

УДК: 330.43, 519.23

Прогнозирование демографических и макроэкономических показателей в распределенной модели мировой динамики

С. А. Махов

Институт прикладной математики им. М. В. Келдыша Российской академии наук,
Россия, 125047, г. Москва, Миусская пл., 4

E-mail: s_makhov@mail.ru

*Получено 15.12.2022, после доработки — 21.12.2022.
Принято к публикации 30.01.2023.*

Разработана динамическая макро модель мировой динамики. В модели мир разбит на 19 регионов по географическому принципу согласно классификации Организации объединенных наций. Внутреннее развитие регионов описывается уравнениями разностного типа для демографических и экономических индикаторов, таких как численность населения, валовой продукт, валовое накопление. Межрегиональные взаимодействия представляют собой агрегированные торговые потоки от региона к региону и описываются регрессионными уравнениями. В качестве регрессоров использовались время, валовой продукт экспортера и валовой продукт импортера. Рассматривалось четыре типа: временная парная регрессия — зависимость торгового потока от времени, экспортная функция — зависимость доли торгового потока в валовом продукте экспортера от валового продукта импортера, импортная функция — зависимость доли торгового потока в валовой продукции импортера от валового продукта экспортера, множественная регрессия — зависимость торгового потока от валовых продуктов экспортера и импортера. Для каждого типа применялось два вида функциональной зависимости: линейная и логарифмически-линейная, всего исследовано восемь вариантов торгового уравнения. Проведено сравнение качества регрессионных моделей по коэффициенту детерминации. Расчеты показывают, что модель удовлетворительно аппроксимирует динамику монотонно меняющихся показателей. Проанализирована динамика немонотонных торговых потоков, для их аппроксимации предложено три вида функциональной зависимости от времени. Показано, что с 10%-й погрешностью множество внешнеторговых рядов может быть приближено пространством семи главных компонент. Построен прогноз автономного развития регионов и глобальной динамики до 2040 года.

Ключевые слова: мировая динамика, математическое моделирование, макроэкономика, внешняя торговля, временные ряды, регрессионный анализ, прогноз

UDC: 330.43, 519.23

Forecasting demographic and macroeconomic indicators in a distributed global model

S. A. Makhov

Keldysh Institute of Applied Mathematics of Russian Academy of Sciences,
4 Miusskaya sq., Moscow, 125047, Russia

E-mail: s_makhov@mail.ru

Received 15.12.2022, after completion – 21.12.2022.

Accepted for publication 30.01.2023.

The paper presents a dynamic macro model of world dynamics. The world is divided into 19 geographic regions in the model. The internal development of the regions is described by regression equations for demographic and economic indicators (Population, Gross Domestic Product, Gross Capital Formation). The bilateral trade flows from region to region describe interregional interactions and represent the trade submodel. Time, the gross product of the exporter and the gross product of the importer were used as regressors. Four types were considered: time pair regression – dependence of trade flow on time, export function – dependence of the share of trade flow in the gross product of the exporter on the gross product of the importer, import function – dependence of the share of trade flow in the gross product of the importer on the gross product of the exporter, multiple regression – dependence of trade flow on the gross products of the exporter and importer. Two types of functional dependence were used for each type: linear and log-linear, in total eight variants of the trading equation were studied. The quality of regression models is compared by the coefficient of determination. By calculations the model satisfactorily approximates the dynamics of monotonically changing indicators. The dynamics of non-monotonic trade flows is analyzed, three types of functional dependence on time are proposed for their approximation. It is shown that the number of foreign trade series can be approximated by the space of seven main components with a 10% error. The forecast of regional development and global dynamics up to 2040 is constructed.

Keywords: world dynamics, mathematical modeling, macroeconomics, foreign trade, time series, regression analysis, forecast

Citation: Computer Research and Modeling, 2023, vol. 15, no. 3, pp. 757–779 (Russian).

Введение

Настоящая работа является логическим продолжением работ, посвященных макромоделированию отдельных стран и их совместной торговли [Махов, 2016; Махов, 2017]. Методика моделирования аналогична той, что использовалась в [Махов, 2013], и опирается на работы [Пирогов и др., 1980; Дубовский, 2005; Дубовский, 2012].

Целью исследования является прогнозирование основных тенденций и построение сценариев глобального развития. В качестве основных методов и подходов используются: а) аппроксимация временных рядов индикаторов статистическими методами; б) построение и анализ динамической модели мировых макропоказателей. Применяется распределенный подход: мир разбивается на регионы по географическому принципу. Для каждого региона строится макромодел, описывающая демографические и экономические процессы, происходящие внутри региона; взаимодействие регионов описывается моделью внешней торговли.

В региональных моделях в качестве основных переменных взяты следующие индикаторы: численность народонаселения, валовой внутренний продукт (ВВП), капитал (стоимость основных производственных фондов), капитальные вложения. Для расчета переменных составляются разностные уравнения регрессионного типа. Параметры регрессий оцениваются по известной в прошлом статистике.

В модели внешней торговли основными показателями являются двусторонние торговые потоки, вычисляемые на основе автономных региональных моделей. То есть основные переменные региональных моделей являются экзогенными величинами для торговой модели. Такая вычислительная схема позволяет разделить модельные блоки и упростить расчеты.

В целях упрощения используется схема разбивки на регионы, применяемая в Организации объединенных наций. Пять континентов, Африка, Америка, Азия, Европа, Океания, подразделяются на 4–5 географических зон каждый (кроме Океании, являющейся единой зоной), всего 19 регионов. Названия регионов, принятые в работе, приведены в табл. 1.

Таблица 1. Названия и сокращения регионов

№	Регион (рус.)	Регион (англ.)	Сокращение	Континент
1	Северная Африка	Northern Africa	NAFR	Африка
2	Восточная Африка	Eastern Africa	EAFR	Африка
3	Центральная Африка	Middle Africa	MAFR	Африка
4	Южная Африка	Southern Africa	SAFR	Африка
5	Западная Африка	Western Africa	WAFR	Африка
6	Северная Америка	Northern America	NAMR	Америка
7	Карибские острова	Caribbean	CRBN	Америка
8	Центральная Америка	Central America	CAMR	Америка
9	Южная Америка	South America	SAMR	Америка
10	Средняя Азия	Central Asia	CAS	Азия
11	Восточная Азия	Eastern Asia	EAS	Азия
12	Юго-Восточная Азия	South-Eastern Asia	SEAS	Азия
13	Южная Азия	Southern Asia	SAS	Азия
14	Западная Азия	Western Asia	WAS	Азия
15	Восточная Европа	Eastern Europe	EEU	Европа
16	Северная Европа	Northern Europe	NEU	Европа
17	Южная Европа	Southern Europe	SEU	Европа
18	Западная Европа	Western Europe	WEU	Европа
19	Океания	Oceania	OCN	Австралия и Океания

Таким образом, объектами исследования являются географически близкие территории, точнее говоря, группы государств, объединенных по географическому признаку. В дополнение к регионам и континентам также отслеживается «суперрегион» — Мир (World, WLD).

Описание модели. Уравнения и параметры

В этом разделе приводятся уравнения автономной региональной модели и модели внешней торговли.

Подмодель автономного развития региона

Регион характеризуется четырьмя основными переменными: численность населения N , основной капитал K , валовой внутренний продукт Y и валовое накопление I .

Демографическое уравнение. Как и в [Махов, 2017], динамика народонаселения происходит по логистическому закону:

$$N(t) = N(t-1) \cdot (1 + d_0 - d_1 \cdot N(t-1)), \quad (1)$$

где $N(t)$ — численность населения в году t , $N(t-1)$ — численность населения в году $(t-1)$, d_0, d_1 — коэффициенты. Форма уравнения (1) предполагает, что темп прироста народонаселения убывает со временем, то есть коэффициент рождаемости снижается и приближается к коэффициенту смертности.

Капитал. Уравнение для капитала является его определением:

$$K(t) = I(t) + I(t-1) + I(t-2), \quad (2)$$

где $I(t), I(t-1), I(t-2)$ — валовое накопление региона в году $t, t-1, t-2$.

Заметим, что ранее в работах [Махов, 2016; Махов, 2017] применялся другой подход, при котором капитал рассчитывался как интеграл от валового накопления с учетом выбытия (в статистических базах выбытие фондов также называется потреблением основного капитала). Отказ от прежнего подхода вызван двумя желаниями: сократить количество разностных уравнений и избежать длинной процедуры оценивания коэффициентов выбытия и восстановления статистических данных капитала. Дело в том, что «капитал» — ненаблюдаемая величина, в отличие, например, от валового продукта, для расчета которого в системе национальных счетов существуют несколько общепринятых методик. Таким образом, (2) задает условную величину $K(t)$ в виде суммы валового накопления за три года. Строгого обоснования, почему именно 3 года, а не, к примеру, 5 лет, нет; можно сказать, что 3 — это минимальное число, когда суммирование дает хорошее сглаживание исходного ряда $I(t)$.

Валовой продукт. Уравнение представляет собой определение производственной функции, для простоты используется линейная зависимость:

$$Y(t) = a_K \cdot K(t) + a_N \cdot N(t), \quad (3)$$

где a_K, a_N — параметры. Разумеется, можно было использовать и нелинейные зависимости, однако, как показало дополнительное исследование, это не повышает точность модели.

Валовое накопление. Под валовым накоплением (в англоязычной литературе — *Gross Capital Formation*) понимается сумма вложений в основные фонды и прироста запасов материальных оборотных средств; уравнение таково:

$$I(t) = c_Y \cdot Y(t-1) + c_K \cdot K(t-1), \quad (4)$$

где c_Y, c_K — коэффициенты. Заметим, что записано динамическое уравнение, поскольку оно связывает искомую величину текущего года с известными величинами прошлого года. При желании уравнению (4) можно придать разностно-дифференциальную форму, но с учетом (2), (3), и так получается авторегрессионное уравнение 3-го порядка.

Кроме четырех основных переменных вычисляются дополнительные:

$$C(t) = c_0 + c_1 \cdot Y(t), \quad (5)$$

$$G(t) = g_0 + g_1 \cdot Y(t), \quad (6)$$

где C — конечное потребление домашних хозяйств, G — государственные расходы (конечное потребление органов государственного управления).

Как видно из (1), численность населения рассчитывается независимо от экономических переменных; сама же она влияет на остальные переменные посредством (3). То есть в модель заложена гипотеза об одностороннем влиянии демографии на экономику.

Помимо параметров $d_0, d_1, a_K, a_N, c_Y, c_K, c_0, c_1, g_0, g_1$, требуется задать год начала моделирования t_0 и начальные условия $N(t_0), I(t_0), I(t_0 - 1), I(t_0 - 2)$, всего 15 констант.

