УДК 314.87

Н. А. Яценко

ФГБОУ ВО «Финансовый университет при Правительстве РФ», Москва,
e-mail: nayaschenko@fa.ru

ЭКОНОМЕТРИЧЕСКИЙ АНАЛИЗ ДЕПОПУЛЯЦИИ НАСЕЛЕНИЯ РОССИИ

Ключевые слова: депопуляция, стационарный временной ряд, независимость процесса остатков, модель авторегрессии, модель скользящего среднего, проверочная статистика.

Проблема воспроизводства населения в нашей стране возникла несколько десятилетий назад. Ещё в 90-ые годы прошлого века распад экономики вызвал резкий рост безработицы и снижение уровня жизни населения. Нестабильная политическая ситуация, дефицит рабочих мест и товаров привели к развалу института семьи, что в свою очередь отразилось на уровне рождаемости. Не смотря на изменившиеся с тех пор экономические условия, принимаемые государством меры по социальной и материальной поддержке семей с детьми, задача сохранения и приумножения человеческого капитала в современной России является актуальной. Представленная работа посвящена анализу годовых демографических показателей за период с 1989 по 2018 год на основе моделей временных рядов. С помощью теста Дики-Фуллера на наличие единичного корня были исследованы свойства стационарности предложенной модели, автокорреляция и гомоскедастичность случайных остатков проверялись по тестам Дарбина-Уотсона, Бройша-Годфри и Голдфелда-Квандта. На основе проведенных статистических тестов предложена смешанная модель авторегрессии и скользящего среднего порядка p и q соответственно $ARMA(p, q)$, как наиболее точная среди рассмотренных моделей стационарных временных рядов, позволяющая состоятельно описывать процесс депопуляции на ближайшие годы.

N. A. Yaschenko

Financial University under the Government of the Russian Federation, Moscow,
e-mail: nayaschenko@fa.ru

ECONOMETRIC ANALYSIS OF DEPOPULATION IN RUSSIA

Keywords: depopulation, standing time series, independence of the remainder process, autoregression model, moving average model, verification statistics.

The problem of population reproduction in our country arose several decades ago. Back in the 90s of the last centuries the collapse of the economy caused a sharp increase in unemployment and a decline in the standard of living of the population. The unstable political situation, lack of jobs and goods led to the collapse of the family institution, which in turn affected the birth rate. Despite the economic conditions that have changed since then and the measures taken by the state to provide social and material support for families with children, the goal of preserving and increasing human capital in modern Russia is topical. Presented work is dedicated to the analysis of annual demographic indicators for the period from 1989 to 2018 based on time series models. The stationarity properties of the proposed model were studied using the Dickey-Fuller test for the presence of a single root, autocorrelation and homoscedasticity of random residues were tested using the Darbin-Watson, Broish-Godfrey, and Goldfeld-Quandt tests. Based on the conducted statistical tests a mixed autoregression and moving average order model p and q respectively $ARMA(p, q)$ is proposed as the most accurate among the considered models of stationary time series which allows us to describe in detail the depopulation process for the coming years.

Экономическое развитие любой страны невозможно без постоянного качественного пополнения трудовых ресурсов, которое можно обеспечить только в условиях устойчивого демографического прироста населения. Однако из-за мировых тенденций по снижению уровня рождаемости все чаще приходится констатировать отсутствие даже

простого воспроизводства, которое характеризуется тем, что коэффициент депопуляции, представляющий собой отношение показателя рождаемости к показателю смертности за определенный период, имеет значение меньше 1.

