

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И ОБРАБОТКА ДАННЫХ

Научная статья

УДК 519.24

DOI: 10.18101/2304-5728-2021-3-62-72

МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ПОСТУПЛЕНИЯ НАЛОГОВ В КОНСОЛИДИРОВАННЫЙ БЮДЖЕТ ПО СУБЪЕКТАМ РОССИЙСКОЙ ФЕДЕРАЦИИ

© **Леонова Ольга Васильевна**

кандидат физико-математических наук, доцент,

Байкальский государственный университет

Россия, 664003, г. Иркутск, ул. Ленина, 11

olga.olgaleonova@yandex.ru

Аннотация. В работе представлена методика использования метода множественной регрессии для исследования зависимости объемов налоговых поступлений от различных факторов. Для исследования были отобраны открытые данные ресурсов федеральных служб за 2020 г. с целью определения факторов, оказывающих значимое влияние на собираемость налогов, сборов и иных обязательных платежей в консолидированный бюджет Российской Федерации. С помощью шагового регрессионного анализа поведен качественный отбор факторов. В процессе анализа поступлений налогов был построен ряд эконометрических моделей с различным типом зависимости. Для выбора наилучшей модели протестированы гипотезы о значимости коэффициентов регрессии и сравнение показателей корреляции, которые дали возможность выделить линейную регрессионную модель, как самую наилучшую в плане аппроксимации исходных данных. Для построения прогнозов и описания количественных зависимостей анализируемых показателей использовалась оцененная множественная линейная регрессия.

Ключевые слова: регрессионный анализ; модель множественной регрессии; коэффициент корреляции; коэффициент эластичности; прогнозирование; эконометрическое исследование.

Для цитирования

Леонова О. В. Моделирование и прогнозирование поступления налогов в консолидированный бюджет по субъектам Российской Федерации // Вестник Бурятского государственного университета. Математика, информатика. 2021. № 3. С. 62–72.

Введение

Порядок формирования государственного бюджета Российской Федерации регламентируется Бюджетным кодексом РФ, в соответствии с которым определяется доходная и расходная части федерального бюджета и бюджетов субъектов (регионов) РФ. Федеральный бюджет направлен на исполнение полномочий государства, а бюджеты регионов обеспечивают полномочия, переданные государством субъектам Российской Федерации.

Федеральный бюджет и бюджеты регионов Российской Федерации вместе составляют общий бюджет страны или, как его сейчас называют, «консолидированный бюджет государства».

Основу доходной части бюджетов, как федерального, так и консолидированного, составляют налоговые поступления. Уровень социально-экономического развития каждого региона и страны в целом напрямую зависит от налогооблагаемой базы, которая, в свою очередь, определяется уровнем развития экономического потенциала каждого субъекта страны.

Сегодня в лучшем положении находятся регионы, где активно ведется добыча газа, нефти, высококалорийных углей и других полезных ископаемых, востребованных на мировом рынке. Не имеют больших проблем с формированием доходной части своего бюджета те регионы страны, на территории которых работают крупные промышленные предприятия по производству алюминия, стали, выпуску широкого ассортимента металлопроката. Регионы, где имеются большие запасы деловой древесины и ведется ее глубокая переработка, а также регионы с высоким уровнем занятости работоспособного населения.

Однако природа неравномерно наделила территории страны богатыми месторождениями, не везде развито промышленное производство и транспортная инфраструктура, а деятельность сферы услуг малоэффективна. Как правило, в таких регионах высокая безработица, низкий уровень заработной платы, слабая налогооблагаемая база и сложности с формированием доходной части регионального бюджета.

Учитывая приведенные аргументы, можно сказать, что определение факторов, непосредственно влияющих на формирование консолидированного бюджета, является актуальной задачей в настоящей момент. В работах [1; 2] авторами предлагается анализ поступлений в консолидированный бюджет РФ с точки зрения состава, структуры и тенденций¹. Однако выявление факторов, построение и изучение зависимостей поступления налоговых сборов еще недостаточно исследованы.

