Technical Documentation

# The German Real-Time Database Gerda:

# Content, Structure, Search and Download

Thomas A. Knetsch

October 2009

# Abstract

The macroeconomic real-time database Gerda maintained by the Deutsche Bundesbank is a comprehensive and permanently up-to-date collection of historical vintages of German national accounts, short-term business and labour market indicators as well as price statistics. The first part of the paper documents the content of the database in terms of indicator coverage and vintage dimension. As useful complementary information, current revision practices and past benchmark revisions are sketched. The real-time database is available on the Bundesbank's website. The second part of the paper is written as a tutorial explaining the various search functionalities and download forms provided by the internet platform. For this purpose, it is useful to introduce the concept of codification used for the real-time database.



**Keywords:** real-time data, revisions, Germany.

**JEL classification:** C82, C88, E01.

# Contents

**1 Introduction** **1**

**2 Data content** **3**
2.1 National accounts (A) . . . . . 3
2.2 Short-term business indicators (I) . . . . . 6
2.3 Short-term labour market indicators (L) . . . . . 9
2.4 Price indices (P) . . . . . 10

**3 Database structure** **10**
3.1 Concept . . . . . 11
3.2 Codification . . . . . 12

**4 Search functionalities** **13**
4.1 Indicator trees . . . . . 14
4.2 Indicator search . . . . . 15
4.3 Hit-list and direct view . . . . . 17

**5 Download forms** **21**
5.1 Matrix format . . . . . 22
5.2 Snapshot format . . . . . 25

**References** **26**

# List of Figures

1 Real-time data matrix of quarterly real output of the German economy . . 11
2 Indicator tree of industrial production . . . . . 14
3 Indicator search of GDP . . . . . 16
4 Codelist of "Accounting system" . . . . . 16
5 Hit-list of a search for GDP . . . . . 18
6 Direct view in table form . . . . . 19
7 Direct view in chart form . . . . . 20
8 Download list of different GDP versions . . . . . 21
9 Download in matrix format . . . . . 22
10 Download in snapshot format . . . . . 25

# The German real-time database Gerda: Content, structure, search and download[0]

## 1 Introduction

Many economic data are subject to revision, which means that the results of economic indicators which statistical agencies publish for the same reporting period may change over time. The main reason for revisions is that early announcements are measured on the basis of incomplete statistical information. However, revisions may also occur when measurement concepts and practices are altered. The time series of economic indicators obviously lack information about past revisions, as they reflect the latest stage taken from available statistical sources. Real-time databases keep this information by storing all issues that were published in the past. In other words, they collect the time-series vintages of economic indicators, each of them identified by its publication date.

Real-time databases make it possible to analyze how prone economic indicators are to revision and to track the "revision history" of the results of individual periods, for example. It also allows a precise reconstruction of the information base at given points in time in the past. With either function, real-time data can be used, for instance, to study the revision processes of economic indicators, to check the performance of forecasting tools under conditions closely conforming to reality, and to evaluate past economic policy strategies and decisions given policymakers' information of that time.

In recent years, macroeconomic real-time databases have been established for a number of industrialized countries. Although they differ with regard to concept and coverage, the major examples share the characteristic of being maintained by a central bank or an international institution, at least if they are sufficiently large in scope. In the 1990s, the Federal Reserve Bank of Philadelphia started reconstructing historical vintages for key macro variables for the U.S. economy (Croushore and Stark, 2001), creating a prototype which has been extensively used by the academic community since then. The Bank of England publishes real-time data for the United Kingdom (Garratt and Vahey, 2006). Real-time data covering a number of European countries are presented on the member-restricted area of the website of the Euro Area Business Cycle Network (EABCN).[1] The OECD also maintains a multi-country real-time database.

[0]Corresponding address: Deutsche Bundesbank, Economics Department,Wilhelm-Epstein-Str. 14, 60431 Frankfurt am Main, Germany, email: thomas.knetsch@bundesbank.de. The real-time database project has been the joint work of the Economics Department, the Statistics Department and the Economic Research Centre of the Deutsche Bundesbank. Matthias Klimpel deserves gratitude for designing the key structure, while Julia Greb has developed the computer application. Permanent conceptual input on data processing has been provided by Reinhold Stahl. Irene Fenzlein, Gerhard Juracka and Mirjam Mathias have been involved in the designation of data and tools. The work has benefited from comments and suggestions by Heinz Herrmann, Robert Kirchner and Patricia Staab. The opinions expressed in this paper are those of the author and do not necessarily reflect the views of the Deutsche Bundesbank.

[1]The EABCN is a research network of European central banks.

In August 2009, the Deutsche Bundesbank made its macroeconmomic real-time database Gerda available online (Deutsche Bundesbank, 2009). The German version of the application can be accessed at http://www.bundesbank.de/statistik/statistik_realtime.php and the English version at http://www.bundesbank.de/statistik/statistik_realtime.en.php. Containing historical vintages of more than 280 economic indicators,[2] Gerda is the first general and comprehensive real-time database for the German economy.[3] It is permanently up-to-date, as newly incoming data are stored immediately after their publication. At the time of its launch, the database contained roughly 18,000 vintages. Owing to the current updating procedure, the content is growing rather rapidly, with an add-on of about 1,000 vintages per quarter. National accounts data at 25 percent and short-term business indicators at 55 per cent constitute the overwhelming part of the vintages stored in the database. Variables stemming from these two accounting systems usually attract most interest in revision analysis and economic studies using real-time data. The remaining part of the database is equally shared by vintages of short-term labour market indicators and price indices.[4]

The aim of this report is to provide a overview on all aspects of the real-time database in a concise form. First, the reader is informed of the set of economic indicators for which real-time data are available. This part includes information on how long vintages go back into the past and enumerates the benchmark revisions of the accounting systems which may be of relevance for data processing across vintages. These lists, however, are designed as brief memos. Detailed information on the nature of benchmark revisions is provided in indicator-specific metadata files which are found on the website. Secondly, the reader is introduced to the codification of the database. Although it presents more technical material, this part may be of interest not only to experts in statistical data processing. Understanding database configuration may also be helpful for users, as it may economize search practices. Finally, the report comprises a tutorial on search and download functions provided by the website application. This explanation is supplemented by a number of examples, not only pointing to the particular utility of the tool considered, but also acquainting the user with special characteristics regarding the publication pattern of German economic statistics.

The data content is outlined in Section 2. Section 3 presents a broad explanation of the formal structure of the database. Sections 4 and 5 provide a tutorial on available search functionalities and download forms.

[2]The indicator counter on the website shows a total number of more than 550 indicators. This discrepancy is explained by the fact that, in the website count, indicators are defined as being separate from each other if they possess different identifiers. This means that, in this formal statistical definition, a new indicator is created if, for instance, the raw data are corrected for seasonal and working-day variations or if real figures are measured in previous year's prices instead of the prices of a constant base year (see also Example 3.1).

[3]Real-time data of German economic indicators had been collected and published earlier (Gerberding et al., 2005; Jacobs and Sturm, 2005, for instance). In contrast to the new database, its predecessors typically referred to specific research projects, meaning that the variables, for which real-time data were collected, were chosen selectively. Thus, such databases are rather limited in scope and have not been kept up to date, as collection stopped after the project had been finalized. A real-time dataset of German national accounts is maintained by the Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI, 2007).

[4]The mentioned proportions do not refer to the current stock but to the quarterly increments. As the scope of the database increases dynamically, actual shares will converge to the shares in increments.