Резкое нарушение воспроизводства населения России относится к началу 1990-х годов. На протяжении всего XX

века в России было несколько демографических кризисов, вызванных первой мировой (1914-1918) и гражданской (1917-1922) войнами, массовым голодом в начале 30-х годов, Великой Отечественной войной (1941-1945), которые унесли жизни огромного количества трудоспособного, по большей части мужского населения [1]. Численность населения после этих событий в стране удалось восстановить, однако рождаемость после указанных событий стала снижаться. Обвальное снижение рождаемости и значительный рост смертности начались с 1989 года и стали следствием разрушения идеологических принципов и ценностей общества, утраты традиционных устоев крепкой семьи, потери мужчиной роли главы семьи и кормильца, отдаления женщины от семьи, ухудшения качества медицинского обслуживания населения. Тяжелейшие экономические и социальные потрясения привели к снижению доходов населения, сокращению рабочих мест по всей стране, жилищным проблемам и т.п., что в итоге привело к резкому увеличению разводов и вызвало рост смертности как среди взрослых, так и молодежи, по причине самоубийств, заболеваниям наркоманией и алкоголизмом. В этом свою отрицательную роль сыграла западная пропаганда раннего сексуального воспитания, свободы выбора пола, однополых браков.

Все это в совокупности привело к тому, что Россия по рождаемости опустилась до уровня развитых стран, при этом индексы интенсивности смертности поднялись до уровня развивающихся стран. Только за период с 1990 по 2010 год население страны уменьшилось на 5 млн. человек [2]. Не смотря на принятый закон о выплате «материнского капитала» на второго и последующего ребенка [3], концепцию о демографической политике на период до 2025 года [4], общая численность населения по данным Росстата на 2015 год составила 146,3 млн и достигла уровня 2001 года только за счет учета проживающих в Республике Крым и городе Севастополь, вошедших в состав РФ [5]. Государственная поддержка института семьи и стимулирование деторожде-

ния [6] несколько улучшило ситуацию по восстановлению популяции в стране начиная с 2015 года, во многих регионах впервые за несколько десятилетий рождаемость превысила смертность. Однако, как скоро демографическая ситуация улучшится настолько, чтобы не представлять собой угрозу постепенного вымирания страны, не могут точно сказать даже эксперты.

Целью данного исследования является проведение эконометрического анализа процесса депопуляции населения России для определения возможных тенденций по сокращению/приращению народонаселения России в ближайшие годы.

Материалом для исследования являются статистические данные с 1989 по 2018 годы, имеющиеся в свободном доступе [7]. **Используемые методы исследования** – статистический анализ временных рядов для построения адекватной модели по прогнозированию будущих значений уровней ряда.

В статье представлены **результаты** проведенного анализа по выбору структуры временного ряда, исследования свойства его стационарности, определения характера влияния случайных факторов на изучаемую эндогенную переменную и вычисления их количественной меры воздействия. Наилучшим образом описывает изменение количества населения в России смешанная модель авторегрессии и скользящего среднего порядка ARMA(1,2).

Рассмотрим динамику годовых показателей в период 1989-2018 гг. количества населения в миллионах человек – обозначим их как y_t (рисунок 1).

Рассматриваемый датированный набор значений y_t представляет собой модель временного ряда, в которой каждое значение уровня y_t объясняется при помощи фактора времени t . Определим вид данного временного ряда.

Поскольку на графике динамики годовых показателей (Рис. 1) явно прослеживается убывающий тренд, а сезонная (циклическая) компонента в значениях уровней ряда не присутствует, спецификацию модели временного ряда представим в простейшем виде как

$$y_t = \alpha + \beta \cdot t + u_t. \quad (1)$$

Рис. 1. Динамика годовых показателей количества населения России (в млн. человек)

Проведем оценку параметров модели (1) методом наименьших квадратов (МНК реализован во встроенной функции «ЛИНЕЙН» – здесь и далее все вычисления будут проводиться в Microsoft Excel) и по оцененной модели

$$\begin{cases} y_t = 149,12 - 0,27t + u_t \\ S_{\alpha} = 0,34; S_{\beta} = 0,019; \tilde{\sigma}_u = 0,92 \end{cases} \quad (2)$$

проверим качество предложенной спецификации модели (1) по F-тесту. В ходе теста выдвигается две гипотезы: основная – все коэффициенты уравнения регрессии (1) равны нулю, то есть ни одна из объясняющих переменных не является значимой, альтернативная – хотя бы один коэффициент не равен нулю:

$$H_0: \alpha = \beta = 0, H_1: \alpha \neq \beta \neq 0.$$

Требуют реализации последующие шаги по проведению F-теста:

1. Найти значение F-статистики для набора данных и определить $(1 - \alpha)$ -квантиль $F_{\text{крит}} = F_{1 - \alpha}$ распределения Фишера.

