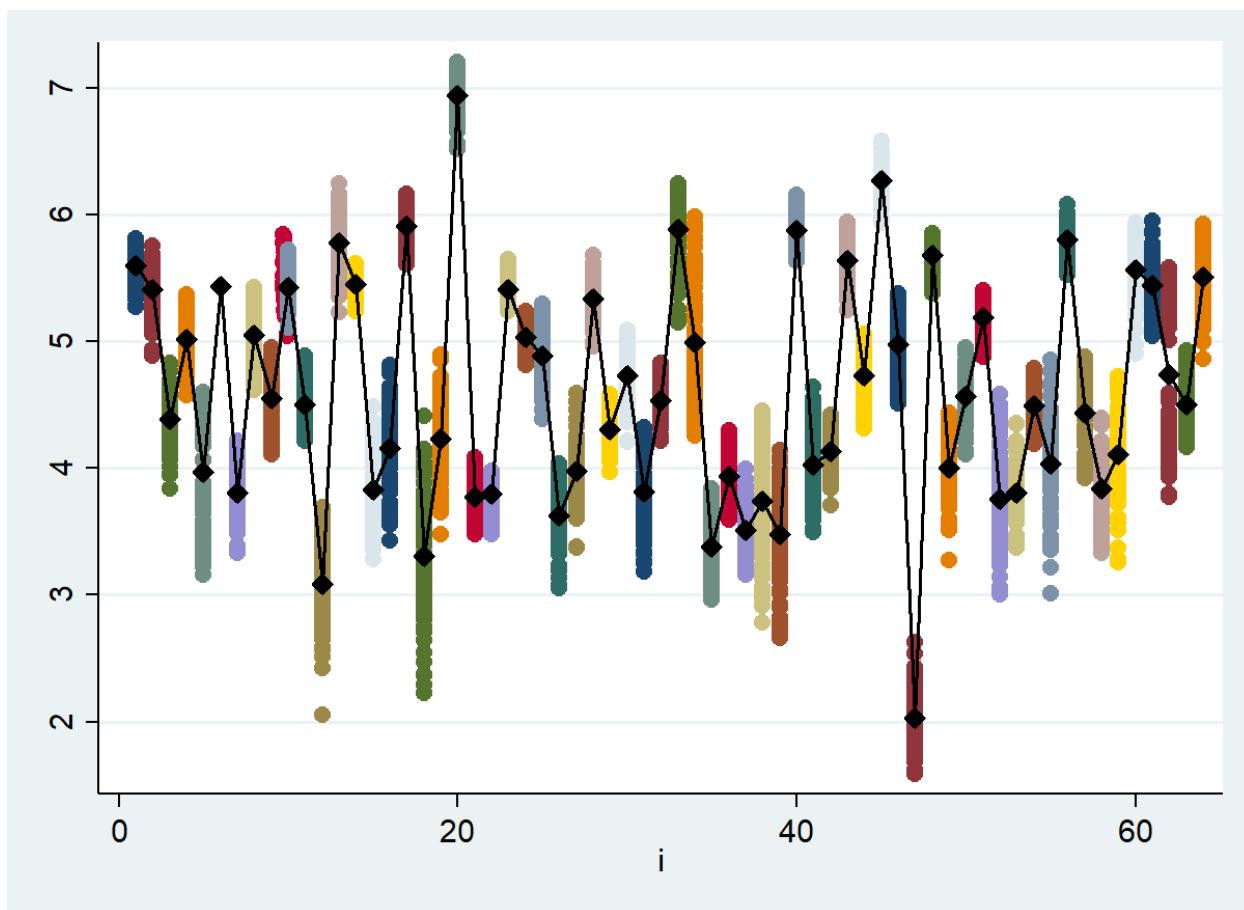
На рисунках и наглядно продемонстрирована гетерогенность наблюдений по времени и по регионам соответственно. На рисунке показаны значения объясняемой переменной в каждый период времени их и среднее (черным квадратом). Таким образом, видно, что среднее значение логарифма среднего потребления бензина в расчете на один автомобиль практически не меняется со временем, однако, несет в себе признаки сезонности.



Источник: составлено авторами

Рисунок 8 – Неоднородность по времени логарифма среднего потребления бензина в расчете на один автомобиль

На рисунке приведены значения логарифма среднего потребления бензина в расчете на один автомобиль в каждом из регионов, среднее по региону за рассматриваемый период обозначено черным квадратом. Видно, что неоднородность значений анализируемой зависимой переменной выше именно в региональном разрезе, чем по времени.



Источник: составлено авторами

Рисунок 9 – Неоднородность по регионам логарифма среднего потребления бензина в расчете на один автомобиль

Далее проводится исследование тесноты связи и характера взаимной зависимости используемых в анализе переменных. И для начала внимание обращено на коэффициенты частной корреляции между переменными модели спроса на бензин с непосредственным учетом количества транспортных средств. Они представлены в таблице .

Таблица 10 – Коэффициенты частной корреляции между переменными модели спроса на бензин с непосредственным учетом количества транспортных средств

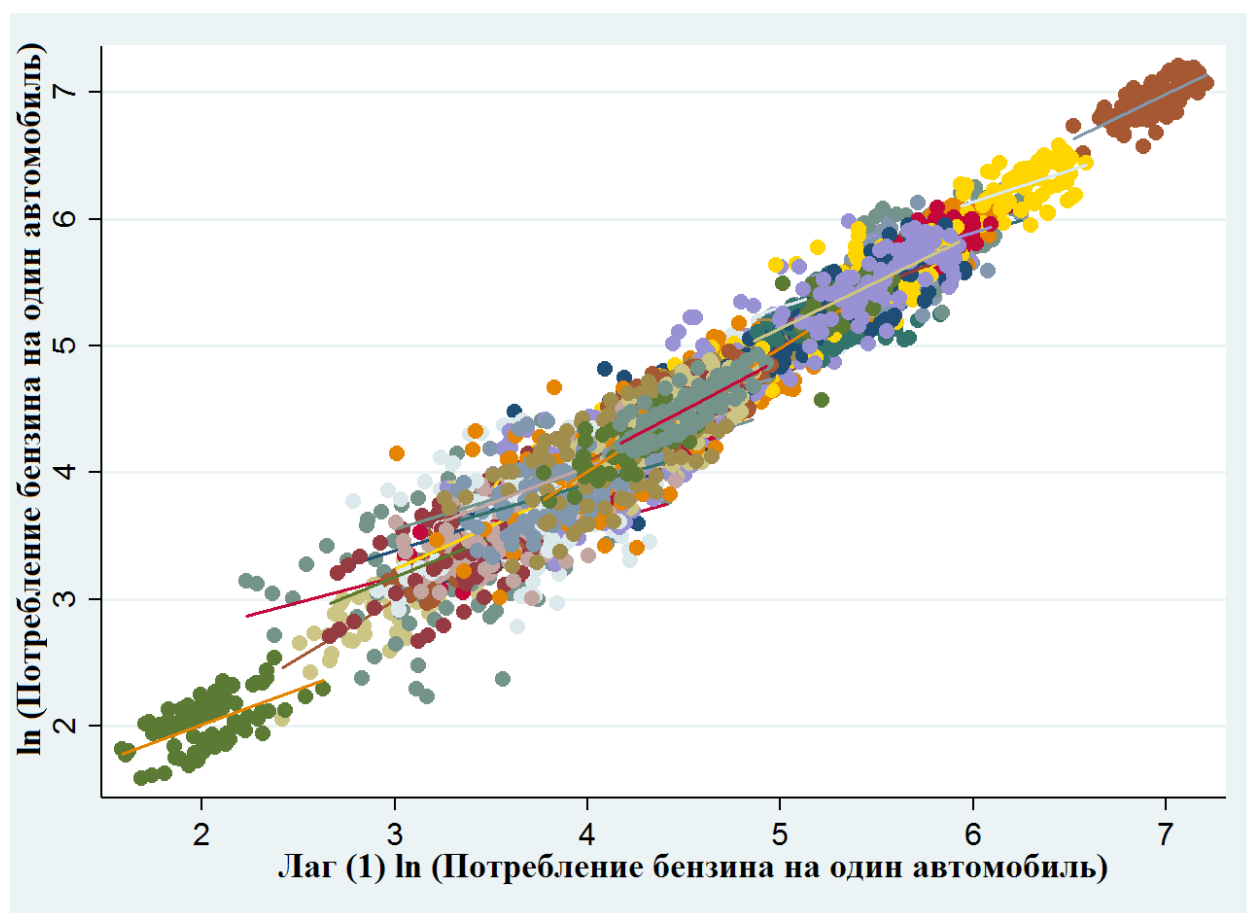
	Логарифм удельного потребления бензина	Логарифм удельного потребления бензина	Логарифм ВРП на душу населения	Логарифм цен на бензин	Логарифм количества авто на душу населения
Логарифм удельного потребления бензина	1.0000	0.9773	0.1010	-0.2545	-0.8305
Логарифм удельного потребления бензина	0.9773	1.0000	0.1011	-0.2531	-0.8309
Логарифм ВРП на душу населения	0.1010	0.1011	1.0000	0.4962	-0.0456
Логарифм цен на бензин	-0.2545	-0.2531	0.4962	1.0000	0.2136

Продолжение таблицы

Логарифм количества авто на одного водителя	-0.8305	-0.8309	-0.0456	0.2136	1.0000
---	---------	---------	---------	--------	--------

Согласно таблице, показывающей коэффициенты частной корреляции между переменными модели спроса на бензин, сильной взаимосвязи между регрессорами не наблюдается, то есть коэффициенты корреляции не превышают 0.5. Исключения составляют лаг логарифма удельного потребления бензина и логарифм количества автомобилей на одного водителя, для которых никак нельзя исключить сильную корреляцию (-0.8309). Однако, этот факт лишь подтверждает выдвинутые предположения об инерционности процесса изменения топливной эффективности автопарка и постоянстве привычек.