# 2 Data content

Gerda merges historical vintages of revision-prone macroeconomic indicators stemming from separate accounting systems. At the current juncture, the database comprises national accounts data, short-term business cycle and labour market indicators as well as price indices.

During the initial collection process, the database was expanding simultaneously in two directions. First, incoming releases of the comprehensive list of economic indicators were continuously recorded by the computerized updating procedure (which is still in action to fill the database at the current end). Second, historical vintages of a limited set of key variables were reconstructed on the basis of printed or electronic documents. The aim was to increase the vintage dimension to an extent necessary for carrying out meaningful real-time data analysis. While the data collected by the computer-automated recording fulfils the highest quality standards, the data recorded manually was double-checked, without making any claims for their accuracy in any case.

As a consequence of the two-sided expansion, Gerda is the real-time data researchers' pet, which is an animal characterized by a short and fat body and a long and lean tail. The body comprises the vintages, which have been recorded automatically for a comprehensive set of indicators. However, as the computerized updating procedure was established in 2005, the vintage dimension of the fat indicator set is short. In contrast, the historical vintages prior to the outset of the updating procedure have had to be recorded from the data sources available in printed and electronic formats. The retrograde recording has been conducted only for a limited number of key indicators, but back into the past as far as possible and reasonable. This subset constitutes the long and lean tail of the animal.

The remainder of this section presents an overview of the content of the real-time database including a description of current publication and revision practices as well as an enumeration of benchmark revisions which occurred in the covered periods of time. The outline is organized according to the accounting systems included in the real-time database.[5]

## 2.1 National accounts (A)

Approximately one-third of Gerda's content is data from the national accounts. The updating procedure took off in May 2005 when the volumes were first published as chain-weighted indices. Since then, the quarterly national accounts statistics are widely reproduced in the real-time database, including the central economic aggregates such as gross domestic product (GDP) and gross national income, the main expenditure components such as private and government consumption, gross fixed capital formation as well as export and imports, the sectoral breakdown of gross value added, and the distribution of national income. Conditional on public availability, these economic entities are found in nominal and real terms

[5]The letters in parentheses provided by the headlines of the subsections denote the GESMES/TS real-time database code of the key component "Accounting system".

as well as with and without seasonal and working-day adjustment. Data on the number and the hours worked of persons employed and employees are also provided—both as total-economy aggregates and in sectoral breakdown. The same structure of data is available for the compensation of employees as well as gross wages and salaries, while net wages and salaries are given only as total-economy aggregates.

The set of economic indicators, for which vintages have been recorded manually, is clearly less comprehensive, with the scope of retrograde recording differing from one indicator to another. Moreover, availability is restricted to seasonal and working-day adjusted data. A natural limit for the vintage dimension is given by the year 1995 when, for the first time (namely in September 1995), the national accounts were published for Germany as a whole. Historical vintages of GDP and the main expenditure aggregates are available in real terms, seasonally and working-day adjusted, back to this date.

Referring to western Germany as a territorial basis, the database comprises vintages back to the publication date of April 1991 for the main expenditure aggregates, once again only in real terms and adjusted format. The GDP vintages start in June 1992, however. Prior to this date, the gross national product (GNP) was seen as the headline measure of aggregate output and, thus, reported as standard in the Bundesbank's economic and statistical publications (Deutsche Bundesbank, 1992, p.27). The GNP vintages even go back to December 1971, with the publication date having been exactly relatable since March 1983. Before, vintages could be assigned only to their publication month.

The current publication practice is to present a first release of quarterly GDP including the total-economy figures of persons employed and employees about six weeks after the end of the reporting period, while the complete national accounts follow approximately ten days later. Just one or two days before the latter date, however, the number and the hours worked of the persons employed as well as the hours worked of employees are released in a sectoral breakdown and, if this has not already happened, also as total-economy aggregates. This multi-segmented schedule was established in May 2003, while there had been only a single publication date prior to this. Starting with the year 2000, the publication lag of the first release was reduced by a fortnight. Traditionally, the first release of the annual national accounts is presented in January, just a few days after the end of the reporting period.

In general, the national accounts of the current and the past three years may be revised at every publication date, as they possess a provisional status. In terms of revision practice, however, the summer release is most important because a comprehensive update of the preliminary data is regularly presented at that date as a consequence of incorporating lagging information about, for instance, the value-added tax statistics and the cost structures of firms. The data which dates back about three years is thereby given a final status. Apart from the summer revision, there might also be revisions at the remaining regular publication dates during the year. These are typically small in magnitude, however, and refer to very recent data only.

Although some benchmark revisions have already been mentioned in the previous description, it is worthwhile surveying them again briefly in chronological order, starting from the latest and going back to the early 1990s:[6]

- **As of May 2005**, the volumes have been published as chain-weighted indices, measured in previous year's prices with reference year 2000. Before that date, the volumes were expressed in constant prices of the base year 1995. Apart from this fundamental conceptual change, the accounting of financial intermediation services has been changed, first, by indirect measurement and, secondly, by assignment to the relevant end users (FISIM). Moreover, improved methods for the measurement of the quality change in price statistics (hedonic approach) has been incorporated (Braakmann et al., 2005; Deutsche Bundesbank, 2005).

- **As of May 1999**, the German national accounts have adapted the accounting standards harmonized within the European Union and set out in the *European System of Accounts 1995* (ESA 1995). This benchmark revision was comprehensive, not only altering concepts, classifications, sources and methods, but also having a substantial bearing on GDP and GNP. The most important change referred to investment, which was conceptually reorganized and extended (the introduction of the category "other investment" comprising intangible assets such as software, for instance). Moreover, public investment in infrastructure was integrated into the basis of fixed assets subject to depreciation, and the sectoral breakdown of gross value added was started being measured at basic prices, while the price concept had previously been market prices less value-added and import taxes. Last but not least, the base year changed from 1991 to 1995 (Strohm et al., 1999).

- **As of September 1995**, the national accounts were published for Germany as a whole and no longer separately for western and eastern Germany (Deutsche Bundesbank, 1995).

- **In September 1993**, the base year switched from 1985 to 1991.

- **In April 1991**, the base year switched from 1980 to 1985. Further conceptual changes affecting the raw data had a rather limited impact on the broad economic aggregates (Lützel, 1991). For GNP and its expenditure components, the Bundesbank introduced a calendar adjustment. Hence, the quarterly data have since been published in seasonal and working-day adjusted terms, while they had only been corrected for seasonal variations previously (Deutsche Bundesbank, 1991).

[6]The following list should not be regarded as a comprehensive description of benchmark revisions. It only highlights the main features appearing to have sufficiently significant qualitative and quantitative effects on the economic indicators contained in the real-time database.

## 2.2 Short-term business indicators (I)

Short-term business indicators give a timely and detailed view of the cyclical stance of the production and trade sectors at monthly frequency. The set of indicators group into production, orders received and turnover indices. Production is measured in real terms, orders received in real and nominal terms. As far as the production sector is concerned, turnover is available only in nominal terms.

**Indicators for the production sector.** The production sector is divided into industry, energy and construction. The industrial sector is further broken down into the main industrial groups, i.e. producers of intermediate goods, producers of capital goods and producers of durable and non-durable consumer goods. Construction indices capture only the main construction industry, which is subdivided into structural and civil engineering. This breakdown is not only found in the output measures but also for the orders received and the turnover indicators. For the industrial sector, the latter two categories, however, are divided into flows from the domestic and the foreign market, with the orders and the turnover from abroad being further differentiated into intra and extra EMU. Orders and turnover of the construction sector are reported separately for housing and in a breakdown by the type of clients, i.e. industrial clients and public sector clients.