2. При справедливости неравенства $F \leq F_{\text{крит}}$ принять основную гипотезу или отвергнуть в противоположном случае.

При принятии нулевой гипотезы спецификация модели признается некачественной.

В результате проведения F-теста получили: $F = 198,09$; $F_{\text{крит}} = 4,196$, то есть качество модели (1) признает-

ся удовлетворительным. Каждое значение уровня y_t объясняется функцией регрессии на уровне 88 %, на что показывает значение коэффициента детерминации модели. Однако, как следует из (2), значение случайного фактора в каждый момент времени t может примерно на 0,92 млн. увеличивать/уменьшать ожидаемое значение численности населения y_p , что является весьма существенным отклонением.

В связи с этим проверим статистическую значимость объясняющих переменных модели по критерию Стьюдента (t-тесту). Выдвигается две гипотезы: основная – коэффициент при интересующей переменной равен нулю, альтернативная – коэффициент при интересующей переменной не равен нулю:

$$H_0: a_i = 0, H_1: a_i \neq 0.$$

Требуется выполнить следующие шаги t-теста:

$$1. \text{ Найти значение дроби } t = \frac{\tilde{a}}{S_{\tilde{a}}}$$

для каждого коэффициента регрессии и определить двусторонний $(1 - \alpha)$ -квантиль $t_{\text{крит}}$ распределения Стьюдента.

2. При справедливости неравенства $t \leq t_{\text{крит}}$ принять основную гипотезу или отвергнуть в противоположном случае.

При принятии нулевой гипотезы переменная признается незначимой, она может быть удалена из модели.

Констатируем, что поскольку для модели (2) все значения t – статистик ($t_a = -14,07$, $t_b = -435,02$ не превышают уровня $t_{\text{крит}} = 2,045$, то все объясняющие факторы в уравнении регрессии (1) являются значимыми.

Определение приближенных значений коэффициентов модели (1) и их среднеквадратических отклонений проводилось по методу наименьших квадратов (МНК). Но является ли данная процедура для (1) самой оптимальной?

Для выяснения оптимальности метода оценивания параметров модели (1) проверим входящие в нее случайные возмущения на гомоскедастичность (отсутствие зависимости между объясняемыми переменными и случайным остатком) по тесту Голдфелда-Квандта и наличие автокорреляции (зависимости случайных остатков предыдущего периода от текущего) с помощью теста Дарбина-Уотсона [8].

В ходе данного теста Голдфелда-Квандта выдвигается две гипотезы: основная гипотеза – случайное возмущение гомоскедастично, альтернативная – случайное возмущение гетероскедастично:

$$H_0 : \text{Var}(u_1) = \text{Var}(u_2) = \dots = \text{Var}(u_n) = \sigma^2,$$

$$H_1 : \text{Var}(u_1) \neq \text{Var}(u_2) \neq \dots \neq \text{Var}(u_n) \neq \sigma^2.$$

Распишем последующие шаги по проведению теста Голдфелда-Квандта:

1. Произведем сортировку всех данных временного ряда по возрастанию значений t .

2. Разделим наши данные на три части: 1 – первые $[0,3n]$ строк (n – количество уравнений наблюдений); 3 – последние $[0,3n]$ строк; 2 – все оставшиеся строки.

3. Оценим коэффициенты методом наименьших квадратов при помощи функции «ЛИНЕЙН» (здесь и далее все вычисления будут проводиться в Microsoft Excel) 1 и 3 части наших данных.

4. Рассчитываем модельную статистику $GQ = \frac{ESS_1}{ESS_2}$ и определяем пороговую статистику как $(1 - \alpha)$ -квантиль $F_{\text{крит}} = F_{1-\alpha}$ распределения Фишера.

