

ОРИГИНАЛЬНАЯ СТАТЬЯ

УДК 330.43(045)

© Карабута А. А., Карабута Е. А., 2023

# Анализ зависимости среднемесячной заработной платы от различных социально-экономических показателей по субъектам России



**Анна Александровна Карабута**, студентка факультета экономики и бизнеса, Финансовый университет, Москва, Россия

**Anna A. Karabuta**, student, Faculty of Economics and Business, Financial University, Moscow

*Russiakarabutaa02@gmail.com*



**Екатерина Александровна Карабута**, студентка факультета экономики и бизнеса, Финансовый университет, Москва, Россия

**Ekaterina A. Karabuta**, student, Faculty of Economics and Business, Financial University, Moscow, Russia

*karabuta020819@gmail.com*

## АННОТАЦИЯ

Статья посвящена анализу влияния различных экономических факторов на уровень среднемесячной заработной платы с помощью построения эконометрической модели. Актуальность работы обусловлена тем, что в России наблюдается большая дифференциация между уровнем жизни в разных регионах, и государственные органы стремятся к максимальному уменьшению этой дифференциации. Такой показатель, как среднемесячная заработная плата, является основным показателем, позволяющим измерить уровень жизни населения, поэтому важно проанализировать факторы, которые оказывают на него наибольшее влияние. Новизна работы состоит в многофакторном анализе заработной платы по всей совокупности регионов России. Целью исследования является выявление основных факторов, влияющих на среднемесячную заработную плату по регионам РФ. В рамках достижения цели авторами проводится подробный анализ факторов различного характера, оказывающих влияние на уровень среднемесячной заработной платы в регионе. Исходя из проведенного теоретического анализа строится эконометрическая модель зависимости среднемесячной заработной платы от экономических показателей регионов, проверяется адекватность данной модели. В заключение дана экономическая интерпретация полученных результатов, определены факторы, оказывающие наиболее значимое влияние на объект исследования.

Научный руководитель: **Ященко Н.А.**, доцент департамента математики, Финансовый университет, Москва, Россия /  
Scientific supervisor: **Yashchenko N.A.**, Associate Professor, Department of Mathematics, Financial University, Moscow, Russia.

**Ключевые слова:** среднемесячная заработная плата; экономические факторы; инвестиции в основной капитал на душу населения; прожиточный минимум; валовой региональный продукт на душу населения; эконометрическая модель; теорема Гаусса-Маркова; уровень жизни населения

**Для цитирования:** Карабута А. А., Карабута Е. А. Анализ зависимости среднемесячной заработной платы от различных социально-экономических показателей по субъектам России. *Научные записки молодых исследователей*. 2023;11(3):28–39.

ORIGINAL PAPER

## Analyzing of Socio-Economic Factors and Average Monthly Wages in Russia

### ABSTRACT

*The paper analyzes the impact of various economic factors on the level of average monthly wages using the construction of an econometric model. The relevance of the work is because in Russia there is a large differentiation between the standard of living in different regions. Government agencies strive to minimize this differentiation as much as possible. The average monthly wage is a key indicator to measure the population's standard of living, so analyzing the factors that impact it is essential. The novelty of the work lies in the multivariate analysis of wages for the entire set of regions of Russia. The purpose of the study is to identify the major factors affecting the average monthly wage in the regions of the Russian Federation. As part of achieving the goal, the authors carry out a detailed analysis of various factors that affect the level of average monthly wages in the region. The research constructed a model to analyze the dependence of wages on economic indicators. Also, the authors identified the most significant factors affecting the object of the study.*

**Keywords:** average monthly wage; economic forces; investment in fixed assets per capita; living wage; gross regional product per capita; econometric model; Gauss-Markov theorem; living standards

**For citation:** Karabuta AA., Karabuta E. A. Analyzing of socio-economic factors and average monthly wages in Russia. *Nauchnye zapiski molodykh issledovatelei = Scientific notes of young researchers*. 2023;11(3):28–39.

### Введение

Заработная плата является основным источником дохода у большинства домохозяйств в России. В рамках существующих различий в степени развития регионов России и среднемесячных заработных плат в регионах возникает вопрос о взаимосвязи этих явлений. Поэтому анализ формирования среднемесячной заработной платы от других экономических показателей по регионам Российской Федерации является актуальным вопросом, имеющим научный и практический интерес.

Объектом исследования выступают экономические показатели регионов России.

Предметом исследования служит влияние экономических показателей регионов России на уровень среднемесячной заработной платы.

В статье использовались такие методы, как эконометрическое моделирование, регресси-

онный и корреляционный анализ, выдвижение и опровержение гипотез, графический метод.

В качестве информационной базы для проведения данного исследования выступают статистические данные официальной статистики Росстат по регионам РФ за 2020 г.

Расчет всех показателей проводился в Microsoft Excel с помощью встроенных функций и инструментов анализа.

### Теоретические предпосылки составления эконометрической модели и анализ существующих эконометрических моделей

Одним из основных показателей уровня жизни населения является средний размер заработной платы как основной источник дохода. Заработная плата — это вознаграждение за труд, которое получает работник в зависимости от различных

факторов (опыт, количество отработанных часов, образование, должность). Средняя зарплата рассчитывается как среднее арифметическое по ограниченной выборке: государства, региона, города или предприятия.

