

УДК 330.55, 311.2, 004.6

DOI: 10.17323/1813-8691-2025-29-2-279-305

## Восстановление истории публикаций российского ВВП и его компонентов<sup>1</sup>

**Горностаев Д.А., Маханькова Н.А., Милютин П.П.,  
Пономаренко А.А., Селезнев С.М.**

Винтажные данные, или ретроспективные, играют важнейшую роль в оценке точности моделей прогнозирования макроэкономических показателей и принимаемых на основе этих моделей решений. Эти данные представляют собой срез информации, доступной на конкретный момент времени в прошлом, и отражают то, как тогда воспринималась экономическая ситуация. Важно понимать, что на основе именно винтажных данных, а не пересмотренных и уточненных показателей, строились прогнозы и принимались решения по вопросам экономической политики. В качестве примера мы показываем, что расхождение траекторий темпов роста ВВП на прогнозном периоде может достигать одного процентного пункта.

В этой работе представлен набор винтажных данных по российскому ВВП и его компонентам по использованию и производству. Датасет состоит из пересмотров номинальных и реальных квартальных показателей в период с декабря 2005 г. по настоящее время. В дополнение к данным в работе описываются характеристики пересмотров реальных и номинальных показате-

---

<sup>1</sup> Настоящая статья отражает личную позицию авторов. Содержание и результаты данного исследования не следует рассматривать, в том числе цитировать в каких-либо изданиях как официальную позицию Банка России или указание на официальную политику или решения регулятора. Любые ошибки в данном материале являются исключительно авторскими.

Авторы выражают благодарность за помощь в формировании набора данных В. Алексеенко, Д. Веневцеву, Н. Горевому, М. Жемкову, М. Иванову, М. Лымарь и Т. Сытиной.

**Горностаев Дмитрий Александрович** – главный экономист Департамента исследований и прогнозирования, Банк России. E-mail: gornostaevda@cbr.ru

**Маханькова Наталья Александровна** – главный экономист Департамента исследований и прогнозирования, Банк России. E-mail: makhankovana01@cbr.ru

**Милютин Петр Павлович** – заместитель начальника отдела Департамента исследований и прогнозирования, Банк России. E-mail: milyutinpp01@cbr.ru

**Пономаренко Алексей Алексеевич** – д.э.н., PhD, начальник Управления экономических исследований Департамента исследований и прогнозирования, Банк России. E-mail: pomomarenkoaa@cbr.ru

**Селезнев Сергей Михайлович** – начальник центра исследования данных и машинного обучения Департамента исследований и прогнозирования, Банк России. E-mail: seleznevsm@cbr.ru

Статья поступила: 13.03.2025/Статья принята: 23.04.2025.

лей ВВП (число, частота, магнитуа) и его компонентов по использованию, а также приводится методология сбора исторических показателей с использованием сервиса Wayback Machine, которая позволяет собирать данные даже в отсутствие истории их релизов.

Мы надеемся, что этот набор данных станет ценным инструментом для исследователей и аналитиков, помогающим лучше понять, как оценивалась экономическая ситуация в прошлом. Он также может стать основой для проведения более корректных экспериментов с моделями в псевдореальном времени.

**Ключевые слова:** ВВП; винтажи данных; пересмотры данных; частота и магнитуа пересмотров; Россия; исторические копии сайтов; Wayback Machine.

**Для цитирования:** Горностаев Д.А., Маханькова Н.А., Милютин П.П., Пономаренко А.А., Селезнев С.М. Восстановление истории публикаций российского ВВП и его компонентов. *Экономический журнал ВШЭ*. 2025; 29(2): 279–305.

**For citation:** Gornostaev D.A., Makhankova N.A., Milyutin P.P., Ponomarenko A.A., Seleznev S.M. Reconstructing the Publication History of Russia's GDP and its Components. *HSE Economic Journal*. 2025; 29(2): 279–305. (In Russ.)

## 1. Введение

Винтажные (ретроспективные) данные являются одним из ключевых ингредиентов для оценки качества моделей прогнозирования макроэкономических показателей и, как следствие, принимаемых на их основе решений. Эти данные представляют собой срез информации, доступной в определенный момент, и отражают то, как воспринималась экономическая ситуация на конкретную дату в прошлом<sup>2</sup>. Именно с использованием винтажных данных, а не пересмотренных/уточненных<sup>3</sup>, делались прогнозы и принимались решения относительно экономической политики.

Чтобы проиллюстрировать важность винтажных данных и их влияние на траекторию прогноза, на рис. 1 показан прогноз ВВП на 12 кварталов для II квартала 2020 г., полученный с использованием простой сезонной авторегрессионной модели (SARIMA<sup>4</sup>) на винтажных данных [Горностаев и др., 2021] и на данных, которые были пересмотрены. Прог-

<sup>2</sup> В этой работе под винтажем понимается именно набор информации, доступной на конкретную дату, а не набор пересмотров значений одного и того же периода.

<sup>3</sup> Пересмотры ранее опубликованных значений экономических показателей являются стандартным явлением в деятельности статистических органов и происходят из-за необходимости соблюдения баланса между оперативностью и точностью предоставляемых статистических показателей (подробнее см.: [Tebrake, 2019]). В большинстве случаев регулярные (запланированные) пересмотры связаны с появлением дополнительной информации, которая может быть использована для расчета статистического показателя, но не была доступна на момент предыдущей публикации этого показателя. Кроме этого, могут происходить разовые пересмотры в связи с методологическими изменениями в методе расчета статистического показателя.

<sup>4</sup> Была взята стандартная реализация SARIMA-модели из пакета statsmodels [Seabold, Perktold, 2010].

ноз был построен в уровнях и для лучшего восприятия переведен в годовые темпы прироста. Из рис. 1 видно, что расхождение траекторий на прогнозном периоде достигает одного процентного пункта.



Рис. 1. Прогноз ВВП (YoY) на основе SARIMA-модели, %

В отсутствие базы данных винтажей для основных макроэкономических показателей исследователям и аналитикам не остается ничего лучше, как использовать пересмотренные данные для оценки формальных метрик качества используемых моделей, однако, как следует из примера выше, это может приводить к некорректной оценке в ситуациях, когда расхождения в прогнозах, вызванные пересмотрами, велики. Во многих случаях метрики, оцененные на пересмотренных данных, лучше, чем на винтажных (см., например: [Mamedli, Shibitov, 2021]), что приводит к завышенным ожиданиям относительно точности модели. Чтобы лучше понимать качество данных и моделей, используемых при принятии решений, а также для стимулирования исследований в этом направлении, центральные банки и статистические агентства во многих странах формируют базы данных винтажных показателей (см. табл. П1 в Приложении).

В этой работе мы завершаем цикл сбора винтажей для основных российских макроэкономических показателей, начатый в работах [Горностаев и др., 2021; Gornostaev et al., 2022], и добавляем к ранее собранным данным по реальному ВВП и показателям, публикуемым Росстатом в сборнике «Краткосрочные экономические показатели Российской Федерации» (далее – КЭП), квартальные данные по номинальному и реальному ВВП и его компонентам по использованию и производству.

Основным вкладом работы является формирование и публикация<sup>5</sup> ретроспективных данных по ВВП и его компонентам. В дополнение к этому мы описываем методоло-

<sup>5</sup> С 2024 г. эти данные и данные из работ [Горностаев и др., 2021; Gornostaev et al., 2022] раз в полгода обновляются на [https://www.cbr.ru/ec\\_research/macroeconomic-vitages/](https://www.cbr.ru/ec_research/macroeconomic-vitages/)

гию сбора данных из открытых источников, которая позволила заполнить около 80% винтажей, и надеемся, будет полезна для исследователей, занимающихся похожими вопросами в других странах.

Дальнейшее изложение построено следующим образом. В разделе 2 описана методология сбора винтажных данных. Раздел 3 посвящен анализу свойств собранных показателей. В разделе 4 обсуждается место нашей работы в научной литературе. Раздел 5 представляет собой заключение.

## 2. Методология восстановления публикаций

Наша методология восстановления публикаций состоит из трех основных шагов, которые мы опишем в этом разделе. На первом этапе с использованием сервиса Wayback Machine собираются файлы<sup>6</sup> с необходимыми показателями с исторических копий сайта Росстата. Несмотря на высокую частоту сохранения копий сайта Росстата сервисом Wayback Machine, нередко они оказываются неполными и не содержат необходимых файлов, поэтому на втором шаге была разработана процедура дозаполнения исходя из информации о том, когда происходили пересмотры, а также дополнительной информации, извлекаемой из новостей о публикации или других открытых источников. На третьем этапе некоторые из данных, которые не удалось восстановить на предыдущих шагах, собирались из различного рода непубличных источников – например, с помощью опроса коллег, которые периодически работают с искомыми данными.

### 2.1. Сбор данных с использованием Wayback Machine

Проект Wayback Machine позволяет получать доступ к историческим копиям страниц интернет-сайтов, которые были собраны в рамках этого проекта в конкретный момент времени в прошлом. Чтобы посмотреть копию определенной страницы, необходимо ввести ее адрес в специальном поле, после чего появляется набор дат с доступными копиями (см. рис. 2). Переходя на страницу с определенной датой, можно увидеть ее копию. К примеру, копия главной страницы сайта Росстата на 10 марта 2012 г. показана на рис. 3. Стоит отметить, что при просмотре страницы, особенно при переходе по внутренним ссылкам сайтов, нужно всегда обращать внимание на дату в верхнем правом углу страницы ввиду того, что Wayback Machine может при нажатии ссылок перенаправлять на копию страницы с другой датой<sup>7</sup>.

Переходя по ссылкам в разделы, связанные с публикацией официальной статистики по ВВП<sup>8</sup>, мы вручную собрали все доступные данные по номинальному и реальному квартальному ВВП и его компонентам по производству и использованию.

---

<sup>6</sup> Приблизительно до конца 2009 г. данные публиковались не в файлах, а в формате htm-страниц.

<sup>7</sup> В некоторых случаях также происходит перенаправление на страницу с другим адресом, но тем не менее обычно это та страница, которая необходима.

<sup>8</sup> Например, в 2012 г. это был раздел «Национальные счета / Внутренний валовой продукт».

## 2.2. Дозаполнение данных

Как упоминалось выше, Wayback Machine содержит файлы с данными далеко не на все даты. Это обуславливается несколькими факторами, среди которых в том числе изменяющаяся структура сайтов, а также более редкое создание копий для более глубоких страниц в иерархии сайтов. Промежутки отсутствия копий страниц могут быть достаточно длинными и составлять до нескольких кварталов, а некоторые страницы вообще могут не иметь исторических копий. В связи с этим возникает необходимость поиска дополнительных методов восстановления исторических данных. В нашем случае достаточно полезной оказалась процедура ручного дозаполнения, которую можно разделить на два вида: дозаполнение слева и дозаполнение справа.