Approximately one-half of the time series collected in Gerda belongs to the group of short-term business indicators. The automatic updating procedure was established in November 2005. From this date onwards, the historical vintages of the above-mentioned indicators are available. An exception is the intra/extra EMU breakdown, which was released for the first time in March 2006. The manual recording of vintages published prior to November 2005 focused on key indicators only. This lean set comprises the production indices for industry (or the manufacturing sector before September 2001), energy and construction as well as the main industrial groups. Orders are reported for industry as a whole and in the broad sectoral and regional breakdowns.

The short-term business statistics of the production sector obey a revision scheme which consists of a quarterly and an annual revision. This was established in March 1999, when a new survey method was introduced with the aim of relieving the enterprises' statistical workload. Under the new concept, the reporting sample of the production survey is divided into mutually exclusive quarterly and monthly reporting parties. As of March 2007, local units with at least 50 employees are obliged to report monthly, while units falling short of this limit but employing more than 20 persons have to report quarterly. This rather simple selection criterion replaced a more complex system at the cost of reducing the coverage.[7] Preliminary figures are released regularly about five weeks after the end of the reference period. The quarterly revision occurs roughly two and a half months after the end of the reporting quarter when the information from the monthly and the quarterly reports are

[7] Up to February 2007, the largest units of the economic sectors in each of the *Bundesländer* are obliged to submit a monthly production report, so that, at the federal state level, at least 75 percent of the output of the firms with 20 or more employees is covered. This provision results in national coverage in excess of 80 percent.

collated. Indices are given final status in the annual revision, which coincides with the quarterly report of the fourth quarter (Herbel and Weisbrod, 1999).[8]

Output indices of the main construction industry are revised regularly when the annual survey of the complete construction sector is available. The revision usually occurs at the beginning of a year and refers to the data of the recent summer semester. The orders received by the main construction industry are published about two weeks later than their industrial counterparts. The raw indices, however, are immediately given final status. Since May 2005, preliminary announcements of construction output have been adjusted in advance by a factor which represents the expected correction of the total annual survey.

Since 1995, there has been further benchmark revisions in the short-term statistics of the production sector, apart from those already mentioned. The list below summarizes them in chronological order, starting with the most recent:

- **In March 2009**, the indices of production, orders inflow and turnover were collected for the first time according to the principles of the pan-European classification of economic sectors NACE Rev 2 and the 2009 edition of the European product classification.[9] This revision had only a marginal impact on the production sector. In addition, the monthly survey in industry and mining were modified, resulting in a reduction of coverage which is substantive in some subcomponents of orders statistics. This benchmark revision also included a base year changeover from 2000 to 2005 (Bald-Herbel, 2009).

- **In March 2003**, the base year changed from 1995 to 2000 for the orders received and turnover indices. The production indices followed **in July 2003**.

- **From September 2001**, a standard EU definition of the main industrial grouping was introduced. First, the existing manufacturing sub-aggregates (intermediate goods, capital goods, durable and non-durable goods producers) were retained under the same names but reclassified. Second, with "energy", a new aggregate was introduced. In addition to electricity, gas, steam and hot water supply, it comprises certain sectors in mining and quarrying sectors as well as a small part of manufacturing (in particular, the manufacture of refined petroleum products). The designation "industry" was chosen for the sum of the production sectors excluding construction and energy. The output of this aggregate, as the successor to manufacturing output, has since been the focus of the Bundesbank's current economic analysis (Deutsche Bundesbank, 2001).[10]

[8] It is worth mentioning that preliminary figures were promptly corrected approximately 20 days after their initial issue in order to incorporate information from delayed monthly reports. However, a methodological change in the treatment of missing reports, which was established in March 2005, significantly improved the quality of the first releases (Statistisches Bundesamt, 2005). As a consequence, the prompt adjustment became superfluous and was eventually abandoned one year later.

[9] NACE Rev 2 is the acronym for *Nomenclature général des activités économiques dans les Communautés européennes revision 2*.

[10] In quantitative terms, the reclassification mainly affected the sub-aggregates whose shares in gross value added of the production sector has shifted by up to 5 percentage points. By contrast, the changeover

- **In March 1998**, the indices of output, orders received and turnover were rebased on the year 1995. This complied with the European economic regulation to introduce a new base year within a period of three years after the end of a year ending in a zero or a five.

- **As of June 1995**, the production statistics in mining and manufacturing were collected for the first time in line with the pan-European NACE and the PRODCOM regulations, requiring a comprehensive reclassification of economic sectors and products (Nowack and Weisbrod, 1995).[11] Their introduction led to major structural changes, which meant that even the results at quite a high level of aggregation were no longer comparable with and without the switch, even though the data of the years 1991 through 1994 were back-calculated.[12] As part of this benchmark revision, the base year was changed from 1985 to 1991. Furthermore, the Bundesbank also published seasonally adjusted economic indicators for Germany as a whole for the first time. **Up to May 1995**, indicators referred to western Germany and were presented in accordance with the German industry classification SYPRO.[13]

**Indicators for the trade sector.** An important part of short-term business statistics are monthly indices of the turnover in the trade sector, where the focus is on retail trade and trade with motor vehicles. The indicators are generally available in current and constant prices as well as with and without seasonal and working-day adjustment. Vintages have been stored from November 2005 onwards. It is worth mentioning that retail turnover indices are published regularly twice a month. Around the turn of the month, there is a first announcement for the reporting period finalized about 30 days ago. A prompt revision within the same (publication) month takes place about 15 days later.[14]

The introduction of NACE Rev 2 in trade statistics in March 2009 called for a redefinition of indicators which include trade with motor vehicles. Through this benchmark revision, only the "pure" retail trade sector has remained more or less unchanged in conceptual terms.[15] The broader aggregate of retail trade including retail trade of motor vehicles and petrol stations, however, could not be maintained after February 2009 because car sales were no longer distinguishable between retail and wholesale. As of March 2009, the database comprises the total sales and repair of motor vehicles in nominal terms as an

from the aggregate "output in the manufacturing sector" to "industrial output" is unlikely to have had a noticeable impact in terms of either quality or quantity on the given cyclical picture as industry, at 79.7 percent compared with 79.5 percent in the case of manufacturing, had only a marginally larger share in the value added of the production sector in 1995.

[11] PRODCOM is the abbreviation of *Products of the European Community*, defining an EU-harmonized classification of products.

[12] The back-calculation could not entirely prevent a statistical break in the time series. See Knetsch and Reimers (2006) for details.

[13] SYPRO is the acronym for *Systematik der Wirtschaftszweige, Fassung für die Statistik im Produzierenden Gewerbe*.

[14] For the handling of this irregularity in real-time data export, see Example 5.2.

[15] The turnover of petrol stations is now part of retail trade, while it was previously imputed to retail trade of motor vehicles.

individual indicator and the merger of this with retail turnover. This benchmark revision also includes a base year changeover from 2003 to 2005.

Substantial revisions in retail turnover data have been caused by the treatment of new firms in the survey. Since existing reporting samples usually shrink rapidly owing to drop-outs, it is current practice to replenish the sample of firms every year. In April 2007, retail turnover indices for the year 2006 including January and February 2007 were published for the first time on the basis of an adjusted sample, creating a statistical break in the time series between December 2005 and January 2006. This break characterizes all vintages issued between April 2007 and February 2008. As of March 2008, retail turnover is presented in the form of a chain index, connecting the levels before and after the adjustment of the reporting sample. As a consequence of the new treatment, the statistical break between December 2005 and January 2006 has disappeared in all vintages published since March 2008.