Самый высокий уровень средней заработной платы в 2020 г. наблюдался в Чукотском АО – 121 193,18 руб., что более чем в 4 раза больше, чем в регионе с самой низкой оплатой труда – Ивановской области, где среднемесячная заработная плата в 2020 г. составляла 28 173,21 руб.<sup>1</sup>

Сравнение заработной платы по субъектам России выдвигает необходимость оценить региональные различия. При сравнении заработной платы необходимо учитывать отраслевую специфику региона и географические, природно-климатические условия, которые влияют на стоимость жизни. Включение в модель такого показателя, как региональный прожиточный минимум для всего населения (среднегодовой), обусловлено необходимостью отражения влияния стоимости жизни в регионе на среднемесячную заработную плату [1].

Оценить степень развития экономики региона, отраслевую специфику можно с помощью нескольких показателей:

1) валовый региональный продукт (ВРП) – представляет собой валовую добавленную стоимость товаров и услуг, созданную резидентами региона, и определяется как разница между выпуском и промежуточным потреблением. ВРП является обобщающим показателем экономической деятельности региона, характеризующим процесс производства товаров и услуг для конечного использования. Включение в эконометрическую модель, объясняющую среднемесячную заработную плату, валового регионального продукта позволяет учесть степень различия экономического развития регионов<sup>2</sup>;

2) инвестиции в основной капитал – совокупность затрат, направленных на строительство, реконструкцию объектов, которые приводят к увели-

чению их первоначальной стоимости, приобретение машин, оборудования, транспортных средств, производственного и хозяйственного инвентаря. Включение этого показателя в модель прогнозирования среднемесячной заработной платы помогает отразить часть инвестированных средств, приходящихся на выплату заработной платы работникам<sup>3</sup>;

3) стоимость основных фондов – первоначальная стоимость основных средств, которая была изменена в ходе достройки, модернизации, дооборудования, реконструкции и частичной ликвидации, а также переоценки и обесценения активов. Этот показатель сильно коррелирует с инвестициями в основной капитал.

Также в рамках исследования зависимости среднемесячной заработной платы особое внимание должно быть уделено численности населения региона, так как различие в этом показателе является существенным и серьезно влияет на величину валового регионального продукта и инвестиций в основной капитал. Поэтому для построения собственной эконометрической модели авторами было принято решение использовать показатели ВРП и инвестиций в основной капитал не в абсолютном значении, а на душу населения, который также рассчитывается Росстатом. Показатель прожиточного минимума, представленный в официальной статистической отчетности, и так рассчитывается на душу населения.

Важным показателем, который может оказать влияние на среднемесячную заработную плату, может быть число предприятий в регионе, так как большое количество предприятий может создавать конкуренцию среди работодателей, что вынуждает предприятия предлагать более выгодные материальные условия труда. Однако стоит отметить, что из анализа статистических данных авторами было замечено, что некоторые регионы с высокими значениями объясняющей переменной (например, Ямало-Ненецкий АО, Ханты-Мансийский АО, Ненецкий АО и т.д.) имеют относительно малое количество предприятий, что связано с отраслевой спецификой региона (основная отрасль – нефтегазовая добыча) и преобладание в структуре предприятий крупных промышленных компаний. Поэтому реальное

<sup>1</sup> Среднемесячная номинальная начисленная заработная плата работников по полному кругу организаций по субъектам Российской Федерации: Официальная статистика. Росстат. М.; 2021. URL: [https://rosstat.gov.ru/storage/mediabank/tab2-zpl\\_09.xlsx](https://rosstat.gov.ru/storage/mediabank/tab2-zpl_09.xlsx) (дата обращения: 12.12.2022).

<sup>2</sup> Методические пояснения к показателю «Валовый региональный продукт». URL: <https://www.gks.ru> 2008. URL: [https://www.gks.ru/bgd/regl/b08\\_14p/isswww.exe/stg/d2/11-08.htm](https://www.gks.ru/bgd/regl/b08_14p/isswww.exe/stg/d2/11-08.htm) (дата обращения: 13.12.2022).

<sup>3</sup> Методические пояснения к показателю «Инвестиции в основной капитал». URL: <https://www.gks.ru> 2008. URL: [https://www.gks.ru/bgd/regl/b08\\_13/IssWWW.exe/Stg/d6/23-23.htm](https://www.gks.ru/bgd/regl/b08_13/IssWWW.exe/Stg/d6/23-23.htm) (дата обращения: 13.12.2022).

влияние данного показателя на объясняющую переменную будет оценено с помощью корреляционной матрицы.

Также на значение эндогенной переменной может оказать влияние такой социально-экономический показатель, как уровень занятости. Низкий уровень занятости в регионе может свидетельствовать о конкуренции между работниками за возможность занимать рабочее место, что снижает требование работников об уровне заработной платы, и, соответственно, низкие значения уровня занятости могут снижать среднемесячную заработную плату в регионе.

Анализ источников литературы по выбранной теме показал, что на данный момент большинство существующих научных исследований по выбранной теме проводятся или как анализ временных рядов в рамках одного субъекта РФ или Российской Федерации в целом, или как пространственные модели, которые исследуют регионы, принадлежащие одному федеральному округу или одной отрасли.