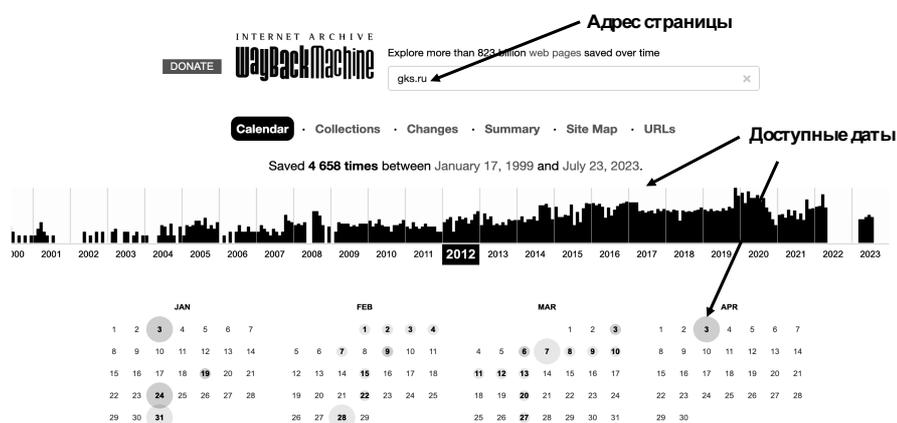


Рис. 2. Визуальное представление доступа к копиям сайта gks.ru с использованием сервиса Wayback Machine



Рис. 3. Копия главной страницы сайта Росстата на 10 марта 2012 г.

Дозаполнение справа проиллюстрировано на рис. 4 и может быть осуществлено в том случае, если известен более поздний винтаж, а также есть дополнительная информация о том, что между неизвестным винтажом и более поздним не было пересмотров. В этом случае, как показано на рис. 4, данные просто копируются из последующих винтажей. При этом информация о наличии пересмотра была получена на основании анализа одного из следующих источников: база данных винтажей КЭП (см.: [Gornostaev et al., 2022]), доклад «Социально-экономическое положение России»<sup>9</sup>, исторические копии страниц с новостями с сайта Росстата, данные винтажей по реальному ВВП (см.: [Горностаев и др., 2021]).

дата выхода в ценах	30.12.2005 2003	03.04.2006 2003	30.06.2006 2003	02.10.2006 2003	29.12.2006 2003	04.04.2007 2003	29.06.2007 2003	28.09.2007 2003	29.12.2007 2003	14.04.2008 2003	01.07.2008 2003	02.10.2008 2003
II квартал 2003	101,2	102,1	102,1	102,1	102,1	102,1		102,1	102,1	102,1	102,1	102,1
III квартал 2003	104,3	104,3	104,3	104,3	104,3	104,3		104,3	104,3	104,3	104,3	104,3
IV квартал 2003	106,2	106,3	106,3	106,3	106,3	106,3		106,3	106,3	106,3	106,3	106,3
I квартал 2004	95,3	95,4	95,4	95,4	95,4	95,5		95,5	95,5	95,5	95,5	95,5
II квартал 2004	102,2	102,9	102,9	102,9	102,9	102,9		102,9	102,9	102,9	102,9	102,9
III квартал 2004	105,7	104,7	104,6	104,6	104,6	105,2		105,2	105,2	105,2	105,2	105,2
IV квартал 2004	105,9	106,3	106,3	106,3	106,3	105,5		105,5	105,5	105,5	105,5	105,5
I квартал 2005	93,5	93,2	93,2	93,2	93,2	93,4		93,4	93,4	93,3	93,3	93,3
II квартал 2005	104,9	105,6	105,7	105,7	105,7	106,4		106,4	106,4	105,2	105,2	105,2
III квартал 2005	105,9	104,3	104,2	104,2	104,2	105,6		105,6	105,6	105,6	105,6	105,6
IV квартал 2005		106,1	106,1	106,1	106,1	105,7		105,7	105,7	106,0	106,0	106,0
I квартал 2006			92,9	92,9	92,9	92,7		92,7	92,7	91,8	91,8	91,8
II квартал 2006				106,7	106,7	104,9		104,9	104,9	105,6	105,6	105,6
III квартал 2006					104,2	106,0		106,0	106,0	106,2	106,2	106,2
IV квартал 2006						106,3		106,3	106,3	106,6	106,6	106,6
I квартал 2007								92,8	92,8	92,3	92,3	92,3
II квартал 2007								105,1	105,1	105,8	105,8	105,8
III квартал 2007									106,3	106,2	106,2	106,2
IV квартал 2007										107,0	107,0	107,0
I квартал 2008											92,0	92,0
II квартал 2008												104,6

Рис. 4. Иллюстративный пример дозаполнения винтажа справа

При дозаполнении слева также используется информация об отсутствии пересмотра. Однако вместо более позднего винтажа при дозаполнении слева используется более ранний винтаж (см. рис. 5). Ключевым отличием от дозаполнения справа является отсутствие в более раннем винтаже последнего значения по восстанавливаемому показателю (выделенная ячейка на рис. 5). Эти данные приходится восстанавливать с использованием альтернативных источников – новостей, докладов «Социально-экономическое положение России» или файлов с показателями квартал к соответствующему кварталу предыдущего года<sup>10</sup>.

<sup>9</sup> Исторические доклады доступны на сайте Росстата.

<sup>10</sup> Стоит также отметить, что если для одного и того же винтажа на любом из этапов было несколько источников данных, то они использовались для дополнительной верификации и поиска ошибок.

дата выхода в ценах	30.12.2005 2003	03.04.2006 2003	30.06.2006 2003	02.10.2006 2003	29.12.2006 2003	04.04.2007 2003	29.06.2007 2003	28.09.2007 2003	29.12.2007 2003	14.04.2008 2003	01.07.2008 2003	02.10.2008 2003
II квартал 2003	101,2	102,1	102,1	102,1	102,1	102,1	102,1	102,1		102,1	102,1	102,1
III квартал 2003	104,3	104,3	104,3	104,3	104,3	104,3	104,3	104,3		104,3	104,3	104,3
IV квартал 2003	106,2	106,3	106,3	106,3	106,3	106,3	106,3	106,3		106,3	106,3	106,3
I квартал 2004	95,3	95,4	95,4	95,4	95,4	95,5	95,5	95,5		95,5	95,5	95,5
II квартал 2004	102,2	102,9	102,9	102,9	102,9	102,9	102,9	102,9		102,9	102,9	102,9
III квартал 2004	105,7	104,7	104,6	104,6	104,6	105,2	105,2	105,2		105,2	105,2	105,2
IV квартал 2004	105,9	106,3	106,3	106,3	106,3	105,5	105,5	105,5		105,5	105,5	105,5
I квартал 2005	93,5	93,2	93,2	93,2	93,2	93,4	93,4	93,4		93,3	93,3	93,3
II квартал 2005	104,9	105,6	105,7	105,7	105,7	106,4	106,4	106,4		105,2	105,2	105,2
III квартал 2005	105,9	104,3	104,2	104,2	104,2	105,6	105,6	105,6		105,6	105,6	105,6
IV квартал 2005		106,1	106,1	106,1	106,1	105,7	105,7	105,7		106,0	106,0	106,0
I квартал 2006			92,9	92,9	92,9	92,7	92,7	92,7		91,8	91,8	91,8
II квартал 2006				106,7	106,7	104,9	104,9	104,9		105,6	105,6	105,6
III квартал 2006					104,2	106,0	106,0	106,0		106,2	106,2	106,2
IV квартал 2006						106,3	106,3	106,3		106,6	106,6	106,6
I квартал 2007							92,8	92,8		92,3	92,3	92,3
II квартал 2007								105,1		105,8	105,8	105,8
III квартал 2007										106,2	106,2	106,2
IV квартал 2007										107,0	107,0	107,0
I квартал 2008											92,0	92,0
II квартал 2008												104,6

Рис. 5. Иллюстративный пример дозаполнения винтажа слева

### 2.3. Сбор из альтернативных источников

С помощью описанных в пунктах 2.1 и 2.2 методов нам удалось восстановить от 80% до 95% винтажей по каждому из показателей номинального и реального<sup>11</sup> ВВП и его компонентов по использованию и производству. Оставшиеся данные мы попытались найти путем опроса коллег и экспертного сообщества.

### 2.4. Итоговый набор данных

В результате описанных в этом разделе шагов нам удалось собрать винтажи по всем компонентам квартального номинального и реального ВВП по производству за период с декабря 2005 г.<sup>12</sup> По компонентам ВВП по использованию нам удалось собрать практически все данные за тот же период. Исключение составили данные по номинальным импорту и экспорту в октябрьском и декабрьском винтажах 2018 г.<sup>13</sup> (см. рис. 6). В эти даты удалось дозаполнить данные слева, однако в альтернативных источниках не было информации по номинальным экспорту и импорту. Также отметим, что по реальным показате-

<sup>11</sup> По реальному ВВП и его компонентам все показатели собирались в темпах роста.

<sup>12</sup> Несмотря на то, что мы также собрали и информацию практически для всех годовых релизов конца января – начала февраля, из которых косвенно могут быть рассчитаны оценки квартальных показателей, здесь и далее мы описываем результаты только для квартальных релизов, так как вопросы перевода годовых оценок в квартальные являются нетривиальными (особенно для номинальных показателей) и лежат вне рамок этой работы. Найденные данные о годовых оценках представлены в дополнительных файлах.

<sup>13</sup> Мы будем благодарны, если кто-то из читателей пришлет нам эти данные.

лям были восстановлены только данные в последних доступных на дату ценах. Единственным исключением стали данные по компонентам реального использованного ВВП за 2017 г. Здесь мы берем данные в базовых ценах 2011 г., хотя в некоторые даты данные в ценах 2016 г. также публиковались. Это связано с достаточно короткой историей данных использованного ВВП в этих релизах и с тем, что на практике в 2017 г. нашими коллегами использовались именно данные в ценах 2011 г.

