Научная статья
УДК 311.3
https://doi.org/10.17586/2713-1874-2025-4-41-51

ИНТЕГРАЦИЯ АЛЬТЕРНАТИВНЫХ ИСТОЧНИКОВ ДАННЫХ В СИСТЕМУ ГОСУДАРСТВЕННОЙ СТАТИСТИКИ

Ирина Николаевна Вейнер

Санкт-Петербургский государственный университет, Санкт-Петербург, Россия, i.veyner@spbu.ru,
https://orcid.org/0000-0002-8557-1667
Язык статьи – русский

Аннотация: В условиях цифровой трансформации происходит изменение источников данных для официальной статистики. Исследование посвящено анализу возможностей интеграции больших данных в традиционную статистическую систему. В данном исследовании проведён обзор существующей практики использования альтернативных источников, таких как мобильные данные (MPD) для учёта миграции и туризма, онлайн-вакансии для мониторинга рынка труда, веб-скрейпинг и данные онлайн-касс для расчёта индекса потребительских цен международными организациями и национальными статистическими службами. Установлено, что от первоначальных идей замены большими данными (альтернативными источниками) данных официальной статистики (традиционных источников) профессиональное сообщество переходит к практике использования больших данных в виде комбинации с данными официальной статистики, эта трансформация обеспечит преимущества в виде большей оперативности, детализации, точности данных, вместе с тем потребует решения множества проблем: методологических, организационных, юридических и т.п.

В рамках исследования была поставлена задача на примере новых статистических административных данных (альтернативный источник) провести анализ агрегированных значений показателей и выяснить возможность их использования в качестве замены соответствующих показателей официальной статистики (традиционный источник). На основе сведений, содержащихся на платформе ФНС России об используемой контрольно-кассовой технике (ККТ), произведён анализ показателей: количества ККТ, выручки, числа чеков с июля 2020 года по октябрь 2025, а также рассчитан размер среднего чека, оценена его динамика, в том числе с учётом инфляции. Полученные результаты сопоставлены с динамикой оборота розничной торговли, представляемой Росстатом. В результате выявлено, что в текущих условиях данные ФНС не могут заменить официальную статистику, но обладают самостоятельной аналитической ценностью и могут выступать оперативным дополнением, обогащающим картину потребительского рынка.

Ключевые слова: административные данные, альтернативные источники данных, большие данные, официальная статистика, Росстат, традиционные источники данных, цифровая трансформация, ФНС России

Ссылка для цитирования: Вейнер И. Н. Интеграция альтернативных источников данных в систему государственной статистики // Экономика. Право. Инновации. – 2025. – Т. 13. – № 4. – С. 41–51. – https://doi.org/10.17586/2713-1874-2025-4-41-51

INTEGRATION OF ALTERNATIVE DATA SOURCES INTO THE STATE STATISTICS SYSTEM

Irina N. Veyner

St. Petersburg State University, St. Petersburg, Russia, i.veyner@spbu.ru, https://orcid.org/0000-0002-8557-1667
Article in Russian

Abstract: In the context of digital transformation, data sources for official statistics are changing.This study analyzes the potential for integrating big data into traditional statistical systems. This study provides an overview of current practices in using alternative sources, such as mobile data (MPD) for recording migration and tourism, online job postings for labor market monitoring, web scraping, and online cash register data for calculating the consumer price index by international organizations and national statistical services. It is established that the professional community is moving from the initial idea of replacing official statistics (traditional sources) with big data (alternative sources) to the practice of using big data in combination with official statistics. This transformation will provide advantages in terms of greater timeliness, detail, and accuracy, but will also require addressing numerous methodological, organizational, and legal challenges.

The study aimed to analyze aggregated indicator values using new administrative statistical data (an alternative source) as an example and explore the feasibility of using them as a substitute for corresponding official statistics (traditional sources). Using information on cash registers stored on the Federal Tax Service platform, we analyzed the following indicators: the number of cash registers, revenue, and receipts from July 2020 to October 2025. We also calculated the average receipt size and assessed its dynamics, including taking inflation into account. The results were compared with retail turnover dynamics reported by Rosstat. It was found that, under current conditions, Federal Tax Service data cannot replace official statistics, but they do have independent analytical value and can serve as a useful supplement, enriching the consumer market picture.

Keywords: administrative data, alternative data sources, big data, official statistics, Rosstat, traditional data sources, digital transformation, Federal Tax Service of Russia

For citation: Veyner I. N. Integration of Alternative Data Sources into the State Statistics System. *Ekonomika. Pravo. Innovacii*. 2025. Vol. 13. No. 4. pp. 41–51. (In Russ.). https://doi.org/10.17586/2713-1874-2025-4-41-51.

Введение. С развитием современных технологий происходит революционный переворот в области данных. Возникли электронные платформы, на которых фиксируются совершённые транзакции, публикуются данные сканирования товаров, регистрации дорожных событий, доступны платформы рекрутинговых агентств и многое другое. В связи с этим профессиональные сообщества активно обсуждают ожидаемые перемены в своей сфере деятельности.

Появившиеся возможности автоматической регистрации различных событий и объектов меняют традиционные подходы статистиков-практиков к сбору исходной информации, появляются альтернативные источники данных. Уже не требуются титанические усилия по сбору и регистрации первичных данных, с которых начинается любое статистическое исследование.

Новые данные (big data) отличаются от имеющихся оперативностью, масштабом, степенью детализации. Сформулированы их характеристики как «три V»: Volume – объем обрабатываемой информации; Velocity – скорость накопления данных и обработки потока данных; Variety – многообразие и недостаточная структурированность данных [1].