Большинство моделей, выявляющие взаимосвязь среднемесячной заработной платы, построены как модели линейной парной регрессии, которая отражает влияние только одного фактора на среднемесячный уровень начисленной номинальной заработной платы. Как правило, таким факторов в рамках существующих моделей выступает величина прожиточного минимума. К другим факторам, которые включают в модель, относятся уровень занятости, расходы регионального бюджета [2].

Также была исследована статья, которая использует кластерный анализ в качестве основного метода анализа региональной дифференциации регионов России. В рамках данного исследования авторы И. В. Орлова и Е. С. Фионова выявили два основных фактора, которые влияют на уровень среднемесячной заработной платы по регионам: стоимость фиксированного набора товаров и услуг и инвестиции в основной капитал. Деление регионов на различные группы и использование фиктивных переменных, с одной стороны, делает эконометрическую модель более универсальной для всех регионов России, но с другой стороны – увеличивает временные издержки на разделение регионов на группы и уменьшает удобство интерпретации построенной эконометрической модели [3].

В рамках исследования авторами будет использоваться модель множественной регрессии, где будут учитываться несколько факторов и их влияние на объясняемую переменную, что повышает научную и практическую значимость работы.

Таким образом, посредством теоретического и экономического смысла объясняющей переменной (среднемесячного уровня заработной платы по регионам РФ), а также из анализа научных исследований на данную тему авторами были сделаны выводы об исследовании влияния на объясняющую переменную таких факторов, как уровень занятости в регионах, число предприятий в регионе, величина основных фондов, прожиточный минимум, ВРП на душу населения, инвестиции в основной капитал на душу населения.

### Построение спецификации эконометрической модели

Построение спецификации модели включало несколько этапов:

1. Построение корреляционной матрицы, использование которой позволило выявить факторы, имеющие наиболее тесную связь с объясняемой переменной. В *табл. 1* приведена полученная корреляционная матрица. Можно заметить, что наибольшие парные коэффициенты корреляции между среднемесячной заработной платой наблюдаются у таких показателей, как прожиточный минимум, ВРП на душу населения и инвестиции в основной капитал на душу населения. Получившиеся результаты корреляционного анализа стали основанием включения перечисленных факторов в конечную регрессионную модель.

2. Проверка оставшихся факторов на мультиколлинеарность. Для этого была построена новая корреляционная матрица оставшихся факторов и оценены полученные коэффициенты корреляции (отсутствие частичной мультиколлинеарности означает неперевышение всех коэффициентов парной корреляции между факторами уровня 0,75). Получившаяся корреляционная матрица приведена в *табл. 2*.

3. Выявление вида связи между объясняемой и объясняющими переменными. Для этого были построены диаграммы рассеивания и выявлены линейные связи между эндогенной переменной и регрессорами.

4. Последовательное включение в уравнение спецификации факторов для анализа влияния

каждого из них на скорректированный коэффициент детерминации. Каждый из включенных факторов увеличивал скорректированный коэффициент детерминации, что еще раз подтвердило значимость каждого из факторов.

5. Построение конечной спецификации модели.

Конечная спецификация модели имеет вид, отраженный в формуле (1):

$$AS_i = a_1 * LW_i + a_2 * GRPPC_i + a_3 * FCIPC_i + u_i, \quad (1)$$

где  $a_i$  – коэффициенты регрессии;

$AS_i$  – среднемесячная начисленная заработная плата;

$LW_i$  – прожиточный минимум;

$GRPPC_i$  – валовой региональный продукт на душу населения;

$FCIPC_i$  – величина инвестиций в основные средства на душу населения;

$u_i$  – случайные остатки.

В спецификации не приведены оценки случайных остатков модели, так как проверка предположений о нулевом математическом ожидании и постоянстве дисперсии случайных остатков проведены в анализе соблюдения предпосылок теоремы Гаусса-Маркова.

Невключение в модель свободного члена имеет две предпосылки: во-первых, при включении в модель свободного члена он оказывается незначим по критерию Стьюдента, что повышает вероятность прийти к выводу о неадекватности модели; во-вторых, включение в модель такого фактора, как прожиточный минимум, который по своим значениям очень схож с величиной минимального размера оплаты труда (невключение в модель МРОТ объяснено тем, что во многих регионах РФ принят единый уровень МРОТ и включение этого фактора понизило бы его «эффективность»), который при возрастающей функции и экономическому смыслу данного показателя может выступать свободным членом в модели прогнозирования средней заработной платы.

Оцененная модель (1) имеет вид формулы (2):

$$\begin{cases} AS_i = 2,18 * LW_i + 0,03 * GRPPC_i + \\ 0,015 * FCIPC_i + u_i, \\ Sa_1 = 0,12 \quad Sa_2 = 0,0037 \quad Sa_3 = 0,006 \\ \sigma_u = 3596,069 \end{cases} \quad (2)$$

где  $Sa_i$  – средние квадратические (стандартные) ошибки коэффициентов;

$\sigma_u$  – среднее квадратическое отклонение случайных остатков.

### Проверка на ошибки спецификации переменных в уравнении регрессии

Для спецификаций эконометрических моделей могут быть характерны несколько видов ошибок. Наличие ошибок первого, второго или третьего рода могут привести к лишению оцененных параметров модели свойств эффективности и несмещенности, что может привести к неверным выводам об адекватности модели и невозможности ее практического использования [4].