Теперь более пристальное внимание будет обращено на то, как каждый из регрессоров влияет на зависимую переменную. Итак, были построены соответствующие диаграммы рассеяния для каждого из регионов на одном графике и предсказаны линейные регрессии логарифма среднего потребления бензина в расчете на один автомобиль на каждую из объясняющих переменных. Наблюдаемые результаты будут проинтерпретированы применительно к краткосрочному периоду. Сначала был рассмотрен лаг зависимой переменной (рисунок).

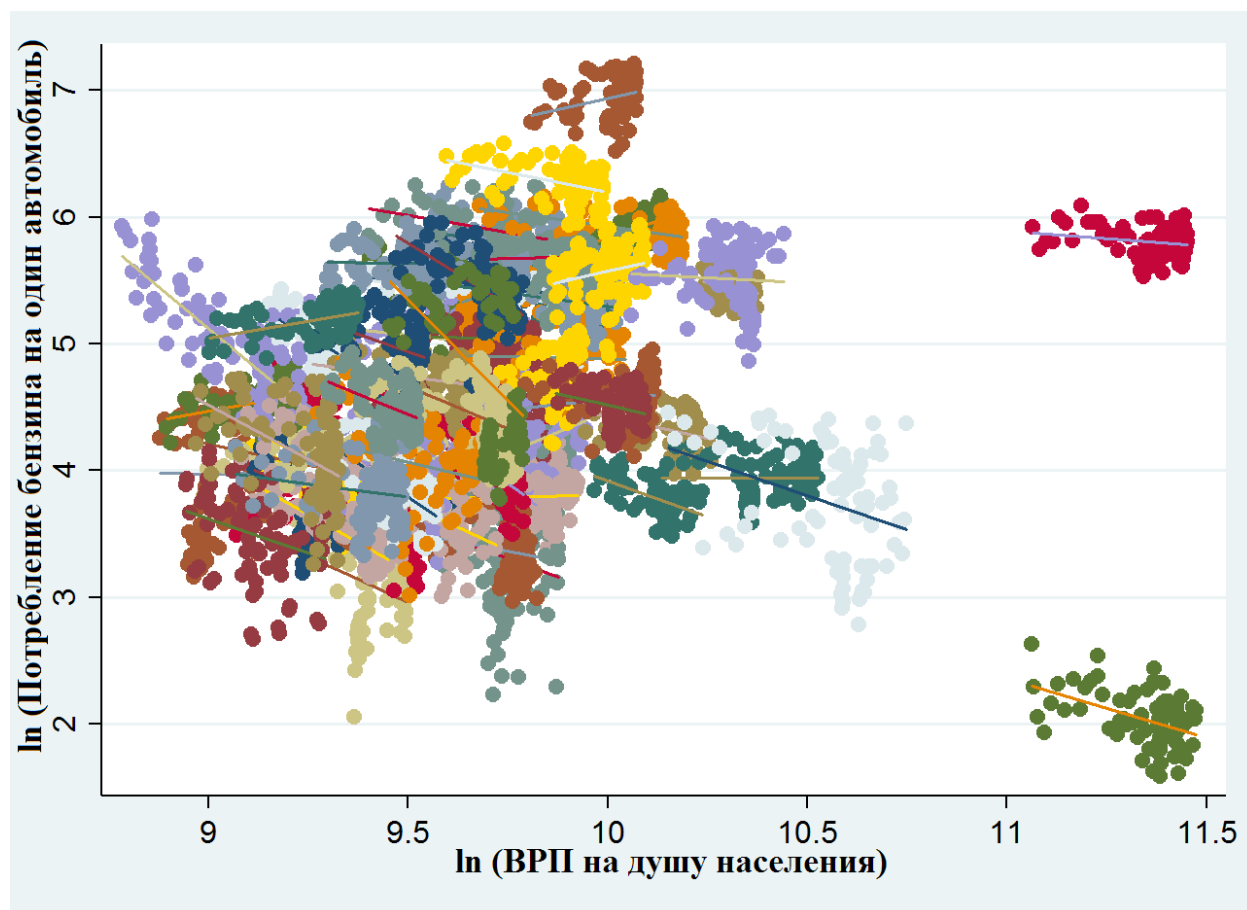


Источник: составлено авторами

Рисунок 10 – Диаграммы рассеяния и графики линейной регрессии логарифма среднего потребления бензина в расчете на один автомобиль на лаг этой переменной

Из рисунка можно заметить, что в данном случае наблюдается явная линейная зависимость. Линии линейной регрессии для всех включенных в анализ регионов имеют положительный наклон, углы наклона этих линий практически совпадают, а уровни отличаются. В целом тангенс угла наклона близок к единице. Таким образом, можем сделать вывод, что среднее потребление бензина в расчете на один автомобиль прямо зависит от значения этого показателя в предыдущем периоде. Более того, его изменение в предыдущем периоде приводит к близкому к равнозначному изменению в текущем периоде, что соответствует выдвинутой нами гипотезе (см. таблица) и подтверждается очень высоким коэффициентом частной корреляции (0.9773) рассматриваемых переменных.

На рисунке представим аналогичные построения для зависимости среднего потребления бензина в расчете на один автомобиль от логарифма реального ВРП на душу населения.

