

А.В.СЕМЕНОВ*

АНАЛИЗ ВРЕМЕННЫХ РЯДОВ В ПОЛИТИЧЕСКОЙ НАУКЕ: ВОЗМОЖНОСТИ И ОГРАНИЧЕНИЯ

Аннотация. В данной статье рассматриваются возможности и ограничения статистического моделирования динамики политических процессов. На примере моделирования рейтинга одобрения президента В. Путина и готовности россиян присоединиться к коллективным действиям с экономическими требованиями демонстрируется аналитический потенциал базовых моделей авторегрессии со скользящим средним и интеграцией (ARIMA), авторегрессии с распределенным лагом (ADL) и модели коррекции регрессионных остатков (ECM). В ходе статистического моделирования временных рядов в политике исследователи сталкиваются с целым рядом аналитических проблем. Данная статья призвана обозначить основные «развилки» в исследовательском процессе и основания для выбора того или иного варианта исследовательского дизайна. Не претендуя на полноту охвата всей дискуссии, данная статья призвана стимулировать использование данного метода применительно к российскому материалу.

Ключевые слова: временные ряды; динамические процессы; политические процессы; статистическое моделирование.

Для цитирования: Семенов А.В. Анализ временных рядов в политической науке: возможности и ограничения // Политическая наука. – 2021. – № 1. – С. 76–97. – Режим доступа: <http://www.doi.org/10.31249/poln/2021.01.03>

Политические процессы по определению динамичны: рейтинги политиков взлетают и падают, интенсивность межгосударственных конфликтов меняется вместе с характером международной

* Семенов Андрей Владимирович, кандидат политических наук, научный сотрудник, Пермский федеральный исследовательский центр Уральского отделения Российской академии наук (Пермь, Россия), e-mail: andreysemenov@comparativestudies.ru

системы, а государственные расходы привязаны к электоральным циклам. Анализ временных рядов (time-series analysis), пришедший в политическую науку из экономической науки, стал важным инструментом развития теорий политических процессов и их эмпирического тестирования, а также прикладного анализа и прогнозирования. Благодаря анализу временных рядов были сформулированы и проверены ключевые модели политических бизнес-циклов, динамики «макрополитической принадлежности» (macropartisanship), одобрения президентской власти, экономического голосования, макрополитической динамики среди прочих [Abramson, Ostrom, 1991; Anderson, 2000; Lebo, Cassino, 2007]. Вместе с расширением предметной области применения анализа временных рядов возникли методологические дискуссии о возможностях и ограничениях метода. Значительную роль сыграли тематические выпуски журнала «Политический анализ»¹ – ведущего издания по количественным методам в политической науке.

В российской политической науке анализ временных рядов и его разновидности пока являются редкостью, что обусловлено как особенностями подготовки российских политологов, так и инфраструктурными проблемами в научной коммуникации. Поиск по словосочетанию «анализ временных рядов» в ведущих журналах по дисциплине за 2000–2020 гг. («Политическая наука», «Полис», «Полития», «Политэкс», «Международные процессы», «Мировая экономика и международные отношения») выдает лишь девять результатов, из которых одна статья – перевод, другая касается анализа политики в области контроля рождаемости, остальные не применяют метод как таковой. Родственный метод панельных данных (dynamic panels или cross-sectional time-series, CSTS) используется несколько шире, но даже в этом направлении российская политическая наука отстает от социологии и экономики по объему накопленных знаний. В то же время российский материал активно используется исследователями, например для тестирования теорий популярности президента в гибридных режимах [Treisman, 2011]. Таким образом, существует явная необходимость представить аналитический потенциал анализа временных рядов российскому академическому сообществу и вводить его в оборот.

¹ Political Analysis: Vol. 4 за 1992 г. и Vol. 24 за 2016 г.

Данный обзор призван начать дискуссию о систематическом применении анализа временных рядов: не претендуя на полноту, я ставлю своей целью, во-первых, обобщить мировую дискуссию об особенностях анализа временных рядов в политике. Теоретизации кратковременной и долговременной памяти процесса через автокорреляцию и скользящее среднее, полной и частичной интеграции, цикличности и «структурных смещений» (structural breaks) представляют собой неотъемлемую часть продвижения научного знания в области динамики политических процессов. Во-вторых, в статье будут рассмотрены два самых распространенных варианта спецификации моделирования временных рядов – модель с распределенным лагом (autoregressive distributed lag, ADL) и модель коррекции регрессионных остатков (error-correction model, ECM). Статистическое моделирование будет проиллюстрировано материалами из российской политики. Данный обзор должен служить своеобразной «картой» для исследователей, которые работают с динамическими рядами в политике и хотят использовать статистические методы для их осмысления и моделирования.

Анализ одномерных временных рядов

Ключевая особенность временного ряда – упорядоченность наблюдений во времени, что создает естественную зависимость наблюдений между собой. Текущая интенсивность какого-либо межгосударственного конфликта определяется его динамикой в предыдущие годы; ключевые макроэкономические показатели также коррелируют между собой во времени; общественное мнение, например рейтинг доверия президенту, имеет значительную инерцию. Зависимость во времени может проявляться в виде тренда (детерминистического и стохастического), памяти процесса (автокорреляции или интеграции) и циклов. Детерминистический тренд – изменение значений временного ряда как функция от времени – в меньшей степени представляет собой интерес для статистического моделирования: редко какой социальный процесс является исключительно линейной функцией от движения во времени.

Гораздо больший интерес для исследователей представляют стохастические компоненты политических процессов, в частности автокорреляция, циклы и интеграция. Автокорреляция отражает

«память процесса» – зависимость текущего значения временного ряда от прошлых значений. В ситуации автокорреляции возмущения, связанные с изменением значений в некоторой переменной в период t , оказывают воздействие на наблюдения в период $t + 1$, $t + 2 \dots t + k$ с экспоненциально убывающим эффектом. Циклы, напротив, связаны с воздействием возмущений в строго определенные числа временных периодов, после чего их эффект исчезает (таким образом, текущие значения связаны с возмущениями в $t \pm k$ периодах). Наконец, эффект возмущений может накапливаться во времени; такие серии получили название интегрированных.