Появились идеи решения многих проблем на базе больших данных. Исходя из скорости и объема данных ожидались новые решения в сельском хозяйстве на основе спутниковых данных, в учёте миграции – на основе мобильной связи, в грузоперевозках – с использованием GPS и многие другие. «В зависимости от источников большие данные могут быть условно разделены на несколько основных категорий:

1) данные, генерируемые человеком (например, сообщения в социальных сетях, поисковые запросы в сети Интернет, текстовые сообщения и т. д.);

2) данные, генерируемые в результате выполнения различных процессов (бизнес-процессы и работа веб-сайтов);

3) данные, генерируемые различными машинами (сенсорными устройствами, датчиками и компьютерными системами)» [2, С. 50].

Изучением возможностей использования больших данных в статистике занялись международные организации, прежде всего Статистическая комиссия ООН. В 2015 г. она организовала Глобальную рабочую группу по большим данным – Big Data UN Global Working Group. Национальные статистические службы стали разрабатывать отдельные направления для потенциального применения в отраслевых статистиках альтернативных источников.

При работе с более узкими задачами выяснилось, что альтернативные источники данных при безусловных преимуществах имеют ограничения, которые не позволяют осуществить мгновенный переход на них или хотя бы частично включить их в традиционные источники официальной статистики.

Однако за последнее десятилетие произошли существенные сдвиги, были осуществлены удачные проекты применения больших данных из альтернативных источников. Например, проведено несколько исследований по оценке миграции с использованием мобильной связи. Анализ перемещений на основе данных мобильных телефонов был использован при организации общественного транспорта, учёта туристических потоков, обеспечения оптимальных действий при чрезвычайных ситуациях. Отдельный интерес исследователей вызывают социальные

сети. На их основе производится оценка различных социальных индикаторов, таких как оценка самочувствия, эмоциональное состояние, отношение к различным явлениям. Поисковые запросы стали новым фундаментом для прогнозирования всплесков заболеваемости.

Литературный обзор. В связи с широким распространение мобильных устройств и возможностью отслеживания их местонахождения появились идеи применения больших данных для учёта миграции населения, высказывались предположения о замене официальных источников статистики миграционного движения. Появился метод сбора исходной информации с помощью данных пассивного мобильного позиционирования (mobile positioning data – MPD), представляющих собой координаты местоположения мобильных телефонов в сотовой сети и записи детализации вызовов. На их основе проведено большое количество исследований. Уже в 2018 году были опубликованы результаты обследования, осуществлённого эстонскими исследователями с целью выявления поколенческих различий в пространственной мобильности [3]. В качестве эмпирической базы для анализа использовались данные крупнейшего эстонского оператора мобильной связи Telia за 2014 год. В результате на основе метода мобильного позиционирования учёными сделаны некоторые значимые выводы: пространственная мобильность линейно снижается с возрастом, представители молодых поколений преимущественно перемещаются внутри Эстонии, а старших – за её пределами и др. При этом сами авторы отмечают, что использование метода мобильного позиционирования накладывало некоторые ограничения, прежде всего – необходимость наличия мобильного телефона. В дальнейшем было проведено множество исследований, в которых использовались мобильные данные для различных целей, в том числе в принятии управленческих решений в сфере транспорта, услуг и т.п. По мнению автора [4, С. 52] «мы можем сделать вывод, что большие данные – это источник для возможных замеров краткосрочной и «текущей» мобильности населения, определение которой пока отсутствует и в перспективе может оказаться достаточно широким». Вместе с тем, автор выделяет ограничения применения больших данных для измерения миграции: фактически происходит учёт мобильных устройств, а не людей, кроме того, статистика миграции оперирует точными определениями и характеристиками, что не всегда может быть известно при использовании мобильных данных.

Близкой по тематике является статистика туризма, для которой также актуально применение данных MPD. Автор [5] сделал обзор исследований, касающихся проблем использования традиционных источников данных и источников больших данных в этой области прикладной статистики. Рассмотрев публикации на основе источников данных обоих типов, автор делает заключение: «комбинация приемов официальной статистики и альтернативных методов должна дать, по мнению многих зарубежных исследователей, всеобъемлющую и последовательную оценку текущих пространственно-временных характеристик туристских поездок [5, С. 70].

Отдельным блоком источников альтернативных данных стали различные рекрутинговые платформы, которые содержат сведения об актуальных вакансиях и потенциальных кандидатах. Некоторые из них публикуют свои аналитические отчёты на основе обработки имеющихся у них данных, другие обеспечивают открытый доступ к агрегированным сведениям. Безусловным преимуществом данных рекрутинговых компаний является их актуальность. В то же время трудно оценивать их репрезентативность, а также они, как правило, сгруппированы в соответствии с критериями, отличающимися от официальных. Подробное сопоставление данных онлайн-вакансий и официальной статистики проведено в [6]. В исследовании использованы данные портала поиска работы hh.ru и соответствующие показатели Росстата. Авторы анализировали сведения о количестве и структуре вакансий и заработных плат. Предварительно были сопоставлены характеристики «вакансий» в обоих источниках, осуществлена перекодировка отраслевых категорий портала hh.ru в соответствии с разделами ОКВЭД. В результате сравнительного анализа квартальных и месячных данных за 2023 год были обнаружены схожие тренды количества вакансий и заработных плат. Отраслевые структуры показали синхронность лишь в некоторых сферах: по вакансиям – в транспорте

и логистике, строительстве и др.; по заработным платам – в торговле, добыче полезных ископаемых и др. В заключение авторы исследования отмечают: «Достоверность и валидность результатов сопоставления официальных статистик и данных онлайн-вакансий зависит от многих факторов и государственных политик стран» [6, С. 241]. Кроме того, осталось не выясненным, «какую именно часть рынка труда представляют онлайн-вакансии: более оперативно отражают рынок труда или дополняют данные официальной статистики.» [6, С. 241]

В настоящее время во многих странах ведутся работы по использованию новых источников данных для расчёта индекса потребительских цен. Размещение ценовой информации на сайтах продавцов и на электронных торговых площадках позволяют собирать информацию о ценах (веб-скрейпинг), а также автоматическая фиксация торговых операций через контрольно-кассовую технику (онлайн-кассы) формирует массивы больших данных о проданных товарах, их ценах. Эти возможности рассматривались первоначально, как замена наблюдений официальной статистики за потребительским рынком. Современные инструментальные средства автоматизируют процесс сбора первичных данных, они позволяют получать оперативную информацию и даже информацию в режиме реального времени, упрощается процедура пересмотра набора учитываемых товаров.