дата выхода	30.06.2017	02.10.2017	29.12.2017	02.04.2018	02.07.2018	02.10.2018	29.12.2018	01.04.2019	01.07.2019	02.10.2019	31.12.2019	01.04.2020
...	...	...	...	...	...	...	...	...	...	...	...	...
I квартал 2017	5575,8	5575,8	5575,8	5583,5	5583,5	5583,5	5583,5	5578,1	5578,1	5578,1	5578,1	5578,1
II квартал 2017		5621,5	5621,5	5636,2	5636,2	5636,2	5636,2	5636,4	5636,4	5636,4	5636,4	5629,8
III квартал 2017			5886,3	5897,5	5897,5	5897,5	5897,5	5890,1	5890,1	5890,1	5890,1	5882,9
IV квартал 2017				6848,9	6848,9	6848,9	6848,9	6889,7	6889,7	6889,7	6889,7	6871,9
I квартал 2018					6599,2	6599,2	6599,2	6578,3	6578,3	6578,3	6578,3	6566,9
II квартал 2018						?	?	7749,3	7749,3	7749,3	7749,3	7748,2
III квартал 2018							?	8376,5	8376,5	8376,5	8376,5	8372,5
IV квартал 2018								9228,5	9228,5	9228,5	9228,5	9242,0
I квартал 2019									7706,6	7706,6	7706,6	7697,2
II квартал 2019										7620,7	7620,7	7565,8
III квартал 2019											7758,4	7772,4
IV квартал 2019												8122,0

Рис. 6. Данные по номинальному экспорту, которые не удалось восстановить

Для исследований и анализа часто необходимо иметь более длинные ряды, чем доступно в ценах последнего базового года. Поэтому аналогично [Горностаев и др., 2021] мы также продлили в прошлое (на отдельном листе в дополнительно публикуемых к этой работе файлах) ряды по реальному использованному ВВП на основе публикуемых файлов с показателями квартал к соответствующему кварталу предыдущего года (индекса физического объема). В дополнение к этому для каждого винтажа на базе собранных файлов графика публикаций Росстата и новостей были приближенно (с точностью до нескольких дней) восстановлены даты появления данных на сайте Росстата.

### 3. Анализ свойств собранных показателей

Для иллюстрации свойств собранных показателей и характеристик их пересмотров в этом разделе мы проводим анализ номинального и реального ВВП по использованию. Мы не показываем результаты для произведенного ВВП ввиду того, что они менее иллюстративны из-за перехода с ОКВЭД на ОКВЭД 2, что привело к несопоставимости многих временных рядов по смыслу.

#### 3.1. Анализ пересмотров

Для реального ВВП по использованию до 2010 г. включительно базовым периодом был 2003 г., с 2011 по 2015 гг. – 2008 г., с 2016 по 2017 гг. – 2011 г.<sup>14</sup>, с 2018 по 2022 гг. – 2016 г.

<sup>14</sup> В нашей выборке. См. комментарии относительно винтажей 2017 г. из раздела 2.4.

Для ясности дальнейшего изложения уточним:

- под наблюдением мы подразумеваем значение ВВП или его компонентов (или их темпы роста, если речь идет о реальных показателях) за некоторый период, представленное в одной из публикаций Росстата;
- публикацией/релизом  $i$ -го квартала мы называем информацию о временном ряде квартального ВВП (его компонентов или темпов роста реальных показателей), опубликованную или в конце  $i$ -го квартала, или в начале  $i + 1$  квартала;
- пересмотр – факт уточнения значения наблюдения в публикации по сравнению со значением из предыдущей публикации;
- частота пересмотра отражает долю наблюдений, для которых произошел пересмотр, среди как пересмотренных, так и непересмотренных наблюдений, и позволяет понять, через какое время от отчетного квартала данные могут уточняться;
- магнитуда пересмотра – процентное отклонение значения наблюдения от впервые опубликованного значения.

На рис. 7–12 представлены описательные статистики базы данных квартального ВВП. Рисунки 7–9 построены по данным номинального ВВП и его компонентов, рис. 10–12 – по данным темпов роста реального ВВП и его компонентов. Последовательно отражены частота пересмотров, магнитуда пересмотров и число пересмотров по релизам.

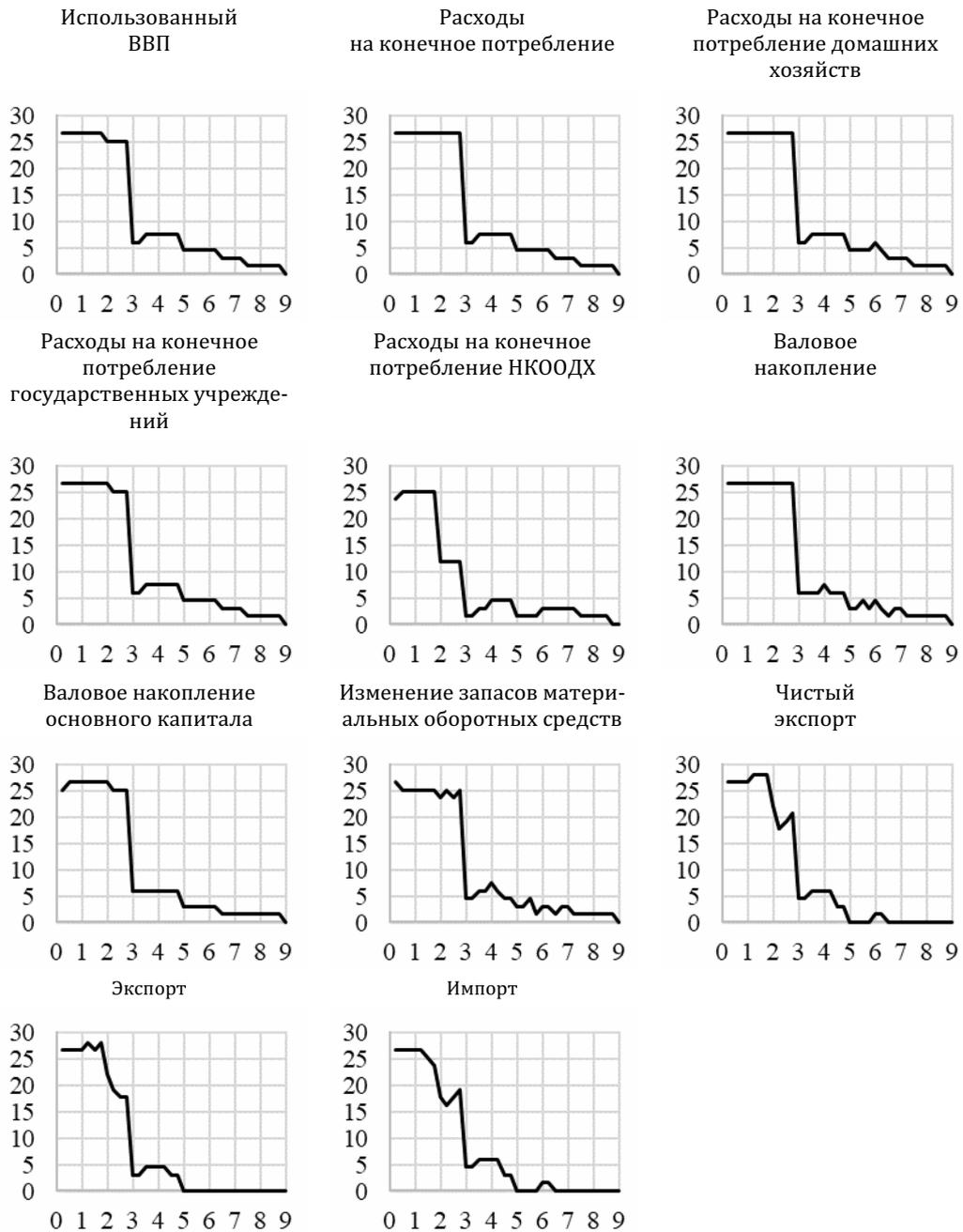
Частота пересмотров номинального ВВП и его компонентов представлена на рис. 7. От 25 до 30% наблюдений пересматриваются спустя один квартал с момента первого релиза. По большому числу показателей в течение 11 кварталов частота пересмотров сохраняется. Спустя 3 года (12 кварталов) частота пересмотров резко снижается и составляет меньше 8% от общего числа наблюдений. Для экспорта и импорта пересмотры прекращаются спустя 5 и 7 лет соответственно, для прочих показателей пересмотры прекращаются спустя 9 лет с момента первого релиза.

Магнитуду пересмотров позволяет оценить рис. 8. Среднее значение магнитуды таких показателей, как использованный ВВП, расходы на конечное потребление домашних хозяйств и государственных учреждений, валовое накопление основного капитала, возрастает от нескольких десятых до примерно 5% (относительно величины первого релиза) на протяжении 5 лет после первого релиза. В меньшей степени пересматриваются показатели экспорта и импорта (значение по модулю не превышает 0,5%), при этом показатели в среднем пересчитываются в сторону снижения (значение магнитуды уходит ниже нуля с увеличением горизонта). В большей степени пересматриваются такие показатели, как изменение запасов материальных оборотных средств (свыше 75% по модулю спустя год) и расходы на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства (НКООДХ) (около 10% по модулю), что связано со спецификой сбора информации о данных показателях<sup>15</sup>.

На количество пересмотров по отношению к публикации в тот или иной календарный период указывает рис. 9. При этом на оси абсцисс указывается публикация, для наблюдений которой подсчитывается число пересмотров спустя квартал<sup>16</sup>.

<sup>15</sup> Отметим, что изменение запасов материальных оборотных средств и расходы на конечное потребление НКООДХ составляют небольшую долю ВВП, поэтому магнитуда пересмотров этих показателей не оказывает большое влияние на магнитуду пересмотра ВВП.

<sup>16</sup> Например, среди наблюдений, опубликованных на рубеже IV квартала 2005 г. и I квартала 2006 г., 11 наблюдений подверглись пересмотрам в конце I квартала 2006 г. (для ясности см. рис. П1 в Приложении).