Базовой моделью в моделировании динамики является понятие стационарного временного ряда. Стационарность определяется через постоянство математического ожидания (среднего значения) и дисперсии реализации значений ряда. Распределение значений стационарного ряда, таким образом, не меняется со временем, его ключевым свойством является тенденция возвращения к среднему после отклонений от него. Любой одномерный временной ряд, описывающий какой-либо процесс, может быть идентифицирован с помощью модели авторегрессии с интеграцией и скользящим средним (autoregressive integrated moving average model), ARIMA (p, d, q), где p – порядок автокорреляции, d – порядок интеграции, а q – порядок скользящего среднего процесса.

Рисунок 1 иллюстрирует различия между основными модификациями стационарного ряда. Верхняя левая панель отражает стохастический процесс («белый шум»), верхняя правая – белый шум с временным положительным трендом (тренд-стационарный процесс). Две нижние панели визуализируют особенности рядов с сильной автокорреляцией и выраженным скользящим средним: несмотря на то что обе серии стремятся к среднему значению, «темпы» их возвращения отличаются от стационарного ряда. При сильной автокорреляции ряд «движется» в определенном направлении некоторое время, прежде чем повернуть назад к среднему уровню. В модели со скользящим средним воздействие шоков существует определенный период, после чего исчезает, что обуславливает циклическую динамику.

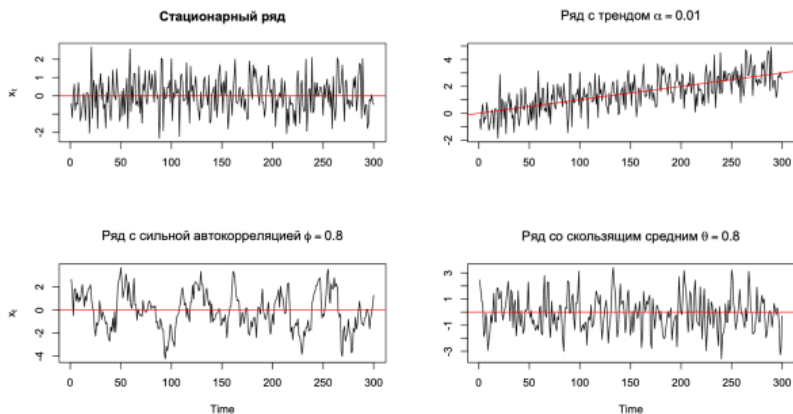


Рис. 1.

Симуляция динамических рядов с различными компонентами временной зависимости, $N = 300$, ошибки являются стохастическим процессом с нормальным распределением, нулевым средним значением и стандартным отклонением в одну единицу

Нестационарный временной ряд отличается тем, что эффекты возмущений, связанных с вариацией значений вокруг среднего, не исчезают со временем, а накапливаются. Иными словами, как среднее, так и дисперсия «интегрированного» ряда может бесконечно меняться. Нестационарный ряд первого порядка интеграции можно обратить в стационарный вычитанием (первой разностью) предыдущих значений из текущего. Такая процедура меняет интерпретацию самого временного ряда: вместо значений переменной, измеренной в уровне, она превращается в изменения этого уровня между соседними наблюдениями. Рисунок 2 отражает симуляцию временного ряда с интеграцией первого порядка и его дифференцированную форму. Как видно из рисунка, оригинальная серия не возвращается к среднему значению, тогда как дифференцированные значения временного ряда превращают его в стационарный.

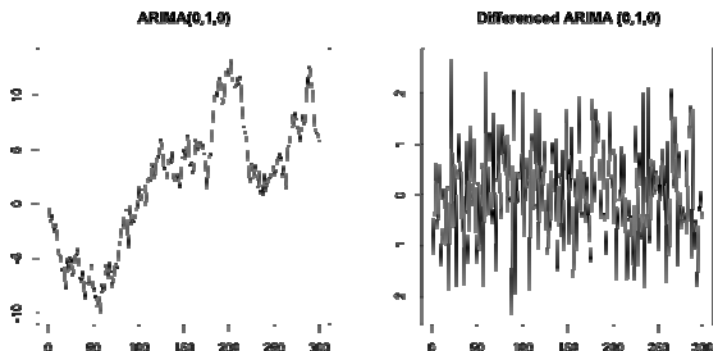


Рис. 2.

Динамические ряды с различными компонентами временной зависимости. $N = 300$, ошибки являются стохастическим процессом с нормальным распределением $N(0,1)$

Ключевой проблемой при моделировании временных рядов в социальных науках является корректная идентификация его динамических свойств. Визуальное обследование временного ряда – первый шаг в данном направлении: наличие временного тренда и циклов, как правило, можно заметить на графике процесса. Однако проблема возникает при разведении рядов с сильной автокорреляцией и интегрированных процессов: в ситуации, когда наблюдений недостаточно, определить, стремится ли процесс к своему «естественному» среднему уровню или нет, сложно. Данную проблему можно решить на уровне теории: поскольку интеграция подразумевает накопление «шоков», предположение о ее наличии, скажем, в рейтинге одобрения президента подразумевает, что текущие оценки зависят от значений 10-летней давности. Сильная автокорреляция же указывает на «длинную память», но не подразумевает накопление эффектов во времени¹.

¹ С математической точки зрения агрегированные показатели общественного мнения в принципе не могут быть интегрированными, поскольку диапазон их значений, выраженный в процентных пунктах, ограничен, выйти за его пределы невозможно, что противоречит асимптотическим свойствам интегрированного временного ряда.