Подмодель внешней торговли

Внешняя торговля региона описывается двусторонними товарными потоками, каждый из которых в общем случае зависит от валовых региональных продуктов и времени:

$$T_{ij} = f(t, Y_i, Y_j), \quad (7)$$

где T_{ij} — товарный поток из региона i в регион j ; Y_i, Y_j — валовые продукты регионов экспортера i и импортера j соответственно; t — время в годах: либо календарный год, либо номер года, начиная с определенного момента (например, с 1970 г. или 1995 г.). Потоки считаются брутто-величинами, то есть внутренняя торговля региона ненулевая: $T_{ii} \neq 0$.

В качестве функции f использовались два типа зависимости, линейная и лог-линейная:

$$T_{ij} = A + \alpha \cdot Y_i + \beta \cdot Y_j + \theta \cdot t, \quad (7')$$

$$\ln T_{ij} = \ln A + \alpha \cdot \ln Y_i + \beta \cdot \ln Y_j + \theta \cdot t, \quad (7'')$$

где A, α, β, θ — параметры регрессии. Однако ни в виде (7'), ни в виде (7'') внешнеторговое уравнение не исследовалось, использовались упрощенные зависимости, либо только от ВВП, либо только от времени. Дополнительно с множественной регрессией $f(Y_i, Y_j)$ рассматривались парные регрессии, получающиеся из (7') и (7'') при $\alpha = 1$ или $\beta = 1$. Итого получилось 8 регрессионных моделей:

$$T_{ij} = A + \alpha \cdot Y_i + \beta \cdot Y_j, \quad (7a)$$

$$\ln T_{ij} = \ln A + \alpha \cdot \ln Y_i + \beta \cdot \ln Y_j, \quad (7b)$$

$$\frac{T_{ij}}{Y_i} = \gamma + \delta \cdot Y_j, \quad (7c)$$

$$\ln \frac{T_{ij}}{Y_i} = \gamma + \delta \cdot \ln Y_j, \quad (7d)$$

$$\frac{T_{ij}}{Y_j} = \gamma + \delta \cdot Y_i, \quad (7e)$$

$$\ln \frac{T_{ij}}{Y_j} = \gamma + \delta \cdot \ln Y_i, \quad (7f)$$

$$T_{ij} = \eta + \theta \cdot t, \quad (7g)$$

$$\ln T_{ij} = \eta + \theta \cdot t. \quad (7h)$$

Уравнения (7a) и (7b) соответствуют множественной линейной и степенной зависимости, уравнения (7c) и (7d) описывают линейную и степенную экспортную функцию, (7e) и (7f) — линейную и степенную импортную функцию. Заметим, что для внутрирегионального потока T_{ii} экспортная и импортная функции совпадают.

Зная все потоки, можно рассчитать суммарный товарный экспорт T_{iW} и суммарный товарный импорт T_{Wi} i -го региона:

$$T_{iW} = \sum_j T_{ij}, \quad (8)$$

$$T_{Wi} = \sum_k T_{ki}. \quad (9)$$

Суммирование происходит по всем торговым партнерам, включая сам регион i . Полный региональный экспорт с учетом услуг X_i связан линейной зависимостью с суммарным товарным экспортом:

$$X_i = e_0 + e_1 \cdot T_{iW}. \quad (10)$$

Аналогично: полный региональный импорт M_i определяется через суммарный товарный импорт:

$$M_i = m_0 + m_1 \cdot T_{Wi}. \quad (11)$$

В уравнениях (10), (11) коэффициенты e_0 , e_1 , m_0 , m_1 — константы для каждого региона.

Региональный экспорт X_i и региональный импорт M_i входят в известное балансное соотношение:

$$Y_i = C_i + G_i + I_i + X_i - M_i. \quad (12)$$

Уравнение (12) описывает обратную связь от подмодели внешней торговли и позволяет оценить погрешность вычисления Y_i .

Общий размер модели: 551 уравнение для 551 переменной, а также 1425 параметров и начальных значений.

Эмпирические данные

Для оценки параметров модели и последующей верификации уравнений использовались статистические данные из следующих источников:

- 1) база данных экономической статистики ООН (United Nations Statistics Division, Economic Statistics) [United Nations Statistic Division, 2020];
- 2) статистика индикаторов мирового развития Всемирного банка (World Bank, World development indicators) [World Bank, 2020];
- 3) база данных ЮНТКАД — Конференция ООН по торговле и развитию (UNCTAD — United Nations Conference on Trade and Development) [UNCTAD, 2020].

Выяснилось, что используемые в автономной подмодели индикаторы повторяются во всех трех базах, поэтому технически оказалось достаточно базы ЮНТКАД, содержащей полный набор данных в нижеперечисленных таблицах.

Демографическая статистика приведена в таблице Total and urban population, annual: оценки численности населения с 1950 г. (самое раннее) по 2019 г., а также прогнозные значения с 2020 по 2050 г.

Макроэкономические агрегаты содержатся в таблице Gross domestic product: GDP by type of expenditure, VA by kind of economic activity, total and shares, annual с 1970 по 2019 г. и измеряются в постоянных долларах США 2015 г.

Двусторонний товарный экспорт и импорт даны в таблице Merchandise trade matrix, annual в текущих долларах США за период 1995–2019 гг., практически, за последние 25 лет¹. Исключение: часть данных об экспорте в Южную Африку за 1995–1999 гг. отсутствует.

При этом полный экспорт и полный импорт (как макроэкономические агрегаты) имеются как в текущих, так и в постоянных долларах. Поэтому необходимо привести все величины к одним единицам измерения.

Также двусторонние потоки не содержат данные о торговле услугами, есть только данные о полном экспорте и импорте услуг данной страны с остальным миром. Кроме того, не совпадают величины экспорта и импорта, относящихся к одному и тому же потоку. Поясним: экспорт из региона А в регион Б — то же самое, что импорт в регион Б из региона А в натуральных показателях; при этом денежное выражение, согласно статистике, различное, импорт, как правило, превышает экспорт. Связано это с тем, что экспорт рассчитывается в ценах по условиям FOB, а импорт — в ценах по условиям CIF². Исходя из сказанного, было принято решение в качестве показателя, отражающего товарные перевозки между странами, взять двусторонний экспорт в постоянных долларах 2015 г. Для перевода из текущих в постоянные доллары использовались данные об инфляции полного экспорта, сам пересчет производился по формуле

$$T_{ij}^{2015}(t) = T_{ij}(t) \frac{X_i^{2015}(t)}{X_i(t)}. \quad (13)$$

Здесь $T_{ij}^{2015}(t)$ — товарный экспорт из региона i в регион j в долларах 2015 г. в году t ; $T_{ij}(t)$ — товарный экспорт региона i в регион j в текущих долларах в году t ; $X_i^{2015}(t)$ — полный экспорт региона i в долларах 2015 г. в году t ; $X_i(t)$ — полный экспорт страны i в текущих долларах в году t .

Статистика товарного импорта теоретически может быть использована для оценки разброса и погрешностей данных, в настоящей работе это не проводилось.

Верификация модели

Проверка проводилась в два этапа: 1) верификация автономной подмодели (уравнения (1)–(6)); 2) верификация подмодели торговых взаимодействий (уравнения (7)–(11)).

Демография региона

Демографическое уравнение не всегда адекватно описывает динамику численности населения региона за последние 25 лет. В частности, плохие результаты показали Африканский и Европейский континенты. Связано это с тем, что уравнение (1) описывает ситуацию устойчивого снижения темпов прироста численности населения, т. е. второй стадии демографического перехода [Капица, 1999; Вишневецкий и др., 2005], в то время как в Европе эта фаза демографического перехода фактически завершилась, а в Африке только-только началась.

¹ Ситуация весны-лета 2021 г.; в ноябре 2021 г. единая таблица разбита на 5 таблиц с 5-летними временными интервалами: 1995–2000, 2001–2005, 2006–2010, 2011–2015, 2016–2020.

² FOB — от английской аббревиатуры FOB, означающей Free on Board (буквально «бесплатно на борт судна»). Условия FOB предусматривают, что продавец обязан доставить товар в порт и погрузить на указанное покупателем судно; расходы по доставке товара на борт судна ложатся на продавца. CIF — от английской аббревиатуры CIF, означающей Cost, Insurance and Freight (буквально «стоимость, страхование и фрахт»). Условия CIF означают, что продажная цена включает в себя стоимость товара, фрахт или транспортные расходы, а также стоимость страховки для морских перевозок. Таким образом, эти условия подходят только для перевозок водным транспортом.

Последнее обстоятельство позволяет все же использовать уравнение (1) для африканских регионов со сдвигом начала расчетов: в Северной и Восточной Африке начало расчетов — 2015 г., в Центральной Африке — 2013 г., в Южной Африке — 2014 г., в Западной Африке — 2012 г. В качестве статистического материала для нахождения параметров уравнений использованы прогнозные данные ООН до 2050 г. Качество регрессий оказалось высоким: коэффициент детерминации всех регрессий попал в диапазон от 0,95 до 0,99. Фактически удалось получить демографический прогноз более простым способом, чем по методике, используемой в отделе народонаселения Департамента по экономическим и социальным вопросам ООН¹ [United Nations, 2019]. В качестве примера на рис. 1 приведены статистические данные и уравнение регрессии для Северной Африки.

Менее удачно с точки зрения применимости уравнения (1) обстоит дело в европейских регионах. Начать с того, что само уравнение не подошло ни для одного региона. Тем не менее в трех регионах удалось связать темп прироста народонаселения с временем t : в Западной и Южной Европе зависимость оказалась линейной начиная с 1971 г., в Северной Европе статистические данные неплохо легли на экспоненту начиная с 2010 г. (рис. 2). Таким образом, вместо уравнения (1) имеют место две его модификации:

$$N(t) = N(t-1) \cdot (1 + d_0 - d_1 \cdot t), \quad (1')$$

$$N(t) = N(t-1) \cdot (1 + d_0 \cdot \exp(-d_1 \cdot t)), \quad (1'')$$

Хуже всего ситуация в Восточной Европе: устойчиво отрицательные с 1993 г. темпы прироста не удалось описать простой зависимостью. Поэтому при дальнейших расчетах пришлось использовать «костыль» — прогноз ООН до 2050 г.

Демография прочих регионов хорошо описывается уравнением (1), более того, на довольно длительном интервале: с 1971 по 2050 г. Исключения: Северная Америка, Средняя Азия, Южная Азия, Западная Азия, которые лучше аппроксимируются уравнением (1) на интервалах 1995–2050, 2012–2050, 1983–2050, 2000–2050 соответственно.