В данной статье исследуется зависимость размеров налоговых сборов и иных обязательных платежей в консолидированный бюджет Российской Федерации в 2020 г. от различных факторов, которые, по мнению экспертов, оказывают влияние на изучаемый показатель. Это дает возможность найти новые рычаги, оказывающие влияние на уровень собираемости налогов.

Для проведения исследования применяются эконометрические методы, которые успешно использовались в различных сферах для моделирования социально-экономических процессов [3–7]. В этом случае предполагается существование определенного типа связей между результирующими

¹ Анализ тенденций в бюджетно-налоговой сфере России // Российский экономический университет им. Г. В. Плеханова. 2020. Вып. 22. С. 1–6. DOI: 10.21686/atbns/22.2021 (дата обращения: 10.07.2021).

щим показателем и объясняющими переменными. Для такого рода исследований требуется грамотный подбор всех величин исходя из анализа структуры поступлений налогов, и к тому же появляется возможность проанализировать динамику социально-экономических процессов в субъектах Российской Федерации.

1 Постановка задачи

Исследуемая величина — поступление налогов, сборов и иных обязательных платежей в консолидированный бюджет Российской Федерации в 2020 г. Для исследования этой величины были отобраны различные факторы, которые, по мнению экспертов, оказывают влияние на объем поступления налогов. В качестве эмпирической базы используются данные по субъектам РФ из открытых источников: Федеральная служба государственной статистики и Федеральная налоговая служба.

Обозначим через Y зависимую переменную (результатирующий показатель) – поступление налогов, сборов и иных обязательных платежей в консолидированный бюджет в 2020 г. по 85 субъектам Российской Федерации (млн руб.). Исследуем зависимость этой переменной от следующих факторов:

x_1 — численность рабочей силы в возрасте от 15 лет и старше (тыс. чел.);

объем отгруженных товаров собственного производства, выполненных работ и услуг собственными силами:

x_2 — добыча полезных ископаемых (млн руб.),

x_3 — обрабатывающие производства (млн руб.),

x_4 — производство электрической энергии, газа и пара, кондиционирование воздуха (млн руб.),

x_5 — водоснабжение, водоотведение, организация сбора и утилизации отходов, деятельность по ликвидации загрязнений (млн руб.);

x_6 — объем работ, выполненных по виду «Строительство» (млн руб.);

x_7 — эксплуатационная длина железнодорожных путей общего пользования (км);

x_8 — оборот розничной торговли (млн руб.);

x_9 — среднедушевые денежные доходы населения (руб./месяц);

x_{10} — грузооборот автомобильного транспорта (млн тонн);

x_{11} — грузооборот железнодорожного транспорта (млн тонн).

В качестве эмпирической базы используются данные за 2020 г. по субъектам РФ из открытых источников: Федеральная служба государственной статистики и Федеральная налоговая служба².

² Федеральная налоговая служба www.analytic.nalog.ru. Федеральная служба государственной статистики www.rosstat.gov.ru (дата обращения: 15.07.2021).

Сами данные в статье не приводятся, поскольку их включение оказывает отрицательное влияние на оригинальность текста.

По двум субъектам — Республика Калмыкия и г. Севастополь — данные неполные, поэтому данные субъекты не будут участвовать в исследовании.

Для решения поставленной задачи построим уравнения множественной регрессии — зависимости поступления налогов, сборов и иных обязательных платежей в консолидированный бюджет страны от отобранных факторов (при расчетах будем использовать прикладной пакет «Анализ данных» MS Excel). Из построенных моделей множественной регрессии выберем ту, которая лучше других аппроксимирует исходные данные, по наилучшей модели построим точечные и интервальные прогнозы по каждому субъекту в отдельности и дадим содержательную интерпретацию найденных количественных взаимосвязей между исследуемыми переменными.