Существует ряд тестов, которые позволяют проверить гипотезу об интеграции: Augmented Dickey-Fuller (ADF), Kwiatkowski-Phillips-Schmidt-Shin (KPSS) и Phillips-Perron (PP) являются параметрическими тестами [Hyndman, Athanasopoulos, 2018]. Эти тесты обладают низкой статистической мощностью (ADF, к примеру, имеет большую долю ложноположительных результатов, т.е. склонен идентифицировать интеграцию там, где ее нет, особенно если ряд обладает «длинной памятью»), общая рекомендация состоит в применении нескольких тестов и сравнении их результатов между собой.

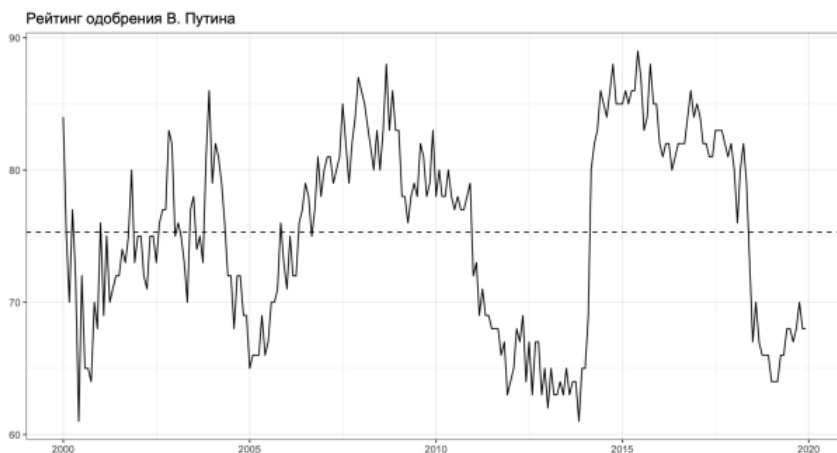


Рис. 3.
Одобрение деятельности В. Путина на посту президента (премьер-министра), % положительных ответов. Среднее значение в 75,3 п.п. обозначено прерывистой горизонтальной линией¹

Рассмотрим указанные выше проблемы в определении свойств одномерного временного ряда на примере рейтинга одобрения В. Путина на посту президента (премьер-министра с 2008 по 2012 г.). Рисунок 3 демонстрирует динамику одобрения с

¹ Данные: Левада-центр. – Mode of access: <https://www.levada.ru/indikatory/> (accessed: 20.11.2020).

2000 по 2019 г. Визуальное обследование указывает на сильную автокорреляцию, наличие циклов и «структурных разрывов» (structural breaks) – существенных изменений в поведении временного ряда в результате шоковых событий¹. В целом движения рейтинга колеблются вокруг среднего значения в 75,3 п.п.

На рис. 4 представлены коррелограммы функции автокорреляции (ACF, левая панель) и частной автокорреляции (PACF, правая панель), которые подтверждают наличие сильного AR-компонента и некоторой цикличности: «пики», пересекающие прерывистую линию, на значениях оси X, отражающей порядок лага, означают наличие корреляции между наблюдениями. ACF позволяет определить наличие AR, а PACF – наличие MA-компонентов. Сильная автокорреляция также может указывать на интеграцию (хотя как указывалось выше, обосновать ее наличие теоретически сложно). Тест-статистики ADF (для шести лагов) и PP, нулевой гипотезой которых является отсутствие стационарности, оказываются статически незначимыми (p-value 0,35 и 0,08 соответственно), что не позволяет принять альтернативную гипотезу о стационарности. Аналогично KPSS-тест с нулевой гипотезой тренд-стационарности оказывается значим на уровне 0,05 (p-value = 0,035), что позволяет принять альтернативную гипотезу о нестационарности. Результаты тестов не согласуются между собой и поэтому выбор в пользу той или иной модели остается за исследователем.

Альтернативной процедурой идентификации параметров модели может быть подбор комбинации значений ARIMA (p, d, q), минимизирующий логарифмическое правдоподобие и информационные критерии (AIC и BIC) модели. Иными словами, задача состоит в поиске модели, наилучшим образом аппроксимирующей набор данных. В случае с рейтингом доверия такой моделью выступает ARIMA (3, 0, 4), т.е. модель с выраженной автокорреляцией и скользящим средним, но без интеграции.

¹ В данном примере структурные разрывы совпадают по времени с монетизацией льгот (начало 2005 г.), российско-украинским конфликтом (март 2014 г.) и пенсионной реформой (июнь 2018 г.). Анализ структурных разрывов находится за пределами данной статьи, но, безусловно, представляет теоретический и практический интерес, см. гл. 7 работы: [Time series analysis for the social sciences, 2014].

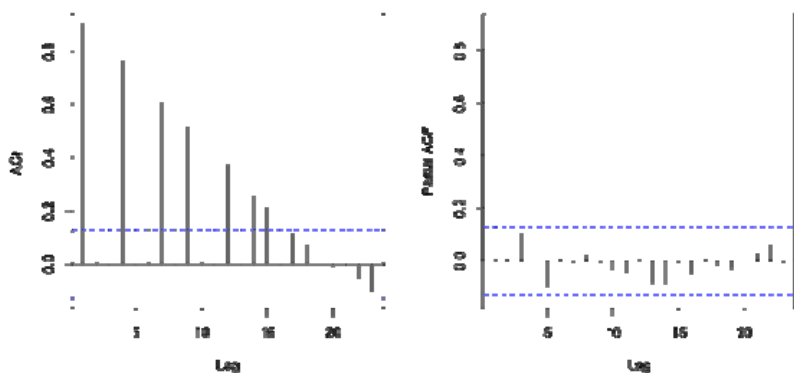


Рис. 4.