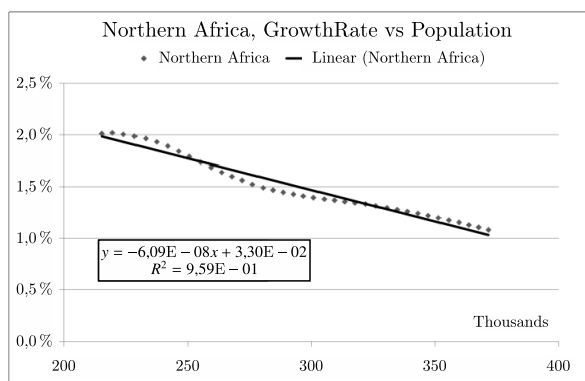


Рис. 1. Темп прироста народонаселения в зависимости от численности населения в Северной Африке. Численность населения приведена в тыс. чел., темп прироста — в процентах. Уравнение регрессии и коэффициент детерминации — в рамке

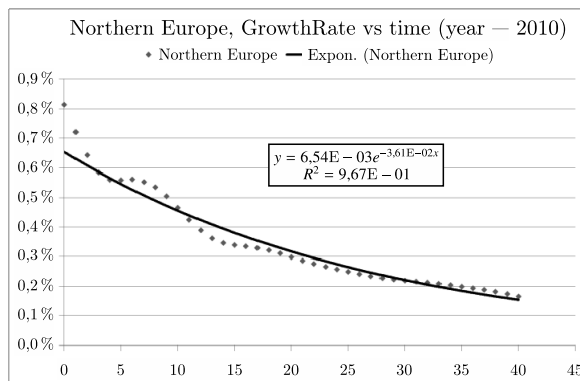


Рис. 2. Темп прироста народонаселения в зависимости от времени в Северной Европе. Время измеряется в годах, отсчет идет с 2010 г. (т. е. 10 на горизонтальной оси соответствует 2020 г.), темп прироста — в процентах

¹ Разумеется, мы не претендуем на большую точность, тем более не пытаемся подменить своей моделью вычисления, проведенные большой группой статистиков. Мы лишь пытаемся обойти сложность, связанную с расчетом огромного количества демографических показателей, построив модель для вычисления единственного индикатора — «численность народонаселения».

Макроэкономика региона

Соотношение (2) не требует верификации, поскольку является определением основного капитала. Проверка уравнений (3), (4) дала в целом удовлетворительные результаты: коэффициенты детерминации всех регрессий оказались выше 0,9 на интервале с 1972 по 2019 г., за исключением четырех регионов. Для трех из них пришлось сдвинуть время начала расчетов: Восточная Африка (с 1992), Средняя Азия (с 1994) и Восточная Европа (с 1999). Для Южной Америки уравнения (3) и (4) подверглись модификации, в правые части были добавлены константы:

$$Y(t) = a_Y + a_K \cdot K(t) + a_N \cdot N(t), \tag{3'}$$

$$I(t) = c_I + c_Y \cdot Y(t - 1) + c_K \cdot K(t - 1). \tag{4'}$$

Даже в модифицированном виде качество регрессии (4') оказалось чуть ниже, чем у немодифицированного уравнения (4) для остальных регионов: коэффициент детерминации с точностью до третьего знака составил 0,806.

Уравнения (5), (6) успешно прошли верификацию на всем диапазоне имеющихся данных с 1970 по 2019 г. Более того, во многих случаях свободные члены (постоянные слагаемые) c_0 , g_0 оказались незначимы, и их можно положить равными нулю. В качестве иллюстрации на рис. 3 и рис. 4 приведены наилучшая регрессия и наихудшая регрессия соответственно.

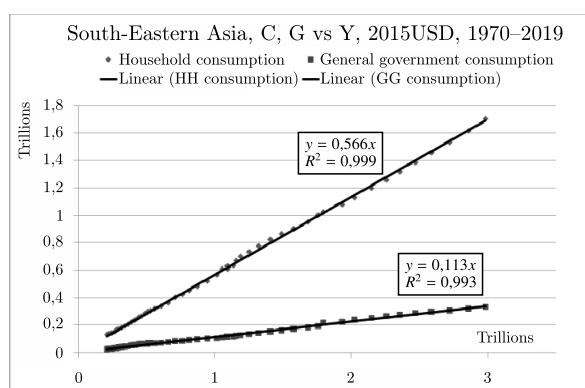


Рис. 3. Потребление домашних хозяйств (ромбики) и государственные расходы (квадратики) в Юго-Восточной Азии. Все величины представлены в триллионах долларов 2015 г. Уравнения регрессии и коэффициенты детерминации — в рамках

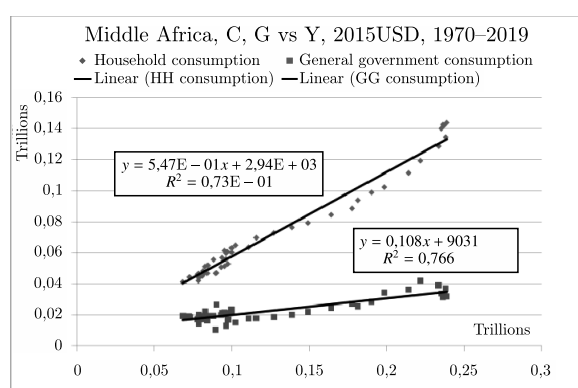


Рис. 4. Потребление домашних хозяйств (ромбики) и государственные расходы (квадратики) в Центральной Африке. Все величины представлены в триллионах долларов 2015 г. Уравнения регрессии и коэффициенты детерминации — в рамках

Внешняя торговля

Помимо верификации, также проводилось сравнение видов торгового уравнения по коэффициенту детерминации. Сравнение показало, что парные регрессии (7с), (7д), (7е), (7ф) в среднем хуже аппроксимируют данные, чем множественные регрессии (7а) и (7б), причем в ряде случаев заметно хуже. Основных причин этому две.

Во-первых, очевидно, что чем больше в модели объясняющих переменных (регрессоров), тем выше качество такой модели в том смысле, что выше доля объясненного разброса. Поэтому при прочих равных условиях коэффициент детерминации R^2 множественной регрессии всегда превышает аналогичный коэффициент парной регрессии. Иначе говоря, зависимость, описываемая соотношением (7а) или (7б), более гибка, чем зависимость вида (7с), (7е) или (7д), (7ф).

Вторая причина связана с видом левой части экспортных и импортных уравнений: здесь в качестве зависимой переменной выступает доля торгового потока в ВВП экспортера или импортера. Проблема в том, что преобразование исходного ряда T_{ij} в ряд $\frac{T_{ij}}{Y_i}$ или в $\frac{T_{ij}}{Y_j}$ «портит» его свойства с точки зрения разброса значений: ряд становится более «зашумленным». Более того, если проделать обратную операцию, т. е. умножить T_{ij} на Y_i или на Y_j , то новый ряд будет «лучше» с точки зрения разброса значений и качества регрессии. При этом итоговая зависимость в обоих вариантах получится одинаковой, но в первом случае коэффициент детерминации окажется меньше, чем во втором.

Кажущееся противоречие здравому смыслу возникает из-за слишком формального подхода к качеству регрессионной модели. На самом деле все несколько проще, если помнить о стохастической природе регрессии. Запись $Z = f(X)$ означает, что выражение $Z - f(X)$ (или $\frac{Z}{f(X)}$, или $\log\left(\frac{Z}{f(X)}\right)$, или что-нибудь подобное) является белым шумом — стационарным случайным процессом [Бокс, Дженкинс, 1974]. Обычно математическое ожидание такого процесса считается нулевым, хотя это необязательно, поскольку ненулевое среднее значение всегда можно вычесть. Дисперсия процесса определяет ошибку прогноза: чем больше дисперсия, тем выше погрешность и ниже точность прогноза.

Можно сделать следующие умозаключения:

- а) малые (близкие к нулю) значения R^2 говорят об отсутствии тренда (либо о его незаметности на фоне колебаний) и большой дисперсии остатков, однако это не значит, что прогноз невозможен; если получится белый шум, то в качестве прогноза можно взять константу — среднее значение ряда;
- б) средние значения R^2 — аргумент в пользу существования тренда, значительной дисперсии остатков и невысоком качестве модели;
- в) большие (близкие к единице) значения R^2 свидетельствуют о наличии тренда, малой дисперсии остатков и, как следствие, большой прогностической силе модели.

Возникает вопрос: как провести границы между случаями а), б) и в)? Простейшее решение — проявить волюнтаризм и проложить их симметричным образом, к примеру: а) [0; 0,3); б) [0,3; 0,7); в) [0,7; 1]. Таким образом можно получить некоторую информацию о качестве применяемой модели, но произвол в выборе границ оставляет чувство неудовлетворенности.

Второй вариант — разбить промежуток [0; 1] не на три, а на большее число интервалов; например, на десять одинаковых с шагом 0,1: [0; 0,1), [0,1; 0,2), [0,2; 0,3), ..., [0,9; 1], и далее посчитать количество регрессий, попавших в каждый интервал. То есть, по сути, построить гистограмму распределения регрессионных моделей по коэффициенту детерминации. Учитывая, что общее число моделей — 361 — довольно велико, можно получить наглядную картину того, насколько удачно та или иная модель справляется с подгонкой данных.

Наконец, еще один способ, исключающий произвол, но менее наглядный, — проранжировать регрессии по коэффициенту R^2 . Причем можно построить как прямую зависимость «ранг-размер», так и обратную — накопленную частоту от значения, т. е. эмпирическую функцию распределения. Последняя зависимость примечательна еще и тем, что связана с гистограммой интегрально (как первообразная с оригинальной функцией).

В целях упорядочения изложения назовем первый способ проведения границ грубым способом, второй — гистограммным или интервальным, третий — интегральным или непрерывным. Описанный метод позволяет сравнить торговые уравнения между собой в целом, не вникая в детали, какой торговый поток оказался лучше описан той или иной моделью.

Продолжая тему сравнения регрессионных моделей, коснемся вопроса, почему предлагаемые модели вообще могут быть применены для прогнозирования. По построению модели предлагают в качестве прогноза условное математическое ожидание (линию регрессии), показывающую тенденцию временного ряда. Если ряд имеет ярко выраженный восходящий или нисходящий тренд, то регрессионная модель четко это демонстрирует. Во всех исследуемых моделях (7a), (7b), (7c), (7d), (7e), (7f) в качестве объясняющих переменных используются региональные ВВП, которые в целом возрастают со временем (рис. 13–16).

Некоторые срывы наблюдаются в 1980-х (Западная Африка, Центральная Африка, Южная Азия, Северная Европа), в 1990-х (Центральная Африка, Южная Африка, Карибы, Средняя Азия, Восточная Европа) и в конце 2000-х (Северная Африка, Америка, Европа). Колебания выражены слабо и носят скорее нерегулярный характер; вообще, на 50-летнем интервале если и возможно определить экономические циклы, то только 3-летние циклы запасов (краткосрочные) и 10-летние инвестиционные циклы (среднесрочные). Выявление циклов, определение амплитуд и различение циклов от волн — дело непростое, требующее специального исследования; в настоящей работе такое исследование не проводилось.

Правые части уравнений (7a), (7b), (7c), (7d), (7e), (7f) так или иначе зависят от региональных ВВП и, таким образом, растут со временем. Следовательно, если конкретный двусторонний торговый поток монотонно изменяется со временем, то любая регрессионная модель из перечисленных покажет наличие тренда. Поэтому имеет смысл рассмотреть и включить в перечень сравнения непосредственную зависимость торгового потока от времени.