2 Отбор факторов и построение линейной модели множественной регрессии

Отбор факторов, которые будут участвовать в описании зависимости — это очень важная процедура для любого эконометрического исследования. Существуют различные методы отбора факторов, среди которых можно выделить шаговый регрессионный анализ [8], который будем использовать в данной работе.

Для выявления наиболее значимого фактора построим таблицу 1, которая содержит значения t -статистики для каждого параметра соответствующего фактора.

Таблица 1

Фактор	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
t -	-0,5	7,29	1,78	3,66	-2,3	0,79	-4	1,8	-0,7	0,49	-0,1

Наибольшее по модулю значение t -статистики имеет фактор x_2 — добыча полезных ископаемых, с этого фактора начнем включение независимых переменных в модель.

Парная линейная зависимость с переменной x_2 имеет вид:

$$Y = 108892,1 + 0,84x_2 + \varepsilon, \quad (1)$$

(2,12)

где ε — случайная составляющая, в скобках t -статистики коэффициентов регрессии.

Далее в модель (1) добавляем новый фактор, учитывая значение его t -статистики. После добавления каждого фактора тестируем новую модель на значимость коэффициентов с помощью критерия Стьюдента [9]. Если параметры при введенных факторах оказываются незначимыми, то дан-

ный фактор из модели исключается и в дальнейшем исследовании участия не принимает.

После нескольких шагов (добавления очередных и исключения незначимых факторов) оценена модель множественной линейной регрессии:

$$Y = -116291 + 0,63x_2 + 0,13x_3 + 0,46x_8 + \varepsilon, \quad (2)$$

$\begin{matrix} (-5,28) & (14,47) & (2,09) & (6,05) \end{matrix}$

$$R^2 = 0,917, \quad F = 293,$$

где R^2 — коэффициент детерминации, F — наблюдаемое значение критерия Фишера.

Качество модели (2) можно оценить как хорошее, поскольку все параметры значимые, коэффициент детерминации близок к 1, F -критерий имеет большое значение.

Из уравнения регрессии (2) можно сделать вывод, что между поступлением налогов в бюджет и факторами x_2 , x_3 , x_8 существует линейная зависимость.

3 Аппроксимация данных нелинейными моделями

Нелинейные модели нередко дают неплохие результаты при исследовании различных зависимостей. В данной работе построим следующие модели: полином второй степени, обратную, степенную, показательную, полулогарифмическую. Для нахождения оценок параметров перечисленных моделей необходима их линейризация. В зависимости от типа нелинейности для линейризации будем использовать замену переменных или логарифмические преобразования.

После линейризации всех нелинейных моделей и применения к ним метода наименьших квадратов [9] была построена совокупность различных типов зависимостей, которая приведена в таблице 2.

Таблица 2

Модели в линейризованном виде

Модели	Стандартный и скорректированный коэффициент детерминации	F -статистика
Линейная модель $Y = -116291 + 0,63x_2 + 0,13x_3 + 0,46x_8 + \varepsilon,$ $\begin{matrix} (-5,28) & (14,47) & (2,09) & (6,05) \end{matrix}$	$R^2 = 0,917$ $\bar{R}^2 = 0,914$	$F = 239$
Полином второй степени $Y = 3661,53 + 0,01x_2 - 0,075x_3 +$ $\begin{matrix} (0,17) & (0,11) & (-0,82) \end{matrix}$ $+ 0,37x_8 + 2,7 \cdot 10^{-7}x_2^2 + 1,63 \cdot 10^{-7}x_3^2 -$ $\begin{matrix} (3,46) & (6,95) & (3,69) \end{matrix}$ $- 1,3 \cdot 10^{-7}x_8^2 + \varepsilon$ $\begin{matrix} (-2,31) \end{matrix}$	$R^2 = 0,96$ $\bar{R}^2 = 0,95$	$F = 320$