Коррелограммы функции (частичной) автокорреляции для рейтинга одобрения Владимира Путина на посту президента

Таким образом, теория и тесты не позволяют сделать однозначных выводов относительно динамических свойств доверия В. Путину, что довольно типично для политических процессов [Time series analysis for the social sciences, 2014]. Решением этой проблемы может быть оценка параметра «частичной» (fractional) интеграции – как указывалось выше, формальные тесты слабо различают интеграцию первого порядка и «околоинтегрированные» (near-integrated) серии. Такой подход использовался в оценке связи оценок экономической ситуации в стране с президентским рейтингом в работе Даниела Трейсмана [Treisman, 2011], однако теоретическое обоснование частичной интеграции еще более затруднительно, а оценка этого параметра затруднена отсутствием конвенциональных тестов¹.

Базовые модели многомерного анализа временных рядов

Исследователи редко заинтересованы исключительно в свойствах одномерного временного ряда. Теории и модели поли-

¹ Из наиболее встречаемых в эконометрике – локальная оценочная функция Уиттла (local Whittle estimator) и Гауссова полупараметрическая оценка частичной интеграции Робинсона [Velasco, 2003].

тических процессов в большей степени фокусируются на связи между различными процессами, например между экономическим ростом и одобрением главы исполнительной власти. Однако, как указывают де Беф и Кили, в социальных науках «теория постулирует только в самом общем виде, что (настоящее и) прошлое имеют значение; она остается достаточно неопределенной в отношении деталей» [De Boef, Keele, 2008, p. 186]. Иными словами, даже если теория предполагает связь между двумя процессами, она редко позволяет точно определить, какой лаг в независимой переменной имеет значение, является ли эта связь в уровнях (levels) или изменениях уровня переменной (changes) и какие экзогенные переменные имеют значения в долговременной перспективе.

Модель авторегрессии с распределенным лагом (autoregressive distributed lag model, ADL / ARDL) выступает в качестве наиболее общей формы для статистического анализа связи стационарных временных рядов. Формула 1 отражает общую спецификацию данной модели:

$$Y_t = \alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{j=1}^n \sum_{i=0}^q \beta_{ij} X_{t-i} + \epsilon_t \quad (1), \quad \text{где } \left| \sum_{i=1}^p \alpha_i \right| < 1,$$

т.е., Y_t является стационарным процессом, набор предикторов X_t является (слабо)экзогенным по отношению к Y_t ($E(\epsilon_t, X_{t'}) = 0 \forall t, s, j$), а ϵ_t является «белым шумом». Модель обозначается ADL (p, q; n), где p указывает на количество лагов в зависимой переменной, q – количество лагов в предикторах, а n – на число последних в модели. Преимуществами данной модели является консистентная оценка параметров методом наименьших квадратов за счет контроля на (возможную) автокорреляцию в Y_t , а также гибкость в дальнейшей спецификации: модель не содержит изначальных допущений о том, какой лаг в регрессорах имеет значение и позволяет в дальнейшем протестировать введение необходимых ограничений с помощью t-теста или F-теста.

Спецификация модели в эмпирическом анализе в первую очередь должна быть связана с теорией: исследователи так или иначе должны обозначить, каковы теоретические ожидания относительно наличия одновременных эффектов и включение какого лага каждой из зависимых переменных является обоснованным. К примеру, в дискуссии об источниках мнения избирателей относительно состояния и перспектив национальной экономики [Elite

economic forecasts ..., 1999] используют модель «частичной коррекции» (partial adjustment) для оценки связи президентского рейтинга, экономических новостей и экспертных оценок, которая включает лишь одновременный эффект независимой переменной без ее лагов. Другие ограничения предполагают только долговременный эффект у предиктора [De Voef, Keele 2008, p. 187].

Модель ADL остается недостаточно распространенной в политической науке [Time series analysis for the social sciences, 2014, p. 88–90], в том числе потому, что исследователи предпочитают более «компактные» модели с предзаданными (и не всегда обоснованными) ограничениями [De Voef, Keele 2008]. Однако данная модель позволяет не только контролировать временную зависимость в измеряемой переменной и эффекты предикторов, но и разделить последние на кратко- и долгосрочные (short-run / impact и long-run multipliers). В ADL долговременный мультипликатор определяется отношением суммы коэффициентов независимой переменной к параметру $1 - \alpha_1$, который определяет «скорость», с которой зависимая переменная возвращается к своему среднему значению после шокового воздействия. Например, для модели с непосредственным и отложенным в один лаг эффектами в независимой переменной долговременный мультипликатор будет определяться по формуле:

$$\frac{\beta_0 + \beta_1}{1 - \alpha_1} .$$

Оценка стандартных ошибок долговременного мультипликатора может быть произведена с помощью дельта-метода или методом трансформации Бьюли (Bewley transformation).

Основной альтернативой ADL выступает модель коррекции регрессионных остатков (error-correction model, ECM), которая получила широкое распространение в политической науке с начала 1990-х годов, что связано в первую очередь с тем, что ее функциональная форма в виде первых разниц переменных позволяет преобразовать интегрированные серии в стационарные. Кроме того, эта форма непосредственно отражает долговременный эффект регрессоров, а механизм коррекции ошибок в случае его наличия – существование долговременного равновесия между зависимой переменной и предиктором.

Математически ЕСМ является преобразованием модели авторегрессии с распределенным лагом. Интуиция, лежащая в основе модели, заключается в том, что два процесса могут быть интегрированы друг с другом: т.е. между ними существует долгосрочное равновесие, которое не позволяет им отойти слишком далеко друг от друга или приблизиться слишком близко. Если такое происходит, срабатывает механизм коррекции и обе серии возвращаются к долгосрочному «расстоянию» между ними в последующие периоды. Существование такого механизма хорошо заметно в случае, если две серии являются ARIMA (p, 1, q) процессами: регрессионные остатки в такой модели будут стационарными, в отличие от своих родительских серий.