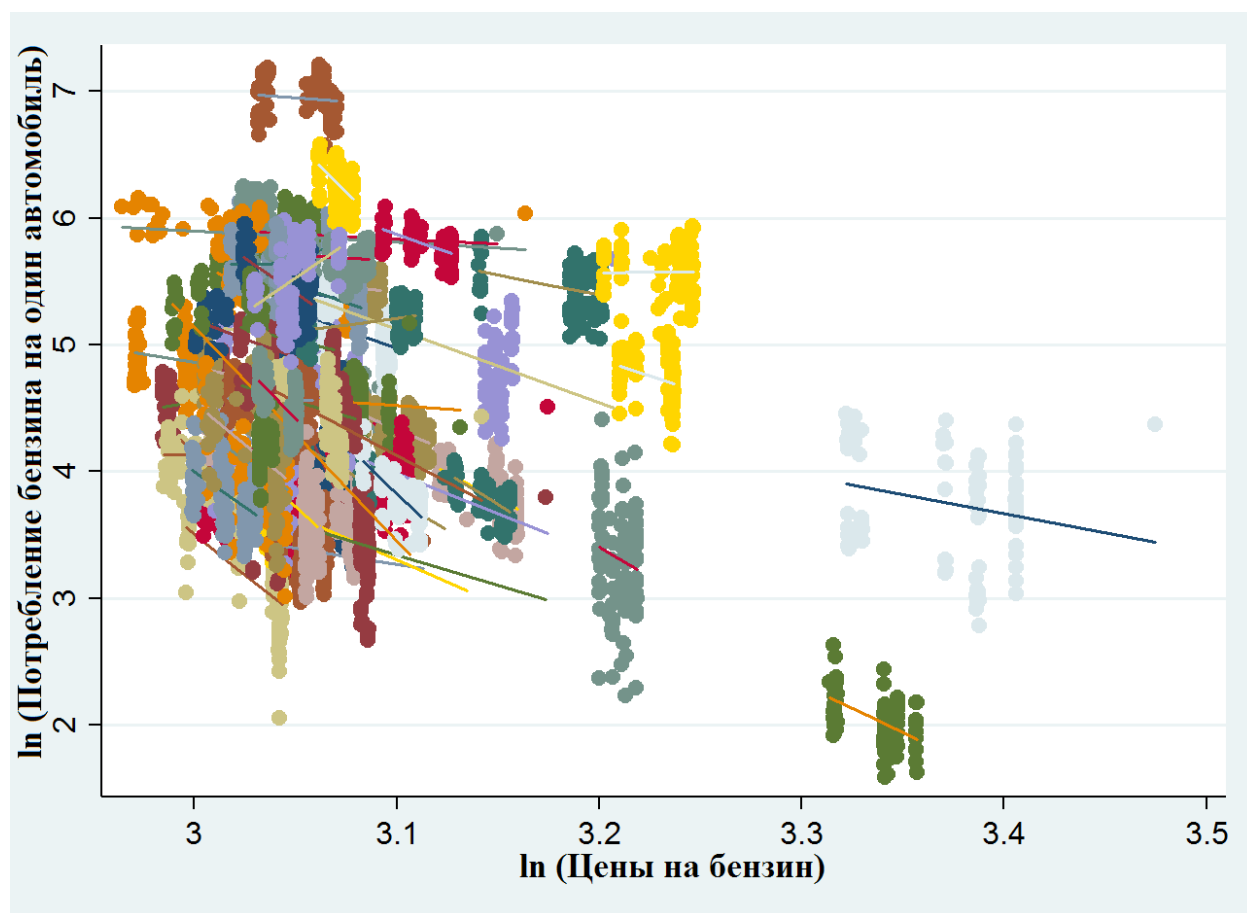


Источник: составлено авторами

Рисунок 11 – Диаграммы рассеяния и графики линейной регрессии логарифма среднего потребления бензина в расчете на один автомобиль на логарифм реального ВРП на душу населения

В данном случае зависимость не такая очевидная – облака точек для разных регионов вытянуты в разные стороны и, соответственно, графики линейной регрессии имеют разный наклон, что согласуется с довольно низким значением коэффициента частной корреляции рассматриваемых переменных. Совокупное облако точек для всех регионов вытянуто в положительном направлении, угол его наклона небольшой. То есть, с увеличением реального ВРП на душу населения на 1% среднее потребление бензина в расчете на один автомобиль возрастет гораздо меньше, чем на 1%. Это соответствовало бы предположению о низкой эластичности спроса на бензин по доходу, поскольку автомобильное топливо является нормальным благом и относится к товарам первой необходимости. Однако, графики линейной регрессии для регионов по отдельности в большинстве своем говорят о наличии обратной зависимости. Поэтому по данному рисунку сложно делать однозначные выводы.

Перейдем к диаграммам рассеяния и графикам линейной регрессии логарифма среднего потребления бензина в расчете на один автомобиль на логарифм реальных цен на бензин (рисунок).

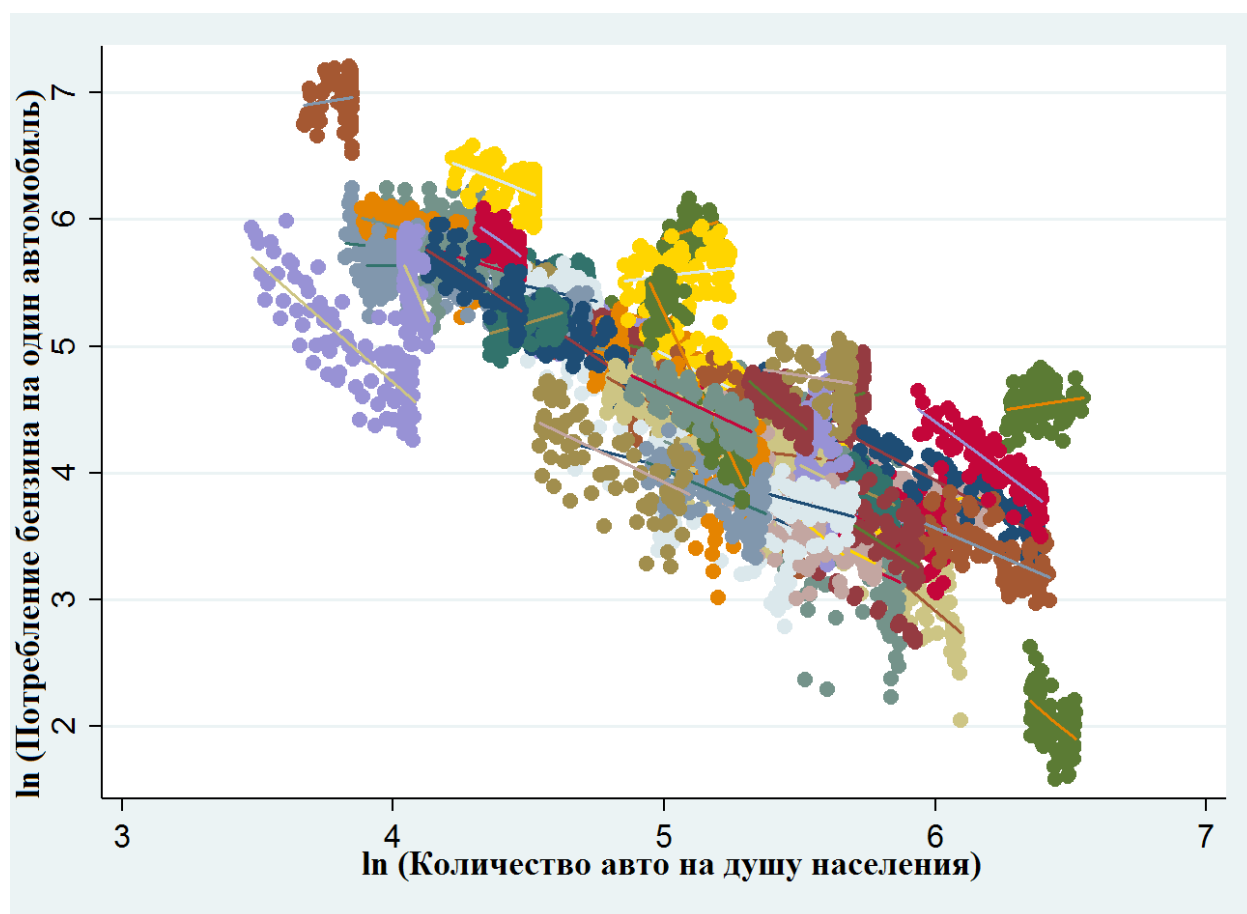


Источник: составлено авторами

Рисунок 12 – Диаграммы рассеяния и графики линейной регрессии логарифма среднего потребления бензина в расчете на один автомобиль на логарифм реальных цен на бензин

Из рисунка можно сделать вывод об обратном характере влияния логарифма реальных цен на бензин на логарифм среднего потребления бензина в расчете на один автомобиль, что отвечает значению коэффициента частной корреляции. Тангенс угла наклона графиков линейной регрессии для многих регионов близок к -1. Из этого следует, что при повышении цен на бензин на 1% спрос на него снизится примерно на 1%. Данные выводы согласуются с законом спроса и дают основания предположить эластичность спроса на бензин по цене близкой к единичной.

На рисунке представлены диаграммы рассеяния и графики линейной регрессии логарифма среднего потребления бензина в расчете на один автомобиль на логарифм количества автомобилей в расчете на душу населения.



Источник: составлено авторами

Рисунок 13 – Диаграммы рассеяния и графики линейной регрессии логарифма среднего потребления бензина в расчете на один автомобиль на логарифм количества автомобилей в расчете на душу населения

На рисунке можно наблюдать явно выраженное обратное направленное влияние логарифма количества автомобилей в расчете на душу населения на логарифм среднего потребления бензина в расчете на один автомобиль, что подтверждается высоким значением коэффициента частной корреляции (-0.8305) этих переменных. Тангенс угла наклона графиков линейной регрессии для многих регионов по модулю меньше 1. Это

дает основания полагать, что при увеличении показателя количества автомобилей на душу населения на 1% среднее потребление бензина в расчете на один автомобиль снизится менее, чем на 1%, что также соответствует выдвинутым гипотезам (см. таблица) и логике модели.

Перейдем теперь к оцениванию уравнения спроса на бензин с непосредственным учетом количества транспортных средств. И в первую очередь необходимо выяснить, какой из основных типов моделей позволяет лучше описать особенности, присущие анализируемой панели данных.

Сначала была рассмотрена модель индивидуальных регрессий. Часто модели индивидуальных регрессий по объектам дают несостоятельные и не робастные оценки при небольших количествах наблюдений по времени, однако имеющаяся база данных охватывает 64 региона, каждый из которых наблюдается достаточно большое количество временных периодов, а именно 84, благодаря чему едва ли придется столкнуться с данной проблемой. Кроме того, данные модели имеют неоспоримое преимущество – модели индивидуальных регрессий наиболее подробно из всех описанных в данном исследовании моделей учитывают существующую гетерогенность объектов. Однако, если требуется получить результаты, которые можно было бы распространять на всю совокупность объектов в условиях панельных данных, оценивание отдельных индивидуальных регрессий не является оправданным. В общем случае уравнение линейной регрессии принимает вид, описанный формулой (12):

$$Y_{it} = \alpha_i + \beta_i X_{it} + \varepsilon_{it} \quad , \quad i=1, \dots, N; \quad t=1, \dots, T \quad (12)$$

где α – свободный член модели,

β – вектор коэффициентов,

X – вектор объясняющих переменных,

ε – случайная ошибка,

i – индикатор (номер) объекта,

t – индикатор (номер) временного такта.