5. При выполнении следующих условий остаток признается гомоскедастичным:

$$\begin{cases} GQ \leq F_{\text{крит}} \\ GQ^{-1} \leq F_{\text{крит}} \end{cases}.$$

В результате проведения теста получили:

$$GQ = 1,74; GQ^{-1} = 0,57; F_{\text{крит}} = 3,44.$$

Оба условия теста Голдфелда-Квандта выполняются, что указывает на гомоскедастичность случайного остатка в модели (1), то есть показатель дисперсии σ_u^2 является константой для каждого уравнения наблюдения.

В ходе теста Дарбина-Уотсона основная и альтернативная гипотеза выглядят так:

$$H_0 : \text{Cov}(u_i, u_j) = 0, H_1 : \text{Cov}(u_i, u_j) > 0.$$

Алгоритм теста Дарбина-Уотсона включает шаги:

1. По оцененной модели (2) рассчитывается статистика $DW = \frac{\sum_{i=2}^n (\tilde{u}_i - \tilde{u}_{i-1})^2}{\sum_{i=1}^n \tilde{u}_i^2}$.

2. По таблице значений статистик d_L и d_U критерия Дарбина-Уотсона по количеству наблюдений и входящих в модель регрессоров находятся критические значения нижней и верхней границы для статистики $DW - d_L$ и d_U .

3. Смотрим в какой интервал попала статистика и делаем соответствующие выводы:

Проверим автокоррелированность случайного остатка в модели (2):

$$DW = 0,7077; d_L = 1,36;$$

$$d_U = 1,5 \text{ при } k = 1, n = 30.$$

Статистика Дарбина-Уотсона получила близкое к 0 значение и попадает в область отклонения гипотезы H_0 – случайный остаток имеет положительную автокорреляцию между соседними уровнями случайных остатков. Из этого следует, что МНК-оценки в модели (2) не являются оптимальными и причиной этому может служить возможная ошибка в спецификации модели (1). Наличие автокорреляции случайных остатков может сигнализировать о пропуске

в уравнении связи важной объясняющей переменной. Так как для объяснения значений y_t в модели (1) никакие факторы, кроме времени, не рассматриваются, возьмем в качестве такой переменной величину y_{t-1} . Тогда вместо спецификации (1) будем рассматривать модель

$$y_t = a \cdot y_{t-1} + u_t, \quad (3)$$

которая называется моделью авторегрессии порядка p (в данном случае $p = 1$) $AR(p)$.

Проверим стационарность временного ряда (3) по тесту Дики-Фуллера. иными словами проверим значение коэффициента a в формуле (3): если $a = 1$, то процесс имеет единичный корень и временной ряд y_t является не стационарным; при $|a| < 1$ y_t является стационарным временным рядом.

Если выражение (3) записать при помощи оператора разности первого порядка $\Delta y_t = y_t - y_{t-1}$, то оно примет вид $\Delta y_t = b \cdot y_{t-1} + u_t$, где $b = a - 1$. Временной ряд имеет единичный корень (является интегрированным первого порядка), если его разности первого порядка образуют стационарный ряд.

Для модели авторегрессии первого порядка суть теста Дики-Фуллера заключается в проверке гипотезы H_0 о наличии единичного корня (или гипотезы о нулевом значении коэффициента b). Принятие/отклонение гипотезы H_0 происходит в результате сравнения DF -статистики для коэффициентов регрессии (3) с проверочными значениями статистики, взятыми из критических таблиц МакКиннона, составленных для конкретных видов тестовых регрессий на заданном уровне значимости.

Для модели (3) без константы и тренда по тесту Дики-Фуллера $DF = 2,58$, что при используемом объеме выборки в 29 наблюдений позволяет отклонить гипотезу H_0 и считать ряд y_t стационарным, поскольку $DF > DF_{kr} = -1,95$ (для 5 % уровня).

Однако предполагая возможную авторегрессию в рамках (3) не первого, а более высокого порядка, был проведен расширенный (пополненный) тест Дики-Фуллера для модели (3), в которую был добавлен лаг первой разности эндогенной переменной, который позволил принять гипотезу H_0 при $L = 1$; для более высоких порядков лага $L = 2$ и $L = 3$ гипотеза отклоняется.