В базовом виде модель ЕСМ с одним лагом в предикторе X выглядит так:

$$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \epsilon_t \quad (2),$$

$$\text{где } \alpha_1^* = \alpha_1 - 1, \beta_0^* = \beta_0, \text{ а } \beta_1^* = \beta_0 + \beta_1,$$

литеры без звездочки указывают на эквиваленты параметров в модели ADL (формула 1).

Коэффициент у ΔX_t указывает на одновременный эффект изменений в предикторе на изменения в зависимой переменной, коэффициент у X_{t-1} отражает долгосрочный эффект и указывает на скорость эквilibрации зависимой переменной после того, как она была выведена из равновесия. Долгосрочный мультипликатор (long-run multiplier, LRM) эффекта X задан отношением

$$LRM = \frac{\beta_1^*}{\alpha_1^*}.$$

Моделирование динамики протестных настроений

Рассмотрим модели авторегрессии с распределенным лагом и коррекции регрессионных остатков на примере индикатора готовности присоединиться к массовым выступлениям с экономическими требованиями, который отслеживается рядом социологических служб. Опросы Левада-центра производятся с регулярными интервалами в два месяца, их визуализация (рис. 5) позволяет предположить наличие временного тренда до весны 2018 г. (наклонная линия) при среднем значении в 20 п.п. (прерывистая горизонтальная

линия). После объявления правительства о начале пенсионной реформы показатели готовности поднимаются до исторических максимумов, сравнимых с кризисными 1997–1998 гг.

Функция автокорреляции указывает на сильную инерцию в серии. Формальные тесты на интеграцию не позволяют сделать однозначного вывода о ее наличии: ADF-тест с пятью лагами предполагает интеграцию, PP-тест – ее отсутствие, KPSS-тесты указывают на тренд-стационарность (результаты приведены в Приложении). Как уже обсуждалось выше, переменные, имеющие ограниченный размах значений, не могут быть интегрированными, более того, их трудно отличить от процессов с «длинной памятью». Использование информационных критериев указывает, что ARIMA (1, 1, 1) является оптимальной, хотя стохастический тренд на снижение уровня готовности протестовать очевидным образом прервался в 2018 г.

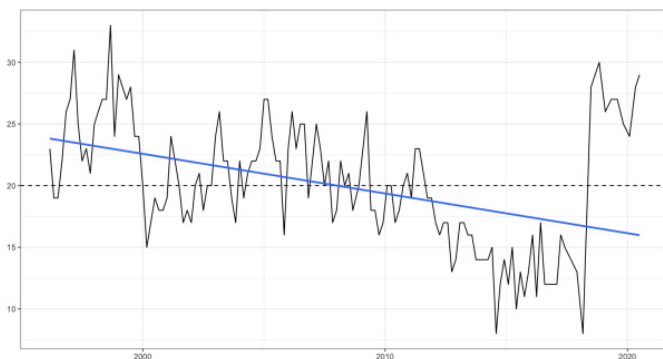


Рис. 5.

Готовность присоединиться к выступлениям с экономическими требованиями, 1996–2020¹

Поскольку готовность выступать с экономическими требованиями может быть связана с ухудшением объективных экономических условий (замедлением экономического роста, падением реаль-

¹ Данные: Левада-центр. – Mode of access: <https://www.levada.ru/indikatory/> (accessed: 20.11.2020).

ных доходов, инфляционными шоками и ростом безработицы), можно предположить связь между этими показателями и готовностью протестовать. Помимо непосредственного эффекта, например от всплеска инфляции, можно предположить наличие отложенного эффекта, поскольку изменения в объективных экономических условиях не сразу инкорпорируются в субъективные оценки состояния экономики и недовольство. Инфляция и безработица являются одними из наиболее существенных источников экономического недовольства, их аддитивный индекс в экономической литературе получил название «индекса невзгод» (misery index, см.: [Welsch, 2007]). Для простоты экспозиции я использую эти два индикатора для предсказания значений готовности присоединиться к коллективным действиям с экономическими требованиями.

Таблица 1

Зависимая переменная: готовность к участию в коллективных выступлениях с экономическими требованиями

	Модель 1	Модель 2	Модель 3
Готовность _{t-1}	-0,03 (0,08)	-0,04 (0,08)	-0,38*** (0,08)
Инфляция _t	0,14 (0,09)	0,12 (0,11)	
Безработица _t	0,63 (0,72)	0,55 (0,76)	
Инфляция _{t-1}	0,63** (0,24)	0,45* (0,27)	0,69** (0,27)
Безработица _{t-1}	1,13** (0,54)	1,07** (0,54)	1,33** (0,58)
Δ Инфляция _t			0,11 (0,09)
Δ Безработица _t			0,64 (0,73)
Свободный член	-0,03 (0,28)	0,22 (0,73)	-0,09 (0,28)
Думту-контроль	Нет	Есть	Нет
Число наблюдений	137	137	137
Скоп. R ²	0,08	0,05	0,18
Остаточная ошибка	3,09 (df = 131)	3,14 (df = 129)	3,14 (df = 122)
F-статистика	3,34*** (df = 5; 131)	1,61*** (df = 12; 128)	6,78*** (df = 14; 122)
Обозначения:	* p < 0,1; ** p < 0,05; *** p < 0,01		

Все три серии в анализе содержат ARMA-компоненты, которые нуждаются в очистке: инфляция через фильтр ARIMA (1, 1, 2) с 12-месячным сезонным компонентом, безработица через фильтр ARIMA (0, 1, 0). Таблица 2 представляет результаты моделирования: первая модель является базовой ADL (1, 1, 2), вторая включает набор фиктивных переменных для различных экзогенных шоков (кризисы 1998 и 2008 гг., реформа монетизации льгот и пенсионная реформа, кампания «За честные выборы!» 2011–2012 гг. и президентская кампания Навального 2017–2018 гг., а также индикаторы президентских сроков Ельцина и Путина). Последняя модель представляет собой ECM с первым лагом независимых переменных.