Однако при более тщательном рассмотрении выясняется, что при безусловных плюсах новых технологий есть некоторые ограничения их применения. Например, в [7] выделяются недостаточность охвата источников ценовой информации; сложности идентификации конкретных товаров и отнесения их к товарным группам; возможные расхождения цен на сайтах организаций и при непосредственной продаже; сложности привязки ценовой информации к конкретной территории. Исследователи Бюро статистики труда США обнаружили сложности фиксации цен при обновлении ассортимента: снижение цен на старые товары при появлении новых. При этом надо отметить, что такие данные применяются не в качестве замены, а в качестве альтернативного или дополнительного источника [8]. Тем не менее, уже есть опыт использования альтернативных источников данных, например, данные из Интернета включены в расчёт индекса потребительских цен в Австралии, в Нидерландах индекс цен на продукты формируется на основании альтернативных источников.

Интересный опыт использования данных о потребительских товарах описывает группа российских исследователей [9]. В нём они используют технологию веб-скрейпинга и альтернативную технологию сбора данных кассовой техники.

Вместе с полученным опытом практических действий пришло и лучшее понимание происходящих процессов. Большие данные приобрели новые характеристики, теперь выделяют «пять V»: объём информации, как количество записей и их характеристик (Volume); скорость прироста и обработки (Velocity); многообразие источников, носителей и форматов данных (Variety), достоверность (Veracity), ценность (Value). В [10] автор проводит подробный анализ современного применения больших данных в официальной статистике, как международной, так и национальной. Он выделил три типа источников данных: официальные статистические наблюдения (в т.ч. ведомственные); ведомственные реестры, регистры и административные записи; большие данные (данные операторов мобильной связи, коммерческих банков, торговых сетей и т.п.) и дал им характеристики. Автор сформулировал потенциальные возможности использования больших данных в официальной статистике и назвал существующие «проблемы: взаимодействие с владельцами данных в части доступа к ним и к метаданным, включая описание методов сбора и обработки первичных сведений; уникальность массивов данных и спонтанность их возникновения/исчезновения; территориальная и динамическая несопоставимость данных; разработка механизма поиска данных, оценки возможности их использования (ценности) и методов преобразования; использование неофициальных классификаций и методов систематизации объектов и случаев наблюдений; неполнота охвата наблюдением изучаемых объектов и возможное наличие смещения в их характеристиках; сложность структур данных и многообразие форматов и архитектуры хранилищ, систем доступа.» [10,

С. 10] Опираясь на изученный опыт, автор делает вывод о необходимости поиска решений прежде всего для экономии ресурсов официальными статистическими структурами.

Таким образом, идёт активная трансформация подходов к традиционным статистическим источникам под влиянием больших данных. Активную работу ведут как исследователи-теоретики, так и национальные статистические службы. Найти возможности гармонизации официальной статистики и новых альтернативных источников информации – задача, решением которой заняты многие.

Постановка задачи (цель исследования). В рамках исследования была поставлена задача на примере новых статистических административных данных (альтернативный источник) провести анализ агрегированных значений показателей и выяснить возможность их использования в качестве замены соответствующих показателей официальной статистики (традиционный источник). Большие данные очень обширны и различаются по многим параметрам (способам сбора, наличием/отсутствием методологии, сферой применения, организаторам и др.). Наиболее близкими к данным официальной статистики могут рассматриваться данные других государственных ведомств.

Методы и материалы исследования. Федеральная налоговая служба Российской Федерации активно развивает свою деятельность в цифровой среде, в том числе она публикует агрегированные деперсонифицированные данные на платформе поставки данных ФНС России, которые автоматически регистрируются на основе контрольно-кассовой техники в соответствии с Федеральным законом от 22 мая 2003 года № 54-ФЗ «О применении контрольно-кассовой техники при осуществлении расчетов в Российской Федерации» [11]. Проведён статистический анализ представленных агрегированных данных, а также сопоставлены данные альтернативного источника и традиционного источника официальной статистики.

Полученные результаты. Данные, представленные на платформе поставки данных ФНС России, публикуются за каждый месяц, начиная с июля 2020 года. На рисунке 1 представлена динамика количества контрольно-кассовой техники (ККТ) с 2020 года (с июля по декабрь) по 2025 год (с января по октябрь) с делением на ККТ у индивидуальных предпринимателей и у юридических лиц. Число аппаратов в июле 2020 года составляло 3,5 млн шт., а в октябре 2025 года – 3,8 млн шт., т.е. за этот период их число выросло на 10%, соотношение между индивидуальными предпринимателями и юридическими лицами оставалось примерно одинаковым.