**Рис. 7.** Частота пересмотров в зависимости от количества лет после первой публикации (номинальный ВВП и его компоненты), %

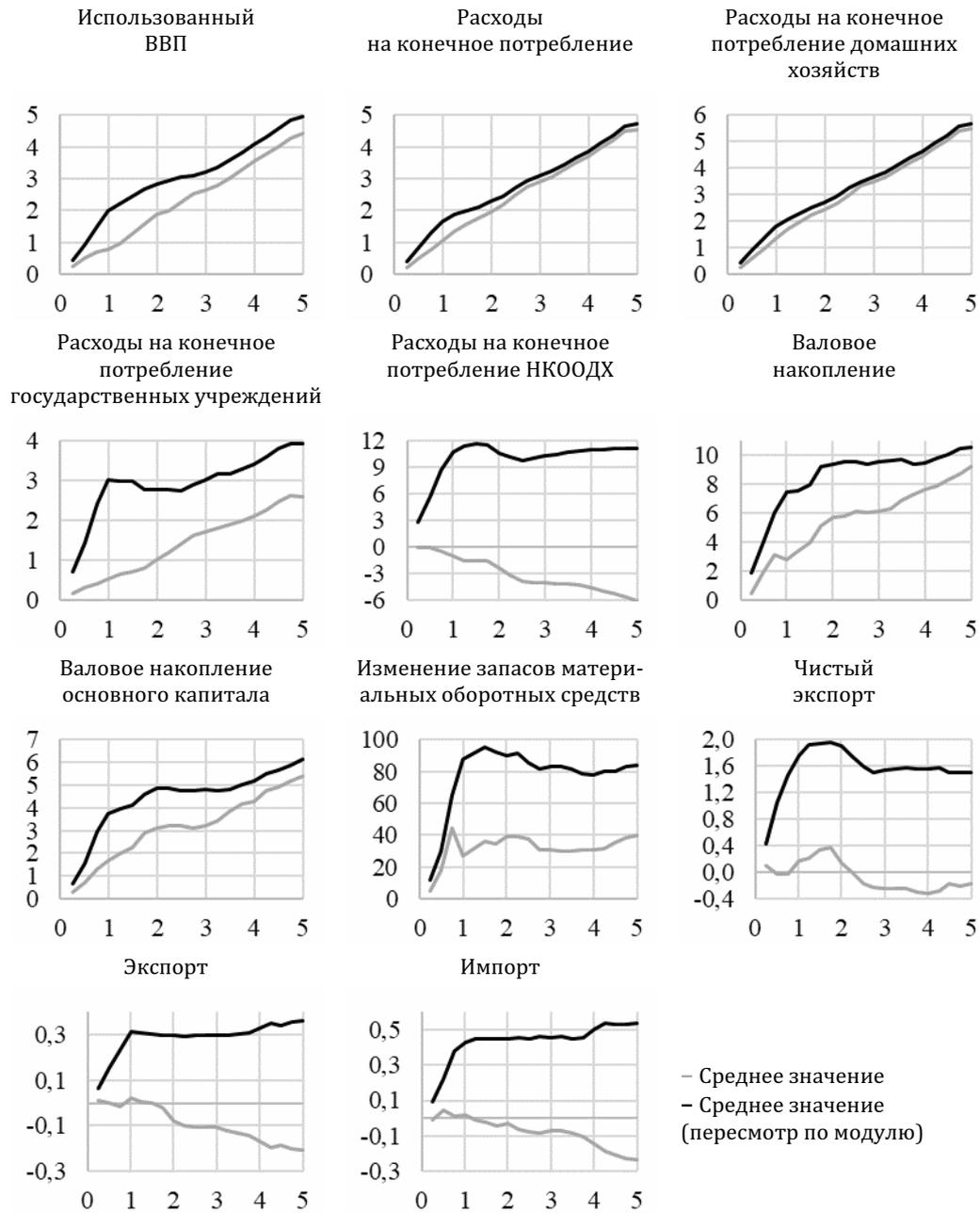
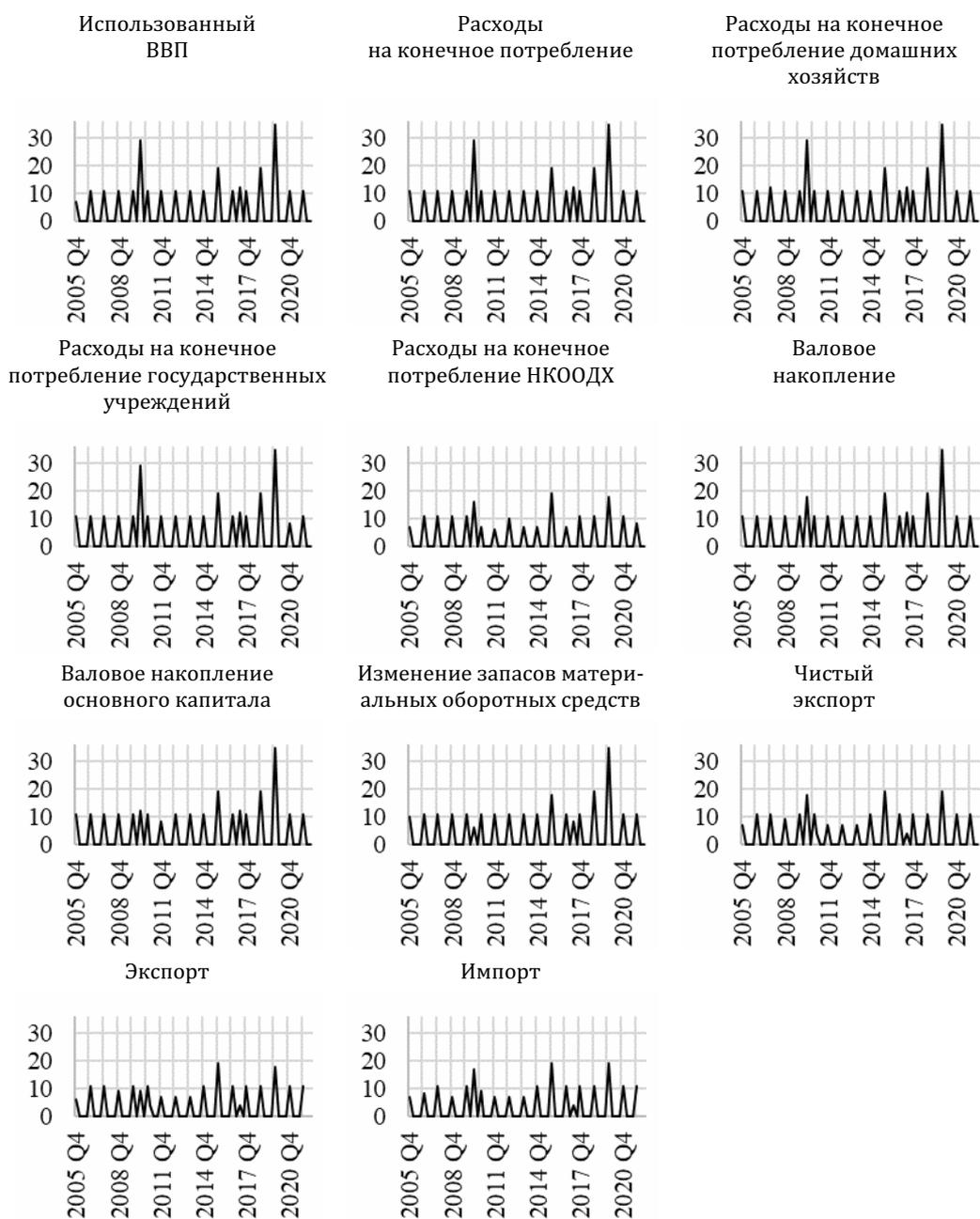


Рис. 8. Магнитуда пересмотров (номинальный ВВП и его компоненты), п.п.



**Рис. 9.** Число пересмотров на момент последнего релиза (номинальный ВВП и его компоненты)

Из рис. 9 видно, что чаще всего пересмотры совершаются в конце I квартала года (т.е. для публикации IV квартала), причем, как правило, пересматриваются данные последних 11 календарных наблюдений (2 года и 3 квартала). То, что чаще всего пересмотры совершаются в конце I квартала, связано с получением более детальных годовых данных при построении годовых национальных счетов.

Публикации I, II, III кварталов практически не подвергаются пересмотрам (исключения составляют публикации III квартала 2009 г., II квартала 2010 г., I квартала 2011 г. (только для экспорта и чистого экспорта), II квартала 2017 г.). В нескольких случаях пересматривалось больше, чем 11 наблюдений (для ВВП это публикации II квартала 2010 г. (29 пересмотров), IV квартала 2015 г., IV квартала 2018 г. (по 19 пересмотров), IV квартала 2019 г. (35 пересмотров)). Такого рода пересмотры связаны с развитием информационной базы и совершенствованием методологии расчетов.

Что касается темпов роста реального ВВП и его компонентов, 25–30% наблюдений пересматривается спустя 1–11 кварталов после релиза, спустя 12 кварталов (3 года) – 5–10% наблюдений, спустя 7 лет информация практически не пересматривается (см. рис. 10).

Магнитуда пересмотров темпов роста реальных показателей представлена на рис. 11. Мы стремимся отразить магнитуду изменений, не связанных со сменой базовых периодов, поэтому принцип построения рис. 11 отличается от принципа построения рис. 8 (где не было базовых периодов и магнитуда рассчитывалась относительно значения из первой публикации). Сначала рассчитывается процентное отклонение относительно публикации, которая впервые использует тот же базовый период. Далее рассчитывается значение последней магнитуды (также в процентах), рассчитанной для предыдущего базового периода. В завершении два рассчитанных отклонения суммируются (для ясности см. рис. П2 в Приложении), результат чего отражен сплошными линиями на рис. 11. Для построения пунктирных линий рис. 11 не учитывались магнитуды, рассчитанные при переходе на новый базовый период.