Результаты анализа однозначно свидетельствуют против одновременного эффекта предикторов на зависимую переменную: инфляционные шоки и рост безработицы хотя и действуют в предсказанном положительном направлении на готовность присоединиться к выступлениям с экономическими требованиями, не достигают конвенционального уровня статистической значимости. Отсутствие эффекта наблюдается как в уровнях (модели 1–2), так и в изменениях (модель 3). Можно также отметить, что размер краткосрочных эффектов в моделях 1 и 3 практически одинаков. Однако первый лаг (в один месяц) в независимых переменных оказывается положительно и значимо связан с готовностью протестовать. Согласно первой модели, рост цен в 1 п.п. связан с ростом готовности протестовать в последующий месяц на 0,62 п.п., а рост безработицы на 1 п.п. – с ростом в 1,13 п.п. Размер эффектов несколько больше в модели ECM, поскольку коэффициент для лага здесь отражает сумму кратко- и долговременных эффектов.

Модель коррекции регрессионных остатков также позволяет легко вычислить долговременный мультипликатор: поскольку она отражает (возможный) долгосрочный эквilibrium между сериями, изменения в предикторах выводят серии из этого равновесия и готовность протестовать должна также «сдвинуться» вверх, вслед за значениями инфляции и безработицы в долгосрочной перспективе. Темп балансировки продиктован коэффициентом лага зависимой переменной (-0,38), LRM для роста цен составляет $0,69/0,38 = 1,8$ п.п. в долгосрочной перспективе, для безработицы – $1,33/0,38 = 3,5$ п.п., из чего можно сделать вывод, что если инфляция и безработица поднимаются к новым постоянным уровням, готовность протестовать бу-

дет «следовать» за ними. Отсутствие немедленного эффекта у обоих предикторов может отражать как необходимость «обработки» сигналов индивидами о состоянии внешней экономической среды, перед тем как включить их в обновленные оценки готовности протестовать, так и асинхронность замеров общественного мнения и фиксации экономических изменений. Стандартные тесты на автокорреляцию в остатках модели добавляют дополнительной уверенности в результатах динамического моделирования, хотя небольшое значение скорректированного R-квадрат указывает на значительную долю необъясненной вариации в зависимой переменной. В этой перспективе изменения инфляции и безработицы, кажется, играют не самую значительную роль.

Рабочая схема анализа временных рядов в политической науке

Как было показано выше, исследователи сталкиваются с целым рядом проблем при моделировании политической динамики. Идентификация временных компонентов, приведение набора переменных в баланс друг с другом, поиск коинтеграции и решение проблемы эндогенности накладываются на слабую мощность тестов, позволяющих корректно определить свойства временных рядов. Более того, структурные смещения либо гетероскедастичность в дисперсии создают дополнительные трудности¹. В результате предложить общую рабочую схему анализа динамики политических процессов затруднительно: социальные процессы сложны и зачастую не похожи друг на друга, что превращает работу с ними в подобие искусства, а не просто науку.

Тем не менее можно выделить некоторые общие шаги в работе с динамическими процессами. Во-первых, статистическое моделирование не заменяет хорошо сформулированную теорию

¹ Генерализованная авторегрессионная модель с условной гетероскедастичностью (generalized autoregressive conditional heteroskedasticity (GARCH)) позволяет учесть неравномерность дисперсии, а динамическая обусловленная корреляция (dynamic conditional correlation, DCC) – учитывать волатильность эффектов в зависимости от внешних условий. Обсуждение этих моделей выходит за рамки данной статьи, заинтересованные читатели могут подробнее узнать о них в: [Engle, 2002; Lebo, Box-Steffensmeier, 2008].

относительно динамических свойств рассматриваемого временного ряда и его связи с экзогенными процессами. Насколько серьезной ожидается инерционная составляющая, есть ли циклы и / или интеграция, временной тренд? Например, общественное мнение практически всегда имеет авторегрессионную составляющую: граждане так или иначе инкорпорируют оценки за предыдущие периоды в свои текущие оценки, что обеспечивается как памятью, так и непрерывностью информационных потоков в медиасреде (последние, например, постоянно напоминают гражданам, какой рейтинг у главы исполнительной власти был в предыдущем месяце или какова общая динамика безработицы). Трудовые протесты могут быть связанными с долговременными экономическими циклами, а коллективные действия в странах с цикличным климатом – с временами года. Наконец, некоторые процессы могут иметь очевидную временную тенденцию – к примеру, явка на выборы со временем падает, а одобрение политиков со временем сходит на нет, как демонстрируется на примере рейтингов глав исполнительной власти США и Великобритании.

Вторым шагом является визуальное обследование и протесты временного ряда. Визуальное обследование позволяет предварительно удостовериться или опровергнуть некоторые теоретические ожидания, а также увидеть детали процесса, не вытекающие из теории непосредственно. Например, рис. 3 позволяет предположить наличие структурных разрывов в рейтинге президента Владимира Путина (в марте 2014 и июне – августе 2018 г.), когда движения оказались слишком резкими. И хотя природа этих разрывов легко идентифицируется (эффект «единения вокруг знамени» после крымских событий и разочарование после начала пенсионной реформы), сами по себе структурные разрывы не могут быть выведены из общей теории процесса. При формулировании теории и визуализации имеет значение уровень агрегирования единиц наблюдения (день, неделя, месяц, год, декада): чем он выше, тем менее заметны колебания. Некоторые политические процессы (например, протесты или электоральные рейтинги) могут быть измерены на уровне недель или даже дней, другие (межгосударственные соглашения, безработица) доступны для измерения лишь на более высоких уровнях агрегации.