Следующий тип моделей – модель сквозной регрессии. Данный тип моделей не учитывает панельную структуру данных и воспринимает их как пространственную выборку. Это часто не позволяет получить состоятельных оценок коэффициентов. Однако, в случае небольшой панели, где из-за малого числа наблюдений выделение гетерогенности не оправдано и ведет к значительному ухудшению качества модели, применение моделей

данного типа целесообразно. Более того, если гетерогенность отсутствует, мы получим не только состоятельные, но и эффективные оценки. Уравнение линейной регрессии в общем случае принимает вид, заданный формулой (13):

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} \quad (13)$$

Представим в таблице результаты оценивания модели сквозной регрессии. Как можно заметить, регрессия значима в целом. Доля дисперсии зависимой переменной, объясняемая с помощью данной модели, составляет 0.9563. Каждый из регрессоров, как и константа, значим на однопроцентном уровне значимости.

Таблица 11 – Результаты оценивания модели сквозной регрессии обычным методом наименьших квадратов

Линейная регрессия			Количество наблюдений = 5,312			
			F(4, 5307)	=	29627.04	
			Prob > F	=	0.0000	
			R-squared	=	0.9563	
			Root MSE	=	0.19475	
lngaspercar	Коэффициент	Стандартная ошибка (робастные)	t	P> t	[95% Доверительный интервал]	
lag1	0.9230501	0.0061021	151.27	0.000	0.9110874	0.9350127
Lnvrp	0.024832	0.0064762	3.83	0.000	0.012136	0.037528
Lngasprice	-0.1832213	0.0647573	-2.83	0.005	-0.3101722	-0.0562704
lncarperdriver	-0.085072	0.0078771	-10.80	0.000	-0.1005143	-0.0696296
_cons	1.110337	0.1826177	6.08	0.000	0.7523314	1.468343

Примечание – указаны робастные стандартные ошибки

Перейдем к модели с детерминированными индивидуальными эффектами. Данный вид моделей представляет собой некий компромисс между моделью индивидуальных регрессий и моделью сквозной регрессии, позволяя учесть пообъектную гетерогенность. Уравнение линейной регрессии в общем случае принимает вид (формула (14)):

$$Y_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it}, \quad E(\alpha_i) = \text{const}, \text{var}(\alpha_i) = 0 \quad (14)$$

В таблице приведены результаты оценивания регрессии с фиксированными эффектами с помощью обычного метода наименьших квадратов.

Таблица 12 – Результаты оценивания регрессии с фиксированными индивидуальными эффектами обычным методом наименьших квадратов

Регрессия с фиксированными эффектами (FE/within)	Количество наблюдений = 5,312
Переменная группировки: i	Количество групп = 64
Коэффициент детерминации:	Количество наблюдений в группе:

within = 0.4747	минимальное = 83.0
between = 0.9633	среднее = 83.0
совокупный = 0.9297	максимальное = 83.0
	(Std. Err. adjusted for 64 clusters in i)
	F(4, 63) = 124.15
corr(u_i, Xb) = 0.6031	Prob > F = 0.0000

Продолжение таблицы

Lngaspcar	Коэффициент	Стандартная ошибка (робастные)	t	P> t	[95% Доверительный интервал]	
lag1	0.5471889	0.0455203	12.02	0.000	0.4562238	0.6381541
Lnvrp	0.132074	0.0460948	2.87	0.006	0.039961	0.2241871
Lngasprice	-1.05299	0.2911469	-3.62	0.001	-1.634801	-0.4711797
Lncarperdriver	-0.3879447	0.0470361	-8.25	0.000	-0.4819388	-0.2939506
_cons	6.025415	0.8935774	6.74	0.000	4.239743	7.811087
sigma_u	0.22110807					
sigma_e	0.17537208					
Rho	0.61384022	(доля дисперсии, обусловленной u_i)				

Данная регрессия значима в целом, совокупный коэффициент детерминации составляет 0.9297. Все включенные в модель регрессоры и константа значимы на однопроцентном уровне значимости. Ценовая эластичность спроса на бензин получилась равной -1.05299, а эластичность по доходу – 0.132074.

Модель со случайными индивидуальными эффектами. Еще один вид моделей, являющийся обобщением модели с детерминированными эффектами – модель со случайными эффектами. В отличие от модели с фиксированными эффектами, данная модель имеет в своей основе достаточно сильные предположения о некоррелированности индивидуальных эффектов с регрессорами. Если данное предположение верно, то полученные оценки можно считать состоятельными и эффективными. Однако, при нарушении этого предположения оценки окажутся несостоятельными. Уравнение регрессии в общем виде представлено формулой (15):

$$Y_{it} = \beta X_{it} + u_{it}, \text{ где } u_{it} = \alpha_i + \varepsilon_{it}, \quad E(\alpha_i) = 0, \text{ var}(\alpha_i) = \sigma_\alpha^2 \quad (15)$$

где α_i и ε_{it} – случайные возмущения between и within, независимые между собой и не зависящие от X_{it} .

Результаты оценивания модели со случайными эффектами приведены в таблице . Согласно этим результатам, регрессия значима в целом; переменная $\ln v_{tr}$, отвечающая за логарифм реального ВРП на душу населения, незначима; переменная $\ln gasprice$, означающая логарифм реальных цен на бензин, значима лишь на 10-процентном уровне значимости. Остальные объясняющие переменные значимы на однопроцентном уровне значимости.

Таблица 13 – Результаты оценивания регрессии со случайными эффектами обобщенным методом наименьших квадратов

Регрессия со случайными эффектами (RE, GLS), оцененная ОМНК			Количество наблюдений = 5,312			
Переменная группировки: i			Количество групп = 64			
Коэффициент детерминации:			Количество наблюдений в группе:			
within = 0.4465			минимальное = 83.0			
between = 0.9987			среднее = 83.0			
совокупный = 0.9563			максимальное = 83.0			
			Wald chi2(4) = 28541.19			
corr(u_i, X) = 0 (assumed)			Prob > chi2 = 0.0000			
lngaspercar	Коэффициент	Стандартная ошибка	z	P> z	[95% Доверительный интервал]	
lag1	0.9230501	0.0185563	49.74	0.000	0.8866804	0.9594198
lnvrp	0.024832	0.0168062	1.48	0.140	-0.0081075	0.0577715
lngasprice	-0.1832213	0.1055905	-1.74	0.083	-0.3901749	0.0237323
lncarperdriver	-0.085072	0.0236924	-3.59	0.000	-0.1315082	-0.0386357
_cons	1.110337	0.3722816	2.98	0.003	0.3806788	1.839996
sigma_u	0.000					
sigma_e	0.17537208					
rho	0.000	(доля дисперсии, обусловленной u_i)				

Для систематизации результатов и выбора оптимального вида модели необходимо провести ряд тестов, позволяющих попарно сравнить модели между собой.

Для сравнения модели сквозной регрессии (Pool) и модели с индивидуальными эффектами (FE) необходимо провести обычный F – тест (тест Вальда) на равенство всех индивидуальных эффектов нулю, нулевая гипотеза в котором будет следующей: H_0 :

$\alpha_i = 0$. Для этого необходимо получить суммы квадратов остатков двух моделей и построить статистику вида (формула ()):

$$F = \frac{(RSS_{pool} - RSS_{FE}) / (N - 1)}{RSS_{FE} / (NT - (N - 1) - (k + 1))}, \quad \sim F(N - 1, NT - N - k) \quad (16)$$

где

RSS_{pool} – сумма квадратов остатков модели сквозной регрессии;

RSS_{FE} – сумма квадратов остатков модели с фиксированными индивидуальными эффектами;

N – количество объектов наблюдения;

T – количество временных тактов;

k – количество оцениваемых в модели параметров.

В таблице представлены результаты F-теста на сравнение модели сквозной регрессии и модели с индивидуальными фиксированными эффектами.