## 2.3 Short-term labour market indicators (L)

Short-term labour market indicators are monthly data on employment, unemployment, hours worked, gross wages and salaries as well as (negotiated) pay rates on a monthly and hourly basis. The data are taken from several statistical sources. While most indicators refer to the total economy, some of them are also available for the industrial sector and the main construction industry. Real-time data are collected for indices with and without seasonal and working-day adjustment except pay rates, for which adjusted data is generally not available.

In particular, Gerda comprises the monthly statistics of persons employed (domestic concept) and employees subject to social security contributions. It also includes the number of unemployed persons and the unemployment rate as defined by the concept of the International Labour Organization (ILO). For the industrial sector and the main construction industry, data on the number and the hours worked of employees are available. Gross wages and salaries are found in the same sectoral structure, while pay rates are provided as total-economy aggregates and for the production sector including construction.

For employment and unemployment, the computer-automated vintage collection started in August and September 2005 respectively. The sectoral data on the number and the hours worked of employees as well as on gross wages and salaries has been recorded since November 2005. The vintages of employees subject to social security contributions are available from March 2006 onwards. Total pay rates, which include ancillary benefits and one-off payments, have been collected since November 2005. Pay rates excluding one-off payments and basic pay rates, which additionally correct for ancillary benefits, are available since they were initially issued in August 2007 (Deutsche Bundesbank, 2007).

As the data summarized under the headline of short-term labour market indicators stem from various primary statistics, publication schemes and revision practices differ from one indicator to another. The number of persons employed in the economy is first published about one month after the reference period on the basis of incomplete information. In the aftermath of this preliminary release, further statistical sources become successively

available, leading to revisions which are substantive at the beginning but tend to decay in magnitude as time goes by. After three years, the indicator can be regarded as final (Fritsch, 2006).

Sectoral data on employees and on gross wages and salaries are taken from the monthly and quarterly surveys in the production sector, implying that the revision schedule is similar to that of the short-term business indicators (see Section 2.2). The employees subject to social security contributions and the statistics on hours worked are edited by the Institut für Arbeitsmarkt- und Berufsforschung (IAB). The indices of negotiated wages and salaries are published in the Monthly Report of the Deutsche Bundesbank. ILO unemployment data are revision-free because they are gathered from a telephone survey.

### 2.4 Price indices (P)

Price indicators condense the information of wide spread and detailed price observations. As surveys are complete prior to the first announcement, regular revisions do not emerge in price statistics. Thus, Gerda is predominantly informative on the timing of vintage publication. Benchmark revisions, however, even alter price indices. Base year changeovers reflect adjustments in weighting schemes because underlying commodity baskets change. Methodological and procedural innovations are often implemented in the context of base year changeovers. The base year was switched to 2005 in February 2008 in the case of consumer prices (Elbel and Egner, 2008; Deutsche Bundesbank, 2008), in September 2008 in the case of construction prices (Vorholt, 2008) and in March 2009 in the case of producer prices as well as export and import prices (Peter, 2009).

The set of price indices available in the real-time database group into consumer prices, producer prices as well as export and import prices. The consumer price index is published with and without the energy component. Further subcomponents are those for food products, consumer durables and non-durables excluding energy and food, services excluding house rents, and house rents. The database contains producer prices of industrial goods with and without energy as well as a construction price index and a price index of farm products. External trade prices capture all products traded.

The price indicators are measured monthly, except the construction price index which is available only quarterly. With the same exception, raw data is complemented by seasonally adjusted indices. The seasonal adjustment of the consumer price index as well as its energy price and services price subcomponent additionally include a correction for working days.

## 3 Database structure

In general, the architecture of a macroeconomic real-time database differs from usual time series databases in two respects. First and most obviously, the time series of an economic indicator issued at distinct points in time are regarded as different entities. Hence, they have to be stored separately, assigning discriminating identifiers to them. Second, a real-time database contains a set of indicators which is heterogenous from a statistical point of view because the proneness to revision is a characteristic shared by many economic

Figure 1: Real-time data matrix of quarterly real output of the German economy

| Statistical indicator definition | Quarterly GDP, seasonally and calendar-adjusted, at constant prices | | | | | | | Quarterly GDP, seasonally and calendar-adjusted, at previous year's prices | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Database key | Q:DE:Y:A:AG1:CA010:C:A | | | | | | | Q:DE:Y:A:AG1:CA010:A:I | | | |
| Publication date | 6.9.1995 | 7.12.1995 | ... | 4.3.1999 | 8.6.1999 | ... | 15.2.2005 | 12.5.2005 | ... | 15.5.2009 | 13.8.2009 |
| Base year | 1991 | 1991 | | 1991 | 1995 | | 1995 | – | | – | – |
| Measurement unit | DM bn | DM bn | | DM bn | € bn | | € bn | 2000=100 | | 2000=100 | 2000=100 |
| 1991 Q1 | 709.4 | 709.4 | | 713.3 | 428.56 | | 429.86 | 86.04 | | 85.95 | 85.94 |
| 1991 Q2 | 717.3 | 717.3 | | 716.4 | 428.97 | | 428.43 | 85.32 | | 85.30 | 85.31 |
| ... | | | | | | | | | | | |
| 1995 Q1 | 750.1 | 750.1 | | 746.2 | 447.89 | | 447.23 | 89.77 | | 89.82 | 89.81 |
| 1995 Q2 | 758.1 | 758.1 | | 751.1 | 451.37 | | 451.26 | 90.77 | | 90.73 | 90.73 |
| 1995 Q3 | | 758.0 | | 748.9 | 451.57 | | 451.16 | 90.92 | | 90.86 | 90.87 |
| 1995 Q4 | | | | 749.3 | 451.01 | | 451.14 | 90.67 | | 90.80 | 90.80 |
| ... | | | | | | | | | | | |
| 1998 Q3 | | | | 800.2 | 472.99 | | 468.18 | 94.88 | | 94.96 | 94.97 |
| 1998 Q4 | | | | 797.2 | 472.33 | | 468.67 | 94.85 | | 94.78 | 94.79 |
| 1999 Q1 | | | | | 474.40 | | 473.56 | 95.84 | | 95.83 | 95.82 |
| 1999 Q2 | | | | | | | 472.39 | 95.71 | | 95.78 | 95.78 |
| ... | | | | | | | | | | | |
| 2004 Q3 | | | | | | | 502.77 | 102.62 | | 101.80 | 101.85 |
| 2004 Q4 | | | | | | | 501.60 | 102.48 | | 101.83 | 101.84 |
| 2005 Q1 | | | | | | | | 103.55 | | 102.00 | 101.98 |
| 2005 Q2 | | | | | | | | | | 102.72 | 102.57 |
| ... | | | | | | | | | | | |
| 2008 Q4 | | | | | | | | | | 107.61 | 107.74 |
| 2009 Q1 | | | | | | | | | | 103.52 | 103.93 |
| 2009 Q2 | | | | | | | | | | | 104.26 |

Source: Deutsche Bundesbank (2009).

statistics. However, not all indicators which are subject to revision face the same attention in a real-time context. A special database for this purpose is thus not a merger of revision-prone economic statistics but an appropriate selection of indicators which are both subject to revision and sufficiently interesting from the standpoint of real-time data research.