Третьим шагом являются диагностические тесты на ARIMA-компоненты, которые могут добавить уверенности в предположе-

ниях, основанных на теории и визуальном обследовании. Функции (частичной) автокорреляции позволяют идентифицировать наличие ARMA-признаков, а панель тестов на интеграцию – наличие стохастического тренда. Если временная структура ряда относительно проста, а сам ряд стационарен, использование авторегрессии с распределенным лагом, с контролем на нужное количество прошлых значений зависимой переменной дает консистентные оценки параметров и позволяет напрямую интерпретировать связи между переменными в модели в оригинальных единицах измерения. Дифференцирование нестационарного ряда меняет интерпретации связи: изменения, а не абсолютный уровень, становятся предметом анализа.

Сложности почти неизбежно возникают, когда тесты демонстрируют комплексную динамику (присутствие автокорреляции, циклов и тренда). Очистка данных становится почти обязательной процедурой, равно как и диагностика итоговой модели на автокорреляцию ошибок с помощью тестов Портмонто (Portmanteau test), мультипликатора Лагранжа (Lagrange multiplier test) и др. Большинство программ содержат специализированные пакеты с соответствующими наборами тестов, однако их корректное применение и демонстрация в результатах остается задачей самих исследователей.

Наконец, исследователям стоит обращать внимание на коинтеграцию, эндогенность и волатильность вариации при анализе динамических процессов. Все три концепта являются значимыми как содержательно, так и технически: коинтеграция (к примеру, в военных расходах между двумя соперничающими державами) может быть предметом формального моделирования или указывать на наличие механизмов стабилизации динамической системы взаимодействий. Учет эндогенности позволяет эмпирически тестировать ключевые гипотезы политических наук, например «закона неизбежной репрессивности» (law of coercive responsiveness). Наконец, неравномерность дисперсии также может указывать на наличие теоретически интересных связей между контекстом и процессом: к примеру, исследователи указывают, что оценки перспектив развития экономики имеют больший эффект на рейтинги одобрения правительства в период избирательной кампании, а не между ними [Lebo, Vox-Steffensmeier, 2008]. В любом случае, политические процессы – слишком комплексны и интересны, чтобы подходить к ним только с определенным набором инструментов;

тщательная работа с альтернативными спецификациями (включая уровень агрегации, количество лагов, векторы коинтеграции, наборы предикторов и т.д.) расширяет знания о динамике политических процессов, даже если результат не соответствует изначальным ожиданиям и теории.

Заключение

Политические процессы не просто динамичны, их отличает комплексность, петли обратной связи и нелинейность. Данная статья демонстрирует возможности статистического моделирования политических процессов и многочисленные «подводные камни», неизбежно сопровождающие работу исследователя в этой области. Так, анализ рейтинга одобрения президента Владимира Путина с помощью модели авторегрессии со скользящим средним показывает наличие долговременной памяти и циклов: возмущения, которые выводят одобрение из состояния равновесия, заставляют рейтинг двигаться в определенном направлении некоторое время, после чего он возвращается к своему среднему значению в 75 п.п. Анализ готовности присоединиться к акциям протеста с экономическими требованиями, в свою очередь, свидетельствует о том, что россияне реагируют на изменения в инфляции и безработице в предсказанном направлении, но с запозданием. Более того, модель коррекции регрессионных остатков указывает на долгосрочное равновесие между источниками экономического недовольства и готовностью к коллективным действиям.

Безусловно, приведенные выше примеры являются всего лишь иллюстрациями и нуждаются в дополнительных исследованиях. Тем не менее они согласуются с существующим знанием относительно динамики общественного мнения, будь то рейтинги одобрения или протестные настроения, и воздействия объективных экономических показателей на них. Применение инструментов анализа временных рядов позволяет не только протестировать ключевые гипотезы, но и выявить ограничения, а также высветить аспекты, которые могут быть изучены в дальнейшем: например, что отражает долговременная память в рейтингах доверия Владимиру Путину – его персональную харизму, компетентность или

действия контекстуальных факторов, таких как экономический рост или политическая стабильность?

Анализ временных рядов в политической науке на данный момент является одним из наиболее динамично развивающихся направлений: будучи заимствованным из экономики и опираясь на схожий концептуальный аппарат, этот метод значительно расширил возможности исследователей. Его значимость не ограничивается возможностью прогнозов, напротив, наиболее важный вклад анализ временных процессов вносит в теоретические дискуссии о макрополитической динамике и ее микрооснованиях. Техническое совершенствование инструментария (например, байесовские модели или динамические панели) зачастую ставят математическую сложность исследований в один ряд с передовыми работами в области экономики, снижая доступность результатов, но это не должно становиться препятствием к его использованию. Сложность политических процессов не должна останавливать исследователей в попытках выявить динамические регулярности и связи.

Список литературы

- Abramson P.R., Ostrom C.W.* Macropartisanship: An empirical reassessment // *American political science review.* – 1991. – Vol. 85, N 1. – P. 181–192. – DOI: <https://doi.org/10.2307/1962884>
- Anderson C.J.* Economic voting and political context: a comparative perspective // *Electoral studies.* – 2000. – Vol. 19, N 2–3. – P. 151–170. – DOI: [https://doi.org/10.1016/s0261-3794\(99\)00045-1](https://doi.org/10.1016/s0261-3794(99)00045-1)
- De Boef S., Keele L.* Taking time seriously // *American journal of political science.* – 2008. – Vol. 52, N 1. – P. 184–200. – DOI: <https://doi.org/10.1111/j.1540-5907.2007.00307.x>
- Elite economic forecasts, economic news, mass economic judgments, and presidential approval / *R. Nadeau, R.G. Niemi, D.P. Fan, T. Amato* // *The Journal of politics.* – 1999. – Vol. 61, N 1. – P. 109–135. – DOI: <https://doi.org/10.2307/2647777>
- Engle R.* New frontiers for ARCH models // *Journal of applied econometrics.* – 2002. – Vol. 17, N 5. – P. 425–446. – DOI: <https://doi.org/10.1002/jae.683>
- Hyndman R.J., Athanasopoulos G.* Forecasting: principles and practice. – 2018. – Mode of access: <https://otexts.com/> (accessed: 29.08.2020).
- Lebo M.J., Box - Steffensmeier J.M.* Dynamic conditional correlations in political science // *American journal of political science.* – 2008. – Vol. 52, N 3. – P. 688–704. – DOI: <https://doi.org/10.1111/j.1540-5907.2008.00337.x>