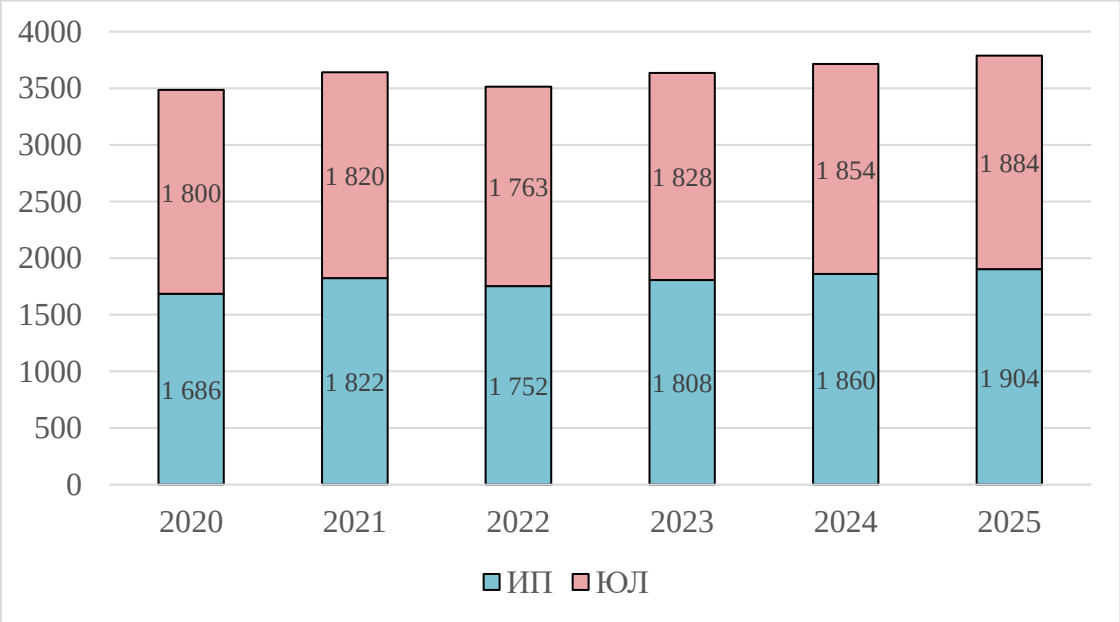

Рисунок 1 – Среднемесячное количество контрольно-кассовой техники, тыс. шт.
Источник: составлено автором на основе [12]

На портале ФНС представлены данные и о составе ККТ. Выделены активные, автономные и неактивные ККТ. В целом, соотношения достаточно устойчивы, однако сравнивая структуру июля 2020 г. и октября 2022 г. (см. таблицу 1), можно отметить увеличение доли активных ККТ на 11 п.п. за счет сокращения двух других видов ККТ.

Таблица 1

Состав контрольно-кассовой техники по видам

Источник: составлено автором на основе [12]

Виды ККТ	Июль 2020 г.	Октябрь 2025 г.
Активные	71,6	82,6
Автономные	5,3	1,5
Неактивные	23,1	15,9
Итого	100,0	100,0

Проанализируем динамику ещё двух показателей, представленных на платформе поставки данных ФНС России: выручка и количество чеков. Ежемесячные изменения выручки представлены в виде графика на рисунке 2.

Следует обратить внимание на два момента. Первый: объем ежемесячной выручки ежегодно увеличивается, что объясняется инфляцией. При этом есть месяцы, в которые значения почти не отличаются от аналогичных прошлогодних (например, февраль и март 2022 и 2023 годов). Второй: во все рассматриваемые годы наблюдается сезонный рост выручки в декабре (15–25%) и обратное сокращение в январе.

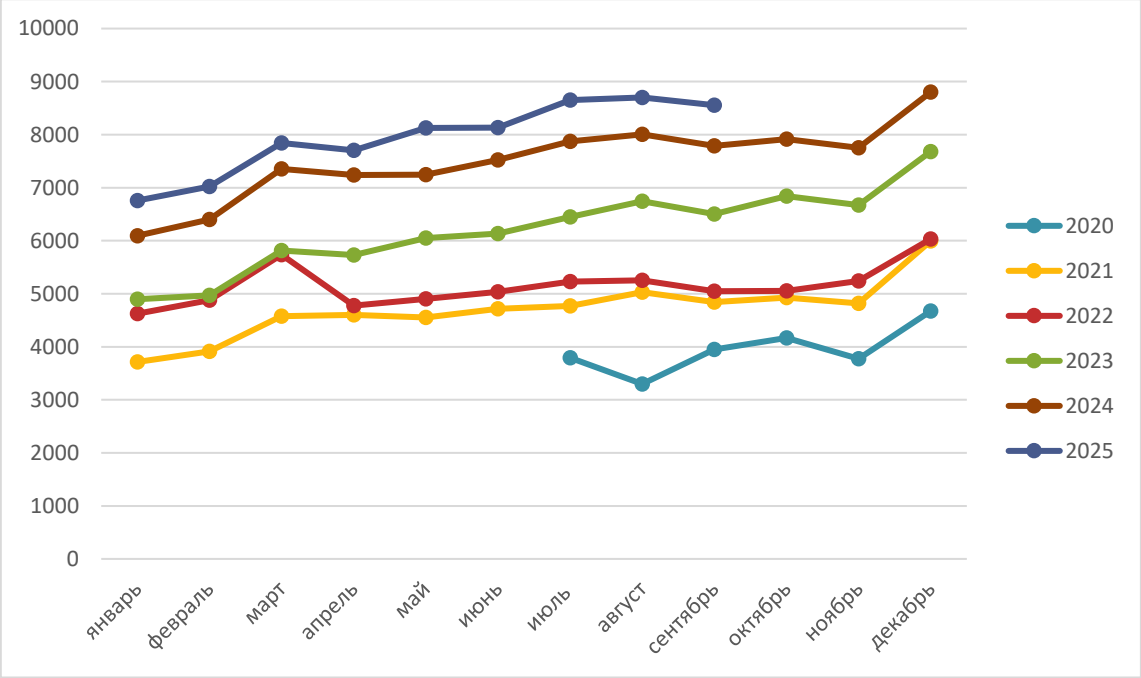

Рисунок 2 – Ежемесячный объём выручки, млрд руб.

Источник: составлено автором на основе [12]

Для оценки годовых изменений рассчитаем среднемесячные размеры выручки за каждый год в номинальном исчислении. Получим приросты 19, 10, 20, 17 и 6% за 2021, 2022, 2023, 2024 и 2025 годы соответственно. В целом выручка выросла в 2,3 раза. Однако для оценки изменения реальных величин исключим инфляцию из имеющихся значений, используя официальные значения ИПЦ (см. рисунок 3).