## 3.1 Concept

Gerda is conceptually quite flexible, allowing for data export which can be tailored to the various purposes for which macroeconomic real-time data are usually used. In general, each vintage of the time series of an economic indicator has a unique identifier, representing the defining characteristics of the stored vector of figures. In the context of a real-time database, the publication date is an identifying criterion, too. The exact publication date is thus part of the key structure.

The remaining part of the key structure comprises components identifying economic indicators. In order to ensure the broadest applicability of the database, economic entities are defined according to their exact statistical convention. In the case of fundamental benchmark revisions concerning key components, this choice implies that vintages regarded as congruent from an economic point of view may differ not only in the publication date but also in other key components (see Example 3.1). The disadvantage is that the full revision history of some economic entities cannot be retrieved by downloading the full

real-time data matrix which is stored for a single (statistically defined) indicator. Under these circumstances, the vintages of two or even more indicators have to be merged after individual downloads. In this respect, it may be useful to consult the metadata sheets which provides supplementary information in order to find economically congruent, but statistically unequal, vintages. Knowledge of fundamental benchmark revisions is often very helpful for real-time data research. In particular, the empirical results of revision analysis may be strongly affected by whether or not benchmark revisions are dealt with appropriately (Knetsch and Reimers, 2009; Lorenz, 2008, for instance). Further useful information, such as base-year shifts which cannot be taken from the identifier, are delivered in the form of vintage-specific attributes.

**Example 3.1** Since May 2005, the economic entity "quarterly real output of the German economy" has been measured by a chain index of GDP in previous year's prices, while, before, it was measured by a volume in constant prices of a specific base year. As the concepts differ from each other in statistical terms, they possess different identifiers in the real-time database. Consequently, two data exports (namely those of the two statistical concepts) are required to create the real-time data matrix of the above-defined economic entity shown in Figure 1.

## 3.2 Codification

Codification is based on GESMES/TS (abbreviating the term *Generic Statistical Message for Time Series*), which has become the preferred method in international data exchange. As the indicators contained in Gerda stem from rather heterogenous statistical sources, a special key family has been developed in order to keep the coding system as concise as possible. In particular, the key structure consists of the following 11 key components:

1. Frequency
2. Reference area
3. Adjustment indicator
4. Accounting system
5. Indicator category
6. Indicator identification
7. Price or unit reference
8. Calculation basis
9. Publication date: year
10. Publication date: month
11. Publication date: day

The first eight key components identify the economic indicator, while the remaining three specify the publication date.[16] The codelists of the first three key components ("Frequency", "Reference area", "Adjustment indicator") are defined in accordance with international conventions. The components "Price or unit reference" and "Calculation basis" are also found in a number of other GESMES/TS key families. However, the codelists of these key components are Gerda-specific.

The components "Accounting system", "Indicator category" and "Indicator identification" are tailored to the purpose at hand. In technical terms, they form a block in which the relational structure, which still exists between all other key components and between these components and the block as a whole, is abandoned. Instead, the three components are organized hierarchically. In this structure, subordinate key components are given codelists which are conditional on the superior key components (see Example 3.2). The hierarchy is introduced in order to keep the key structure sufficiently concise despite the heterogeneity amongst economic indicators.

**Example 3.2** In the national accounts section of the database (set A in "Accounting system"), value added, employment, hours worked and so on, which can be picked in the "Indicator category", are available both as total-economy aggregates and in a sectoral breakdown. The latter can be specified in the "Indicator identification". However, when "Consumption" is chosen in the "Indicator category", for instance, the codelist of the "Indicator identification" comprises "Private households including non-profit institutions serving households" and "General government". After having set I in "Accounting system", short-term business cycle indicators are found in the codelist of the "Indicator category", e.g. IP1 for "Production", IO1 for "Orders received", IT1 for "Turnover". In the case of "Production" and the various forms of "Orders received", the codelist of the "Indicator identification" contain the main industrial groups and components of the construction sector. In the case of "Turnover", several aggregates of retail trade are additional members in the codelist of the subordinate component.

# 4 Search functionalities

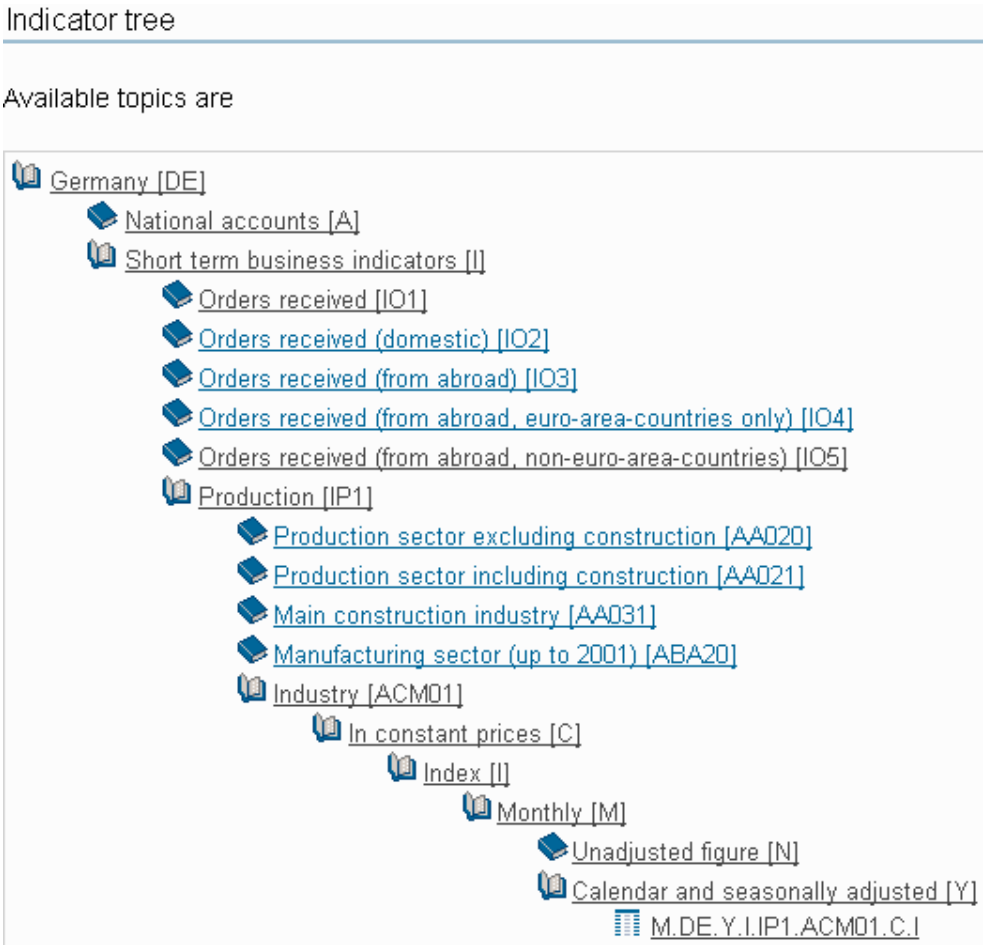
For convenience, the search for time series, which are individual vintages of an economic indicator in a real-time database, is separated into two stages. The search functionalities as such refer to the selection of economic indicators only, while vintages are chosen within download options (see Section 5).[17]

Indicators are found via predetermined indicator trees, free text search or an indicator search form. In general, the first two options may be more appropriate for less experienced users. At the cost of requiring some knowledge in general economic statistics, the indicator search offers more flexibility, as it allows users to search along each dimension of the

[16] Owing to the separation of search functionalities and download forms in the website application, users are never confronted with full identifiers. In hit-lists and download lists as well as the direct view and data output, indicators are generally abbreviated by their eight-key-components identifier.