Таблица 14 – Результаты F-теста на равенство индивидуальных эффектов нулю (сравнение модели сквозной регрессии и модели с индивидуальными фиксированными эффектами)

Регрессия с фиксированными эффектами (FE/within)			Количество наблюдений = 5,312			
Переменная группировки: i			Количество групп = 64			
Коэффициент детерминации:			Количество наблюдений в группе:			
within = 0.4747			минимальное = 83.0			
between = 0.9633			среднее = 83.0			
совокупный = 0.9297			максимальное = 83.0			
			F(4, 5244) = 1184.53			
corr(u_i, Xb) = 0.6031			Prob > F = 0.0000			
Lngaspcar	Коэффициент	Стандартная ошибка	t	P> t	[95% Доверительный интервал]	
lag1	0.5471889	0.0115181	47.51	0.000	0.5246086	0.5697692
Lnvrp	0.132074	0.0347219	3.80	0.000	0.0640046	0.2001435
Lngasprice	-1.05299	0.2102449	-5.01	0.000	-1.465158	-0.6408226
Lncarperdriver	-0.3879447	0.0277246	-13.99	0.000	-0.4422965	-0.3335929
_cons	6.025415	0.5563235	10.83	0.000	4.93479	7.116041
sigma_u	0.22110807					
sigma_e	0.17537208					
Rho	0.61384022	(доля дисперсии, обусловленной u_i)				
F-тест на равенство всех u_i=0:			F(63, 5244) = 20.65		Prob > F = 0.0000	

Согласно F – тесту на наличие фиксированных эффектов, модель с фиксированными эффектами лучше описывает данные по сравнению с моделью сквозной регрессии.

Для сравнения модели сквозной регрессии (Pool) и модели со случайными эффектами (RE) был проведен тест Бройша – Пагана (тест множителей Лагранжа), нулевая

гипотеза которого говорит об отсутствии индивидуальных эффектов: $H_0: \sigma_u^2 = 0$.
Статистика для теста Бройша – Пагана выглядит следующим образом (формула (17)):

$$LM = \left(\frac{\partial l}{\partial \theta} \right)^T \left(-E \left[\frac{\partial^2 l}{\partial \theta \partial \theta'} \right] \right)^{-1} \left(\frac{\partial l}{\partial \theta} \right), \quad \sim \chi_1^2 \quad (17)$$

Результаты теста Бройша – Пагана на сравнение модели сквозной регрессии и модели со случайными эффектами представлены в таблице.

Таблица 15 – Результаты теста Бройша – Пагана на сравнение модели сквозной регрессии и модели со случайными эффектами

Тест Бройша – Пагана (тест множителей Лагранжа) на наличие случайных эффектов		
lngaspercar [i, t] = Xb + u [i] + e [i, t]		
Результаты оценивания:		
	Var	sd = sqrt(Var)
Lngaspercar	0.8681487	0.931745
E	0.0307554	0.1753721
U	0	0
Тестируемая гипотеза:		
Var(u) = 0		
chibar2(01) = 0.00		
Prob > chibar2 = 1.0000		

Согласно тесту Бройша-Пагана, модель со случайными эффектами лучше описывает данные по сравнению с моделью сквозной регрессии.

Для того, чтобы сравнить модели с фиксированными эффектами (FE) и со случайными эффектами (RE), необходимо провести тест Хаусмана, в ходе которого проверяется наличие систематической разницы в оценках двух методов. При принятии нулевой гипотезы была выбрана модель со случайными эффектами ввиду эффективности ее оценок в данном случае. Для проведения данного теста, необходимо рассчитать статистику равную (формула (18)):

$$\left(\hat{\beta}_{IV} - \hat{\beta}_{OLS} \right)^T \left(\hat{V}_{\hat{\beta}_{IV}} - \hat{V}_{\hat{\beta}_{OLS}} \right)^{-1} \left(\hat{\beta}_{IV} - \hat{\beta}_{OLS} \right), \quad \sim \chi^2(rank(\hat{V}_{\hat{\beta}_{IV}} - \hat{V}_{\hat{\beta}_{OLS}})) \quad (18)$$

В таблице продемонстрированы результаты теста Хаусмана на сравнение модели с фиксированными эффектами и модели со случайными эффектами.

Таблица 16 – Результаты теста Хаусмана на сравнение модели с фиксированными эффектами и модели со случайными эффектами

	Коэффициенты			
	(b)	(B)	(b-B)	sqrt (diag (V_b-V_B))
	FE	RE	Разность	Стандартная ошибка
lag1	0.5471889	0.9230501	-0.3758611	0.0102039

Lnvtp	0.132074	0.024832	0.107242	0.0339841
Lngasprice	-1.05299	-0.1832213	-0.869769	0.205171
Lncarperdrive r	-0.3879447	-0.085072	-0.3028728	0.0267771
b = состоятельны и при нулевой гипотезе (H ₀), и при альтернативной (H _a)				
B = несостоятельны при альтернативной гипотезе (H _a), состоятельные и эффективные при нулевой гипотезе (H ₀)				
Тестируемая гипотеза:				
H ₀ : разница в оценках коэффициентов не систематическая				

Продолжение таблицы

	$\chi^2(4) = (b-B)' [(V_b - V_B)^{-1}] (b-B)$
	= 1365.34
	Prob > χ^2 = 0.0000

Согласно тесту Хаусмана, модель со случайными эффектами дает несостоятельные оценки по сравнению с моделью с детерминированными эффектами.

Суммируя полученные результаты, можно сделать вывод, что оптимальной моделью является модель с детерминированными индивидуальными эффектами, которая позволяет учитывать гетерогенность объектов наблюдения. То есть, как и предполагалось при изучении описательных статистик данных, потребление бензина по регионам отличается друг от друга довольно значительно.

Далее спецификация с фиксированными индивидуальными эффектами была оценена различными методами. Оценки параметров спроса на бензин в уравнении () представлены в таблице .

Таблица 17– Эмпирические результаты оценивания модели спроса на бензин с непосредственным учетом количества транспортных средств

	Краткосрочный период					Долгосрочный период		
	1-ый лаг зависимой переменной	ВРП на душу населения	Цена на бензин	Кол-во автомобилей на человека	R-квадрат	ВРП на душу населения	Цена на бензин	Кол-во автомобилей на человека
OLS	0.923*** (0.006)	0.025*** (0.006)	-0.183*** (0.064)	-0.085*** (0.008)	0.956	0.323 (0.084)	-2.381 (0.842)	-1.106 (0.102)
Within	0.547*** (0.046)	0.132*** (0.046)	-1.053*** (0.291)	-0.388*** (0.047)	0.929	0.292 (0.102)	-2.325 (0.643)	-0.857 (0.104)
LSDVc	0.565*** (0.011)	0.124*** (0.032)	-1.013*** (0.165)	-0.372*** (0.028)		0.285 (0.072)	-2.328 (0.379)	-0.855 (0.064)
GLS	0.923*** (0.019)	0.025 (0.017)	-0.183* (0.106)	-0.085*** (0.024)	0.956	0.323 (0.218)	-2.381 (1.372)	-1.106 (0.308)
GLS-HC	0.600*** (0.005)	0.084*** (0.007)	-0.768*** (0.046)	-0.316*** (0.007)		0.212 (0.092)	-1.921 (0.596)	-0.791 (0.093)
Within-2SLS	0.376* (0.200)	0.184** (0.088)	-1.244*** (0.386)	-0.552*** (0.188)	0.889	0.295 (0.141)	-1.994 (0.620)	-0.884 (0.301)
FD-2SLS	0.147 (0.124)	1.441*** (0.271)	-1.589*** (0.505)	-1.555*** (0.278)	0.649	1.689 (0.318)	-1.863 (0.592)	-1.824 (0.326)
FD-GMM	0.483*** (0.044)	0.236*** (0.066)	-1.648*** (0.428)	-0.433*** (0.064)		0.457 (0.127)	-3.188 (0.828)	-0.837 (0.123)
sys-GMM	0.501*** (0.040)	0.373*** (0.098)	-1.733*** (0.442)	-0.496*** (0.075)		0.748 (0.197)	-3.474 (0.887)	-0.994 (0.151)

Примечание

1 Уровни значимости *** – 1%; ** – 5%; * – 10%.

- 2 OLS – сквозная регрессия, оцененная методом наименьших квадратов (МНК);
- 3 Within – регрессия с фиксированными эффектами, оцененная методом наименьших квадратов;
- 4 LSDVc – регрессия с фиксированными эффектами и поправкой на смещение;
- 5 GLS – регрессия со случайными эффектами, оцененная обобщенным методом наименьших квадратов (ОМНК);
- 6 GLS-HC – регрессия с фиксированными эффектами, оцененная доступным ОМНК;
- 7 Within-2SLS – регрессия с фиксированными эффектами, оцененная двухшаговым МНК;
- 8 FD-2SLS – метод Андерсона – Хсiao;
- 9 FD-GMM – метод Ареллано – Бонда;
- 10 sys-GMM – метод Бланделла – Бонда.

Таким образом, предполагая экзогенность объясняющих переменных, для проведения исследования использовались следующие методы оценивания: сквозная регрессия, оцененная обычным методом наименьших квадратов, которая не учитывает какие-либо эффекты; регрессия с фиксированными эффектами, оцененная обычным методом наименьших квадратов; регрессия с фиксированными эффектами и поправкой на смещение; а также обобщённый метод наименьших квадратов, при котором предполагается, что специфичные для объектов эффекты случайны.