Поскольку множественные модели (7a) и (7b) по причине большего числа регрессоров в целом лучше подгоняют статистические данные, чем модели (7c), (7d), (7e) и (7f), то необходим эталон, о котором заранее нельзя сказать, лучше или хуже он множественной модели. По этой причине регрессионные модели, описываемые уравнениями (7g) и (7h), мы назовем эталонными.

Результаты сравнений, включая таблицы и графики, приведены в приложении 1. Общие выводы:

- 1) нелинейные модели (7b), (7d), (7f), (7h) лучше подгоняют данные, чем аналогичные линейные модели (7a), (7c), (7e), (7g);
- 2) модель (7b) лучше справляется с подгонкой данных, чем остальные модели;
- 3) в моделях (7c), (7d), (7e), (7f) случаев с близким к нулю R^2 больше, чем в эталонных регрессиях.

Если условно упорядочить модели по качеству подгонки, то выглядеть такой ряд будет так: (7a), (7b) > (7g), (7h) > (7d), (7f) > (7c), (7e). Здесь знак «>» означает переход в следующий качественный класс, через запятую перечислены модели одного класса. Соответственно, уравнение (7b) показало свою состоятельность, последующие расчеты и прогнозы сделаны на его основе.

Ложку дегтя в предыдущие рассуждения относительно сравнения регрессионных моделей добавляет тот факт, что судить о качестве модели на основании распределения коэффициента детерминации не вполне корректно. Академически правильнее сравнивать уравнения для каждого торгового потока поименно, применяя соответствующие статистические тесты. Например, информационный критерий Акаике, либо критерий Шварца, либо подходящий тест для невлоченных моделей [Вербик, 2008]. Такого рода исследование мы надеемся провести в ближайшем будущем.

Остается обсудить верификацию уравнений (10) и (11), поскольку соотношения (8) и (9) определяют суммарный товарный экспорт и суммарный товарный импорт региона, но не являются уравнениями связи.

Результаты оказались схожи с результатами проверки уравнений (5), (6): успешная верификация на всем диапазоне имеющихся данных и незначимость во многих случаях постоянных слагаемых e_0 , m_0 . Иллюстративный материал включает в себя рис. 5–8: на рис. 5 приведена наилучшая (с наибольшим R^2) регрессия вида (10), на рис. 6 — наихудшая (с наименьшим R^2) регрессия вида (10); на рис. 7, 8 приведены соответственно наилучшая и наихудшая регрессия вида (11).

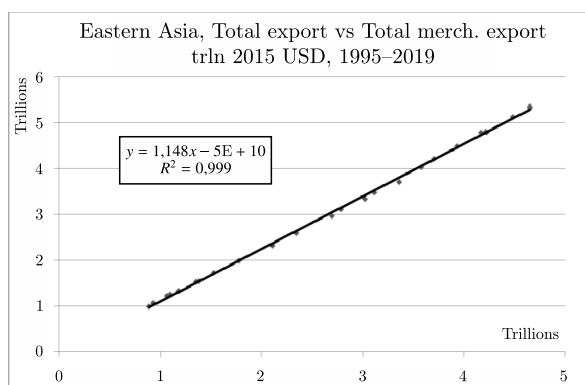


Рис. 5. Зависимость полного экспорта (ось ординат) от товарного экспорта (ось абсцисс) Восточной Азии. Величины представлены в трлн долл. 2015 г. Уравнение регрессии и коэффициент детерминации — в рамке

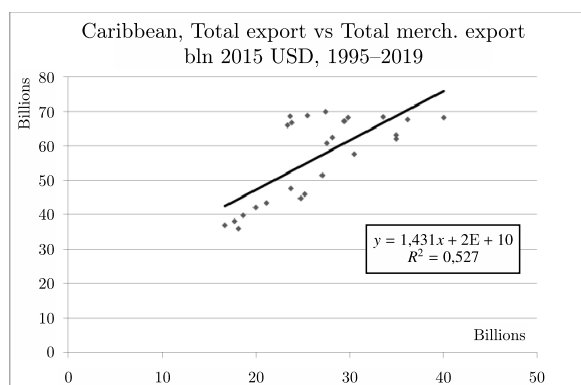


Рис. 6. Зависимость полного экспорта от товарного экспорта Карибского региона. Величины представлены в млрд долл. 2015 г. Уравнение регрессии и коэффициент детерминации — в рамке

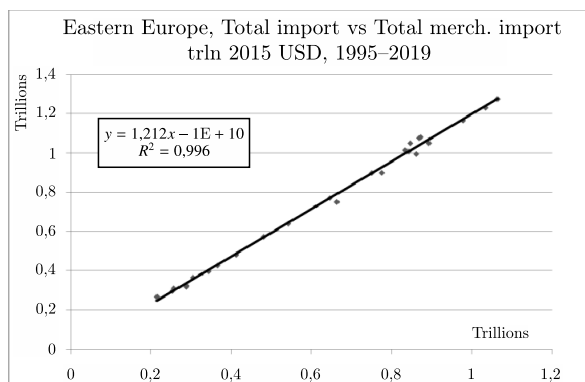


Рис. 7. Зависимость полного импорта от товарного импорта Восточной Европы. Величины представлены в трлн долл. 2015 г. Уравнение регрессии и коэффициент детерминации — в рамке

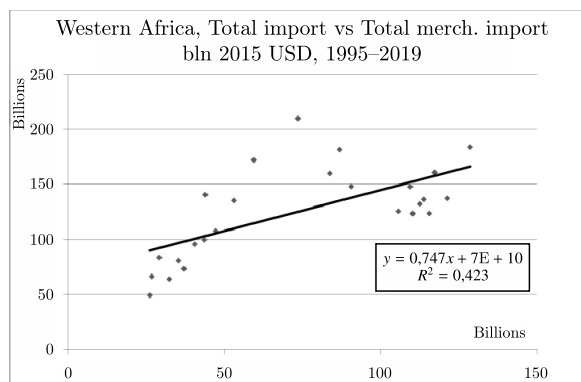


Рис. 8. Зависимость полного импорта от товарного импорта Западной Африки. Величины представлены в млрд долл. 2015 г. Уравнение регрессии и коэффициент детерминации — в рамке

Результаты расчетов

Учитывая размерность модели, мы не имеем возможности показать результаты всех вычислений, в противном случае объем работы существенно выйдет за допустимые рамки. Поэтому

графики будут приведены для части расчетных величин и лишь при одном фиксированном наборе параметров. В первую очередь демонстрируются результаты расчетов автономных моделей регионов. Ради экономии на одном рисунке 4–5 расчетных кривых, относящихся преимущественно к одному континенту.

На рис. 9–12 приведены графики численности населения с 1970 по 2040 г.: данные ООН, включая прогноз, и модельные расчеты. На рис. 9 отражена демография регионов Африки, на рис. 10 – Америки, на рис. 11 – Азии, на рис. 12 – Европы и Океании.

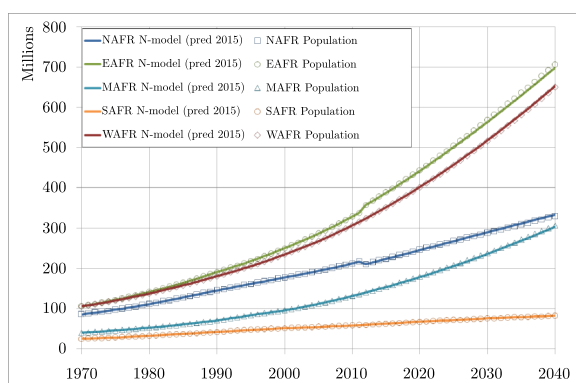


Рис. 9. Динамика численности населения регионов Африки в млн чел. Маркеры – исторические оценки с 1970 г. и прогноз ООН с 2020 г., сплошные линии – модельные расчеты, год начала расчетов указан в скобках рядом с подписью линии. Например, «NAFR N-model (pred 2015)» обозначает численность населения Северной Африки, рассчитанную по модели, начиная с 2015 г.

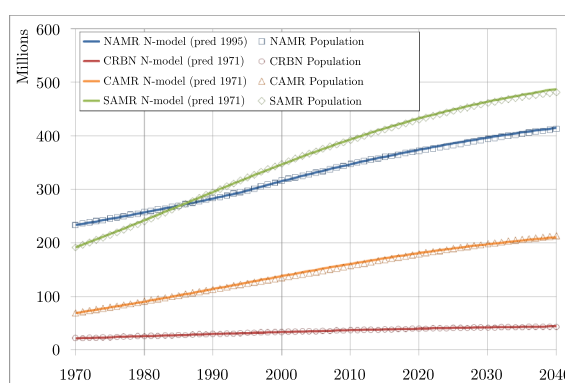


Рис. 10. Динамика численности населения американских регионов в млн чел.

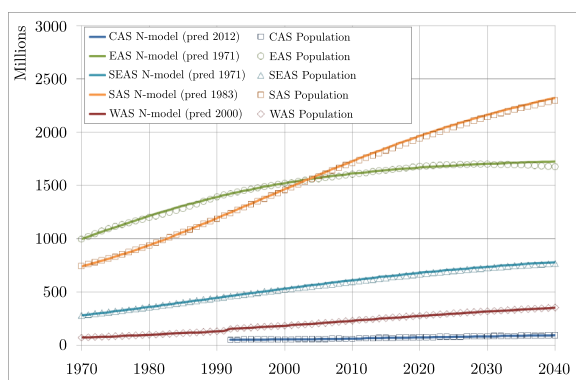


Рис. 11. Динамика численности населения регионов Азии в млн чел.

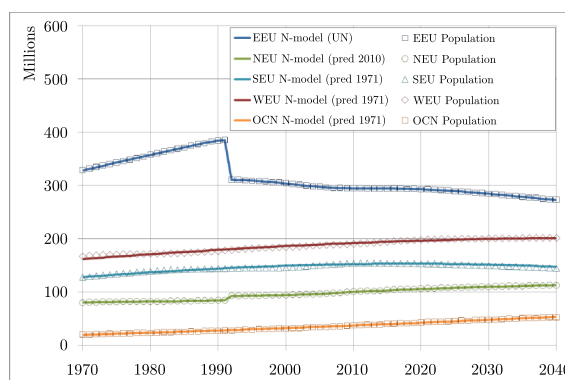


Рис. 12. Динамика численности населения европейских регионов и Океании в млн чел.

Наблюдается хорошее совпадение вычислений с прогнозом ООН, что неудивительно, так как параметры модели подбирались именно под эти статистические данные. К тому же, за редким исключением, демографические величины ведут себя непрерывно и монотонно, что идеально подходит для регрессионной модели.