[17] Technically speaking, only the first eight key components of the full GESMES/TS code are determined at maximum after the search process has been completed.

Figure 2: Indicator tree of industrial production




key structure, allocating either one or more codes to key components and/or keeping key components unspecified.[18] Even in this case, however, the platform provides enough assistance to be self-explanatory, for instance, by offering codelists for each key component which presents the eligible members as a formal code with a verbal description.

Apart from the option of performing the systematic search, the indicator search form also provides a free text field where a key word search can be carried out.

## 4.1 Indicator trees

An intuitive access to the content of the real-time database is provided by a search along indicator trees. Users are guided through the (indicator-related) key components step-by-

[18] It is necessary to fill in just one key component for the system to present a hit-list of the time series available in the database.

step and in a pre-specified way. At each node, available codes are presented as branches and twigs, with the option of selecting one of them. The tree structure deviates from the ordering of components in the key structure in order to economize search and to enhance the readability of trees. In particular, the simplest choices are located at the very beginning which, in turn, usually restrict the set of available indicators the most. Search via indicator trees avoids the risk of being disturbed by acronyms.

The tree's root is the question on the reference area. After opting for either "Germany" or "West Germany", the user is then asked to broadly define the economic indicator by specifying the three hierarchical key components "Accounting system", "Indicator category" and "Indicator identification", where available codes are often numerous but easily understandable. Finally, the platform guides the user through the four remaining key components, asking for indicator characteristics which, at first glance, may be less familiar to an unexperienced user. However, as eligible nodes are usually confined only to a few members at these stages, selection should not be too complicated. Figure 2 displays the indicator tree that has to be climbed when searching for the monthly index of German industrial production in constant prices, seasonally and working-day adjusted.

## 4.2 Indicator search

More experienced users may prefer to choose the indicator search form (see Figure 3) because, in this environment, key components can be filled by codes in quite a flexible manner. In contrast to indicator trees, search is constrained neither by the pre-specified ordering of key components nor by the restriction of selecting precisely one code. By pressing "start key component search", the user can always make the system create a hit-list of indicators which meet the determined search criteria. Of course, the more key components are determined, the shorter is the resulting hit-list. The number of hits for the given specification is presented on top of the search form. With every new entry, this piece of information is updated.

In general, it is possible to fill the fields directly by typing in the desired codelist members and pressing "apply" to save the entries. This practice, however, requires familiarity with the codification. The standard approach is therefore to select codes via codelists which can be displayed by pressing "show list" for the key component of interest. The list contains the eligible members as acronyms accompanied by a verbal description (see Figure 4). Selection means marking the squares to the left of the entries and ultimately pressing "submit" to transfer the chosen codes to the search form. In general, more than one code can be selected per key component.

An effective help in the search process is that the codelist of the key component, which is currently being filled in, is conditioned on the predetermination of other key components. Hence, users are never disturbed by codelist elements which are not relevant at the given stage of the search process. Note that irrelevance refers not only to conceptual impossibility (recall Example 3.2) but also to data non-availability (see Example 4.1).

Figure 3: Indicator search of GDP

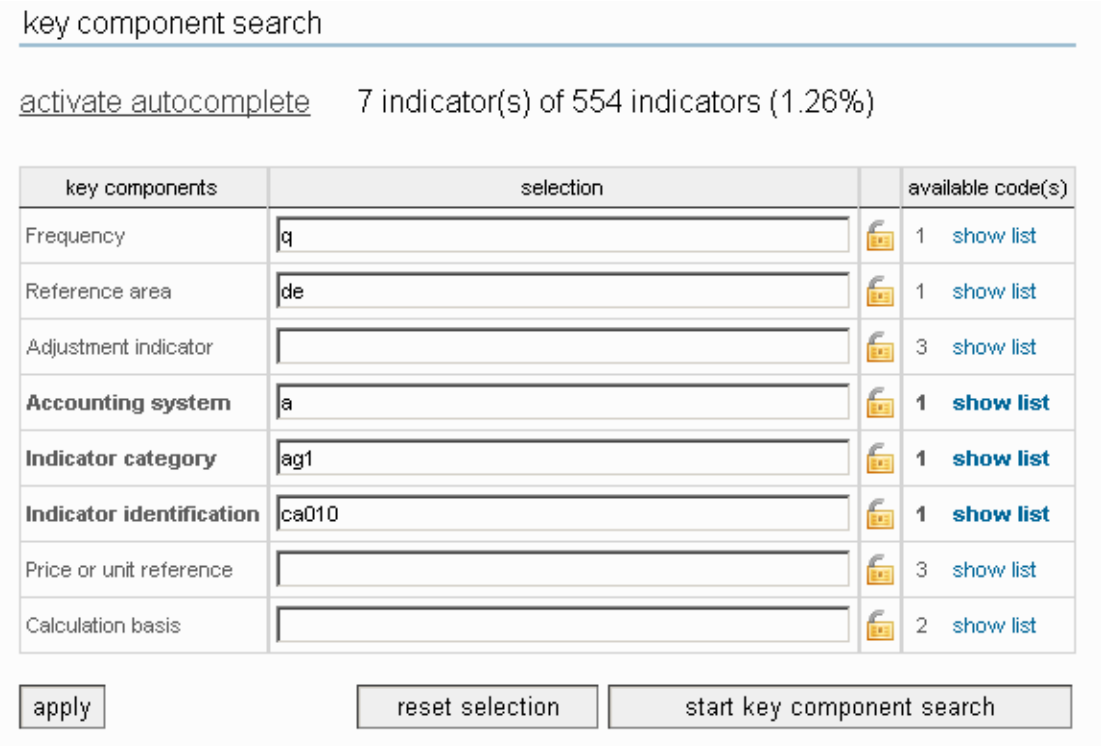
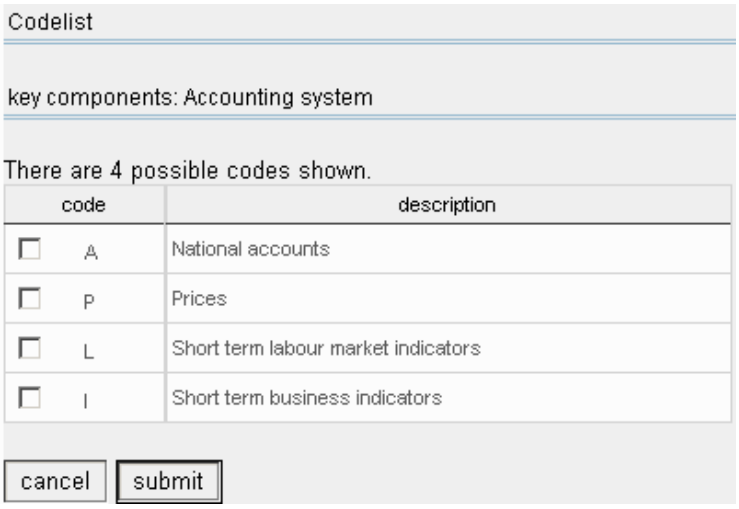



Figure 4: Codelist of "Accounting system"



**Example 4.1** The codelist of the reference area is based on the ISO country codes coding countries, groups of countries (such as the euro area), regions, continents and so on. The real-time database comprises only indicators for Germany with country code DE and West Germany with country code R1. The vast number of irrelevant elements which are actually not eligible are not displayed in the codelist.