Для оценки параметров спроса на бензин в уравнении () были использованы робастные оценки дисперсионно-ковариационной матрицы для всех методов, за исключением регрессии с фиксированными эффектами и поправкой на смещение. В данном случае применялись бутстрапированные оценки.

Состоятельность применяемых стандартных методов оценивания панельных данных зависит от отсутствия автокорреляции вектора ошибок. Проведенные тесты показывают отсутствие автокорреляции остатков первого порядка.

Сквозная регрессия, оцененная обычным методом наименьших квадратов, включена в анализ для целей сравнения. Она дает самый большой коэффициент при лаге зависимой переменной, что соответствует результатам работ Baltagi и Griffin (1997) [CITATION Bal97 \l 1049] и Baltagi et al. (2003) [CITATION Bal031 \l 1049], поскольку имеет место смещение из-за упущенных региональных эффектов (результаты оценивания представлены в таблице). То же самое верно для регрессии со случайными эффектами, оцененной обобщенным методом наименьших квадратов (результаты содержатся в таблице).

При оценивании регрессии со случайными эффектами обобщенным методом наименьших квадратов влияние валового регионального продукта на душу населения оказывается незначимым, а при оценивании сквозной регрессии обычным методом наименьших квадратов – значимым на 1-процентном уровне значимости. Кроме того, краткосрочная эластичность спроса в зависимости от количества автомобилей на душу населения, найденная с помощью этих двух методов, кажется недооцененной по сравнению с оценками других методов, а долгосрочные эластичности завышены из-за более низкой скорости адаптации к долгосрочному равновесию.

Оценка методом Within учитывает специфические для регионов фиксированные эффекты и нейтрализует их возможную корреляцию с объясняющими переменными, в связи с чем данный метод находит широкое применение. Результаты оценивания модели посредством данного метода представлены в таблице . Все коэффициенты являются разумными по величине, знаку и значимости (однако необходимо учитывать тот факт, что значение t -статистик искажено для всех рассмотренных оценок.). Краткосрочная эластичность по цене равна -1.053, а долгосрочная – (- 2.325). Краткосрочный коэффициент эластичности по количеству автомобилей на душу населения – (-0.388). Для сравнения значение этого показателя, полученное Baltagi et. al. (2003) [CITATION Bal031 \l 1049], равно -0.240; Baltagi и Griffin (1997) [CITATION Bal97 \l 1049] – (-0.74), Рокк (2010) [CITATION Рок10 \l 1049] – (-0.696) и (-0.228) в модели, учитывающей количество дизельных автомобилей. Поскольку количество автомобилей входит в уравнение в качестве как зависимой, так и независимой переменной, можно рассчитать долгосрочную трансформированную эластичность потребления бензина по количеству автомобилей – $(1+\delta)$, то есть, когда домохозяйство покупает второй автомобиль, при прочих равных условиях, общее потребление бензина увеличится на 14.33%, что приблизительно в три раза меньше, чем результаты, полученные в работах Dahl и Sterner (1991) [CITATION Dah91 \l 1049] и Baltagi и Griffin (1997) [CITATION Bal97 \l 1049] (47%), Рокк (2010) [CITATION Рок10 \l 1049] (23%). Однако, в работе Baltagi et al. (2003) [CITATION Bal031 \l 1049] было показано, что общее потребление бензина и вовсе снижается на 12.68%, что исследователи связывают с неучетом количества дизельных автомобилей.

Регрессия с индивидуальными фиксированными эффектами, оцененная доступным (реализуемым) обобщенным методом наименьших квадратов, учитывает гетероскедастичность и пространственную корреляцию, а также характерную для панельных данных структуру ошибок AR (1). Полученные доступным (реализуемым) обобщенным методом наименьших квадратов оценки статистически значимы. Коэффициент при лаге зависимой переменной равен 0.600, что больше, чем оценки других

методов, кроме сквозной регрессии, оцененной обычным методом наименьших квадратов, и регрессии со случайными эффектами, оцененной обобщенным методом наименьших квадратов. Ввиду очень маленьких стандартных ошибок, регрессия с индивидуальными фиксированными эффектами, оцененная доступным (реализуемым) обобщенным методом наименьших квадратов, кажется подходящей для нашей панельной модели, что подтверждается выводами Baltagi и Griffin (1997)[CITATION Bal97 \l 1049].

Однако, поскольку наша модель динамична и учитывает индивидуальные региональные эффекты, лаг зависимой переменной коррелирует с вектором ошибок и, следовательно, приводит к несостоятельности оценок МНК, Within и ОМНК. С целью учета данной особенности модели в научных исследованиях используют метод инструментальных переменных, который обеспечивает состоятельность, но не обязательно эффективность оценок параметров модели.

Таким образом, с целью учета проблемы эндогенности лага переменной удельного расхода бензина GAS_{t-1} были применены четыре наиболее популярных метода, основанных на использовании инструментальных переменных. Регрессия с фиксированными эффектами, оцененная двухшаговым МНК, дает довольно низкую оценку коэффициента при лаге зависимой переменной, в то время как краткосрочные эластичности несколько переоценены по сравнению с регрессией с фиксированными эффектами, оцененной обычным МНК.

Это верно и для оценок, полученных с помощью метода Андерсона – Хсiao (FD-2SLS). Полученная данным методом скорость адаптации к долгосрочному равновесию невелика и незначима.

Далее был применен одношаговый метод Ареллано – Бонда (FD-GMM). Данный метод позволяет получить хорошие оценки. Коэффициент при лаге спроса на бензин составляет 0.483, что несколько ниже, чем аналогичный коэффициент, полученный с помощью LSDVc, тогда как оценки краткосрочных эластичностей выше по модулю. Это приводит к тому, что в долгосрочном периоде эластичности, хоть и остаются на разумном уровне, но выше, чем соответствующие значения, полученные при оценивании LSDVc и доступным (реализуемым) ОМНК (GLS-HC). Обилие используемых FD-GMM инструментов дает возможность получить более эффективные оценки, чем, например, FD-2SLS.

Обращаясь к одношаговому методу Бланделла – Бонда (sys-GMM), мы наблюдаем среднюю скорость адаптации к долгосрочному равновесию при применении полного набора условий на моменты. На небольших выборках это может вызвать сверхподгонку

инструментируемой переменной, в результате чего оценки смещаются в сторону оценок МНК. Закономерно, что на используемой статистической базе этот эффект не наблюдается.

Преимущество метода Бланделла – Бонда заключается в нивелировании смещенности оценок при работе с малыми выборками, которая имеет место при применении методов FD-2SLS и FD-GMM вследствие недостаточного количества инструментов. Однако с увеличением числа регрессоров, количество условий на моменты приближается к количеству наблюдений в малых выборках, что может привести к сверхподгонке инструментируемой переменной, и полученные оценки окажутся смещенными (см., например, Baltagi (2005) [CITATION Bal05 \l 1049]).

Следует также принимать во внимание тот факт, что в работе Bun и Kiviet (2006) [CITATION Bun06 \l 1049] авторы показали, что желаемые свойства оценок, полученных методами FD-2SLS, Within-2SLS, FD-GMM и sys-GMM, достигаются асимптотически при большом количестве объектов, а иначе могут оказаться смещенными и неэффективными.

Как показано в работе Bruno (2005) [CITATION Bru05 \l 1049], регрессия с фиксированными эффектами и поправкой на смещение (метод LSDV corrected) на малых выборках работает заметно лучше, чем метод инструментальных переменных и метод моментов. Значения поправок для коррекции смещения были получены на основании оценок метода Ареллано - Бонда, обеспечивающего состоятельные оценки.

Проведенные тесты на отсутствие автокорреляции остатков свидетельствуют о состоятельности соответствующих оценок.

Итак, для оценки уравнения спроса на бензин в Российской Федерации в общей сложности было применено девять методов. Качественные результаты оценки уравнения спроса на бензин в целом соответствуют результатам, полученным в предыдущих исследованиях, таких как Baltagi и Griffin (1983, 1997) [CITATION Bal83 \l 1049], [CITATION Bal97 \l 1049], Baltagi et al. (2003) [CITATION Bal03 \l 1049] и Johansson и Schipper (1997) [CITATION Joh97 \l 1049]. Коэффициенты имеют ожидаемые знаки: ВРП на душу населения – положительный, цена на бензин и среднее количество автомобилей на душу населения – отрицательный. Абсолютные значения полученных в ходе выполнения данной работы оценок эластичности цен на бензин (см. таблицу) в основном выше, чем принятый в литературе диапазон от -0,2 до -0,5 в краткосрочном периоде и от -0,6 до -0,8 в долгосрочном (см. Graham and Glaister (2002) [CITATION Gra02 \l 1049]). Данное обстоятельство может быть обосновано, в частности, с тем фактом, что в представленном анализе не учтено количество дизельных автомобилей. Использование в уравнении спроса на бензин общего количества автомобилей, а не отдельно автомобилей с бензиновыми и дизельными двигателями, приводит к завышенным в абсолютном выражении коэффициентам, поскольку в данном случае игнорируется растущая доля автомобилей с дизельными двигателями (Роск, (2010) [CITATION Рос10 \l 1049]).