<p>Полулогарифмическая модель</p> $Y = -3132777 + 65817 \ln x_2 + 2527 \ln x_3 + 219877 \ln x_8 + \varepsilon$ <p style="text-align: center;"> <small>(-5,22) (3,13) (0,05)</small> <small>(2,81)</small> </p>	$R^2 = 0,34$ $\bar{R}^2 = 0,32$	$F = 14$
<p>Обратная пропорциональность</p> $\frac{1}{Y} = 2,89 \cdot 10^{-5} - 1,2 \cdot 10^{-11} x_2 - 3,7 \cdot 10^{-11} x_3 + 3,2 \cdot 10^{-11} x_8 + \varepsilon$ <p style="text-align: center;"> <small>(5,26) (-1,13)</small> <small>(-2,38) (1,67)</small> </p>	$R^2 = 0,12$ $\bar{R}^2 = 0,08$	$F = 3,64$
<p>Степенная модель</p> $\ln Y = 0,31 + 0,21 \ln x_2 + 0,32 \ln x_3 + 0,41 \ln x_8 + \varepsilon$ <p style="text-align: center;"> <small>(0,41) (8) (5,13)</small> <small>(4,12)</small> </p>	$R^2 = 0,819$ $\bar{R}^2 = 0,813$	$F = 118$
<p>Показательная модель</p> $\ln Y = 10,73 + 1,3 \cdot 10^{-6} x_2 + 1,3 \cdot 10^{-6} x_3 - 4,5 \cdot 10^{-7} x_8 + \varepsilon$ <p style="text-align: center;"> <small>(95,5) (5,8) (4,04)</small> <small>(-1,12)</small> </p>	$R^2 = 0,63$ $\bar{R}^2 = 0,61$	$F = 44$

Для дальнейшего исследования следует оставить только те модели, у которых значимые параметры. Проверку значимости параметров моделей можно провести с помощью критерия Стьюдента. Сравнивая по модулю t -статистики параметров моделей с критической точкой $t_{кр} = 1,99$, можно сделать вывод, что только линейная и степенная модели имеют значимые параметры при входящих в модели факторах.

Теперь предстоит из этих двух моделей выбрать наилучшую. Для этого оценим тесноту связи переменных в каждой модели.

В случае линейной модели в качестве меры линейной зависимости используем коэффициент множественной корреляции $r = 0,96$. Для степенной модели этот показатель называется индекс корреляции $\eta = 0,57$.

Коэффициент корреляции ближе к 1, чем индекс корреляции, это означает, что линейная модель лучше аппроксимирует исходные данные по переменным x_2 , x_3 , x_8 . Поэтому линейную модель будем использовать для анализа и прогноза.

4 Прогноз на основе множественной линейной регрессии

По линейной модели (2) спрогнозируем поступление налогов, в качестве прогнозируемых значений переменных будем использовать среднее поступление налогов по каждому федеральному округу. Такие благополучные регионы, как г. Москва и Московская область, г. Санкт-Петербург и Ленинградская область выделим отдельно, учитывая большое отклонение значений для этих регионов.

Таблица 3

Среднее прогнозируемое значение количества поступления налогов, сборов и иных обязательных платежей в консолидированный бюджет РФ по федеральным округам

№	Субъект Российской Федерации	Прогнозное значение			Прогнозируемое поступление налогов (точечный прогноз)	Прогнозируемое поступление налогов (интервальный прогноз)
		x_2	x_3	x_8		
1	Центральный федеральный округ	22909,15	451726	274788,33	85235,259	(52568,45; 117902,07)
2	Московская область и г. Москва	619903,9	4926873,1	4318952,7	2932308,3	(2893311,7; 2971304,8)
3	Северо-Западный федеральный округ	106279,43	357375,33	172165,61	77749,65	(45076,98; 110422,32)
4	Ленинградская область и г. Санкт-Петербург	27478,55	1857641,5	1105546,9	658861,03	(625339,54; 692382,52)
5	Южный федеральный округ	64177,483	503365,32	631220,76	284341,62	(251469,67; 317213,56)
6	Северокавказский федеральный округ	3723,1857	68840,986	252294,37	12742,307	(-20111,64; 45596,25)
7	Приволжский федеральный округ	159318,11	703070,81	454621,04	288081,62	(255425,75; 320736,79)
8	Уральский федеральный округ	981092,93	931714,3	518538,7	866998,32	(833542,41; 900454,22)