The search process is further accelerated if the autocomplete option is activated. This is done by pressing "activate autocomplete". Under this mode, the system will automatically fill in all remaining key components for which the intermediate selection results in codelists with only a single member (see Example 4.2). However, after any system-induced determination of key components, it should nonetheless be examined whether or not the automatic entries are desired in the sense that they do not contradict the properties of the indicators which are intended to be found.

**Example 4.2** In developing the identifier for GDP, the user may first enter A in "Accounting system" and, second, AG1 in "Indicator category". If autocomplete is activated, the system will automatically set CA010 in "Indicator identification", meaning "Overall economy (domestic concept)", because this is the only member of the codelist which is eligible.

Users may often be interested in downloading real-time data of more than one indicator. If they share commonalities in terms of key components, iterated selection need not always restart from the very beginning if entries common to some or all indicators have been fixed in the first iteration. For each key component, this can be done by activating the lock symbol (through a click on the open-lock symbol ). A fixed entry is not erased when "reset selection" is pressed.

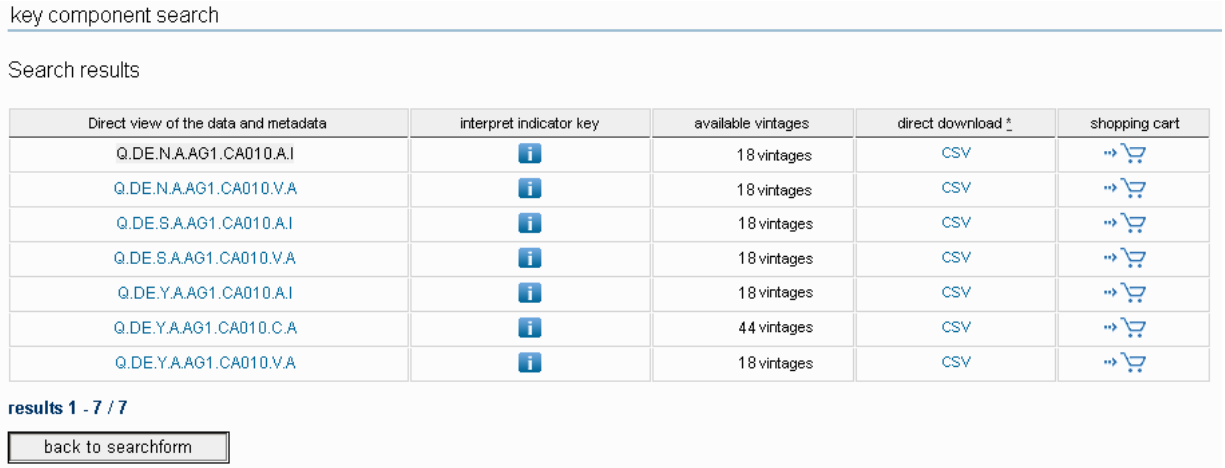
**Example 4.3** The aim is to snapshot the time series of quarterly real output, consumption and investment of the German economy (in seasonal and working-day adjusted terms), as they stood on 1 July 2006. The identifiers of the three indicators differ only in "Indicator category" and "Indicator identification". During the first search iteration, say for GDP, the user should fix the entries of all key components except those mentioned above. After GDP has been fully specified and written in the indicator list, "reset selection" can be pressed to erase all entries except those which has been kept locked. In this case, refilling "Indicator category" and "Indicator identification" is sufficient to specify the codes of consumption and investment.

## 4.3 Hit-list and direct view

After triggering "Start key component search", the results of indicator search are summarized in a hit-list (see Figure 5). This lists all indicators meeting the criteria determined in the search form.[19] For the sake of readability, displayed tables are restricted to 20 entries per page. In the case of longer hit-lists, tables are separated. Apart from providing information on the vintages available for the selected indicators, hit-lists offer direct view and CSV download as well as a shopping cart function.

[19]The search along the indicator tree also ends up with a hit-list. Of course, the hit-list contains only a single entry in this case.

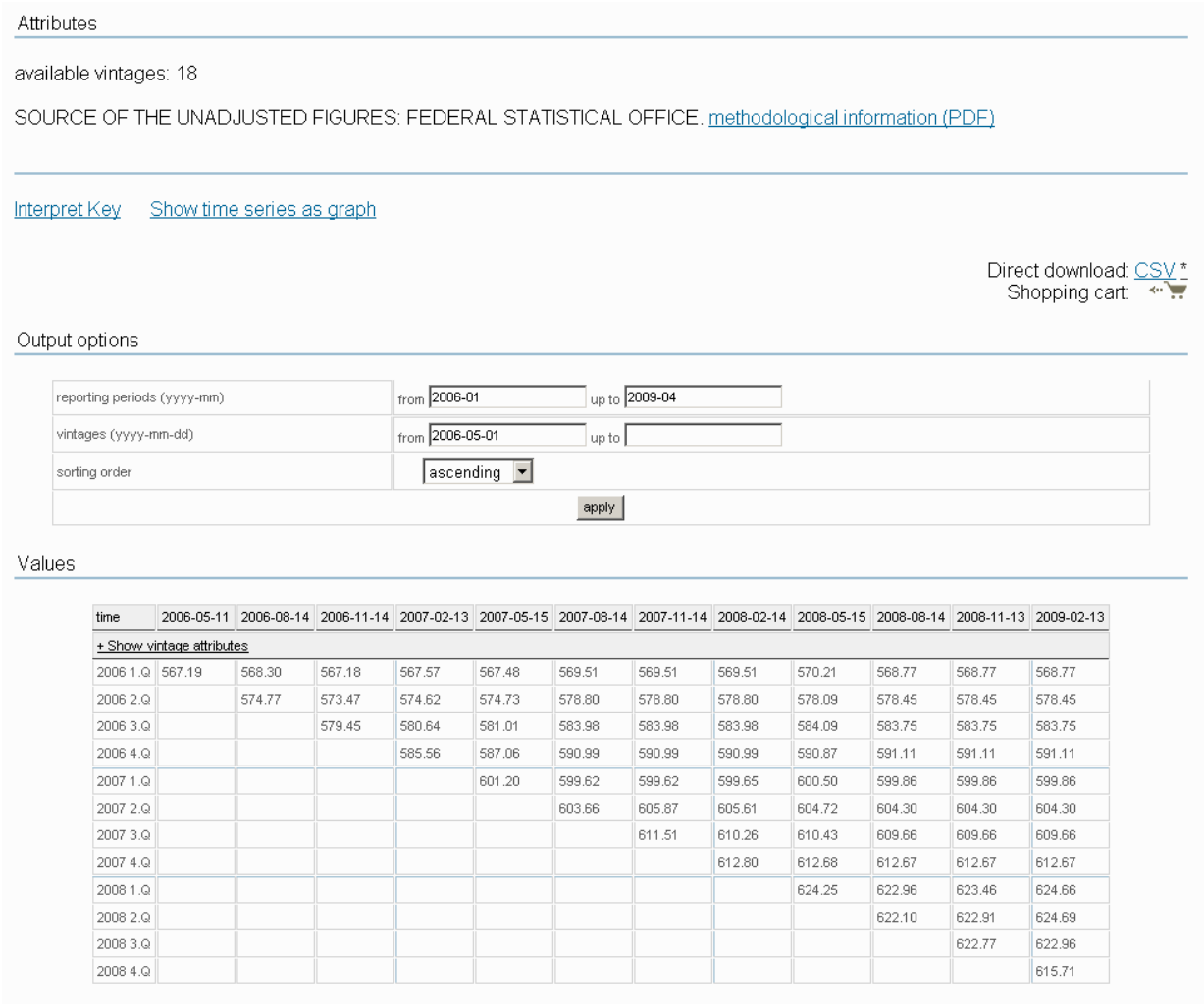
Figure 5: Hit-list of a search for GDP




In order to consider the scope and structure of the real-time data of an economic indicator, users may click on its identifier sorted in the left column of the hit-list. As a standard, the display then shows the full real-time data matrix. However, there are options for altering the range of displayed reporting periods and vintages as well as the ordering of the matrix. In addition, information on indicator and vintage attributes, data sources and benchmark revisions are provided. Detailed statistical metadata are tabulated in a separate text file. Figure 6 illustrates the direct view where the real-time data matrix is strongly tailored in terms of reporting periods and vintages. Figure 7 shows the corresponding display in chart form which can be produced by clicking "Show time series as graph" on the direct view display.