На рис. 13–16 приведены графики ВВП с 1970 по 2040 г.: исторические данные и модельные расчеты. Рис. 13 отражает экономику регионов Африки, рис. 14 – Америки, рис. 15 – Азии,

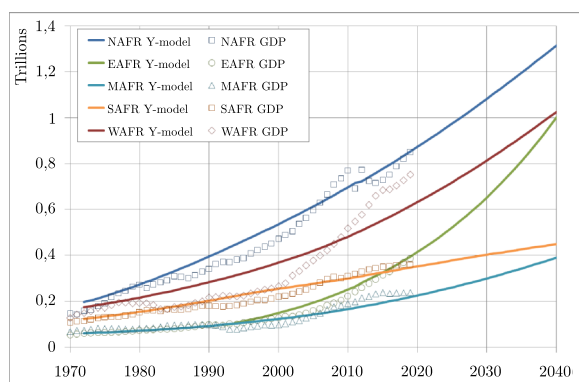


Рис. 13. Динамика ВВП регионов Африки в трлн долл. 2015 г.

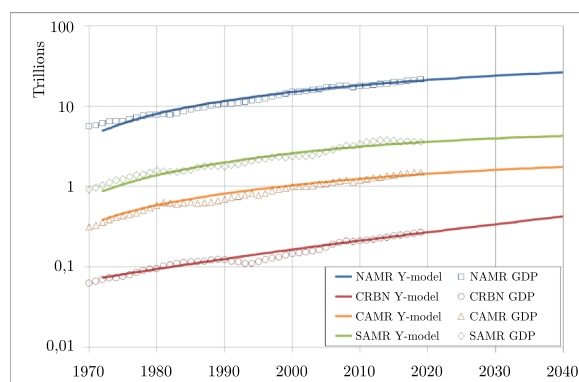


Рис. 14. Динамика ВВП регионов Америки в трлн долл. 2015 г. Масштаб логарифмический

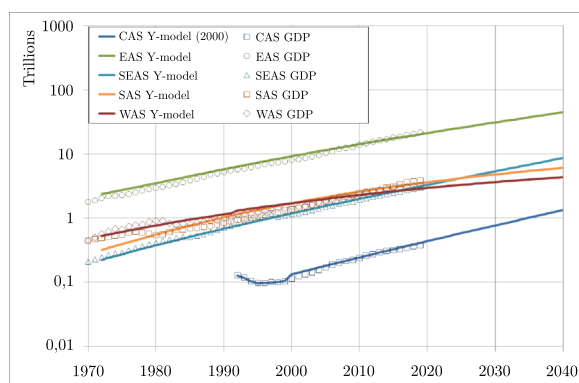


Рис. 15. Динамика ВВП регионов Азии в трлн долл. 2015 г. Масштаб логарифмический

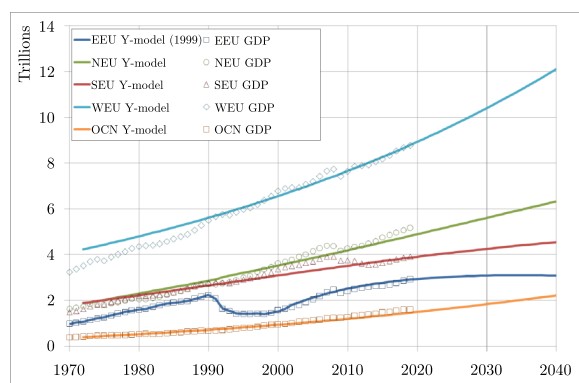


Рис. 16. Динамика ВВП европейских регионов и Океании в трлн долл. 2015 г.

рис. 16 — Европы и Океании. В отличие от демографии макроэкономика региона хуже описывается предложенной моделью, что заметно в регионах Африканского континента (рис. 13). Более или менее «прилично» ведут себя Центральная Африка и Южная Африка, в прочих регионах расхождение с моделью сильнее. В Америке и Азии расхождение не столь заметно, однако картину «замазывает» логарифмический масштаб (рис. 14, 15). В Европе и Океании наблюдается неплохое соответствие между данными и расчетом (рис. 16).

Также были подсчитаны глобальные величины (как сумма региональных агрегатов). На рис. 17 представлена численность мирового населения, на рис. 18 — мировой ВВП.

Визуально между эмпирикой и расчетом неплохое соответствие, что для демографии неудивительно, учитывая результаты региональных моделей, для макроэкономики, видимо, сработал эффект «больших чисел». Основной вклад внесли регионы с большими значениями ВВП: Северная Америка, Восточная Азия, Западная Европа; к тому же при суммировании отклонения от расчетных линий частично компенсировались, что снизило относительный разброс.

Вычисления двусторонних торговых потоков по уравнению (7b) дали ожидаемые результаты: в тех случаях, где верификация была успешной, расчеты более или менее близки к статистике; там же, где в ходе проверки возникли трудности, численный эксперимент продемонстрировал несоответствие между моделью и данными.

Лучше всего множественная лог-модель подошла для экспортных потоков десяти регионов: четырех европейских, четырех азиатских (за исключением Средней Азии), Северной Аме-

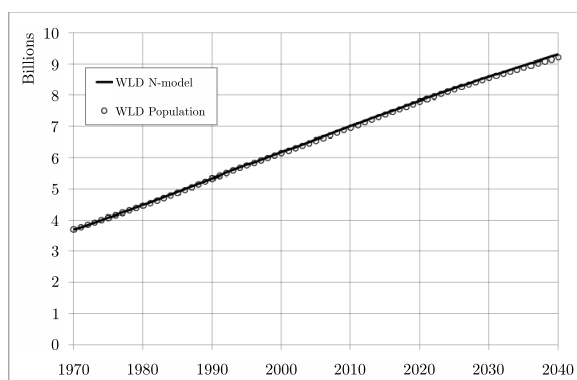


Рис. 17. Динамика мирового народонаселения в млрд чел.

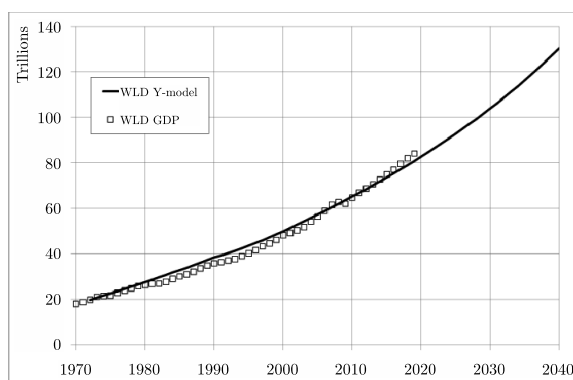


Рис. 18. Динамика мирового ВВП в трлн долл. 2015 г.

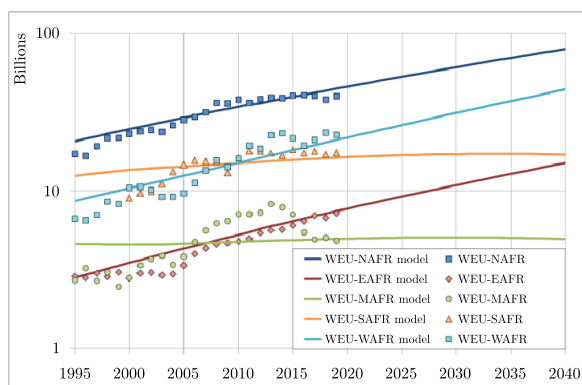


Рис. 19. Двусторонние товарные потоки из Западной Европы в регионы Африки в долл. 2015 г. в логарифмическом масштабе

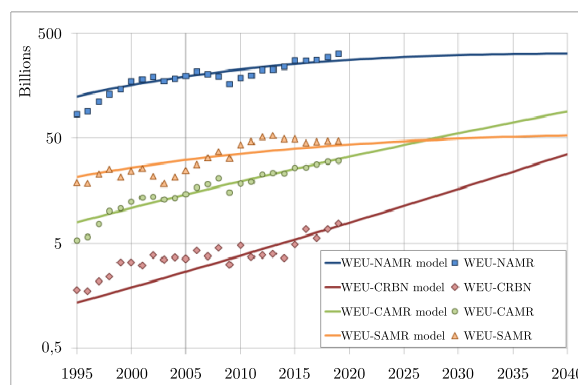


Рис. 20. Двусторонние товарные потоки из Западной Европы в регионы Америки в долл. 2015 г. в логарифмическом масштабе

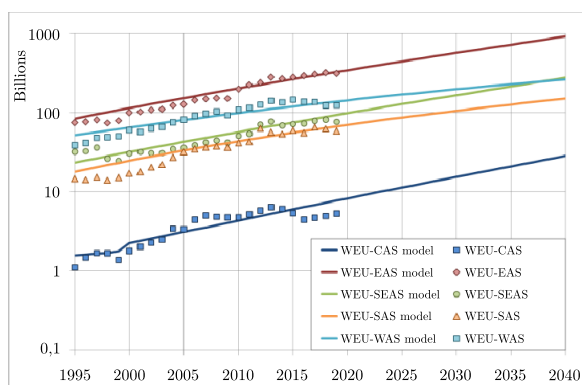


Рис. 21. Двусторонние товарные потоки из Западной Европы в регионы Азии в долл. 2015 г. в логарифмическом масштабе

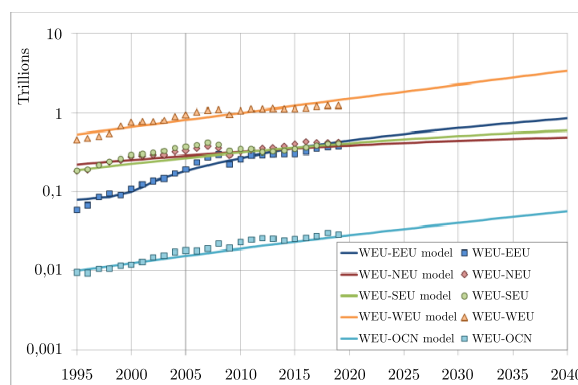


Рис. 22. Двусторонние товарные потоки из Западной Европы в Европу и Океанию в долл. 2015 г. в логарифмическом масштабе

рики и Восточной Африки (см. приложение 2). В качестве образца на рис. 19–22 приведены все экспортные потоки Западной Европы.