**Ошибка первого рода** заключается в неверном выборе типа функции. Данной ошибке характерны 3 симптома:

1. Несоответствие диаграммы рассеяния, построенной по выборке, графику функции. Для модели множественной регрессии невозможно проверить данный симптом.

2. Длительное постоянство знака оценок случайных остатков в упорядоченных уравнениях наблюдений по возрастанию значений объясняющей переменной.

На *рисунке* представлено распределение случайных остатков по возрастанию суммы модулей значений объясняющей переменной в уравнениях наблюдений. На диаграмме можно заметить, что не наблюдается длительное постоянство знаков случайных остатков (максимальная задержка знака – 5 наблюдений).

Разительное отличие одноименных коэффициентов в двух оцененных моделях (моделях, оцененных по двум разным выборкам, с примерно одинаковым объемом наблюдений).

В *табл. 3* приведены оцененные коэффициенты ( $\tilde{a}'_i$ ) и их среднеквадратической ошибки по первой части выборки ( $S_{\tilde{a}'_i}$ ), рассчитаны доверительные интервалы для коэффициентов из второй части выборки (доверительный интервал – интервал, нижней границей которого является разница оцененного коэффициента и его среднеквадратической ошибки ( $\tilde{a}'_i - S_{\tilde{a}'_i}$ ), а верхней границей – сумма этих показателей ( $\tilde{a}''_i + S_{\tilde{a}'_i}$ ), а также приведены одноименные оцененные коэффициенты из второй части выборки ( $S_{\tilde{a}''_i}$ ) и их доверительные интервалы.

Таблица 1

Корреляционная матрица

	Средняя ЗП за 2020	Прожиточный минимум	ВРП на душу населения	Инвестиции в основной капитал на душу населения	Уровень занятости	Число предприятий	Величина основных фондов
Средняя ЗП за 2020	1						
Прожиточный минимум	0,87	1					
ВРП на душу населения	0,81	0,68	1				
Инвестиции в основной капитал на душу населения	0,77	0,28	0,31	1			
Уровень занятости	0,47	0,64	0,59	0,35	1		
Число предприятий	0,26	0,09	0,06	0,84	0,26	1	
Величина основных фондов	0,41	0,20	0,24	0,90	0,38	0,95	1

Источник: составлено авторами.

Таблица 2

Корреляционная матрица

	Средняя ЗП за 2020 г.	Прожиточный минимум	ВРП на душу населения	Инвестиции в основной капитал на душу населения
Средняя ЗП за 2020 г.	1			
Прожиточный минимум	0,78	1		
ВРП на душу населения	0,87	0,5	1	
Инвестиции в основной капитал на душу населения	0,71	0,35	0,65	1

Источник: составлено авторами.

Все коэффициенты регрессии принадлежат интервалам оценок коэффициентов или один из интервалов входит в другой, что говорит об отсутствии 3-го симптома первого типа ошибки.

**Ошибка второго рода** заключается во включении в модель незначущей объясняющей пере-

менной. Проверка наличия данной ошибки проводится через проведение *t*-критерия Стьюдента для каждого оцененного значения коэффициента регрессора.

Согласно критерию Стьюдента: если  $\left| \frac{\tilde{a}_t}{S_{\tilde{a}_t}} \right| \leq t_{\text{крит}}$  =>  $x_t$  не значим.

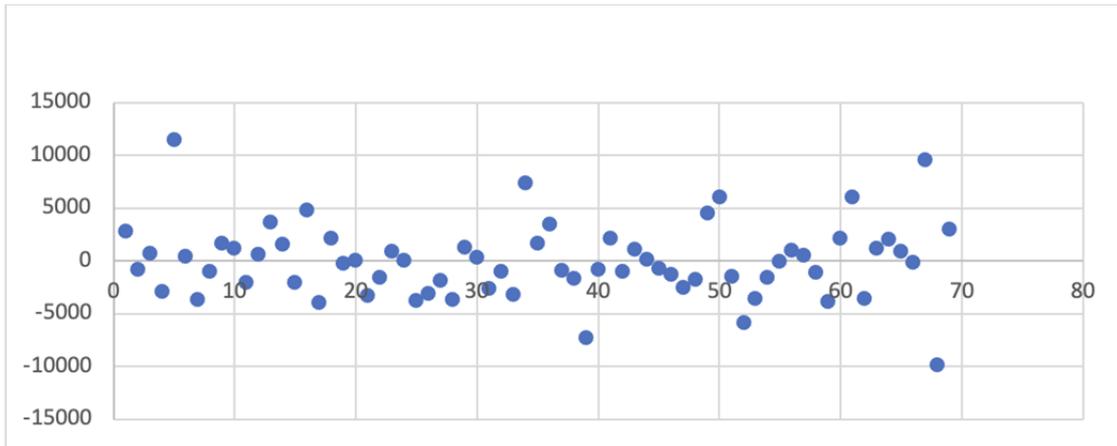


Рис. Распределение случайных остатков  $u_i$

Источник: составлено авторами.