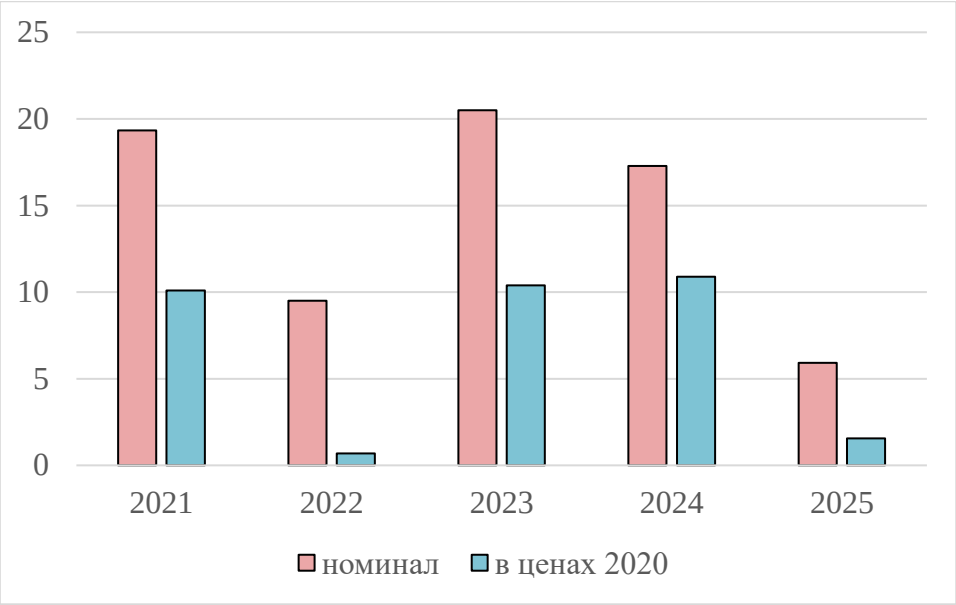

Рисунок 3 – Ежегодные относительные приросты среднемесячного объёма выручки, %
Источник: составлено автором на основе [12, 13]

После исключения инфляции видно, что рост выручки наблюдался в 2021, 2023 и 2024 годах на уровне 10%, а в 2022 и 2025 годах – лишь 0,7 и 1,6%. Данные по 2025 году включают месячные данные лишь до сентября включительно, традиционный рост выручки в конце года увеличит имеющийся результат.

Распределение количества чеков по месяцам рассматриваемого периода выглядит следующим образом (рисунок 4):

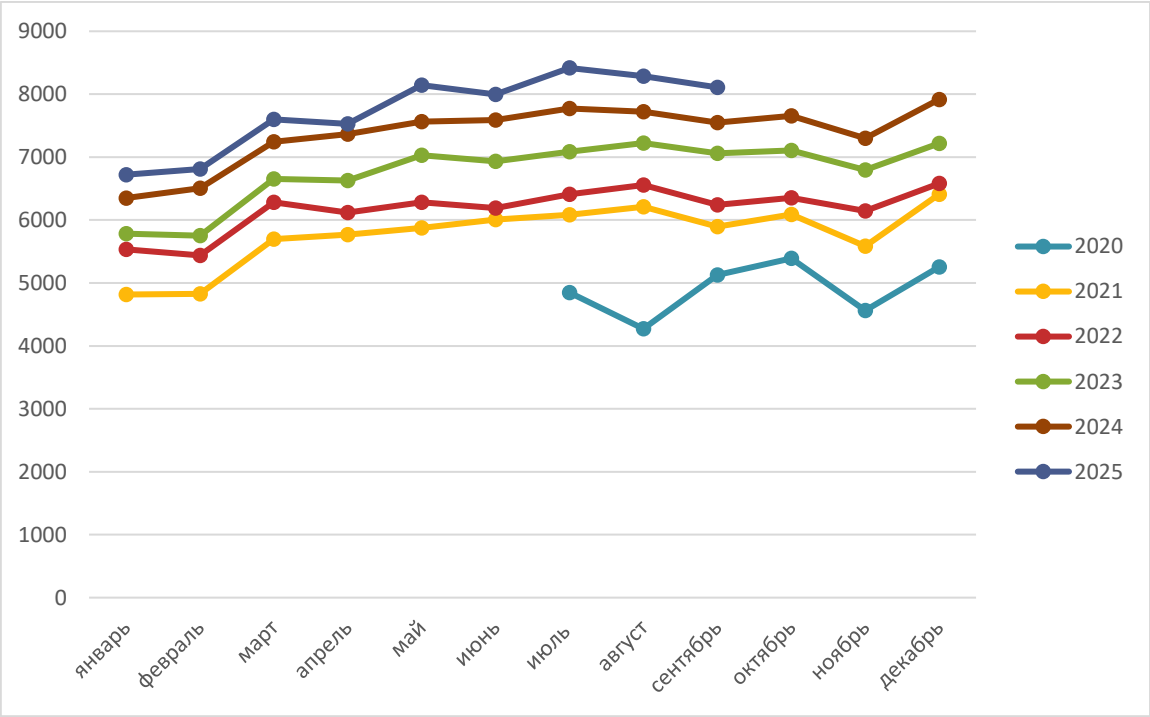

Рисунок 4 – Количество чеков за каждый месяц, млн шт.
Источник: составлено автором на основе [12]

График наглядно демонстрирует довольно равномерный ежегодный рост числа чеков и небольшую сезонную волну в конце-начале года. Наблюдается значительный рост числа чеков в 2021 году по сравнению с 2020 годом, далее происходит замедление этого роста, и приросты составляют примерно +10% по сравнению с аналогичным месяцем

прошлого года. А рост их количества за весь рассматриваемый временной промежуток составил 67 %.