- Lebo M.J., Cassino D.* The aggregated consequences of motivated reasoning and the dynamics of partisan presidential approval // *Political psychology*. – 2007. – Vol. 28, N 6. – P. 719–746. – DOI: <https://doi.org/10.1111/j.1467-9221.2007.00601.x>
- Time series analysis for the social sciences / J.M. Box-Steffensmeier, J.R. Freeman, M.P. Hitt, J.C. Pevehouse.* – Cambridge : Cambridge university press, 2014. – 280 p.
- Treisman D.* Presidential popularity in a hybrid regime: Russia under Yeltsin and Putin // *American journal of political science*. – 2011. – Vol. 55, N 3. – P. 590–609. – DOI: <https://doi.org/10.1111/j.1540-5907.2010.00500.x>
- Velasco C.* Gaussian semi - parametric estimation of fractional cointegration // *Journal of time series analysis*. – 2003. – Vol. 24, N 3. – P. 345–378. – DOI: <https://doi.org/10.1111/1467-9892.00311>
- Welsch H.* Macroeconomics and life satisfaction: Revisiting the “Misery Index” // *Journal of applied economics*. – 2007. – Vol. 10, N 2. – P. 237–251. – DOI: <https://doi.org/10.1080/15140326.2007.12040489>

A.V. Semenov*

Time-series analysis in political sciences: opportunities and limitations

Abstract. In this paper, I consider opportunities and limitations of modelling the political dynamics with the time-series instruments. Using the examples of the president Putin’s approval rating and readiness to join the collective actions with economic demands I demonstrate the analytical potential of autoregressive integrated moving average model (ARIMA), autoregressive distributed lag model (ADL), and error correction model (ECM). Modelling the political dynamics faces a string of analytical dilemmas. This paper aims at identifying the basic choices in application of the statistical instruments to dynamic processes and helping the other researchers to navigate through them. While it is hard to account in a single paper for all the developments in the discipline, which has been substantially advanced substantially and technically in the last three decades, this text also aims at stimulating the discussion on the opportunities and limitations when applied to Russian politics.

Keywords: presidentialism; responsibility attribution; economic crisis; economic hardships; political institutions; accountability.

For citation: Semenov A.V. Time-series analysis in political sciences: opportunities and limitations. *Political science (RU)*. 2021, N 1, P. 76–97. DOI: <http://www.doi.org/10.31249/poln/2021.01.03>

* **Semenov Andrei**, Perm Federal Research Center of the Ural Branch of the Russian Academy of Sciences (Perm, Russia), e-mail: andreysemenov@comparativestudies.ru

References

- Abramson P.R., Ostrom C.W. Macropartisanship: An empirical reassessment. *American political science review*. 1991, Vol. 85, N 1, P. 181–192. DOI: <https://doi.org/10.2307/1962884>
- Anderson C.J. Economic voting and political context: a comparative perspective // *Electoral studies*. 2000, Vol. 19, N 2–3, P. 151–170. DOI: [https://doi.org/10.1016/s0261-3794\(99\)00045-1](https://doi.org/10.1016/s0261-3794(99)00045-1)
- Box-Steffensmeier J.M., Freeman J.R., Hitt M.P., et al. *Time series analysis for the social sciences*. Cambridge : Cambridge university press, 2014, 280 p.
- De Boef S., Keele L. Taking time seriously. *American journal of political science*. 2008, Vol. 52, N 1, P. 184–200. DOI: <https://doi.org/10.1111/j.1540-5907.2007.00307.x>
- Engle R. New frontiers for ARCH models. *Journal of applied econometrics*. 2002, Vol. 17, N 5, P. 425–446. DOI: <https://doi.org/10.1002/jae.683>
- Hyndman R.J., Athanasopoulos G. *Forecasting: principles and practice*. 2018. Mode of access: <https://otexts.com/> (accessed: 29.08.2020).
- Lebo M.J., Box - Steffensmeier J.M. Dynamic conditional correlations in political science. *American journal of political science*. 2008, Vol. 52, N 3, P. 688–704. DOI: <https://doi.org/10.1111/j.1540-5907.2008.00337.x>
- Lebo M.J., Cassino D. The aggregated consequences of motivated reasoning and the dynamics of partisan presidential approval. *Political psychology*. 2007, Vol. 28, N 6, P. 719–746. DOI: <https://doi.org/10.1111/j.1467-9221.2007.00601.x>
- Nadeau R., Niemi R.G., Fan D.P., et al. Elite economic forecasts, economic news, mass economic judgments, and presidential approval. *The Journal of politics*. 1999, Vol. 61, N 1, P. 109–135. DOI: <https://doi.org/10.2307/2647777>
- Treisman D. Presidential popularity in a hybrid regime: Russia under Yeltsin and Putin. *American journal of political science*. 2011, Vol. 55, N 3, P. 590–609. DOI: <https://doi.org/10.1111/j.1540-5907.2010.00500.x>
- Velasco C. Gaussian semi - parametric estimation of fractional cointegration. *Journal of time series analysis*. 2003, Vol. 24, N 3, P. 345–378. DOI: <https://doi.org/10.1111/1467-9892.00311>
- Welsch H. Macroeconomics and life satisfaction: Revisiting the “Misery Index”. *Journal of applied economics*. 2007, Vol. 10, N 2, P. 237–251. DOI: <https://doi.org/10.1080/15140326.2007.12040489>