Таблица 3

Числовые параметры для оценки одноименных коэффициентов модели по двум частям выборки

$a_i$	$\tilde{a}'_i$	$S_{\tilde{a}'_i}$	$\tilde{a}'_i - S_{\tilde{a}'_i}$	$\tilde{a}'_i + S_{\tilde{a}'_i}$	$\tilde{a}''_i$	$S_{\tilde{a}''_i}$	$\tilde{a}''_i - S_{\tilde{a}''_i}$	$\tilde{a}''_i + S_{\tilde{a}''_i}$
$a_1$	2,42	0,25	2,17	2,66	2,024	0,26	1,77	2,28
$a_2$	0,019	0,0089	0,01	0,028	0,033	0,006	0,027	0,04
$a_3$	0,028	0,032	-0,0041	0,06	0,012	0,007	0,0047	0,019

Источник: составлено авторами.

где:  $\tilde{a}_i$  – оцененные коэффициенты регрессии;  
 $S_{\tilde{a}_i}$  – среднеквадратическая ошибка оцененного коэффициента.

$T_{крит}$  – двусторонняя  $(1-\alpha)$  квантиль распределения Стьюдента. В модели (2)  $\left| \frac{\tilde{a}_1}{S_{\tilde{a}_1}} \right| = 17,59$ ,  $\left| \frac{\tilde{a}_2}{S_{\tilde{a}_2}} \right| = 8,067$ ,  $\left| \frac{\tilde{a}_3}{S_{\tilde{a}_3}} \right| = 2,45$ , при  $t_{крит} = 1,9965$ . Таким образом, все переменные регрессии являются значимыми.

**Ошибка третьего рода** заключается в пропуске в линейной регрессионной модели значащей объясняющей переменной. В модели (2) данная ошибка не может быть проверена с помощью теста Дарбина-Уотсона, однако отсутствие этой ошибки можно констатировать через отсутствие симптомов первой ошибки и диаграмму рассеивания случайных остатков (см. рисунок).

Таким образом, проведенный анализ типичных ошибок первого, второго и третьего рода позволяет сделать вывод об их отсутствии в оцениваемой эконометрической модели и ее пригодности для дальнейшего использования.

**Проверка выполнения предпосылок теоремы Гаусса-Маркова**

Примененный метод МНК дает наилучшие оценки параметров при выполнении 4 условий теоремы Гаусса-Маркова [4]:

$$E(u_1) = E(u_2) = \dots = E(u_n) = 0;$$

$$Var(u_1) = Var(u_2) = \dots = Var(u_n) = \sigma_u^2;$$

$$Cov(u_i, u_j) = 0 \text{ при } i \neq j;$$

$$Cov(u_i, x_{1j}) = 0.$$

**Первую предпосылку** теоремы Гаусса-Маркова можно признать выполненной при отклонении нулевой гипотезы  $H_0 : a_1 = a_2 = a_3 = 0$ , где  $a_i$  – коэффициенты перед регрессорами, и принятии альтернативной гипотезы  $H_1$  (о неравенстве нулю хотя бы одного из коэффициентов перед регрессорами) [5].

Для принятия одной из выдвинутых гипотез проведем  $F$ -тест и используем  $R$ -критерий.

Для оценки способности регрессоров объяснить эндогенную переменную и исследования качества регрессии вычислим коэффициент детерминации ( $R^2$ ).

Для модели (2) коэффициент детерминации равен 0,9924. Полученный нецентрированный коэффициент детерминации свидетельствует о сильной связи между выбранной эндогенной переменной и преопределенными переменными, что свидетельствует о высоком качестве регрессии.

Чтобы доверять полученному коэффициенту детерминации, необходимо провести  $F$ -тест. Ситуация совершенно плохой спецификации модели равносильна справедливости статистической гипотезы  $H_0 : a_1 = a_2 = a_3 = 0$ . Статистикой данной гипотезы против альтернативы  $H_1 = \overline{H_0}$  служит случайная переменная  $F$ . Алгоритм выполнения  $F$ -теста:

$$1. F = \frac{R^2 / k}{(1 - R^2) / (n - k)}, \text{ где } R^2 \text{ – коэффициент}$$

детерминации;  $k$  – количество регрессоров в модели;  $n$  – объем обучающей выборки. Для модели (2):  $F = 2864,95$ .

2. Вычисление  $F_{\text{крит}}$  – значение, обратное (правостороннему)  $F$ -распределению вероятности со степенями свободы с уровнем значимости  $\alpha = 0,05$ ,  $v_1 = k; v_2 = n - k$ .  $F_{\text{крит}} = 2,74$ , с уровнем значимости  $\alpha = 0,05$ ,  $v_1 = 3; v_2 = 66$ .

3. Сравнения полученных показателей  $F$  и  $F_{\text{крит}}$ . Если  $F > F_{\text{крит}}$ , т.е. гипотеза  $H_0$  отвергается и принимается альтернативная гипотеза  $H_1$ . Таким образом, мы можем сделать вывод об удовлетворительном качестве регрессии, т.е. о способности регрессоров в рамках линейной модели объяснять значение эндогенной переменной.

**Вторая предпосылка** теоремы Гаусса Маркова предполагает выполнение условия гомоскедастичности случайных остатков. Для ее выполнения необходимо подтверждение гипотезы  $H_0 : Var(u_1) = Var(u_2) = \dots = Var(u_n) = \sigma_u^2$ . Данная пред-

посылка проверяется с помощью проведения теста Голдфелда-Квандта. Данный тест проходит по следующему алгоритму:

1. Упорядочить данные по возрастанию  $z_i$  – суммы модулей значений объясняющих переменных.