Похожая ситуация с вычисленными по формулам (8), (9) региональным товарным экспортом и региональным товарным импортом: чем точнее смоделированы слагаемые, тем ближе

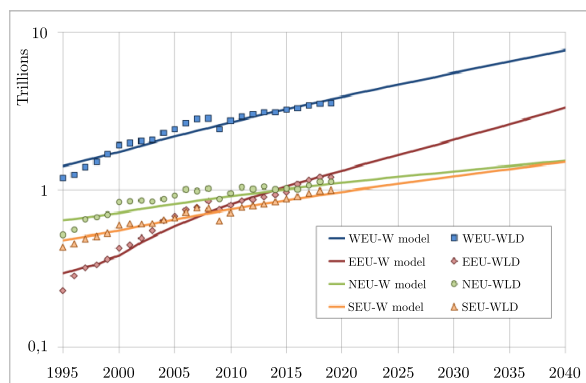


Рис. 23. Суммарный товарный экспорт регионов Европы в логарифмическом масштабе

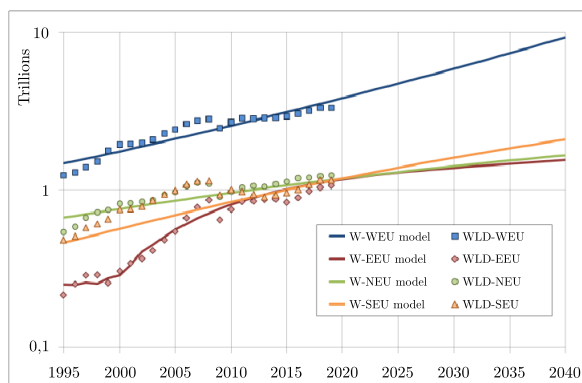


Рис. 24. Суммарный товарный импорт регионов Европы в логарифмическом масштабе

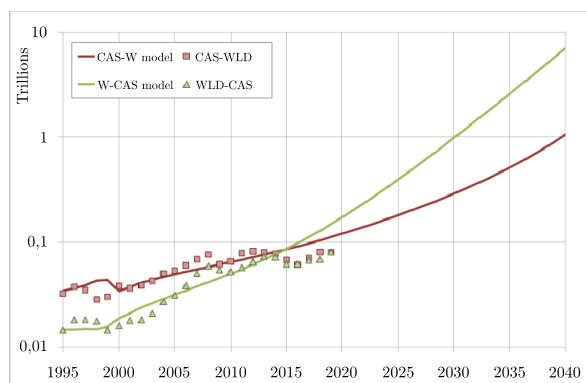


Рис. 25. Суммарный товарный экспорт-импорт Средней Азии в логарифмическом масштабе

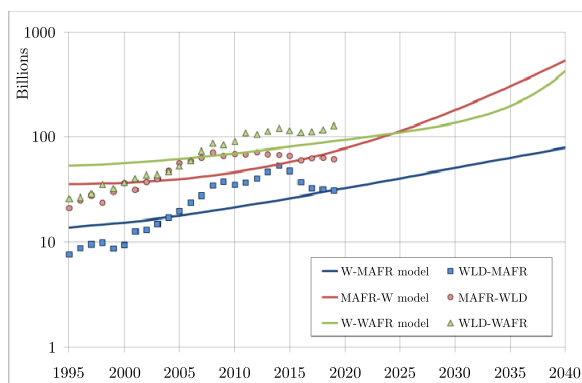


Рис. 26. Суммарный товарный экспорт-импорт Центральной Африки и импорт Западной Африки

общая сумма к расчетной. Рис. 23, 24 иллюстрируют этот факт. На рис. 23 показана динамика суммарного товарного экспорта европейских регионов, на рис. 24 — их товарного импорта.

Региональный товарный экспорт-импорт позволяет оценить интегральную степень расхождения торговой модели с данными. Даже если отдельные двусторонние потоки плохо аппроксимированы, в силу пресловутого закона больших чисел их сумма может быть хорошо подогнана регрессией. Оказалось, что в большинстве случаев так и есть. Исключения: экспорт и импорт Средней Азии (рис. 25), экспорт и импорт Центральной Африки, импорт Западной Африки (рис. 26).

Рис. 25 иллюстрирует ситуацию вычислительной неустойчивости: расчетные линии пошли «вразнос», разбежавшись друг от друга почти в 10 раз. Расчетные значения явно завышены: экспорт на протяжении 25 лет, с 1995 по 2019 г., вырос с 34 до 112 млрд долл., т. е. чуть больше чем в 3 раза, а за 21 год, с 2020 по 2040 г., — со 112 млрд до 1 трлн долл., т. е. почти в 10 раз; импорт с 1995 по 2019 г. вырос с 15 до 150 млрд долл., т. е. в 10 раз, а за расчетный период — со 150 млрд до 7 трлн долл., т. е. почти в 50 раз. Абсурдность этого результата очевидна.

На рис. 26 неустойчивость также имеет место для суммарного экспорта Центральной Африки и суммарного импорта Западной Африки. Второй случай примечателен тем, что расчетная кривая после примерного экспоненциального роста в 2035 г. неожиданно меняет скорость и начинает возрастать сверхэкспоненциально. Третья величина — суммарный импорт Центральной Африки — имеет большую дисперсию остатков, что говорит скорее о неверном выборе регрессионной зависимости, чем об артефакте численного эксперимента.

Причина неустойчивости заключается в неверной спецификации отдельных двусторонних потоков, повлиявших на итоговую сумму. Конкретно для Средней Азии проблемными являются экспортные потоки в Западную Африку и Западную Азию (рис. 27). Показателен поток в Западную Африку: невооруженным глазом видно, что тренда нет, а расчетная величина возрастает по экспоненте. Ошибка вызвана эффектом *мультиколлинеарности*, т. е., по сути, избыточностью множественной модели, парная регрессия лучше аппроксимировала бы данные.

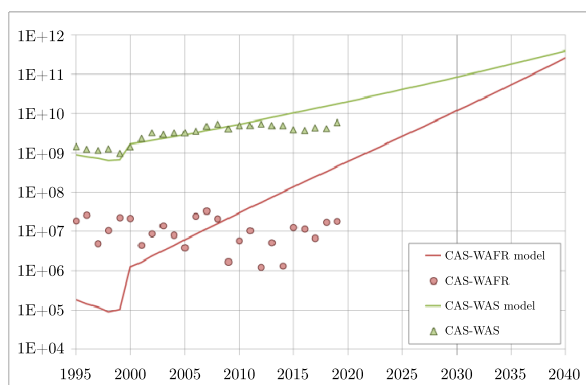


Рис. 27. Товарный экспорт из Средней Азии в Западную Африку (CAS-WAFR) и в Западную Азию (CAS-WAS). Логарифмический масштаб

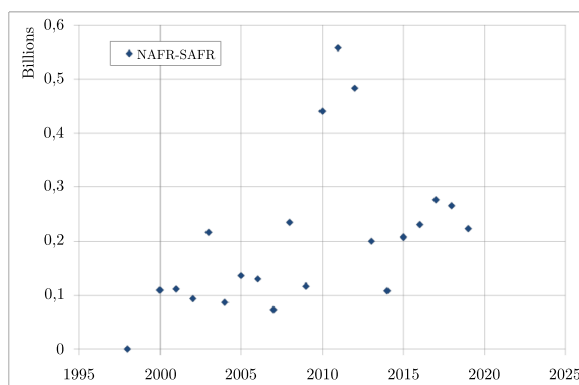


Рис. 28. Стационарный ряд — товарный экспорт из Северной Африки в Южную Африку, млрд долл. 2015 г.

Плохую подгонку экспорта из Средней Азии в Западную Азию отчасти можно приписать мультиколлинеарности, отчасти — неподходящему виду регрессионной зависимости. Возможно, перспективнее использовать S-образную функцию, наподобие логистической кривой.

Сравнительный анализ позволяет выделить три типа зависимостей, описывающих динамику двусторонних потоков, не считая возрастающего тренда.

1. Стационарный ряд: отсутствие линии тренда (либо очень слабый тренд), графически множество данных похоже на «облако». Хорошо аппроксимируется константой — средним значением или близким к нему числом. Пример — на рис. 28.

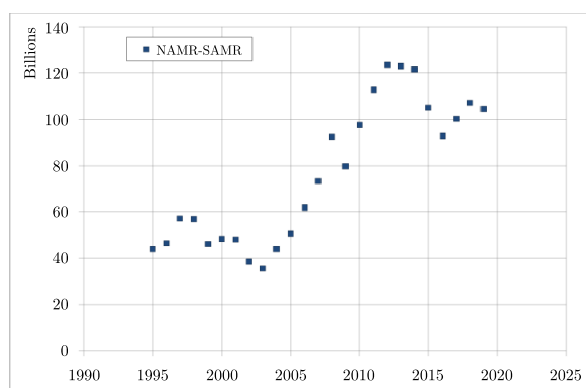


Рис. 29. S-образная зависимость: товарный поток из Северной Америки в Южную Америку

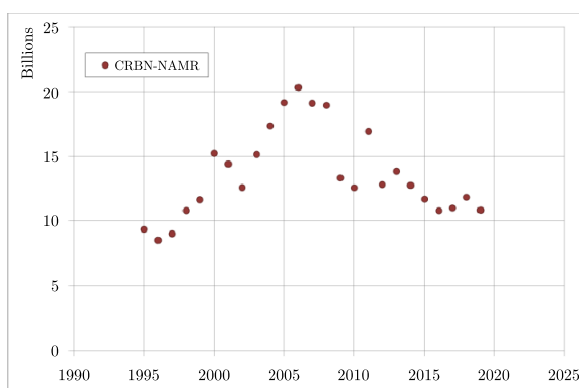


Рис. 30. Одиночная волна: товарный экспорт Карибского региона в Северную Америку

2. S-образная зависимость: колебания в начале и в конце промежутка, характерный подъем в середине участка, как на рис. 29. В отдельных случаях можно подогнать константой, отбросив начало и середину интервала.

3. Волновая динамика: довольно отчетливо виден контур одинокой волны (сначала подъем, затем падение); характерный образец приведен на рис. 30. Самый сложный случай для прогнозирования: даже точная аппроксимация не позволяет с уверенностью сказать, как поведет себя величина в будущем. Может быть, наблюдаемый волновой фронт является случайным всплеском, возможно, это часть периодических колебаний. К сожалению, сделать выводы по одиночной волне невозможно, необходимо хотя бы два-три полноценных цикла. Соответственно, простейший метод — считать аппроксимируемую величину константой, взяв среднее значение либо на всем интервале, либо на конечном участке. Другой способ — повторить профиль волны, приложив начальную точку к конечной. Вариативность удлинняет расчеты и усложняет представление результатов, взамен появляются альтернативность прогноза и дополнительная возможность оценить его точность.

Можно дать более формальное обоснование небольшого числа типов зависимостей торговых потоков от времени, используя методы снижения размерности пространства исследуемых признаков. Относительно простым в использовании является метод главных компонент, позволяющий построить в выборочном пространстве систему некоррелированных рядов — главных компонент — на основе ортонормированных линейных комбинаций исходных признаков [Айвазян и др., 1989]. Также метод дает возможность оценить суммарную дисперсию главных компонент, т. е. долю объясненного разброса при использовании первых k компонент и отбрасывании остальных. Иными словами, можно оценить точность «натягивания» выборочного пространства на первые k главных компонент, заменяя все пространство k -мерной проекцией.

Результаты применения метода главных компонент к торговым потокам и рядам ВВП приведены в приложении 2. Ряды региональных ВВП описываются первой главной компонентой с точностью 0,954, а региональные торговые потоки — тремя главными компонентами с точностью 0,802 и семью компонентами с точностью 0,907.