На основе выручки и числа чеков был рассчитан показатель – средний размер чека за каждый месяц рассматриваемого периода. Результаты представлены на рисунке 5. За весь период в текущих ценах показатель вырос на 35%. Однако уровень цен поднялся на 52%. Как видно на диаграмме, динамика среднего чека в ценах июля 2020 года демонстрирует снижение показателя. При неизменных ценах июля 2020 года размер среднего чека с 782 рублей снизился до 696, снижение составило 11%. Таким образом, при опережающем росте числа чеков размера выручки мы наблюдаем картину уменьшения среднего чека, т.е. объем покупок стал меньше.

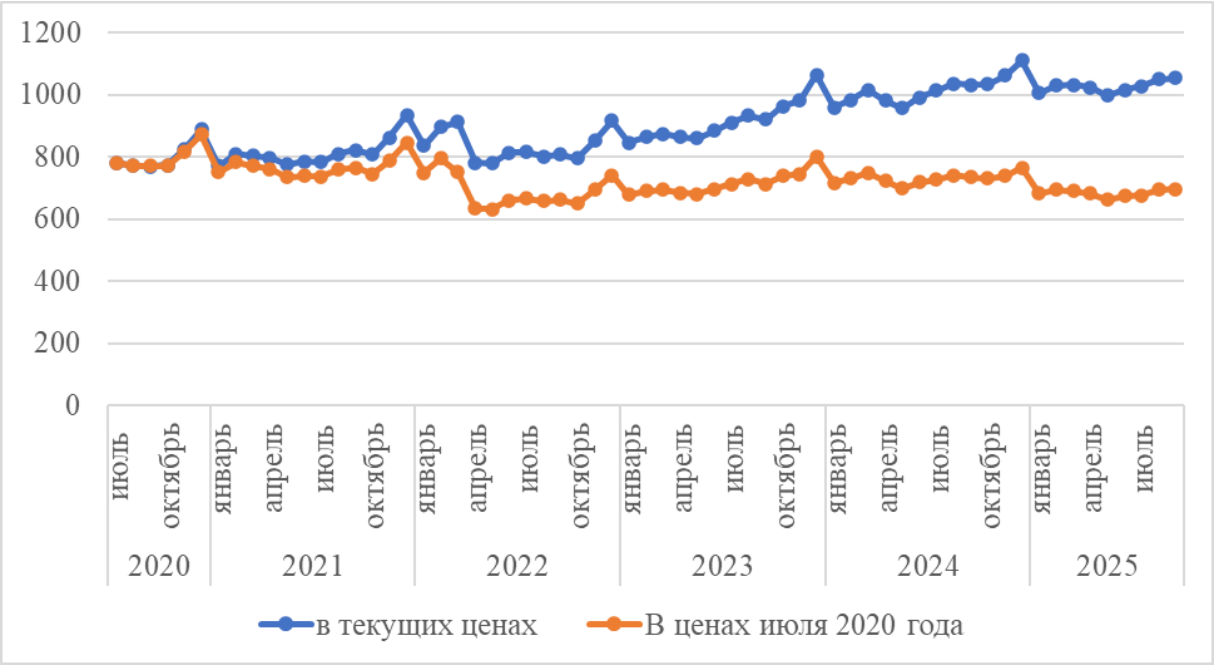

Рисунок 5 – Ежемесячная динамика среднего чека, руб.
Источник: составлено автором на основе [12, 13]

Среди официальных показателей, характеризующих потребление населения, широко применяется оборот розничной торговли. «Оборот розничной торговли – стоимостной показатель, представляющий собой выручку от продажи товаров населению для личного потребления или использования в домашнем хозяйстве за наличный или безналичный расчет» [14]. Сопоставим его динамику с изменениями выручки, зафиксированной через контрольно-кассовую технику (ККТ) (рисунок 6).

Диаграмма наглядно демонстрирует синхронность изменения показателей. Поскольку оба измеряются в текущих ценах, то их рост во многом объясняет одновременность этих движений. Различие значений в начале ряда (выручка 3,79 трлн руб; а РТО – 2,9 трлн руб.) объясняет растущее расхождение, т.е. если в июле 2020 года выручка была больше соответствующего РТО на 30%, то в сентябре 2025 эта величина составила 65% (8,55 и 5,17 трлн руб. соответственно).

Таким образом, наличие данных о цифровой фиксации всех производимых оплат населения на основе ККТ позволяет анализировать эту сферу. Были подтверждены увеличивающиеся объёмы как аппаратов ККТ, выручки, так и количества чеков. При этом установлено:

– доля активных аппаратов превысила 83%;

– соотношение между ККТ у индивидуальных предпринимателей и юридических лиц достаточно стабильно (50 на 50);

– ежегодно наблюдается сезонная волна в конце-начале года;

– размер среднего чека растёт на протяжении рассматриваемого периода.