2. Разделить данные на 3 части, где  $n' = n/3$  при выполнении условия минимальности ( $n' > k$ ).

3. По 1 части  $n'$  найти МНК-оценки,  $ESS_1$  – объясненная с помощью регрессии сумма квадратов отклонений.

4. По последней части  $n'$  найти МНК-оценки,  $ESS_2$ .

$$5. \text{Вычислить статистику } GQ = \frac{ESS_1}{ESS_2}; GQ^{-1} = \frac{1}{GQ}.$$

6. Вычислить  $F_{\text{крит}}$  со степенями свободы  $v_1 = v_2 = n' - k$ .

7. Если  $GQ \leq F_{\text{крит}}$  и  $GQ^{-1} \leq F_{\text{крит}}$ , то 2-я предпосылка теоремы Гаусса-Маркова выполняется, т.е. случайные остатки гомоскедастичны.

Проведем данный тест для модели (2):

$$GQ = \frac{ESS_1}{ESS_2} = 0,61;$$

$$GQ^{-1} = \frac{1}{GQ} = 1,64;$$

$$F_{\text{крит}} = 2,12, \text{ со степенями свободы } v_1 = v_2 = n' - k = 20.$$

Оба неравенства ( $GQ \leq F_{\text{крит}}$  и  $GQ^{-1} \leq F_{\text{крит}}$ ) верные, соответственно, мы можем принять гипотезу  $H_0$  и сделать вывод о гомоскедастичности случайных остатков.

**Третья предпосылка** теоремы Гаусса-Маркова предполагает некоррелированность случайных остатков. Для исследования предпосылки об отсутствии автокорреляции случайных остатков авторы воспользуются графическим методом и тестом Бройша-Годфри.

В рамках графического метода выборка была отсортирована по возрастанию  $z_i$  и построена диаграмма рассеивания значений  $u_i$ . На рисунке видно, что случайные остатки  $u_i$  распределены вокруг 0 и часто меняют знак, что является признаками некоррелированности случайных остатков.

Тест Бройша-Годфри исследует наличие автокорреляции высших порядков при помощи вспомогательной эконометрической модели, в кото-

рой в роли зависимой переменной выступают случайные остатки первоначальной модели, а в качестве объясняющих переменных – значения случайных остатков в предыдущих наблюдениях. Преимуществом данного теста является то, что он может проводиться для однородных моделей (моделей, в рамках которых параметр  $a_0$  не подлежит определению) [6].

Вспомогательная модель для проведения теста Бройша-Годфри, при порядке автокорреляции случайных остатков  $p = 2$ , имеет вид:

$$u_i = g_0 + g_1 * u_{i-1} + g_2 * u_{i-2} + e_i, \quad (3)$$

где  $u_i$  – случайный остаток  $i$ -го наблюдения;

$g_i$  – коэффициенты перед регрессорами вспомогательной модели;

$u_{i-1}$  – случайный остаток  $(i - 1)$ -го наблюдения;

$u_{i-2}$  – случайный остаток  $(i - 2)$ -го наблюдения;

$e_i$  – случайные остатки.

Выдвигается нулевая гипотеза

$$H_0 : g_0 = g_1 = g_2 = 0.$$

Алгоритм теста Бройша-Годфри следующий:

1. Оценивание МНК-модели (1) и нахождение  $\hat{u}_i$ .

2. Определение предполагаемого порядка автокорреляции случайных остатков  $p$  [для модели (2) предполагаемый порядок автокорреляции – 2].

3. Оценивание МНК вспомогательной модели (3).

4. Определение значения статистики Бройша-Годфри по формуле:

$$BG = TR_{aux}^2, \quad (4)$$

где  $BG$  – значение статистики Бройша-Годфри;  $T$  – количество наблюдений вспомогательной модели ( $T = n - p$ );  $R_{aux}^2$  – коэффициент детерминации вспомогательной модели.

$$BG = TR_{aux}^2 = 1,85.$$

5. Определения критического уровня  $\chi^2$  статистики с уровнем значимости равным 0,05 и количеством степеней свободы  $p$ .  $\chi_{крит}^2 = 5,99$ .

6. Сравнение полученных значений  $BG$  и  $t_{крит}$ : при соблюдении неравенства  $BG < \chi_{крит}^2$ , гипо-

теза  $H_0$  принимается и делается вывод об отсутствии автокорреляции случайных остатков.  $1,85 < 5,99$ , соответственно неравенство верное для данной модели, гипотеза  $H_0$  принимается и делается вывод об отсутствии автокорреляции случайных остатков и соблюдении третьей предпосылки теоремы Гаусса-Маркова.

**Четвертая предпосылка** теоремы Гаусса-Маркова предполагает, что ковариация между объясняющими переменными и случайными возмущениями равна нулю.

$$H0 : Cov(X_{mi}, u_j) = 0.$$

Мы можем принять эту гипотезу, так как все данные в матрице  $X$  (матрице объясняющих переменных) определены, а ковариация между детерминированной матрицей и случайной величиной  $u_i$  равна нулю.

Таким образом, примененный МНК дал наилучшие оценки параметров регрессии для модели (1).