Кроме того, с учетом того факта, что в качестве знаменателя переменной CAR используется общая численность населения, а не количество водителей, не представляется возможным учесть демографические эффекты: в работе Schmalensee и Stoker (1999) [CITATION Sch99 \l 1049] авторы на данных США показали, что в результате использования количества водителей значительно снижается расчетная эластичность спроса по доходу. В дальнейшем в работе Роск (2010) [CITATION Рос10 \l 1049] продемонстрировано, что в результате использования общей численности населения оценки эластичностей по цене и по доходу оказались на 20 % выше, чем при учете в качестве прокси для количества водителей количества населения в возрасте от 18 до 69 лет.

Было проведено попарное сравнение моделей, учитывающих фиксированный временной эффект. Для того, чтобы сделать выбор между моделью сквозной регрессии по регионам и моделью со случайными региональными эффектами, которые обе включают в себя фиксированные временные характеристики, использовался тест Бройша – Пагана (множителей Лагранжа).

Из проведенного теста Бройша-Пагана можно заключить, что модель со случайными эффектами лучше подходит для описания имеющихся данных по сравнению с моделью сквозной регрессии.

Интерпретируя результаты теста Хаусмана можно заключить, что модель со случайными индивидуальными эффектами дает несостоятельные оценки по сравнению с моделью с детерминированными индивидуальными эффектами, которые обе включают в себя фиксированные временные эффекты. Таким образом, можно сделать вывод, что оптимальной моделью является модель с детерминированными эффектами.

Тест Вальда на значимость дамми-переменных для фиксирования временных эффектов свидетельствует о том, что они статистически значимы, то есть каждый временной период оказывает отличное от первого временного периода влияние на зависимую переменную.

3 Разработка рекомендаций по экономической политике на основании полученных результатов

В предыдущей главе был проведен количественный анализ эластичностей потребления бензина в Российской Федерации на региональном уровне. Согласно результатам анализа, в большинстве исследуемых спецификаций наблюдается эластичный спрос на бензин с увеличением цены. В регрессии с фиксированными эффектами, оцененной как методом наименьших квадратов, так и с поправкой на смещение, значения краткосрочной эластичности потребления бензина составили -1.053 и -1.013 соответственно. Для этих же спецификаций значения долгосрочной эластичности потребления бензина составили -2.325 и -2.328 соответственно. При использовании инструментальных переменных для учета эндогенности между зависимой переменной и ее первым лагом, результаты получились сопоставимыми. Диапазон оцененных значений краткосрочной эластичности составляет от -1.733 до -1.589, а для долгосрочной эластичности – от -3.474 до -1.863. С учетом того факта, что в работе был проведен всесторонний анализ применяемых спецификаций, который не выявил критичных недостатков модели, можно сделать вывод, что полученные оценки являются состоятельными и разумными.

Из полученных оценок можно сделать вывод, что проводимая в данный момент Правительством политика, направленная на пополнение региональных бюджетов за счет увеличения ставки акцизов на нефтепродукты, является неэффективной и контрпродуктивной. С учетом эластичного спроса, в краткосрочном периоде будет наблюдаться нейтральность бюджета при увеличении ставки акцизов или незначительное сокращение бюджета. В долгосрочной перспективе следует ожидать существенного сокращения поступлений в бюджет от акцизов на нефтепродукты, так как с увеличением розничной цены на 1% потребление сократится на 2-3%. Помимо изменений бюджета, вызванных акцизными платежами, следует ожидать сокращения поступлений и от других видов налогов, которые применяются в топливно-энергетическом секторе. Сокращение потребления нефтепродуктов приведет к сокращению спроса на нефть, что повлечет за собой снижение поступлений от таких налогов, как НДС, налог на прибыль. Помимо сокращения поступлений от налогов, связанных с топливно-энергетическим сектором, следует ожидать и прочих негативных эффектов для экономики, так как с ростом цен на нефтепродукты увеличится цена на все остальные товары и услуги, поскольку в них увеличится стоимость транспортной составляющей.

На наш взгляд, увеличения ставки акцизов при проводимом налоговом маневре приведет к существенному сдерживанию экономического роста. При изначальном сценарии налогового маневра предполагалось, что увеличение внутренней цены на нефть при росте

ставки НДС и сокращении экспортной пошлины на нефть не приведет к росту цен на нефтепродукты, так как увеличение затрат производителей бензина и дизельного топлива будет компенсироваться снижением ставки акцизов. Однако в данный момент происходит увеличение как стоимости нефти, так и ставки акцизов, что приводит к значительному росту оптовых и розничных цен на нефтепродукты. В сложившейся экономической ситуации покупательная способность потребителей нефтепродуктов снижается, а затраты на транспорт увеличиваются, что негативно влияет на экономику Российской Федерации, ограничивая потенциал экономического роста и разгоняя инфляцию. Дальнейшее повышение ставки акцизов или розничной цены на нефтепродукты будет увеличивать стагнацию экономики РФ, так как приведет к удорожанию всех товаров и услуг. Иными словами, при текущем уровне средней заработной платы в РФ потенциал для дальнейшего роста цен на нефтепродукты исчерпан.

Для увеличения темпов экономического роста и повышения конкурентоспособности отечественных товаров с высокой добавленной стоимостью необходимо разработать механизмы пополнения региональных бюджетов, которые бы не приводили к росту розничных цен на нефтепродукты. Для оперативного пополнения «бюджетных дыр» необходимо использование иных инструментов, так как повышение акцизов приводит к сокращению экономической активности, как следствие снижаются налоговые поступления, а для их пополнения государство, согласно текущей модели поведения, увеличит ставку акцизов на нефтепродукты. В данный момент необходимо снижать ставки акцизов для стимулирования развития экономики, так как, несмотря на тот факт, что акцизы уплачиваются нефтеперерабатывающими компаниями, их платежи перекладываются преимущественно на розничных потребителей. Такая практика позволяет компенсировать неэффективность отечественной нефтепереработки за счет потребителей, не вкладывая средства в модернизацию технологий.

Согласно нашим расчетам, в данный момент повышение акцизов на бензин на 1% приводит к сокращению суммарных поступлений в региональные и федеральный бюджеты от таких налогов, как НДС и акцизы на бензин на 0.3%¹. Согласно данному расчету, запланированное повышение акциза на бензин пятого класса², которое произойдет в январе 2020 года до уровня в 12752 рубля за тонну (на 3.6% от уровня 2019 года), приведет к сокращению поступлений примерно на 1.1%. Дальнейшие этапы увеличения акцизов на

1 Исходя из предпосылок, что ориентировочная доля акцизов в оптовой цене на бензин 5 класса составляет 20% от цены, а из тонны нефти производится 600 килограмм светлых нефтепродуктов.

2 Федеральный закон от 29.09.2019 № 326-ФЗ "О внесении изменений в часть вторую Налогового кодекса Российской Федерации и статью 1 Федерального закона "О внесении изменений в часть вторую Налогового кодекса Российской Федерации"

нефтепродукты, запланированные на 2021 (до 13262 рублей за тонну) и 2022 (до 13793 рублей за тонну) приведут к еще большему торможению экономического роста и сокращению поступлений в бюджет. Однако наибольшие риски для экономики возникают не в краткосрочном периоде, а в долгосрочном. Потери для бюджета будут более значительными, так как вследствие снижения мобильности населения сократится спрос на нефтепродукты. Увеличение цен на нефтепродукты приведет к снижению спроса на легковые транспортные средства, увеличив давление на автопроизводителей, которые и так в данный момент испытывают сложности с реализацией готовой продукции.

Повышение ставки акцизов, а не снижение, которое было обозначено при разработке налогового маневра, во многом влияет на экономический рост в Российской Федерации. На наш взгляд, необходимо вернуться к изначальному сценарию реализации налогового маневра, снижая величину ставок акцизов, что позволит снизить цены на нефтепродукты, создав стимулы для развития нефтеперерабатывающей промышленности. Российский бизнес должен видеть, что декларируемые обязательства согласуются с фактическими результатами, а не корректируются в угоду временным потребностям. Стабильные ценовые сигналы позволят снизить риски инвесторов, что создаст дополнительные условия для экономического роста. Сокращение поступлений от акцизов повлияет на формирование региональных дорожных фондов, однако для их обеспечения возможно скорректировать налоговое законодательство. Например, перераспределить часть поступлений от налога на добычу полезных ископаемых, которым облагается нефть, в пользу региональных бюджетов.