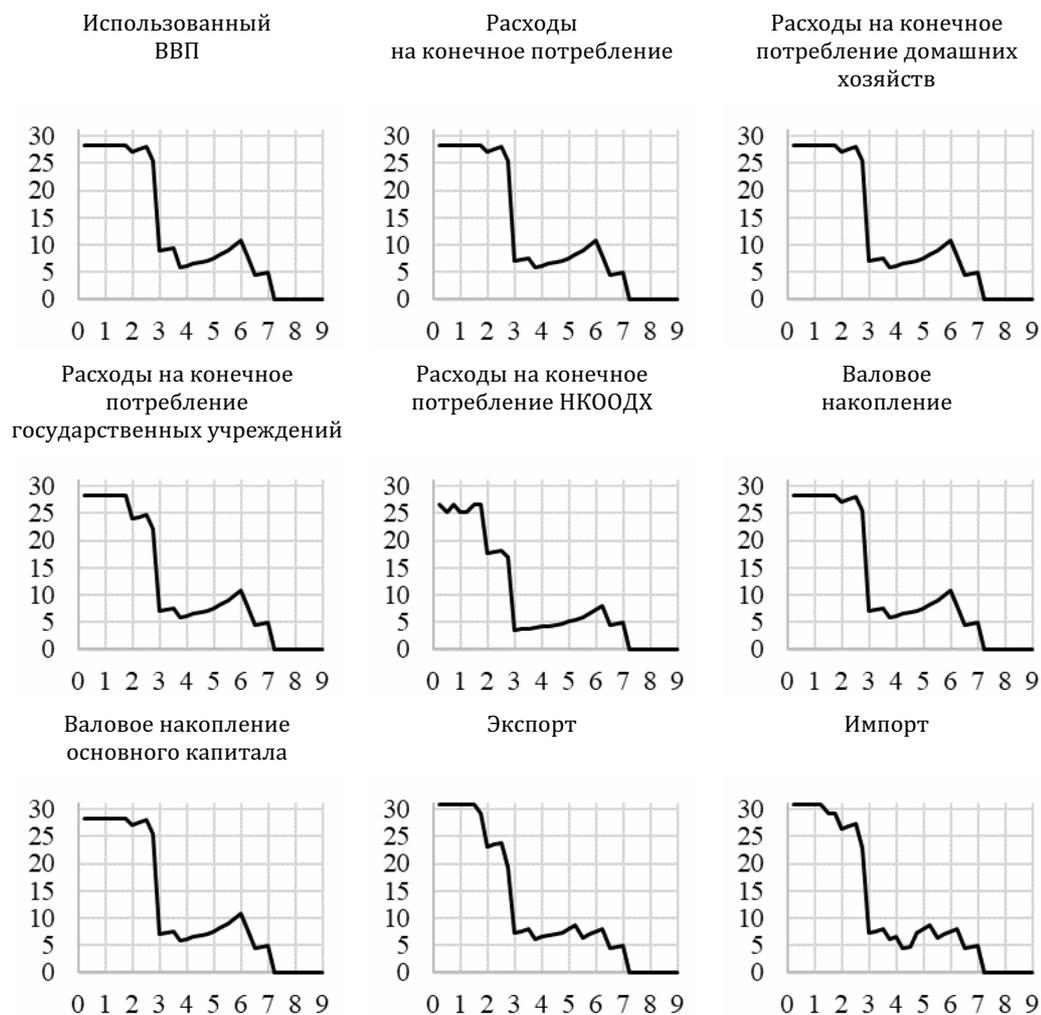
Магнитуда пересмотров по модулю (черные линии на рис. 11) колеблется около 1% для темпов роста расходов на конечное потребление домашних хозяйств и государственных учреждений, экспорта и импорта и также для использованного ВВП. Большая магнитуда (по модулю) наблюдается для темпов роста расходов на конечное потребление НКООДХ (около 1,5%), валового накопления основного капитала (около 3%) и для валового накопления в целом (до 6%). Средние значения магнитуды (серые линии на рис. 11) многих показателей колеблются около нуля, что говорит о том, что темпы роста показателей в реальном выражении пересматриваются как вверх, так и вниз. Анализ альтернативных значений магнитуды (пунктирные линии на рис. 11) приводят к аналогичным выводам.

На рис. 12 так же, как и на рис. 9, видно, что больше всего пересмотров совершается для публикаций IV квартала года. Как правило, совершается 11 пересмотров, но бывает и больше (III квартал 2009 г. (25 пересмотров), II квартал 2010 г. (28 пересмотров), II квартал 2017 г. (24 пересмотра), IV квартал 2019 г. (22 пересмотра)). Большее число пересмотров связано с изменением методологии расчетов, в том числе с уточнением значений базовых периодов.

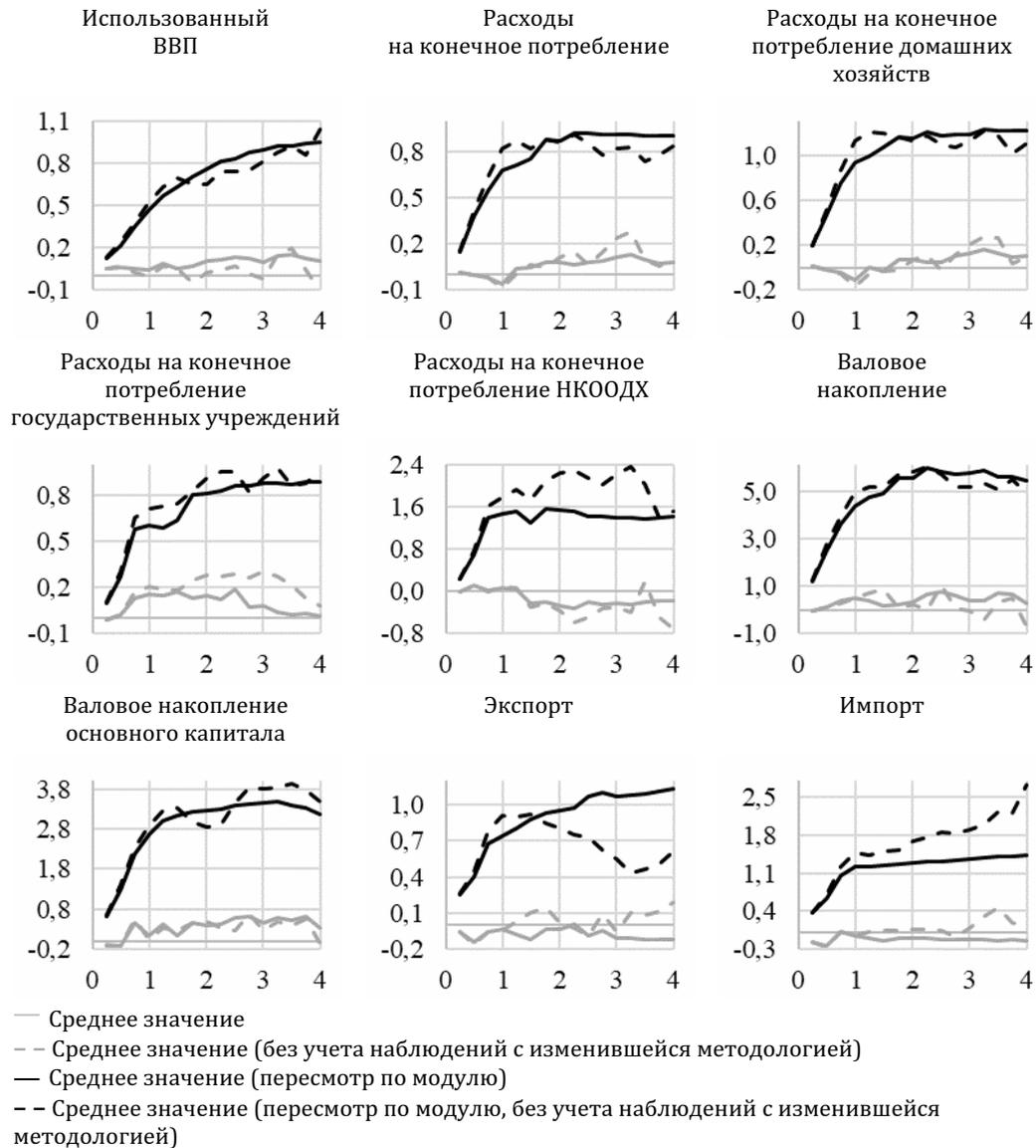
На рис. 12 мы приняли равными нулю число пересмотров для публикаций IV квартала 2010, 2015 и 2017 гг., что связано с тем, что это были последние релизы с использованием некоторого базового периода<sup>17</sup>, который в дальнейшем менялся. Мы не можем с

<sup>17</sup> В нашей выборке. См. комментарии относительно винтажей 2017 г. из раздела 2.4.

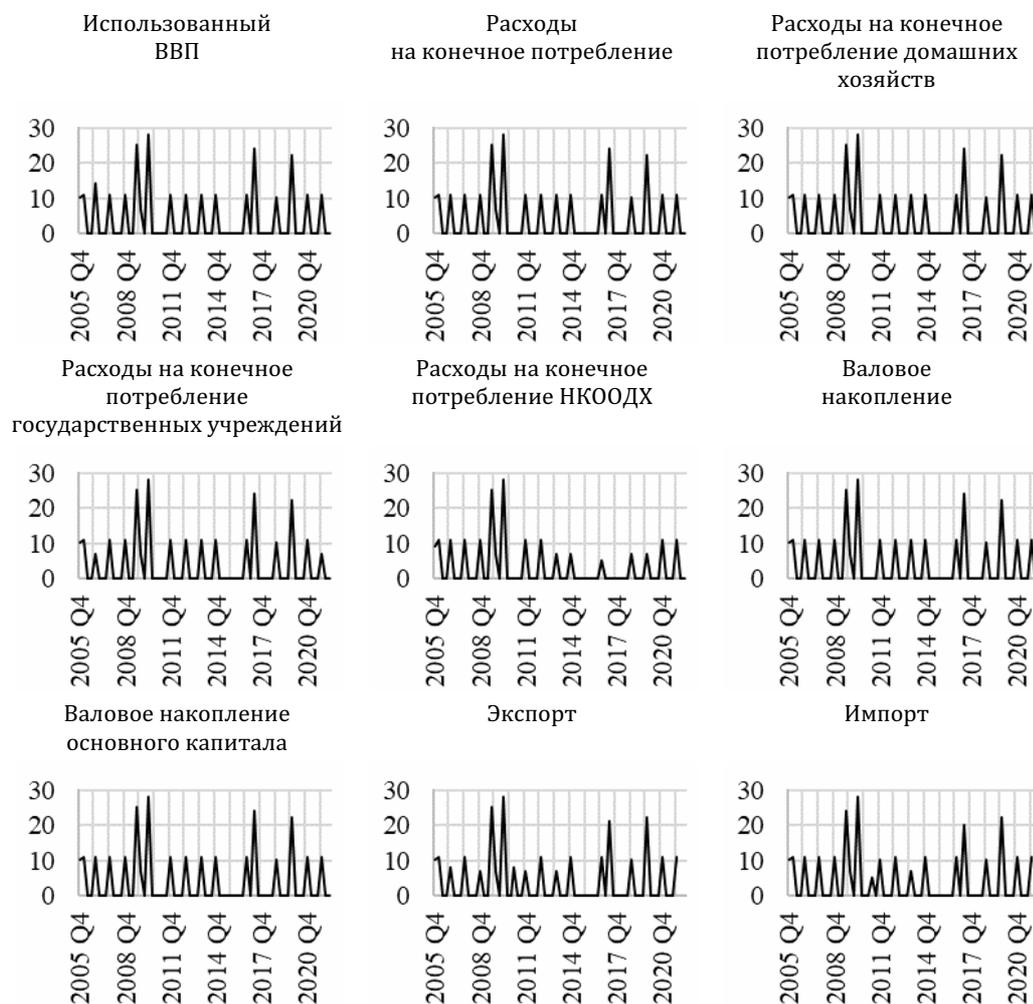
точно сказать, в какой степени изменение значений связано с изменением базы, а в какой – с уточнением показателя, поэтому мы обнулили число пересмотров на рис. 12.