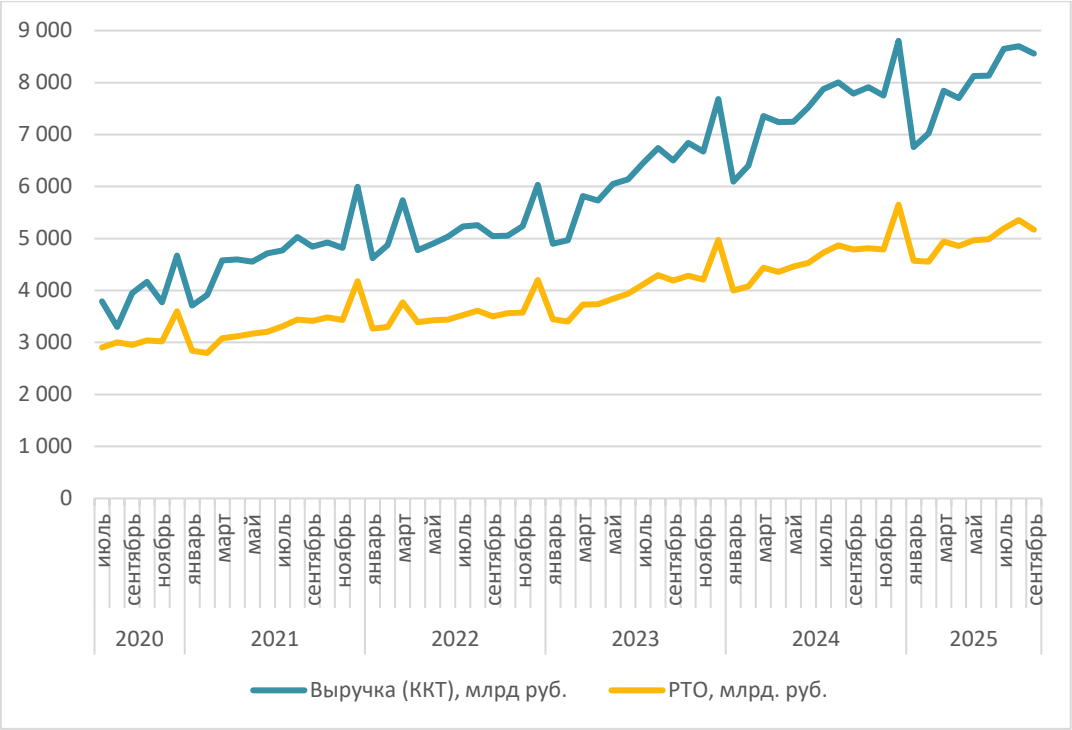

Рисунок 6 – Ежемесячная динамика выручки и розничного товарооборота, млрд руб.
Источник: составлено автором на основе [12, 13]

Таким образом, на основе данных, полученных с помощью новых технологий, был проведён анализ и сделаны содержательные выводы.

Было проведено сопоставление динамики оборота розничной торговли и выручки с платформы поставки данных ФНС с июля 2020 года по сентябрь 2025 года. В результате установлено, что выручка, фиксируемая ККТ, больше розничного товарооборота на протяжении всего периода, их движения происходят синхронно. Следует отметить, что превышение выручки над РТО с течением с времени увеличивается, что следует объяснить тем, что рост стоимостного показателя в существенной степени происходит из-за инфляции, и увеличение от большей базы даёт больший эффект.

Выводы. Изменяющиеся способы сбора статистических данных требуют постоянного наблюдения за этим процессом и внесения корректив в существующие нормы. С момента появления больших данных вопрос о возможности их использовании в официальной статистике заменен вопросом о том, как их использовать. Фактически профессиональным сообществом признано, что к традиционным источникам данных добавились нетрадиционные или альтернативные источники данных. Идёт сложный процесс трансформации отлаженных процессов в новые форматы. На данный момент понятно, что альтернативные источники могут выступать в качестве дополнительных, обогащающих общую картину изучаемого явления. В некоторых ситуациях они могут обеспечить повышенную точность, детализацию в сравнении с традиционными источниками. Кроме того, значимая их особенность – оперативность, так как большие данные могут фиксировать события даже в режиме реального времени.

Однако совместить эти разные источники непросто. Данные официальной статистики имеют чёткую структуру, опираются на принятые классификации, собираются на основе установленной методологии, согласованы с другими источниками. Вместе с тем, статистические службы заинтересованы в сокращении затрат на сбор первичных данных и нагрузки на респондентов.

По-видимому, идёт формирование некоторого сочетания традиционных и альтернативных источников данных, где традиционные данные будут обеспечивать репрезентативность и методологическую базу, а альтернативные – скорость и детализацию.

Проанализированные данные платформы ФНС очевидно не могут заменить данные

официальной статистики, но они имеют собственную аналитическую значимость и уже сейчас могут использоваться для некоторого контроля состояния и динамики соответствующих процессов в области потребления. Многие авторы предвосхищают повышение роли этих сведений после того, как технологии смогут не только фиксировать сумму чека, но и распределение этих сумм по разным категориям товаров, прежде всего, при исчислении инфляции.

Список источников

1. Гладченко В. А. Использование глобальных технологий «bigdata» в качестве эффективного инструмента по осуществлению управления рисками в таможенных органах // Экономика. Право. Инновации. – 2019. – № 2. – С. 36–41.
2. Плеханов Д. А. Большие данные и официальная статистика: обзор международной практики внедрения новых источников данных // Вопросы статистики. – 2017. – № 12. – С. 49–60.
3. Masso A, Silm S, Ahas R. Generational Differences in Spatial Mobility: A Study with Mobile Phone Data // Population, Space and Place. – 2019. – No. 25 (2). – P. e2210. – DOI: 10.1002/psp.2210. – Текст: электронный. (In Eng.).
4. Чудиновских О. С. Большие данные и статистика миграции // Вопросы статистики. – 2018. – Т. 25. – № 2. – С. 48–56.
5. Овчаров А. О. К вопросу о совершенствовании статистического учета туристских поездок в условиях кризиса // Вопросы статистики. – 2021. – Т. 28. – № 2. – С. 67–79. – DOI 10.34023/2313-6383-2021-28-2-67-79. – Текст: электронный.
6. Алтухов В. В., Кудрявцев А. Д. Использование открытых данных онлайн-вакансий в сравнении с данными официальной статистики для мониторинга и прогнозирования динамики рынка труда // Уровень жизни населения регионов России. – 2025. – Т. 21. – № 2. – С. 233–244. – DOI: 10.52180/1999-9836_2025_21_2_5_233_244. – Текст: электронный.
7. Калинин А. М., Волин И. А. Информационные источники для расчета индекса потребительских цен: большие данные сети Интернет и систем ФНС России // Вопросы статистики. – 2022. – Т. 29. – № 1. – С. 44–51. – DOI: 10.34023/2313-6383-2022-29-1-44-51. – Текст: электронный.
8. Konny C. G., Williams B. K., Fried D. M. Big Data in the U.S. Consumer Price Index: Experiences & Plans (Draft) // Big Data for Twenty-First Century Economic Statistics. University of Chicago Press, 2020 [Электронный ресурс]. – URL: https://www.nber.org/books-and-chapters/big-data-twenty-first-century-economic-statistics/big-data-us-consumer-price-index-experiences-and-plans (дата обращения 09.11.2025). – Текст: электронный. (In Eng.).