Анализируя первые семь компонент, можно сделать следующие выводы. Первая главная компонента g_1 возрастает со временем, описывая трендовую составляющую рядов взаимной торговли. Вторая компонента g_2 состоит из двух частей: нисходящий тренд до 2008 г. и восходящий тренд с выходом на константу к 2016 г. Третья компонента g_3 представляет собой фактически одиночную положительную волну, начинающуюся в 2008 г. и заканчивающуюся в 2016 г., вне этой волны имеет место слабый нисходящий тренд. Четвертую компоненту можно описать как колебание из двух волн: положительной волны с 2003 по 2011 г. и отрицательной волны с 2012 по 2017 г. Главные компоненты g_5, g_6, g_7 в той или иной степени можно считать циклическими с нечетко выраженными периодами от 5 до 9 лет. У следующих главных компонент колебания постепенно становятся менее регулярными и более похожими на случайные всплески стационарного случайного процесса.

Сравнивая графики главных компонент с тремя вышеприведенными типами немонотонной динамики двусторонних потоков, заключаем следующее:

- а) линейная комбинация компонент g_1 и g_2 может дать S-образную зависимость;
- б) линейная комбинация компонент g_3 и g_4 описывают ситуацию одиночной волны;
- в) компоненты g_5, g_6, g_7 и все последующие дают циклическую динамику, которую с некоторой точностью можно считать белым шумом.

Заключение

Настоящая работа, с одной стороны, подводит промежуточный итог, с другой — прокладывает путь новым исследованиям в области моделирования региональной и мировой динамики.

Построена динамическая макро модель мировой динамики. В модели мир разбит на 19 географических зон — регионов, каждый регион представляет собой формальное объединение соседних государств. Внутреннее развитие регионов описывается уравнениями разностного типа для демографических и экономических индикаторов. Межрегиональные взаимодействия представляют собой торговые связи, подмодель торговли описывает перетекание валовой продукции от региона к региону. Проведенные расчеты показывают, что модель удовлетворительно описывает динамику монотонно меняющихся показателей. Наилучшее соответствие наблюдается у демографических показателей, наихудшее — у индикаторов внешней торговли.

У ряда двусторонних торговых потоков обнаруживаются выраженные регулярные колебания циклической природы, однако вид и спецификация регрессионных уравнений не позволяют эти колебания выделить и учесть в рамках трендовой составляющей. Кроме того, несмотря на верификацию восьми вариантов торгового уравнения, сравнительный анализ этих разновидностей носит несколько поверхностный характер. Также минусом является отсутствие погрешностей прогнозных значений.

Планы будущих исследований:

- провести более детальное сравнение торговых уравнений при помощи статистических тестов;
- учесть колебания посредством включения циклической составляющей в регрессионные уравнения;
- оценить точность прогноза.

Список литературы (References)

- Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика. Классификация и снижение размерности: справочное издание / под ред. С. А. Айвазяна. — М.: Финансы и статистика, 1989.
- Aivazyan S. A., Buchshaber V. M., Yenyukov I. S., Meshalkin L. D. Prikladnaya statistika. Klassifikatsiya i snizhenie razmernosti: spravochnoe izdanie [Applied statistics: classification and reduction of dimensionality. Reference edition] / ed. by prof. S. A. Aivazyan. — Moscow: Finansy i statistika, 1989 (in Russian).
- Бокс Дж., Дженкинс Г. Анализ временных рядов, прогноз и управление. — М.: Мир, 1974. — Кн. 1.
- Box G., Jenkins G. Time series analysis; forecasting and control. — San Francisco: Holden-Day, 1970. (Russ. ed.: Boks Dzh., Dzhenkins G. Analiz vremennyh ryadov, prognoz i upravlenie. — Moscow: Mir, 1974. — Kn. 1.)
- Вербик М. Путеводитель по современной эконометрии. — М.: Научная книга, 2008.
- Verbeek M. A Guide to modern econometrics. — Chichester: John Wiley and Sons, LTD, 2000. (Russ. ed.: Verbiik M. Putevoditel' po sovremennoi ekonometrii. — Moscow: Nauchnaya kniga, 2008.)
- Вишневецкий А. Г. Избранные демографические труды. 1: Демографическая теория и демографическая история. — М.: Наука, 2005.
- Vishnevskii A. G. Izbrannye demograficheskie trudy. 1: Demograficheskaya teoriya i demograficheskaya istoriya [Selected demographic works. 1: Demographic theory and demographic history]. — Moscow: Nauka, 2005 (in Russian).
- Дубовский С. В. Моделирование циклов Кондратьева и прогнозирование кризисов // Кондратьевские волны: аспекты и перспективы. — Волгоград: Учитель, 2012. — С. 179–188.
- Dubovskii S. V. Modelirovanie ciklov Kondrat'eva i prognozirovanie krizisov [Modeling of Kondratiev cycles and forecasting of crises] // Kondrat'evskie volny: aspekty i perspektivy [Kondratiev waves: aspects and prospects]. — Volgograd: Uchitel, 2012. — P. 179–188 (in Russian).
- Дубовский С. В. Прогнозирование российского экономического роста и финансовой динамики в условиях глобализации и нестабильности // Общественные науки и современность. — 2005. — № 3. — С. 129–136.
- Dubovskii S. V. Prognozirovanie rossiiskogo ekonomicheskogo rosta i finansovoi dinamiki v usloviiakh globalizatsii i nestabilnosti [Forecasting Russian economic growth and financial dynamics in the context of globalization and instability] // Obshchestvennye nauki i sovremennost. — 2005. — No. 3. — P. 129–136 (in Russian).

- Катица С. П.* Сколько людей жило, живет и будет жить на Земле. Очерк теории роста человечества. — М.: Международная программа образования, 1999.
Kapica S. P. Skol'ko lyudei zhilo, zhivet i budet zhit' na Zemle. Ocherk teorii rosta chelovechestva [How many people have lived, live and will live on Earth. An essay on the human growth theory]. — Moscow: Mezhdunarodnaya programma obrazovaniya, 1999 (in Russian).
- Махов С. А.* Динамическая модель внешней торговли стран БРИКС. — М.: Препринты ИПМ им. М. В. Келдыша РАН, 2016. — № 128.
Makhov S. A. Dinamicheskaya model' vneshnej trgovli stran BRIKS [BRICS trade dynamic model]. — Moscow: RAS Keldysh Institute of Applied Mathematics, 2016. — Preprint No. 128 (in Russian).
- Махов С. А.* Динамическая модель стран БРИКС с учетом торговли. — М.: Препринты ИПМ им. М. В. Келдыша РАН, 2017. — № 139.
Makhov S. A. Dinamicheskaya model' stran BRIKS s uchetom trgovli [A dynamic macromodel of the BRICS countries taking into account trade]. — Moscow: RAS Keldysh Institute of Applied Mathematics, 2017. — Preprint No. 139 (in Russian).
- Махов С. А.* Долгосрочная макро модель мировой динамики на основе эмпирических данных // Компьютерные исследования и моделирование. — 2013. — Т. 5, № 5. — С. 883–891.
Makhov S. A. Dolgosrochnaya makromodel' mirovoj dinamiki na osnove empiricheskikh dannyh [The long-term empirical macro model of world dynamics] // Computer Research and Modeling. — 2013. — Vol. 5, No. 5. — P. 883–891 (in Russian).
- Пирогов Г. Г., Богородицкий В. К., Бороздин И. И., Татарникова М. В., Федоровский Ю. П.* Моделирование международных внешнеэкономических связей в моделях глобального развития. — М.: Препринт ВНИИСИ, 1980.
Pirogov G. G., Bogoroditskii V. K., Borozdin I. I., Tatarnikova M. V., Fedorovskii Yu. P. Modelirovanie mezhdunarodnykh vneshneekonomicheskikh svyazei v modeliakh globalnogo razvitiia [Modeling of international foreign economic relations in world dynamics models]. — Moscow: VNIISI, 1980. — Preprint (in Russian).
- UNCTADstat Data Center online. — 2020. — [Electronic resource]. — URL: http://unctadstat.unctad.org/wds/ReportFolders/reportFolders.aspx?sCS_ChosenLang=en (accessed: 18.12.2022).
- United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects 2019: Methodology of the United Nations Population Estimates and Projections (ST/ESA/SER.A/425). — New York: United Nations, 2019. — [Electronic resource]. — URL: https://population.un.org/wpp/Publications/Files/WPP2019_Methodology.pdf (accessed: 18.12.2022).
- United Nations Statistic Division. National accounts main aggregates database. — 2020. — [Electronic resource]. — URL: <https://unstats.un.org/unsd/snaama/Downloads> (accessed: 18.12.2022).
- World Bank. World development indicators online. — 2020. — [Electronic resource]. — URL: https://databank.worldbank.org/data/download/WDI_excel.zip (accessed: 18.12.2022).

Приложение 1

В этом разделе приведены результаты сравнения торговых уравнений (7а), (7b), (7с), (7d), (7е), (7f) и эталонной зависимости (7g), (7h) по коэффициенту детерминации. Используемые краткие названия:

- Etalon-1 — линейная эталонная модель (7g);
- Etalon-2 — лог-линейная (экспоненциальная) эталонная модель (7h);
- Multi-1 — множественная линейная модель (7а);
- Multi-2 — множественная лог-линейная (степенная) модель (7b);
- Export-1 — линейная экспортная функция (7с);
- Export-2 — лог-линейная (степенная) экспортная функция (7d);

- Import-1 — линейная импортная функция (7e);
- Import-2 — лог-линейная (степенная) импортная функция (7f).

1. Грубая прикидка

В табл. 2 приведено 3-интервальное распределение регрессионных моделей по коэффициенту детерминации R^2 .

Таблица 2. 3-интервальное распределение регрессионных моделей по R^2

Модели	Etalon-1	Etalon-2	Multi-1	Multi-2	Export-1	Export-2	Import-1	Import-2
Диапазон	Количество моделей, попавших в диапазон							
0,0–0,3	77	64	41	33	147	129	158	144
0,3–0,7	95	99	85	78	114	116	112	122
0,7–1,0	189	198	235	250	100	116	91	95

2. Гистограммный метод

На рис. 31 и рис. 32 изображены гистограммы распределений восьми моделей по коэффициенту R^2 с шагом 0,1. Ради экономии места и удобства сравнения все линейные модели показаны на рис. 31, все нелинейные — на рис. 32. Конкретные частоты, на основании которых построены гистограммы, даются в табл. 3.