Оцененная модель (1) имеет вид:

$$\begin{cases} AS_i = 2,18 * LW_i + 0,03 * GRPPC_i + \\ + 0,015 * FCIPC_i + u_i, \\ Sa_1 = 0,12 \quad Sa_2 = 0,0037 \quad Sa_3 = 0,006 \\ \sigma_u = 3596,069. \end{cases} \quad (4)$$

По коэффициентам регрессоров оцененной модели (4) можно сделать следующие выводы: все объясняющие переменные модели имеют прямую зависимость с объясняемой переменной. Увеличение прожиточного минимума в регионе на 1000 руб. увеличивает среднемесячную начисленную заработную плату на 2180 руб. Увеличение валового регионального продукта на душу населения на 1 млн руб. увеличивает среднемесячную начисленную заработную плату на 30000 тыс. руб. Увеличение инвестиций в основной капитал на душу населения на 1 млн руб. увеличивает среднемесячную начисленную заработную плату на 15000 тыс. руб.

### Проверка адекватности оцененной модели

Следующим шагом исследования является осуществление проверки адекватности полученной оцененной эконометрической модели. Для выполнения данного этапа необходимо провести процедуру ин-

Таблица 4

**Контролирующая выборка**

Контролирующая выборка					
№	Регион	AS	LW	GRPPC	FCIPC
1	Кировская область	32 360,98	10 604,0	315 153,9	54 031,39
2	Республика Крым	33 617,37	10 851,0	270 363,8	102 673,068
3	Свердловская область	43 139,23	10 828,0	588 270,0	95 039,62
4	Липецкая область	36 793,13	9 761,5	536 151,3	151 531,94

Источник: официальная статистика Росстат. Федеральная служба государственной статистики. URL: <https://rosstat.gov.ru/> (дата обращения: 10.12.2022).

Таблица 5

**Рассчитанные значения для проверки адекватности модели**

№	$\widetilde{AS}_0$	q0	$S_{\widetilde{AS}_0}$	$t_{\text{крит}}$	$\widetilde{AS}_0^-$	$\widetilde{AS}_0^+$
1	33 230,46	0,022	3 607,72	1,9965	26 027,41	40 433,5
2	33 150,086	0,045	3 647,71		25 867,21	40 432,97
3	42 415,017	0,041	3 641,087		35 145,35	49 684,68
4	39 667,38	0,022	3 608,38		32 463,015	46 871,74

Источник: составлено авторами.

тервального прогнозирования средней заработной платы, используя контролируемую выборку.

Проверка адекватности оцененной модели (2) осуществляется по следующему алгоритму:

1. Вычисление точечных прогнозных значений по оцененной модели для каждого набора данных контролирующей выборки.

2. Вычисление  $q_0 = \overline{x_0} (X^T X)^{-1} \overline{x_0}$ ,  $\overline{x_0} = (1, x_0)^T$  для каждого набора данных контролирующей выборки, где  $q_0$  – параметр точности прогноза;  $X$  – матрица с данными, по которым производится оценка модели;  $X^T$  – транспонированная матрица  $X$ ;  $\overline{x_0} = (1, x_0)^T$ .

3. Вычисление ошибки прогноза:

$$S_{\widetilde{AS}_0} = \widetilde{\sigma}_u \times \sqrt{q_0 + 1}$$

для каждого набора данных контролирующей выборки, где  $S_{\widetilde{AS}_0}$  – ошибка прогноза;  $\widetilde{\sigma}_u$  – оцененное среднее квадратическое отклонение модели;  $q_0$  – параметр точности прогноза.

4. Определение нормированной ошибки прогноза ( $t_{\text{крит}}$ ) с уровнем значимости  $\alpha = 0,05$  и количеством степеней свободы  $\nu_2 = n - k$ , где  $\nu_2$  – количество степеней свободы;  $n$  – объем

обучающей выборки;  $k$  – количество регрессоров. Вычисление нормированной ошибки прогноза производится с помощью функции Excel = СТЬЮДЕНТ.ОБР.2Х( $\alpha$ ;  $\nu_2$ ). Так как адекватность модели проверяется по одному уровню значимости и по одному набору обучающей выборки, нормированная ошибка прогноза будет одинаковой для каждого набора данных контролирующей выборки.

5. Определение доверительных интервалов для каждого набора данных контролирующей выборки:

$$\widetilde{AS}_0^- = \widetilde{AS}_0 - S_{\widetilde{AS}_0} \times t_{\text{крит}}$$

$$\widetilde{AS}_0^+ = \widetilde{AS}_0 + S_{\widetilde{AS}_0} \times t_{\text{крит}}$$

где  $\widetilde{AS}_0^-$  – нижняя граница доверительного интервала;

$\widetilde{AS}_0^+$  – верхняя граница доверительного интервала;

$\widetilde{AS}_0$  – оцененное значение эндогенной переменной контролирующей выборки;

$S_{\widetilde{AS}_0}$  – ошибка прогноза.

6. Определение принадлежности фактических значений эндогенной переменной контролирующей выборки рассчитанным доверительным интервалам, определение доли контролирующей выборки, попавшей в доверительные интервалы и формулирование вывода об адекватности модели.

Проверим оцененную модель (2) на адекватность по данному алгоритму. Первоначальные данные контролирующей выборки приведены в *табл. 4*. Все полученные расчетные показатели для контролирующей выборки приведены в *табл. 5*.