**Рис. 10.** Частота пересмотров (темпы роста реального ВВП и его компонентов), %



**Рис. 11.** Магнитуда пересмотров  
 (темпы роста реального ВВП и его компонентов), п.п.



**Рис. 12.** Число пересмотров на момент последнего релиза (темпы роста реального ВВП и его компонентов)

#### 4. Похожие направления в литературе

Наша работа связана с несколькими направлениями исследований, которые можно разделить на три группы: посвященные источникам винтажных данных во всем мире, источниками винтажных данных для России и способам сбора винтажных данных. Ниже мы коснемся каждого из этих направлений чуть подробнее.

#### 4.1. Источники пересмотров российских показателей

Данные по пересмотрам российских показателей представлены в нескольких источниках. За рубежом данные собирались для американской базы данных ALFRED. В ней содержится более 2000 пересмотров показателей, найденных в поиске по ключевому слову «Russia». Эти данные отражают динамику различных переменных, включая реальный ВВП и некоторые его компоненты по использованию. Однако, в отличие от собранных в этой статье, ALFRED содержит лишь отрывочные данные по реальному ВВП и его компонентам. Более того, в основном в базе хранятся сезонно сглаженные, а не исходные данные, что может быть ограничивающим фактором во многих исследованиях.

Насколько нам известно, первой работой российских исследователей, которая представляет винтажные данные, была статья [Астафьева, Турунцева, 2021], в которой авторы исследовали свойства пересмотров индекса физического объема и номинального ВВП. Несмотря на то, что публикация данных не была основной целью, авторы в своей статье представили данные по годовому номинальному ВВП, а также написали о возможности предоставления по запросу информации о квартальных данных для номинального ВВП и индексе физического объема год к году.

Сотрудники Банка России в работах [Gornostaev et al., 2022; Горностаев и др., 2021] представили винтажи по данным из сборника «Краткосрочные экономические показатели Российской Федерации» (более 400 показателей с января 2001 г.) и по темпам квартального роста реального ВВП (с декабря 2005 г.). Данные опубликованы как приложения к соответствующим материалам на сайте Банка России.

#### 4.2. Источники пересмотров показателей других стран

Пионером в сборе винтажных данных является США. Наиболее полными и постоянно обновляемыми базами являются Real Time Data Set for Macroeconomists (RTDSM) и ALFRED, причем базы содержат как данные США, так и данные других стран. Организация экономического сотрудничества и развития создала базу Original Release Data and Revisions Database (ORDRD), которая содержит информацию по странам, входящим в ОЭСР, и по ряду стран, не входящих в организацию. Среди международных баз пересмотров также стоит отметить The Euro Area Real-Time Database (RTDB, покрывает страны еврозоны, Японию и США) и International Real-Time Dataset (покрывает Германию, Канаду, США, Швейцарию и Японию). На страновом уровне винтажные данные собирают Великобритания, Германия, Канада, Новая Зеландия и Австралия. Более детальная информация по базам представлена в табл. П1 Приложения.

На постсоветском пространстве немногие страны собирают винтажные данные. Помимо опыта Банка России, о котором написано выше, можно выделить винтажные данные Банка Литвы, доступные по небольшому числу показателей с 2018 г. В рамках отдельных исследовательских проектов [Benkovskis, 2008; Lewis, 2013] были представлены пересмотры показателей Латвии, Литвы и Эстонии, но после выхода публикаций обновление данных не производилось.

#### 4.3. Подходы к сбору данных по пересмотрам

Базы данных пересмотров состояются на основе статистических сборников, обзоров и иных информационных материалов, которые хранятся в самых разных форматах.

В разные периоды времени это могли быть печатные экземпляры (hard copies, были актуальны на рубеже XX–XXI вв.), микрофиши, CD-диски, pdf-файлы, excel-файлы. Как правило, исследователи вручную составляли базы данных пересмотров. Одну из первых баз данных сформировали Диболд и Рудебуш [Diebold, Rudebusch, 1991] на основе ежемесячного Дайджеста условий ведения бизнеса (Business Condition Digest), подготавливаемого Бюро экономического анализа США. База данных позволила авторам использовать пересмотры сводного опережающего индекса (Composite Leading Index) для прогнозирования выпуска.

Изначально сбор винтажных данных авторы проводили самостоятельно для целей своих исследований, не всегда продолжая обновление рядов в дальнейшем (например, [Diebold, Rudebusch, 1991] в США, [Gerberding et al., 2005] в Германии, [Stone, Wardrop, 2002] в Австралии). Последующие авторы расширяли базы данных предшественников, добавляя актуальные данные. В дальнейшем исследователи формировали базы, которые систематизированно актуализируются до сих пор и доступны в открытых источниках (подробнее см. табл. П1 в Приложении).

В частности, авторы работы [Croushore, Stark, 2001] создали большую базу пересмотров американских показателей Real-Time Data Set for Macroeconomists. Они работали с опубликованными статистическими сборниками, обзорами и иными материалами. Сбор данных, особенно исторических, происходил с распечатанных версий документов. В своей работе авторы благодарят студентов-стажеров, которые под их руководством вручную собирали материалы этих сборников в единую электронную базу данных. Авторы опубликовали собранную базу пересмотров на сайте Федерального резервного банка (ФРБ) Филадельфии. Это была одна из первых баз данных, представленных в широком доступе. Интерес к базе проявился достаточно быстро, и в скором времени появились работы на основе этих данных (например: [Orphanides, Norden, 2002; Giordani, Söderlind, 2003]). В настоящий момент база винтажных данных ФРБ Филадельфии более обширная и содержит ряд дополнительных показателей, в том числе данные Зеленой книги ФРС США.

База данных пересмотров по странам еврозоны создавалась на основе экономических бюллетеней ЕЦБ и ежемесячных бюллетеней ЕЦБ (см.: [Giannone et al., 2012]). Опыт постсоветских стран схожий. Бенковскис [Benkovskis, 2008] использовал винтажные данные по ВВП Латвии, собирая их из публикации «Макроэкономические показатели Латвии». Левис [Lewis, 2013] вручную собирал данные по ежегодным отчетам о процессе трансформации (Transition Report), которые публикует Европейский банк реконструкции и развития по странам Центральной и Восточной Европы (из постсоветского пространства в выборку вошли Латвия, Литва и Эстония). Анализируя данные бюджетно-налоговой политики, Чимадомо [Cimadomo, 2016] отмечает, что существующих баз данных (RTDSM, ALFRED, ECB-EABCN, ORDRD) часто бывает недостаточно, и многие исследователи (особенно на заре использования винтажных данных) собирали данные вручную из таких официальных публикаций, как OECD Economic Outlook, Excessive Deficit Procedure Notifications, Stability and Convergence Programmes

## 5. Заключение

В настоящей работе мы собрали наборы винтажных данных по номинальному и реальному ВВП и его компонентам, а также описали методологию сбора данных. Мы на-

деемся, что собранные нами показатели вкупе с показателями, опубликованными в работах [Горностаев и др., 2021; Gornostaev et al., 2022], помогут исследователям и аналитикам понять, как оценивалась экономическая ситуация в определенный момент в прошлом, а также станут отправной точкой при проведении более корректных экспериментов с моделями в псевдореальном времени.

## Приложение.

### Пояснения к расчету описательной статистики

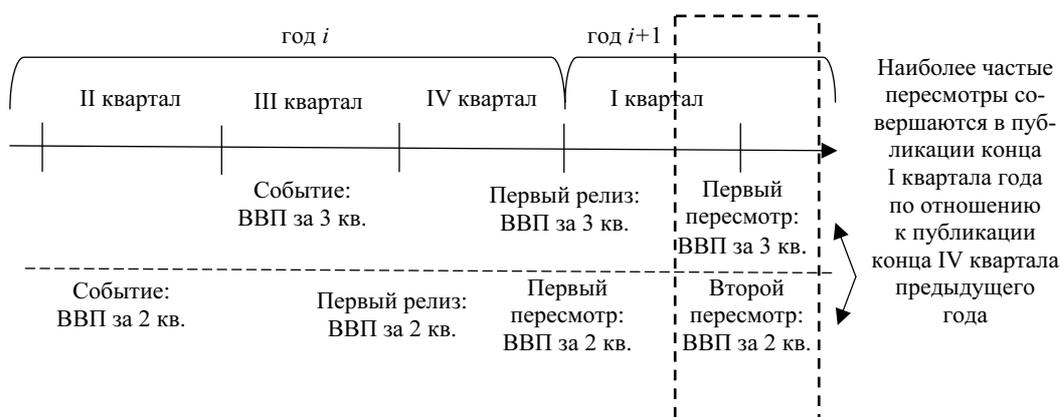


Рис. П1. Схема публикации квартальных данных<sup>18</sup>



$\frac{a_m^{2008}}{a_0^{2008}} - 1$  — отклонение относительно первой публикации с тем же базовым периодом

$\frac{a_k^{2003}}{a_0^{2003}} - 1$  — накопленное отклонение для наблюдений с прошлым базовым периодом

$\frac{a_k^{2003}}{a_0^{2003}} - 1 + \frac{a_m^{2008}}{a_0^{2008}} - 1$  — значение магнитуды для наблюдения  $a_m^{2008}$

Рис. П2. Схема подсчета магнитуды пересмотров

<sup>18</sup> Подразумевается, что оба релиза пересматриваются, хотя в общем случае релиз не всегда пересматривается сразу.

Таблица П1.