9	Сибирский федеральный округ	233309,67	521565,66	323904,24	422989,39	(390340,84; 455637,94)
10	Дальневосточный федеральный округ	203812,3	127160,9	190671,91	118031,93	(85313,73; 150750,13)

5 Количественный анализ исследуемой зависимости

Поскольку параметры линейной модели являются абсолютными показателями тесноты связи, можно дать следующую интерпретацию модели (2):

- при увеличении объемов добычи полезных ископаемых на 1 млн руб. поступление налогов увеличится на 630 тыс. руб. при неизменном уровне объемов обрабатывающих производств и оборота розничной торговли;
- при увеличении объемов обрабатывающих производств на 1 млн руб. поступление налогов увеличится на 130 тыс. руб. при неизменном уровне объемов добычи полезных ископаемых и оборота розничной торговли;
- при увеличении объемов оборота розничной торговли на 1 млн руб. поступление налогов увеличится на 460 тыс. руб. при неизменном уровне объемов добычи полезных ископаемых и обрабатывающих производств.

Для анализа силы влияния независимых переменных на результат можно использовать частные коэффициенты эластичности, которые являются относительными показателями силы связи.

Вычисление частных коэффициентов эластичности привело к следующим результатам:

$$E_{x_2} = 0,43, E_{x_3} = 0,27, E_{x_8} = 0,75.$$

Учитывая эти показатели, можно сделать вывод об относительном влиянии изучаемых показателей:

- если добыча полезных ископаемых изменится на 1% от среднего уровня, это приведет к изменению поступления налогов на 0,43% при неизменном уровне объемов обрабатывающих производств и оборота розничной торговли;
- если объемы обрабатывающих производств изменятся на 1% от среднего уровня, это приведет к изменению поступления налогов на 0,27% при неизменном уровне добычи полезных ископаемых и оборота розничной торговли;
- если объем оборота розничной торговли изменится на 1% от среднего уровня, это приведет к изменению поступления налогов на 0,75% при неизменном уровне объемов добычи полезных ископаемых и обрабатывающих производств.

Показатели частных коэффициентов эластичности дают возможность ранжировать факторы по силе их воздействия на результирующий показатель. Таким образом, наиболее сильное влияние на поступление налогов оказывает оборот розничной торговли, затем добыча полезных ископаемых, и самое слабое влияние оказывает объем обрабатываемых производств.

Заключение

В статье для оценивания и верификации моделей использовался регрессионный анализ и метод множественной регрессии для исследования такого экономического процесса, как поступление налогов в консолидированный бюджет Российской Федерации. Поскольку консолидированный бюджет — это свод бюджетов соответствующих территорий, то для исследования использовались данные по всем субъектам РФ (за исключением Республики Калмыкия и г. Севастополь). Для более полного эконометрического исследования необходимо провести анализ мультиколлинеарности данных и тестирование случайных остатков на гетероскедастичность. Учитывая большой объем материала, результаты такого исследования составят основу для следующей работы.

Статья будет полезной для руководителей регионов страны при определении ими основных направлений экономического развития территорий. Финансовые и экономические ведомства субъектов Российской Федерации, специалисты, занимающиеся вопросами перспективного планирования, могут использовать изложенную методику при разработке дополнительных мер по увеличению доходной части региональных бюджетов за счет создания дополнительных производств и увеличения численности работающих в тех сферах, где может быть достигнут наибольший рост налогооблагаемой базы. Результаты работы могут быть полезны законодательным и исполнительным органам дотационных регионов в плане работы над увеличением налогооблагаемой базы (создание новых производственных объектов, развитие дополнительной сферы услуг, предоставление определенных льгот юридическим и физическим лицам, создающим дополнительные рабочие места).