Определение принадлежности фактических значений эндогенной переменной контролирующей выборки рассчитанным доверительным интервалам:

1.  $32360,98 \in (26027,41; 40433,5)$ ;
2.  $33617,37 \in (25867,21; 40432,97)$ ;
3.  $43139,23 \in (35145,35; 49684,68)$ ;
4.  $36793,13 \in (32463,015; 46871,74)$ .

Все фактические значения контролирующей выборки попали в доверительные интервалы, поэтому с вероятностью 95% модель (2) можно признать адекватной.

## Выводы

Результатом проделанной работы стало получение оцененной эконометрической модели зависимости среднемесячной начисленной номинальной заработной платы от величины прожиточного минимума в регионе, валового регионального продукта на душу населения, величины региональных инвестиций в основной капитал на душу населения.

Из проведенного исследования и построенной эконометрической модели можно сделать вывод о том, что на уровень среднемесячной заработной платы наибольшее влияние оказывают такие региональные экономические факторы как: прожиточный минимум, валовой региональный продукт на душу населения и инвестиции в основной капитал на душу населения.

Модель была проверена на все предпосылки теоремы Гаусса-Маркова, что позволяет говорить об эффективности полученных с помощью метода наименьших квадратов оценок параметров модели.

Также оцененная модель была проверена на адекватность с помощью метода интервального прогнозирования: в рамках этой процедуры подтвердилась адекватность полученной модели с вероятностью 95%.

В ходе проверки на ошибки спецификации переменных в уравнениях регрессии ошибки первого, второго и третьего рода не были обнаружены, что позволяет говорить о возможности практического использования полученной модели для прогнозирования среднемесячной начисленной номинальной заработной платы в регионах Российской Федерации.

Полученная модель может быть использована на практике для прогнозирования величины среднемесячной заработной платы по регионам России, а также проведенное исследование, выводы по нему могут быть использованы органами власти регионов России для разработки региональных стратегий по увеличению уровня жизни населения регионов.

## Список источников

1. Барашов Н.Г., Латков А.В., Толмачев М.Н. Дифференциация муниципальных образований по уровню среднемесячной заработной платы как результат социально-экономического развития. *Промышленность: экономика, управление, технологии*. 2019;3(77). URL: <https://cyberleninka.ru/article/n/differentsiatsiya-munitsipalnyh-obrazovaniy-po-urovnyu-srednemesyachnoy-zarabotnoy-platy-kak-rezultat-sotsialno-ekonomicheskogo> (дата обращения: 11.12.2022).
2. Репринцева Е.В. О соотношении номинальной заработной платы и прожиточного минимума в России. *Экономические исследования*. 2019;(2). URL: <https://cyberleninka.ru/article/n/o-sootnoshenii-nominalnoy-zarabotnoy-platy-i-prozhitochnogo-minimuma-v-rossii> (дата обращения: 13.12.2022).
3. Орлова И.В., Филонова Е.С. Анализ и прогнозирование номинальной заработной платы населения Российской Федерации с учетом ее региональной дифференциации. *Фундаментальные исследования*. 2021;(5):67–74. URL: <https://fundamental-research.ru/ru/article/view?id=43040> (дата обращения: 11.12.2022).
4. Бывшев В.А. Эконометрика. М.: Финансы и статистика; 2008.

5. Kissell R., Poserina J. Regression Models. *Optimal Sports Math, Statistics, and Fantasy*. 2017:39–67.
6. Andrew F. Siegel, Michael R. Wagner Correlation and Regression: Measuring and Predicting Relationships. *Practical Business Statistics* (Eighth Edition). 2022:313–370.

## References

1. Barashov N.G., Latkov A.V., Tolmachev M.N. Differentiation of municipalities in terms of average monthly wages as a result of socio-economic development. *Promyshlennost': ekonomika, upravlenie, tekhnologii = Industry: Economics, management, technology*. 2019;3(77). URL: <https://cyberleninka.ru/article/n/differentsiatsiya-munitsipalnyh-obrazovaniy-po-urovnyu-srednemesyachnoy-zarabotnoy-platy-kak-rezultat-sotsialno-ekonomicheskogo> (accessed on 11.12.2022). (In Russ.).
2. Reprintseva E.V. On the ratio of nominal wages and the subsistence minimum in Russia. *Ekonomicheskije issledovaniya = Economic Research*. 2019;(2). URL: <https://cyberleninka.ru/article/n/o-sootnoshenii-nominalnoy-zarabotnoy-platy-i-prozhitochnogo-minimuma-v-rossii> (accessed on 13.12.2022). (In Russ.).
3. Orlova I.V., Filonova E.S. Analysis and forecasting of the nominal wages of the population of the Russian Federation, taking into account its regional differentiation. *Fundamental'nye issledovaniya = Basic research*. 2021;(5):67–74. URL: <https://fundamental-research.ru/ru/article/view?id=43040> (accessed on 11.12.2022). (In Russ.).
4. Byvshev V.A. Econometrics. Moscow: Finance and statistics; 2008. (In Russ.).
5. Kissell R., Poserina J. Regression Models. *Optimal Sports Math, Statistics, and Fantasy*. 2017:39–67.
6. Andrew F. Siegel, Michael R. Wagner Correlation and Regression: Measuring and Predicting Relationships. *Practical Business Statistics* (Eighth Edition). 2022:313–370.