## Базы пересмотров по странам

Страна	База	Показатели/ периодичность	Сбор данных	Ссылки на статьи
США	Real Time Data Set for Macroeconomists (RTDSM) <sup>19</sup>	Месячные и квартальные данные (около 100 показателей)	Первоначальный сбор данных продолжался восемь лет, авторы и студенты-стажеры вручную собирали базу данных	Croushore, Stark, 2001
США	ALFRED <sup>20</sup>	Месячные, квартальные и годовые (730 тыс. рядов)	Базы данных FRED и ALFRED обладают схожей архитектурой. Благодаря общему дизайну пересмотр данных в базе FRED автоматически отражается в качестве винтажа в базе ALFRED	Anderson, 2006
США	FRED-MD <sup>21</sup> , FRED-QD	Месячные (134 показателя), квартальные (248 показателей)	Соединили ряды Stock, Watson (1996), Jurado et al. (2015), Stock, Watson (2012). Добавили данные из <i>Global Insights Basic Economics Database</i> . Данные автоматически обновляются из системы FRED	McCracken, Ng, 2016; McCracken, Ng, 2020
США	Bureau of Economic Analysis previously published estimates <sup>22</sup>	Квартальные и годовые: ВВП и его компоненты, региональный выпуск, выпуск по отраслям, таблицы «затраты-выпуск», данные платежного баланса	–	–
ОЭСР + Аргентина, Бразилия, Индия, Индонезия, Китай, Россия, Южная Африка <sup>23</sup>	Original Release Data and Revisions Database (ORDRD) <sup>24</sup>	Месячные и квартальные (24 показателя)	Собирали данные из архивных CD-ROM с данными <i>Статистического справочника основных экономических показателей ОЭСР (OECD Main Economic Indicators)</i> . Пересмотры формировались с 1999 г.	–

<sup>19</sup> <https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/real-time-data-set-for-macroeconomists>

<sup>20</sup> <http://alfred.stlouisfed.org>

<sup>21</sup> <https://research.stlouisfed.org/econ/mccracken/fred-databases/>

<sup>22</sup> <https://apps.bea.gov/histdata/>

<sup>23</sup> Не по всем странам доступен каждый из показателей.

<sup>24</sup> [https://stats.oecd.org/Index.aspx?DataSetCode=MEI\\_ARCHIVE](https://stats.oecd.org/Index.aspx?DataSetCode=MEI_ARCHIVE)

Продолжение табл. П1.

Страна	База	Показатели/ периодичность	Сбор данных	Ссылки на статьи
ОЭСР (26 стран)	Real-Time Historical Dataset for the OECD <sup>25</sup>	13 показателей: реальный и номинальный ВВП, дефлятор ВВП, промышленное производство, обрабатывающее производство, коэффициент использования производственных мощностей, уровень безработицы, ИПЦ, денежное предложение, капиталовложения, импорт, экспорт, чистые потоки капитала	Собирали данные из печатных копий <i>Статистического справочника основных экономических показателей ОЭСР</i> . Это позволило получить более ранние пересмотры данных, чем в ORDRD	Fernandez et al., 2011
Великобритания	Gross Domestic Product Real-Time Database <sup>26</sup>	Квартальные: реальный и номинальный ВВП (в том числе его компоненты), дефлятор ВВП, выпуск по отраслям, номинальный доход по секторам экономики, чистое кредитование по секторам, показатели рынка труда	Данные вносятся из Голубой книги, которая публикуется ежегодно Национальной статистической службой Великобритании. В начале 2000-х годов ряды по ВВП и его компонентам получали из публикаций <i>Экономические тренды (Economic Trends)</i> , <i>Финансовая статистика (Financial Statistics)</i> и приложений к ним	Castle, Ellis, 2002; Egginton et al., 2002; Garratt, Vahey, 2006
Великобритания	Revisions triangles for monthly GDP <sup>27</sup>	Месячные: темп роста ВВП, месяц к месяцу	–	–
Зона евро, Япония, США	The Euro Area Real-Time Database (RTDB) под инициативой Euro Area Business Cycle Network (EABCN) <sup>28</sup>	Месячные, квартальные, годовые (около 230 рядов)	Винтажные ряды получали из Экономического бюллетеня ЕЦБ (ранее Ежемесячный бюллетень ЕЦБ). Данные обновляются раз в полгода (в начале января и начале июля)	Giannone et al., 2012

<sup>25</sup> <https://www.dallasfed.org/research/international/oecd>

<sup>26</sup> <https://www.bankofengland.co.uk/statistics/gdp-real-time-database>

<sup>27</sup> <https://www.ons.gov.uk/economy/grossdomesticproductgdp/datasets/revisionstrianglesformonthlygdp>

<sup>28</sup> <https://sdw.ecb.europa.eu/browseExplanation.do?node=9689716>

Продолжение табл. П1.

Страна	База	Показатели/ периодичность	Сбор данных	Ссылки на статьи
Германия	Macroeconomic real-time database <sup>29</sup>	Месячные, квартальные, годовые (около 280 рядов)	База собиралась на основе печатных и электронных документов, она пополняется с помощью автоматизированной системы обновления	Knetsch, 2009
Германия, Канада, США, Швейцария, Япония	International real-time dataset	Квартальные: реальный и номинальный ВВП, инфляция, денежное предложение, уровень безработицы	Использовались данные дискет и CD	Faust et al., 2003
Канада	Real-time Canadian Socio-Economic Information Management System (CANSIM) tables <sup>30</sup>	Месячные и квартальные (21 показатель)	База автоматически пополняется	–
Канада	The real-time database of Bank of Canada staff projections <sup>31</sup>	Квартальные (22 показателя)	Данные были получены из архивных публикаций <i>Экономических прогнозов сотрудников (Staff Economic Projections)</i> Банка Канады.	Champagne et al., 2018
Новая Зеландия	A Real-Time Database for GDP <sup>32</sup>	Месячный ВВП и его компоненты	–	Sleeman, 2006
Австралия	Australian Real-Time Macroeconomic Database <sup>33</sup>	Месячные и квартальные: ВВП и его компоненты, цены, ключевые денежные показатели, показатели рынка труда, 21 показатель бюджетно-налоговой статистики	Пересмотры составлялись на основе рядов, опубликованных статистическим бюро Австралии в форме печатных изданий, микрофишей, CD ROM и режис электронных файлов (PDF, Excel, Lotus 123). Денежно-кредитная статистика публиковалась Резервным банком Австралии в Ежемесячном бюллетене, доступном онлайн	Lee et al., 2012; Lee et al., 2019

<sup>29</sup> [https://www.bundesbank.de/dynamic/action/en/statistics/time-series-databases/time-series-databases/745590/real-time-data?statisticType=BBK\\_RTD&treeId=113205500](https://www.bundesbank.de/dynamic/action/en/statistics/time-series-databases/time-series-databases/745590/real-time-data?statisticType=BBK_RTD&treeId=113205500)

<sup>30</sup> <https://www.statcan.gc.ca/en/dai/btd/rct>

<sup>31</sup> <https://www.bankofcanada.ca/rates/staff-economic-projections/>

<sup>32</sup> <https://www.rbnz.govt.nz/research-and-publications/research-programme/additional-research/a-real-time-database-for-gdp>

<sup>33</sup> <https://fbe.unimelb.edu.au/economics/macrocentre/artmdatabase>

\* \*  
\*

### СПИСОК ЛИТЕРАТУРЫ

- Астафьева Е.В., Турунцева М.Ю.* Пересмотры ВВП: данные и оценка статистических свойств // Экономический журнал ВШЭ. 2021. Т. 25. № 1. С. 65–101.
- Горностаев Д., Поршаков А., Селезнев С., Стерхова А., Чернядьев Д.* Пересмотры ВВП: измерение и последствия // Аналитическая записка. Банк России, 2021.
- Anderson R.G.* Replicability, Real-Time Data, and the Science of Economic Research: FRED, AL-FRED, and VDC // Federal Reserve Bank of Saint Louis Review. 2006. Vol. 88. № 1. P. 81–93.
- Benkovskis K.* Short-term Forecasts of Latvia's Real Gross Domestic Product Growth Using Monthly Indicators // Bank of Latvia Working Paper Series. 2008. № 2008/05.
- Castle J., Ellis C.* Building a Real-Time Database for GDP (E) // Quarterly Bulletin – Bank of England. 2002. Vol. 42. № 1. P. 42–49.
- Champagne J., Poulin-Bellisle G., Sekkel R.* Evaluating the Bank of Canada Staff Economic Projections Using a New Database of Real-Time Data and Forecasts // Bank of Canada Staff Working Paper. 2018. № 2018-52.
- Cimadomo J.* Real-Time Data and Fiscal Policy Analysis: A Survey of the Literature // Journal of Economic Surveys. 2016. Vol. 30. № 2. P. 302–326.
- Croushore D., Stark T.* A Real-Time Data Set for Macroeconomists // Journal of Econometrics. 2001. Vol. 105. № 1. P. 111–130.
- Diebold F.X., Rudebusch G.D.* Forecasting Output with the Composite Leading Index: A Real-Time Analysis // Journal of the American Statistical Association. 1991. Vol. 86. № 415. P. 603–610.
- Egginton D.M., Pick A., Vahey S.P.* «Keep it Real!»: A Real-Time UK Macro Data Set // Economics Letters. 2002. Vol. 77. № 1. P. 15–20.
- Faust J., Rogers J.H., Wright J.H.* Exchange Rate Forecasting: The Errors We've Really Made // Journal of International Economics. 2003. Vol. 60. P. 35–59.
- Fernandez A.Z., Koenig E.F., Nikolsko-Rzhevskyy A.* The Real-Time Historical Database for the OECD // Federal Reserve Bank of Dallas, Globalization and Monetary Policy Institute Working Papers. 2011. № 96.
- Garratt A., Vahey S.P.* UK Real-Time Macro Data Characteristics // Economic Journal. 2006. Vol. 116. P. 119–135.
- Gerberding C., Seitz F., Worms A.* How the Bundesbank Really Conducted Monetary Policy // The North American Journal of Economics and Finance. 2005. Vol. 16. № 3. P. 277–292.
- Giannone D., Henry J., Lalik M., Modugno M.* An Area-Wide Real-Time Database for the Euro Area // Review of Economics and Statistics. 2012. Vol. 94. № 4. P. 1000–1013.
- Giordani P., Söderlind P.* Inflation Forecast Uncertainty // European Economic Review. 2003. Vol. 47. № 6. P. 1037–1059.
- Gornostaev D., Ponomarenko A., Seleznev S., Sterkhova A.* A Real-Time Historical Database of Macroeconomic Indicators for Russia // Russian Journal of Money and Finance. 2022. Vol. 81. № 1. P. 88–103.
- Jurado K., Ludvigson S.C., Ng S.* Measuring Uncertainty // The American Economic Review. 2015. Vol. 105. № 3. P. 1177–1216.
- Knetsch T.A.* The German Real-Time Database Gerda: Content, Structure, Search and Download. Technical Documentation. Deutsche Bundesbank, 2009.
- Lee K., Morley J., Shields K., Tan M.S.* The Australian Real-Time Fiscal Database: An Overview and an Illustration of Its Use in Analysing Planned and Realised Fiscal Policies // CAMA Working Papers. 2019. № 14.
- Lee K., Olekalns N., Shields K., Wang Z.* Australian Real-Time Database: An Overview and an Illustration of Its Use in Business Cycle Analysis // Economic Record. 2012. Vol. 88. № 283. P. 495–516.