ЗАКЛЮЧЕНИЕ

Рынок нефтепродуктов является одним из ключевых рынков, как в Российской Федерации, так и в мире. Изменения условий функционирования данного рынка напрямую оказывают влияние на микроэкономические и макроэкономические показатели государства. Увеличение цен на нефтепродукты приводит к удорожанию всей производственной цепочки, так как увеличиваются затраты на доставку сырья и компонентов, необходимых для производства, а также затраты, связанные с доставкой готовой продукции. Согласно оценкам, на текущий момент в розничной цене бензина около 60-65% составляют различные виды налогов, акцизов и сборов, поэтому любые изменения в налоговом законодательстве оказывают существенное влияние на уровень цен. Акцизы на нефтепродукты используются как довольно быстрый и эффективный способ для пополнения бюджета в краткосрочном периоде, однако влияние повышения ставки акцизов на экономику в долгосрочном периоде изучено недостаточно. В краткосрочном периоде потребление нефтепродуктов практически неэластично по цене, однако в долгосрочном периоде у потребителей существует множество возможностей для сокращения потребления. Сокращение потребления нефтепродуктов может быть вызвано как снижением мобильности населения, так и переходом на более энергоэффективные транспортные средства или альтернативные виды транспорта. Снижения мобильности населения приведет не только к сокращению потребления нефтепродуктов, так как средний годовой пробег на одного жителя уменьшится, но и к снижению спроса на новые транспортные средства, так как часть потребителей откажутся от приобретения легковых автомобилей в условиях роста цен на нефтепродукты. В результате снизится добавленная стоимость машиностроительной отрасли экономики. Учитывая вышесказанное, принятие решений, которые оказывают влияние на цены нефтепродуктов, должно основываться на всестороннем качественном и количественном анализе, так как цена ошибки в данной сфере довольно высока.

С этой целью развитые и развивающиеся страны строят модели, которые позволяют оценить влияния микроэкономических и макроэкономических факторов на потребление нефтепродуктов автомобильным транспортом. Данные модели позволяют снизить риски при принятии решений и прогнозировать реакцию рынка в ответ на изменения условий. Проведенный обзор современной научной литературы показал, что на данном этапе исследователями преимущественно используются две модели: оценка панельных данных и методология коинтеграции временных рядов. Данные подходы, при корректном использовании, дают сравнимые результаты оценок. В качестве ключевых переменных, которые применяются в моделях, выступают удельное потребление топлива на одно транспортное средство, количество транспортных средств, располагаемый доход

потребителей или ВВП на душу населения, цены на нефтепродукты и общая численность населения или количество водителей. Большинство исследований в применяемых методологиях не используют уравнение предложения в моделировании, что может приводить к смещению получаемых результатов. Это обстоятельство связано со сложностью подбора переменных, которые бы оказывали влияние только на спрос (предложение) и не влияли при этом на предложение (спрос). Кроме того, при нахождении переменных, которые оказывают только влияние на одно из уравнений, данные переменные, как правило, являются национальными, что не позволяет проводить оценки на региональном уровне.

Отдельные исследования представляют мета-анализы, проведенные на основе сопоставлений различных исследований. Они позволили выявить ряд закономерностей относительно получаемых результатов. Благодаря данным анализам повышается качество трактовки результатов в исследованиях. На основании данных мета-анализов исследователи могут выбрать более предпочтительные спецификации моделей в зависимости от имеющихся в наличии статистических данных.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Sweeny J. The demand for gasoline in the United States: a vintage capital model // Workshop on Energy and Supply and Demand, International Energy Agency. Paris. 1978. pp. 240–277.
2. Baltagi B.H., Griffin J.M. Gasoline demand in the OECD: an application of pooling and testing procedures // *European Economic Review*. 1983. No. 22. pp. 117–137.
3. Baltagi B.H., Griffin J.M. Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline // *Journal of Econometrics*. 1997. No. 77. pp. 303–327.
4. Schmalensee R., Stoker T.M. Household gasoline demand in the United State // *Econometrica*. 1999. No. 67 (3). pp. 412–418.
5. Baltagi B.H., Bresson G., Griffin J.M., and Pirotte A. Homogeneous, heterogeneous or shrinkage estimators? Some empirical evidence from French regional gasoline consumption // *Empirical Economics*. 2003. No. 28. pp. 795–811.
6. Houthakker H.S., Verleger P.K., and Sheehan D. Dynamic demand for gasoline and residential electricity // *American Journal of Agricultural Economics*. 1974. No. 56. pp. 412–418.
7. Pock M. Gasoline demand in Europe: New insights // *Energy Economics*. 2010. Vol. 32. pp. 54–62.
8. Dahl C.A., Sterner T. Analyzing gasoline demand elasticities: a survey // *Energy Economics*. 1991. No. 11. pp. 203–210.
9. Dargay J., Gately D., and Sommer M. Vehicle ownership and income growth, worldwide: 1960–2030 // *Energy Journal*. 2007. No. 28.
10. Johansson O., Schipper L. Measuring the long-run fuel demand of cars // *Journal of Transport Economic and Policy*. 1997. No. 31. pp. 277–292.
11. Anderson T.W., Hsiao C. Formulation and estimation of dynamic models using panel data // *Journal of Econometrics*. 1982. No. 18.
12. Arellano M., Bond S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations // *Review of Economic Studies*. 1991. No. 58.
13. Blundell R., Bond S. Initial conditions and moment restrictions in dynamic panel data models // *Journal of Econometrics*. 1998. No. 87.
14. Baltagi B.H. *Econometric analysis of panel data* // John Wiley & Sons Ltd. 2005.
15. Bun M.J.G., Kiviet J.F. The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models // *Journal of Econometrics*. 2006. No. 132.
16. Nickell S. Biases in dynamic models with fixed effects // *Econometrica*. 1981. No. 49.
17. Kiviet J.F. On bias, inconsistency and efficiency of some estimators in dynamic panel data models // *Journal of Econometrics*. 1995. No. 68.
18. Bruno G.S.F. Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models // *Economics Letters*. 2005. No. 87.
19. Bun M.J.G., Kiviet J.F. On the diminishing returns of higher-order terms in asymptotic expansions of bias // *Economics Letters*. 2003. No. 79.
20. Blanchard P., Matyas L. Robustness of tests for error component models to non-normality // *Economics Letters*. 1996. No. 51.
21. Montes-Rojas G., Sosa-Escudero W. Robust tests for heteroskedasticity in the one-way error components model // *Journal of Econometrics*. 2011. No. 160.
22. Holly A., Gardiol L. A score test for individual heteroskedasticity in a one-way error components model // *Panel Data Econometrics: Future Directions*. 2000.

23. Baltagi B., Bresson B., and Pirotte A. Joint LM test for homoskedasticity in a one-way error component model // *Journal of Econometrics*. 2006. No. 134.
24. Baltagi B., S. S., and Koh W. Testing panel data regression models with spatial error correlation // *Journal of Econometrics*. 2003. No. 68.
25. Baltagi B., Song S., Jung B., and Koh W. Testing for serial correlation, spatial autocorrelation and random effects using panel data // *Journal of Econometrics*. 2007. No. 140.
26. Erickson T., Whited T. Measurement Error and the Relationship between Investment and Q // *Journal of Political Economy*. 2000. No. 108.
27. Bai J., Ng S. Tests for skewness, kurtosis, and normality for time series data // *Journal of Business and Economic Statistics*. 2005. No. 23.
28. Bera A., Premaratne G. Adjusting the tests for skewness and kurtosis for distributional misspecifications // www.researchgate.net. 2015. URL: https://www.researchgate.net/publication/311536088_Adjusting_the_Tests_for_Skewness_and_Kurtosis_for_Distributional_Misspecifications
29. Gilbert S. Testing the distribution of error components in panel data models // *Economics Letters*. 2002. No. 77.
30. Meintanis S. Testing for normality with panel data // *Journal of Statistical Computation and Simulation*. 2011. No. 81.
31. Galvao A., Montes-Rojas G., Sosa-Escudero W., and Wang L. Tests for skewness and kurtosis in the one-way error component model // *Journal of Multivariate Analysis*. 2013. No. 122.
32. Belsley D.A., Kuh E., and Welsch R.E. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: Wiley, 1980.
33. Baltagi B., Griffin J. Gasoline demand in OECD: An Application of Pooling and Testing Procedures // *European Economic Review* 22 (1983) pp. 117-137. 1983.
34. Graham D.J., Glaister S. The demand for automobile fuel: a survey of elasticities // *Journal of Transport Economics and Policy*. 2002. No. 36.