Заключение

В виду полученных противоречивых результатов теста Дики-Фуллера для модели (3), а также установленной по тесту Дарбина-Уотсона авторегрессии первого порядка остатков в модели (1), можно предположить, что остатки u_t в модели (3) являются не «белым шумом», а некоторым стационарным ARMA-процессом. Использование для описания временного ряда y_t смешанной модели авторегрессии и скользящего среднего ARMA(p, q) вида

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_p y_{t-p} + \lambda_1 z_{t-1} + \lambda_2 z_{t-2} + \dots + \lambda_q z_{t-q} + z_t \quad (4)$$

подтвердили это допущение.

При помощи теста Бройша-Годфри была принята альтернативная гипотеза H_1 о $y_t \sim AR(1)$ и $y_t \sim MA(2)$. Построенная модель ARMA(1,2) содержит случайные влияния в виде «белого шума» (Рис. 3).

Рис. 2. Критические интервалы статистики Дарбина-Уотсона

Остатки модели ARMA(1,2)

Рис. 3. График «белого шума» для смешанной модели

Наблюдаемое резкое отклонение случайного остатка при $t = 20$ от полученного облака точек следует рассматривать как выброс в значениях «белого шума». Анализ данных временного ряда до и после 2010 года, в котором показатели y_t улучшились по сравнению с предыдущими, показал, что никаких структурных сдвигов это не вызвало, поэтому модель вида (4) можно использовать с единым набором параметров, вычисленным по всему имеющемуся набору статистики.

Состоятельные оценки параметров модели (4) получены как оценки метода наименьших квадратов, при этом $\beta_0 = 0$; $\tilde{\beta}_1 = 0,998$; $\tilde{\lambda}_1 = 0,00019$; $\lambda_2 = 0,365$. Но

исследованный на значимость коэффициент λ_1 указал на незначимость $\tilde{z}_{t-1}$ в модели. Это говорит о том, что с учетом обратимости временного ряда (4), все происходящие события в текущем периоде отражаются на значении показателя численности населения только через два года, причем это является более существенным влиянием на эндогенный показатель, чем текущее воздействие случайных факторов, которое составляет 0,352 млн. Этот показатель почти в три раза ниже, чем в модели (2), что говорит о более высокой точности модели ARMA(1,2) и возможности её применения для автопрогнозирования в краткосрочном и среднесрочном периоде.

Библиографический список

1. Олейников В.В. Демографическая проблема и пути ее решения на основе системного подхода // Сборник статей участников VII Международного научного студенческого конгресса «Проект для России», 6–16 апреля 2016 года. Под ред. О.В. Карамовой и А.П. Бувеч / ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации». М.: Финуниверситет, 2016. 1776 с.
2. Тоичкина В.П. Государственная демографическая политика: итоги первого этапа (на примере Мурманской области) // ЭКО. 2011. № 12. С. 108–114.
3. Федеральный закон от 29.12.2006 N 256-ФЗ (ред. от 30.12.2015) «О дополнительных мерах государственной поддержки семей, имеющих детей» (с изм. и доп., вступ. в силу с 12.03.2020) [Электронный ресурс]. URL: <http://base.garant.ru/12151286/> (дата обращения: 03.07.2020).
4. Указ Президента РФ от 09.10.2007 N 1351 «Об утверждении Концепции демографической политики Российской Федерации на период до 2025 года» // «Парламентская газета», 16 октября 2007 г., N 134.
5. Федеральная служба государственной статистики: Численность и состав населения [Электронный ресурс]. URL: <https://www.gks.ru/folder/12781#> (дата обращения: 25.06.2020).
6. Вирановский Н.А., Вирановская Е.В. Решение проблемы депопуляции в России путем мер демографической политики // Научно-методический электронный журнал «Концепт». 2016. Т. 12. С. 56–60.
7. Мировая экономика EREPORT.RU: Экономика России: экономические показатели [Электронный ресурс]. URL: <http://www.ereport.ru/stat.php?razdel=country&count=russia&table=ggecia&time=1> (дата обращения: 25.06.2020).
8. Бабешко Л.О., Бич М.Г., Орлова И.В. Эконометрика и эконометрическое моделирование: учебник. М.: Вузовский учебник: ИНФРА-М, 2017. 400 с.