Литература

1. Кисель А. И. Анализ поступления налогов, сборов и иных обязательных платежей в консолидированный бюджет России // Молодой ученый. 2020. № 7 (297). С. 161–165. URL: <https://moluch.ru/archive/297/67314/> (дата обращения: 08.07.2021). Текст: электронный.
2. Деметрюк Э. Э. Анализ налоговых доходов региональных бюджетов РФ // О некоторых вопросах и проблемах экономики и менеджмента : сборник научных трудов по итогам международной научно-практической конференции. Красноярск. 2014. 304 с. Текст: непосредственный.
3. Аксенюшкина Е. В., Леонова О. В. Моделирование криминогенной обстановки и прогнозирование количества преступлений в регионах Российской Федерации

рации // Вестник Бурятского государственного университета. Математика, информатика. 2020. № 2. С. 36–51. Текст: непосредственный.

4. Сорокина П. Г., Леонова О. В., Волченко Л. Ю. Моделирование налоговой базы по налогу на имущество организаций и прогнозирование поступлений на примере Иркутской области // Вестник УрФУ. Серия экономика и управление. 2018. Т. 17, № 2. С. 310–328. Текст: непосредственный.

5. Волченко Л. Ю., Мамонова Н. В., Завьялова Е. О. Моделирование влияния деятельности таможенных органов на социально-экономическое развитие и инвестиционную активность регионов // Инновационное развитие экономики. 2017. № 6 (42). С. 16–26. Текст: непосредственный.

6. Мамонова Н. В. Анализ нарушения гарантий независимости адвокатов сотрудниками правоохранительных органов при защите личности в уголовном судопроизводстве // Адвокатская практика. 2019. № 2. С. 45–51. Текст: непосредственный.

7. Антипина Н. В. Регрессионный анализ динамики экспорта нефти Российской Федерации // Интеллектуальный и ресурсный потенциалы регионов: активизация и повышение эффективности использования : материалы V Всероссийской научно-практической конференции, Иркутск, 16 мая 2019 г. Иркутск, 2019. С. 15–21. Текст: непосредственный.

8. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Москва: Вильямс, 2007. 392 с. Текст: непосредственный.

9. Елисеева И. И. Эконометрика. Москва: Юрайт, 2012. 449 с. Текст: непосредственный.

Статья поступила в редакцию 15.09.2021; одобрена после рецензирования 15.10.2021; принята к публикации 29.10.2021.

MODELING AND FORECASTING OF TAX RECEIPTS
TO THE CONSOLIDATED BUDGET FOR THE SUBJECTS
OF THE RUSSIAN FEDERATION

Olga V. Leonova
Cand. Sci. (Phys. and Math.), A/Prof.,
Baikal State University
11 Lenina St., Irkutsk 664003, Russia
olga.olgaleonova@yandex.ru

Abstract. The paper presents a method of using the multiple regression method to study the dependence of tax revenues on various factors. For the study, open data from the resources of Federal Services for 2020 were selected to determine the factors that have a significant impact on the collection of taxes, fees, and other mandatory payments to the consolidated budget of the Russian Federation. Qualitative selection of factors is carried out using step-by-step regression analysis. In the process of analyzing tax revenues, some econometric models with different types of dependence were built. To select the best model, hypotheses about the significance of regression coefficients and a comparison of correlation indicators were tested, which made it possible to identify the linear regression model as the best in terms of approximation of the initial data.

Keywords: regression analysis; multiple regression model; correlation coefficient; elasticity coefficient; forecasting; econometric research.

For citation

Leonova O. V. Modeling and Forecasting of Tax Receipts to the Consolidated Budget for the Subjects of the Russian Federation // Bulletin of Buryat State University. Mathematics, Informatics. 2021. N. 3. Pp. 62–72.

The article was submitted 15.09.2021; approved after reviewing 15.10.2021; accepted for publication 29.10.2021.