References

1. Gladchenko V. A. The Use of Global Technology Big Data as An Effective Tool of Risk Management Implementation in Customs. *Ekonomika. Pravo. Innovacii*. 2019. No. 2. pp. 36–41. (In Russ.).

2. Plekhanov D. A. Big Data and Official Statistics: and Overview of International Practices for Implementing New Data Sources. *Voprosy Statistiki*. 2017. No. 12. pp. 49–60. (In Russ.).
3. Masso A., Silm S., Ahas, R. Generational Differences in Spatial Mobility: A Study with Mobile Phone Data. *Population, Space and Place*. 2019. No. 25 (2). pp. e2210. DOI: 10.1002/psp.2210.

4. Chudinovskikh O. S. Big Data and Migration Statistics. *Voprosy Statistiki*. 2018. Vol. 25. No. 2. pp. 48–56. (In Russ.).
5. Ovcharov A. O. On Improving the Statistical Accounting of Tourist Trips in a Crisis. *Voprosy Statistiki*. 2021. Vol. 28. No. 2. pp. 67–79. (In Russ.). DOI: 10.34023/2313-6383-2021-28-2-67-79.

6. Altukhov V. V., Kudryavtsev A. D. The Use of Open Data from Online Vacancies Compared with Official Statistics for Monitoring and Forecasting the Labor Market. *Uroven' Zhizni Naseleniia Regionov Rossii*. 2025. Vol. 21. No. 2. pp. 233–244. (In Russ.). DOI: 10.52180/1999-9836_2025_21_2_5_233_244.

7. Kalinin A. M., Volin I. A. Information Sources for Calculating the Consumer Price Index: Big Data from the Internet and the Systems of the Federal Tax Service of Russia. *Voprosy Statistiki*, 2022. No. 1. pp. 44–51. (In Russ.). DOI: 10.34023/2313-6383-2022-29-1-44-51.
8. Konny C. G., Williams, B. K., Fried D. M. Big Data in the U.S. Consumer Price Index: Experiences & Plans (Draft). Big *Data for Twenty-First Century Economic Statistics. University of Chicago Press*. Available at: https://www.nber.org/books-and-chapters/big-data-twenty-first-century-economic-statistics/big-data-us-consumer-price-index-experiences-and-plans (Accessed: 09.11.2025).

9. Твердые цифры: открытые микроданные о потребительских ценах / А. Исаков, Р. Латыпов, А. Репин [и др.] // Деньги и кредит. – 2021. – Т. 80. – № 1. – С. 104–119. – DOI 10.31477/rjmf.202101.104. – Текст: электронный.
10. Суринов А. Е. Большие данные в официальной статистике: взгляд на проблему // Вопросы статистики. – 2023. – Т. 30. – № 2. – С. 5–22. – DOI: 10.34023/2313-6383-2023-30-2-5-22. – Текст: электронный.
11. Федеральный закон от 22 мая 2003 года № 54-ФЗ «О применении контрольно-кассовой техники при осуществлении расчетов в Российской Федерации» [Электронный ресурс]. – URL: http://www.kremlin.ru/acts/bank/19505 (дата обращения 09.11.2025). – Текст: электронный.
12. Платформа поставки данных ФНС России [Электронный ресурс]. – URL: https://geocheckivpd.nalog.gov.ru/kkt (дата обращения 09.11.2025). – Текст: электронный.
13. Официальный сайт Федеральной службы государственной статистики [Электронный ресурс]. – URL: https://rosstat.gov.ru (дата обращения 09.11.2025). – Текст: электронный.
14. Приказ Росстата от 29.05.2025 № 259 «Об утверждении Официальных статистических методологий по определению обобщающих показателей по статистике оптовой и розничной торговли» [Электронный ресурс]. – URL: http://ssl.rosstat.gov.ru/storage/mediabank/met_259-29052025.pdf (дата обращения 09.11.2025). – Текст: электронный.

9. Isakov A., Latypov, R., Repin, A., et al. Hard numbers: Open Microdata on Consumer Prices. *Den'gi i Kredit*. 2021. Vol. 80 No. 1. pp. 104–119. (In Russ.). DOI: 10.31477/rjmf.202101.104.

10. Surinov A. E. Big Data in Official Statistics: A Look at the Problem. *Voprosy Statistiki*. 2023. No. 2. pp. 5–22. (In Russ.). DOI: 10.34023/2313-6383-2023-30-2-5-22.

11. Federal Law No. 54-FZ of May 22, 2003 «On the Use of Cash Register Equipment for Payments in the Russian Federation». Available at: http://www.kremlin.ru/acts/bank/19505 (Accessed: 09.11.2025). (In Russ.).

12. The data delivery platform of the Federal Tax Service of Russia. Available at: https://geocheckivpd.nalog.gov.ru/kkt (Accessed: 09.11.2025). (In Russ.).
13. The Official Website of the Federal State Statistics Service. Available at: https://rosstat.gov.ru (Accessed: 09.11.2025). (In Russ.).

14. Rosstat Order No. 259 dated 05.29.2025 «On Approval of Official Statistical Methodologies for determining Summary Indicators for Wholesale and Retail Trade Statistics». Available at: http://ssl.rosstat.gov.ru/storage/mediabank/met_259-29052025.pdf (Accessed: 09.11.2025). (In Russ.).

Статья поступила в редакцию 20.11.2025; одобрена после рецензирования 08.12.2025; принята к публикации 12.12.2025.
The article was submitted 20.11.2025; approved after reviewing 08.12.2025; accepted for publication 12.12.2025.