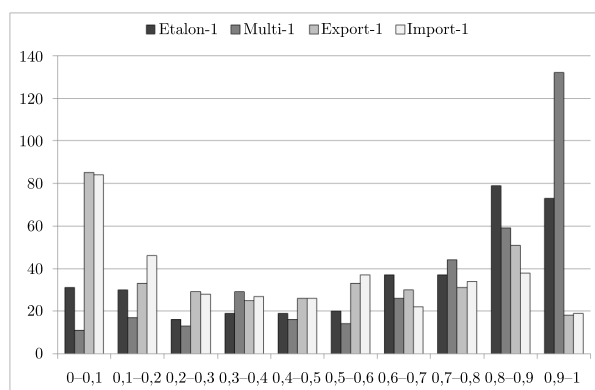


Рис. 31. Гистограмма распределения линейных моделей (7g), (7a), (7c), (7e)

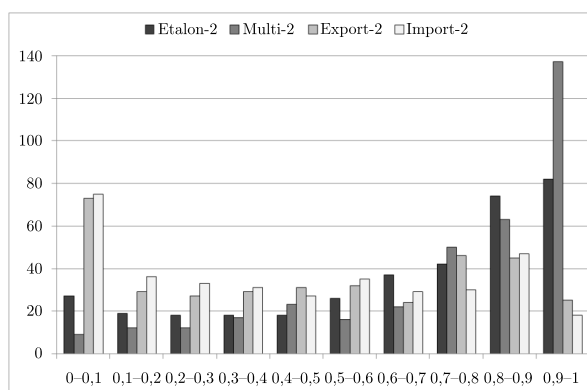


Рис. 32. Гистограмма распределения нелинейных моделей (7h), (7b), (7d), (7f)

3. Непрерывный способ

В силу желания уместить все кривые на двух графиках, сильного несовпадения значений коэффициента R^2 и одинакового количества регрессий для каждого уравнения, проще использовать ранговые зависимости, чем эмпирические функции распределения. На рис. 33 приведены графики для линейных моделей, на рис. 34 — для нелинейных моделей.

По оси ординат отложены значения R^2 , по оси абсцисс — ранг, при этом ранг 1 присвоен минимальному значению, ранг 361 — максимальному. Ранг-размерная зависимость показывает зависимость исследуемой величины от накопленной частоты, по сути — обратную эмпирическую функцию распределения. Чтобы обратная функция получилась в чистом виде, следует перейти от абсолютных частот к относительным, разделив ранг на 361. Пологость конечного участка графика говорит о количестве примеров удачной аппроксимации: чем более полог график, тем выше качество модели.

Таблица 3. 10-интервальное распределение регрессионных моделей по R^2

Модели	Etalon-1	Etalon-2	Multi-1	Multi-2	Export-1	Export-2	Import-1	Import-2
Диапазон	Количество моделей, попавших в диапазон							
0,0–0,1	31	27	11	9	85	73	84	75
0,1–0,2	30	19	17	12	33	29	46	36
0,2–0,3	16	18	13	12	29	27	28	33
0,3–0,4	19	18	29	17	25	29	27	31
0,4–0,5	19	18	16	23	26	31	26	27
0,5–0,6	20	26	14	16	33	32	37	35
0,6–0,7	37	37	26	22	30	24	22	29
0,7–0,8	37	42	44	50	31	46	34	30
0,8–0,9	79	74	59	63	51	45	38	47
0,9–1,0	73	82	132	137	18	25	19	18

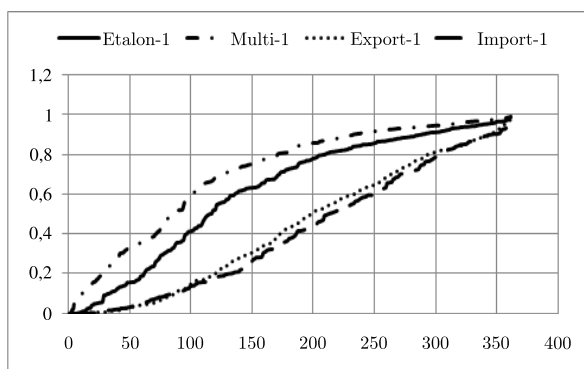


Рис. 33. Ранг-размерная зависимость для моделей (7g), (7a), (7c), (7e)

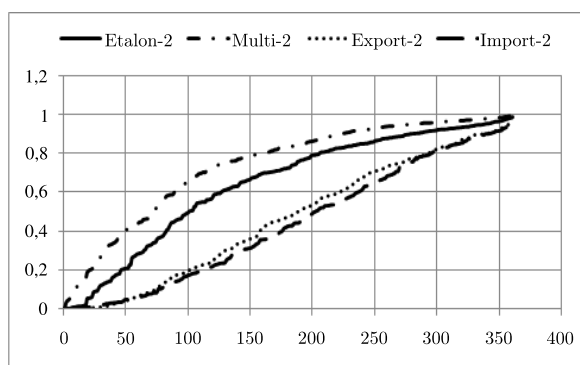


Рис. 34. Ранг-размерная зависимость для моделей (7h), (7b), (7d), (7f)

Приложение 2

Здесь приведены собственные значения корреляционных матриц региональных ВВП и региональных торговых потоков, а также графики нескольких первых главных компонент.

Метод главных компонент использует либо ковариационную, либо корреляционную матрицу исследуемых признаков. В первом случае величины должны быть отцентрированы (средние равны нулю), во втором случае — нормализованы (нулевые средние, единичные дисперсии). Для ковариационной (корреляционной) матрицы вычисляются собственные значения λ_k и собственные векторы u_k , координаты последних задают коэффициенты линейных комбинаций центрированных (нормализованных) признаков в главные компоненты g_k по формулам $g_k = u_k^1 \cdot z_1 + u_k^2 \cdot z_2 + \dots + u_k^n \cdot z_n$, u_k^i — координаты собственного вектора u_k , z_i — центрированные (нормализованные) признаки, n — размер ковариационной (корреляционной) матрицы. Ввиду того что ковариации и дисперсии как региональных ВВП, так торговых потоков численно велики (порядка 10^{24} долл. 2015 г.), удобнее оказалось использовать корреляционную матрицу.

1. Главные компоненты для региональных ВВП

Корреляционная матрица региональных ВВП RY имеет размеры 19×19 , поэтому, хотя выборочное пространство 25-мерное (ежегодные данные за 25 лет с 1995 по 2019 г.), ранг матрицы RY не превышает 19. Критерий информативности первых k главных компонент, показывающий их относительную долю в суммарной дисперсии, рассчитывается по формуле $G_k = \frac{\lambda_1 + \dots + \lambda_k}{\lambda_1 + \dots + \lambda_n}$, $\lambda_1, \dots, \lambda_n$ — собственные значения матрицы RY в порядке убывания. Очевидно, что сумма всех

собственных значений равна 19, т. е. $G_k = \frac{\lambda_1 + \dots + \lambda_k}{19}$. В табл. 4 приведены собственные значения и критерий информативности с точностью до 10^{-4} .

Таблица 4. Собственные значения матрицы RU и критерий информативности 19 главных компонент

k	1	2	3	4	5	6	7	8	9	
λ_k	18,12	0,605	0,193	0,045	$1,51 \times 10^{-2}$	$9,15 \times 10^{-3}$	$3,93 \times 10^{-3}$	$3,65 \times 10^{-3}$	$1,98 \times 10^{-3}$	
G_k	0,9535	0,9853	0,9955	0,9980	0,9988	0,9993	0,9995	0,9997	0,9998	
k	11	12	13	14	15	16	17	18	19	
λ_k	$1,60 \times 10^{-3}$	$8,84 \times 10^{-4}$	$5,62 \times 10^{-4}$	$2,82 \times 10^{-4}$	$1,94 \times 10^{-4}$	$1,24 \times 10^{-4}$	$1,07 \times 10^{-4}$	$2,47 \times 10^{-5}$	$1,71 \times 10^{-5}$	$3,00 \times 10^{-6}$
G_k	0,9999	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1	

2. Главные компоненты для торговых потоков

Корреляционная матрица торговых потоков RT имеет размеры 361×361 , ранг не превышает размерности выборочного пространства. Как отмечалось ранее, потоки имеют лакуны в 1995–1999 гг., из-за чего результаты получаются вычислительно неустойчивы. По этой причине расчеты проведены на интервале 2000–2019 гг., т. е. ранг матрицы RT не превышает 20. Таким образом, у матрицы RT может быть максимум 20 ненулевых (т. е. положительных) собственных значений, в сумме дающих 361.

В табл. 5 приведены собственные значения и критерий информативности 20 главных компонент с точностью до 10^{-4} . Ввиду погрешности вычислительного алгоритма $G_{20} < 1$. Графики первых семи главных компонент показаны на рис. 35.

Таблица 5. Собственные значения матрицы RT и критерий информативности 20 главных компонент

k	1	2	3	4	5	6	7	8	9	10
λ_k	209,9565	47,2331	32,4538	15,0965	10,4513	6,4716	5,9259	5,1194	4,6699	3,4799
G_k	0,5816	0,7124	0,8023	0,8442	0,8731	0,8910	0,9074	0,9216	0,9346	0,9442
k	11	12	13	14	15	16	17	18	19	20
λ_k	3,0285	2,8947	2,5916	2,5242	2,0266	1,8877	1,4455	1,3530	1,1380	0,9729
G_k	0,9526	0,9606	0,9678	0,9748	0,9804	0,9856	0,9896	0,9934	0,9965	0,9992

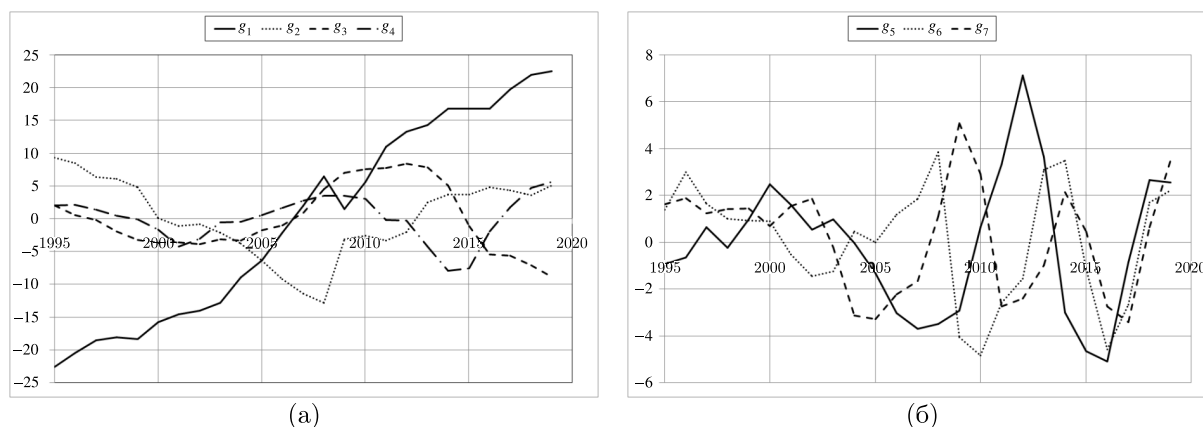


Рис. 35. Первые семь главных компонент для торговых потоков: (а) g_1, g_2, g_3, g_4 ; (б) g_5, g_6, g_7