- Lewis J.* Fiscal Policy in Central and Eastern Europe with Real Time Data: Cyclical, Inertia and the Role of EU Accession // *Applied Economics*. 2013. Vol. 45. № 23. P. 3347–3359.
- Mamedli M., Shibitov D.* Forecasting Russian CPI with Data Vintages and Machine Learning Techniques // *Bank of Russia Working Paper Series*. 2021. № 70.
- McCracken M., Ng S.* FRED-QD: A Quarterly Database for Macroeconomic Research // *National Bureau of Economic Research*. 2020. № 26872.
- McCracken M., Ng S.* FRED-MD: A Monthly Database for Macroeconomic Research // *Journal of Business and Economic Statistics*. 2016. Vol. 34. № 4. P. 574–589.
- Orphanides A., Norden S.V.* The Unreliability of Output-Gap Estimates in Real Time // *Review of Economics and Statistics*. 2002. Vol. 84. № 4. P. 569–583.
- Seabold S., Perktold J.* Statsmodels: Econometric and Modeling with Python // *Proceedings of the 9th Python in Science Conference*. 2010. P. 57–61.
- Sleeman C.* Analysis of Revisions to Quarterly GDP – A Real-Time Database // *Reserve Bank of New Zealand Bulletin*. 2006. Vol. 69. № 1. P. 31–44.
- Stock J.H., Watson M.* Evidence on Structural Instability in Macroeconomic Time Series Relations // *Journal of Business and Economic Statistics*. 1996. Vol. 14. № 1. P. 11–30.
- Stock J.H., Watson M.* Disentangling the Channels of the 2007–2009 Recession // *Brookings Papers on Economic Activity*. 2012. P. 81–156.
- Stone A., Wardrop S.* Real-Time National Accounts Data // *Reserve Bank of Australia Research Discussion Papers*. 2002. № 5.
- Tebrake J.* A Framework for Recording and Communicating Revisions // *Working Party on Financial Statistics*. OECD, 2019.

## Reconstructing the Publication History of Russia's GDP and its Components

**Dmitrii Gornostaev<sup>1</sup>, Natalia Makhankova<sup>2</sup>, Petr Milyutin<sup>3</sup>,  
Alexey Ponomarenko<sup>4</sup>, Sergey Seleznev<sup>5</sup>**

<sup>1</sup> Bank of Russia,  
12, Neglinnaya Street, Moscow, 107016, Russian Federation.  
E-mail: gornostaevda@cbr.ru

<sup>2</sup> Bank of Russia,  
12, Neglinnaya Street, Moscow, 107016, Russian Federation.  
E-mail: makhankovana01@cbr.ru

<sup>3</sup> Bank of Russia,  
12, Neglinnaya Street, Moscow, 107016, Russian Federation.  
E-mail: milyutinpp01@cbr.ru

<sup>4</sup> Bank of Russia,  
12, Neglinnaya Street, Moscow, 107016, Russian Federation.  
E-mail: ponomarenkoaa@cbr.ru

<sup>5</sup> Bank of Russia,  
12, Neglinnaya Street, Moscow, 107016, Russian Federation.  
E-mail: seleznevsm@cbr.ru

Vintage data, or retrospective data, play a crucial role in assessing the accuracy of macroeconomic forecasting models and the decisions based on these models. These data represent a snapshot of information available at a particular point in time in the past and reflect how the economic situation was perceived at that time. It is important to realise that it is on the basis of vintage data, not revised and refined indicators, that forecasts and economic policy decisions were made. As an example, we show that the GDP growth paths in the projection period may diverge by as much as 1pp.

---

**Dmitrii Gornostaev** – chief Economist, Research and Forecasting Department.

**Natalia Makhankova** – chief Economist, Research and Forecasting Department.

**Petr Milyutin** – deputy head of Section, Research and Forecasting Department.

**Alexey Ponomarenko** – Phd, head of Economic Research Division, Research and Forecasting Department.

**Sergey Seleznev** – head of the Center for Data Science and Machine Learning, Research and Forecasting Department.

This paper presents a set of vintage data on Russian GDP and its components by the expenditure and production approaches. The dataset consists of revisions of nominal and real quarterly data for the period from December 2005 to the present. In addition to such data, the paper describes some properties of real and nominal GDP indicator revisions (number, frequency, magnitude) and its expenditure components, as well as the methodology for collecting historical indicators using the Wayback Machine, which enables data collection even in the absence of a saved history of their releases.

We hope that this dataset will become a valuable tool for researchers and analysts, helping to better understand how the economic situation was assessed in the past. It can also provide a basis for conducting more correct experiments with models in pseudo-real time.

**Key words:** GDP; data vintages; data revisions; frequency and magnitude of revisions; Russia; historical copies of websites; Wayback Machine.

**JEL Classification:** C82, E01, E2.

\* \*

\*

### References

- Anderson R.G. (2006) Replicability, Real-Time Data, and the Science of Economic Research: FRED, ALFRED, and VDC. *Federal Reserve Bank of Saint Louis Review*, 88, 1, pp. 81–93.
- Astafieva E., Turuntseva M. (2021) Revisions of GDP: Data and Assessment of Statistical Properties. *HSE Economic Journal*, 25, 1, pp. 65–101. (In Russ.)
- Benkovskis K. (2008) *Short-term Forecasts of Latvia's Real Gross Domestic Product Growth Using Monthly Indicators*. Bank of Latvia Working Paper Series, 2008/05.
- Castle J., Ellis C. (2002) Building a Real-Time Database for GDP (E). *Quarterly Bulletin – Bank of England*, 42, 1, pp. 42–49.
- Champagne J., Poulin-Bellisle G., Sekkel R. (2018) *Evaluating the Bank of Canada Staff Economic Projections Using a New Database of Real-Time Data and Forecasts*. Bank of Canada Staff Working Paper, 2018-52.
- Cimadomo J. (2016) Real-Time Data and Fiscal Policy Analysis: A Survey of the Literature. *Journal of Economic Surveys*, 30, 2, pp. 302–326.
- Croushore D., Stark T. (2001) A Real-Time Data Set for Macroeconomists. *Journal of Econometrics*, 105, 1, pp. 111–130.
- Diebold F.X., Rudebusch G.D. (1991) Forecasting Output with the Composite Leading Index: A Real-Time Analysis. *Journal of the American Statistical Association*, 86, 415, pp. 603–610.
- Egginton D.M., Pick A., Vahey S.P. (2002) «Keep it Real!»: A Real-Time UK Macro Data Set. *Economics Letters*, 77, 1, pp. 15–20.
- Faust J., Rogers J.H., Wright J.H. (2003) Exchange Rate Forecasting: The Errors We've Really Made. *Journal of International Economics*, 60, pp. 35–59.
- Fernandez A.Z., Koenig E.F., Nikolsko-Rzhevskyy A. (2011) *The Real-Time Historical Database for the OECD*. Federal Reserve Bank of Dallas, Globalization and Monetary Policy Institute Working Papers, 96.

- Garratt A., Vahey S.P. (2006) UK Real-Time Macro Data Characteristics. *Economic Journal*, 116, pp. 119–135.
- Gerberding C., Seitz F., Worms A. (2005) How the Bundesbank Really Conducted Monetary Policy. *The North American Journal of Economics and Finance*, 16, 3, pp. 277–292.
- Giannone D., Henry J., Lalik M., Modugno M. (2012) An Area-Wide Real-Time Database for the Euro Area. *Review of Economics and Statistics*, 94, 4, pp. 1000–1013.
- Giordani P., Söderlind P. (2003) Inflation Forecast Uncertainty. *European Economic Review*, 47, 6, pp. 1037–1059.
- Gornostaev D., Ponomarenko A., Seleznev S., Sterkhova A. (2022) A Real-Time Historical Database of Macroeconomic Indicators for Russia. *Russian Journal of Money and Finance*, 81, 1, pp. 88–103.
- Gornostaev D., Porshakov A., Seleznev S., Sterkhova A., Chernyadyev D. (2021) Revisions of GDP: Measurement and Consequences. *Bank of Russia Analytical Notes Series*. (In Russ.)
- Jurado K., Ludvigson S.C., Ng S. (2015) Measuring Uncertainty. *The American Economic Review*, 105, 3, pp. 1177–1216.
- Knetsch T.A. (2009) *The German Real-Time Database Gerda: Content, Structure, Search and Download*. Technical Documentation. Deutsche Bundesbank.
- Lee K., Morley J., Shields K., Tan M.S. (2019) *The Australian Real-Time Fiscal Database: An Overview and an Illustration of Its Use in Analysing Planned and Realised Fiscal Policies*. CAMA Working Papers, 14.
- Lee K., Olekalns N., Shields K., Wang Z. (2012) Australian Real-Time Database: An Overview and an Illustration of Its Use in Business Cycle Analysis. *Economic Record*, 88, 283, pp. 495–516.
- Lewis J. (2013) Fiscal Policy in Central and Eastern Europe with Real Time Data: Cyclical, Inertia and the Role of EU Accession. *Applied Economics*, 45, 23, pp. 3347–3359.
- Mamedli M., Shibaev D. (2021) *Forecasting Russian CPI with Data Vintages and Machine Learning Techniques*. Bank of Russia Working Paper Series, 70.
- McCracken M., Ng S. (2016) FRED-MD: A Monthly Database for Macroeconomic Research. *Journal of Business and Economic Statistics*, 34, 4, pp. 574–589.
- McCracken M., Ng S. (2020) FRED-QD: A Quarterly Database for Macroeconomic Research. *National Bureau of Economic Research*, 26872.
- Orphanides A., Norden S.V. (2002) The Unreliability of Output-Gap Estimates in Real Time. *Review of Economics and Statistics*, 84, 4, pp. 569–583.
- Seabold S., Perktold J. (2010) Statsmodels: Econometric and Modeling with Python. *Proceedings of the 9th Python in Science Conference*, pp. 57–61.
- Sleeman C. (2006) Analysis of Revisions to Quarterly GDP – A Real-Time Database. *Reserve Bank of New Zealand Bulletin*, 69, 1, pp. 31–44.
- Stock J.H., Watson M. (1996) Evidence on Structural Instability in Macroeconomic Time Series Relations. *Journal of Business and Economic Statistics*, 14, 1, pp. 11–30.
- Stock J.H., Watson M. (2012) Disentangling the Channels of the 2007–2009 Recession. *Brookings Papers on Economic Activity*, pp. 81–156.
- Stone A., Wardrop S. (2002) *Real-Time National Accounts Data*. Reserve Bank of Australia Research Discussion Papers, 5.
- Tebrake J. (2019) A Framework for Recording and Communicating Revisions. *Working Party on Financial Statistics*, OECD.