From both the hit-list and the direct view, users are able to initiate the direct download or use the shopping cart function. While the direct download allows users to export the real-time data available for the selected indicator in CSV format, the shopping cart function has to be applied in order to exploit the more sophisticated download options. In particular, an active shopping cart symbol signals that the indicator in question is selected for further processing in the download form section. Deactivation, which is symbolized by , can be turned into activation and vice versa by clicking on these symbols. The current number of indicators collected in the shopping cart is shown in the right upper corner of the screen.

Figure 6: Direct view in table form



Attributes

available vintages: 18

SOURCE OF THE UNADJUSTED FIGURES: FEDERAL STATISTICAL OFFICE. methodological information (PDF)

Interpret Key  Show time series as graph

Direct download: CSV *
Shopping cart:

Output options

| | |
|---|---|
| reporting periods (yyyy-mm) | from 2006-01 up to 2009-04 |
| vintages (yyyy-mm-dd) | from 2006-05-01 up to |
| sorting order | ascending |
| | apply |

Values

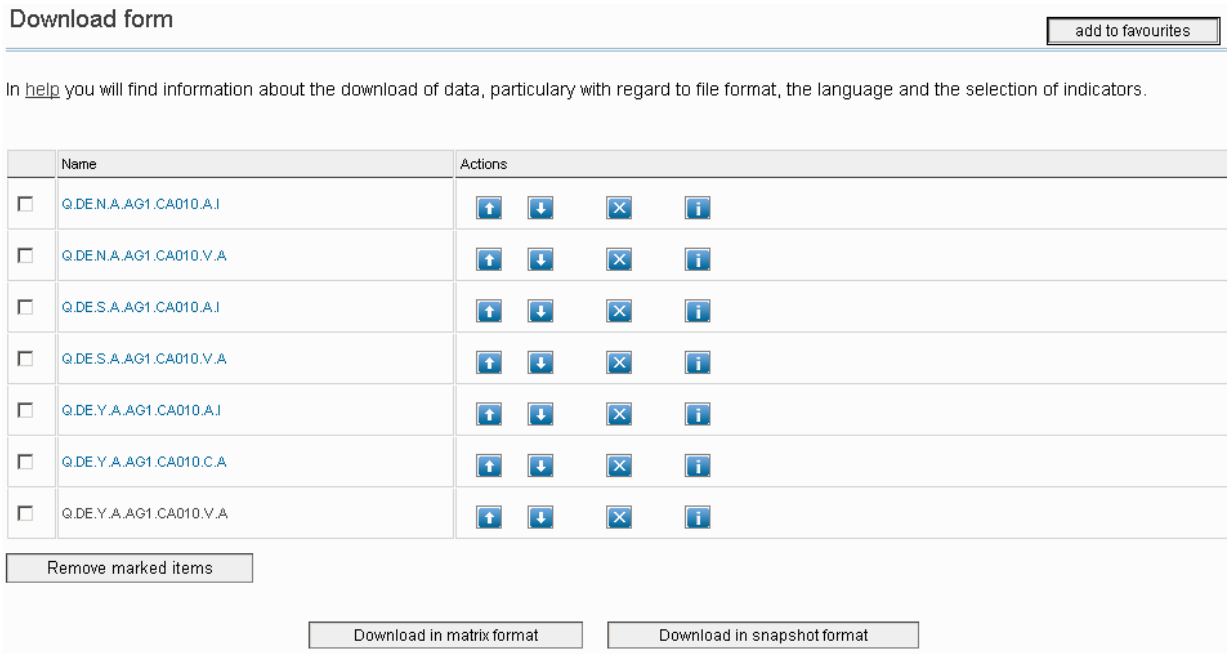
| time | 2006-05-11 | 2006-08-14 | 2006-11-14 | 2007-02-13 | 2007-05-15 | 2007-08-14 | 2007-11-14 | 2008-02-14 | 2008-05-15 | 2008-08-14 | 2008-11-13 | 2009-02-13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| + Show vintage attributes | | | | | | | | | | | | |
| 2006 1.Q | 567.19 | 568.30 | 567.18 | 567.57 | 567.48 | 569.51 | 569.51 | 569.51 | 570.21 | 568.77 | 568.77 | 568.77 |
| 2006 2.Q | | 574.77 | 573.47 | 574.62 | 574.73 | 578.80 | 578.80 | 578.80 | 578.09 | 578.45 | 578.45 | 578.45 |
| 2006 3.Q | | | 579.45 | 580.64 | 581.01 | 583.98 | 583.98 | 583.98 | 584.09 | 583.75 | 583.75 | 583.75 |
| 2006 4.Q | | | | 585.56 | 587.06 | 590.99 | 590.99 | 590.99 | 590.87 | 591.11 | 591.11 | 591.11 |
| 2007 1.Q | | | | | 601.20 | 599.62 | 599.62 | 599.65 | 600.50 | 599.86 | 599.86 | 599.86 |
| 2007 2.Q | | | | | | 603.66 | 605.87 | 605.61 | 604.72 | 604.30 | 604.30 | 604.30 |
| 2007 3.Q | | | | | | | 611.51 | 610.26 | 610.43 | 609.66 | 609.66 | 609.66 |
| 2007 4.Q | | | | | | | | 612.80 | 612.68 | 612.67 | 612.67 | 612.67 |
| 2008 1.Q | | | | | | | | | 624.25 | 622.96 | 623.46 | 624.66 |
| 2008 2.Q | | | | | | | | | | 622.10 | 622.91 | 624.69 |
| 2008 3.Q | | | | | | | | | | | 622.77 | 622.96 |
| 2008 4.Q | | | | | | | | | | | | 615.71 |

Figure 7: Direct view in chart form

# 5 Download forms

The platform offers two download forms which are aligned to the main uses of real-time data. First, it can be presented as a collection of spreadsheets, each of which comprises the real-time data of an indicator in the standard matrix representation. Second, spreadsheets may be ordered according to cut-off dates, containing the data status of the chosen set of indicators at the specific points in time. With the snapshot export, users are able to reconstruct the information set which was available to policymakers at the time when decisions were made.

Figure 8: Download list of different GDP versions




By entering the section "Download form", the download list is shown first (see Figure 8). This comprises those indicators for which the shopping cart were triggered during the search process. The buttons in the right-hand column can be used to manipulate the download list. The placement of an indicator in the list may be controlled by pressing ⬆ ("move identifier up") and ⬇ ("move identifier down"). An indicator can be deleted from the list with ✕ ("Delete identifier"). Indicator information can be recalled by clicking ℹ ("interpret indicator key"). Fine-tuning seems to be a sensible operation, as download lists can be stored for iterative use by the bookmark option. The procedure for storing download lists can be triggered with "Add to favourites".

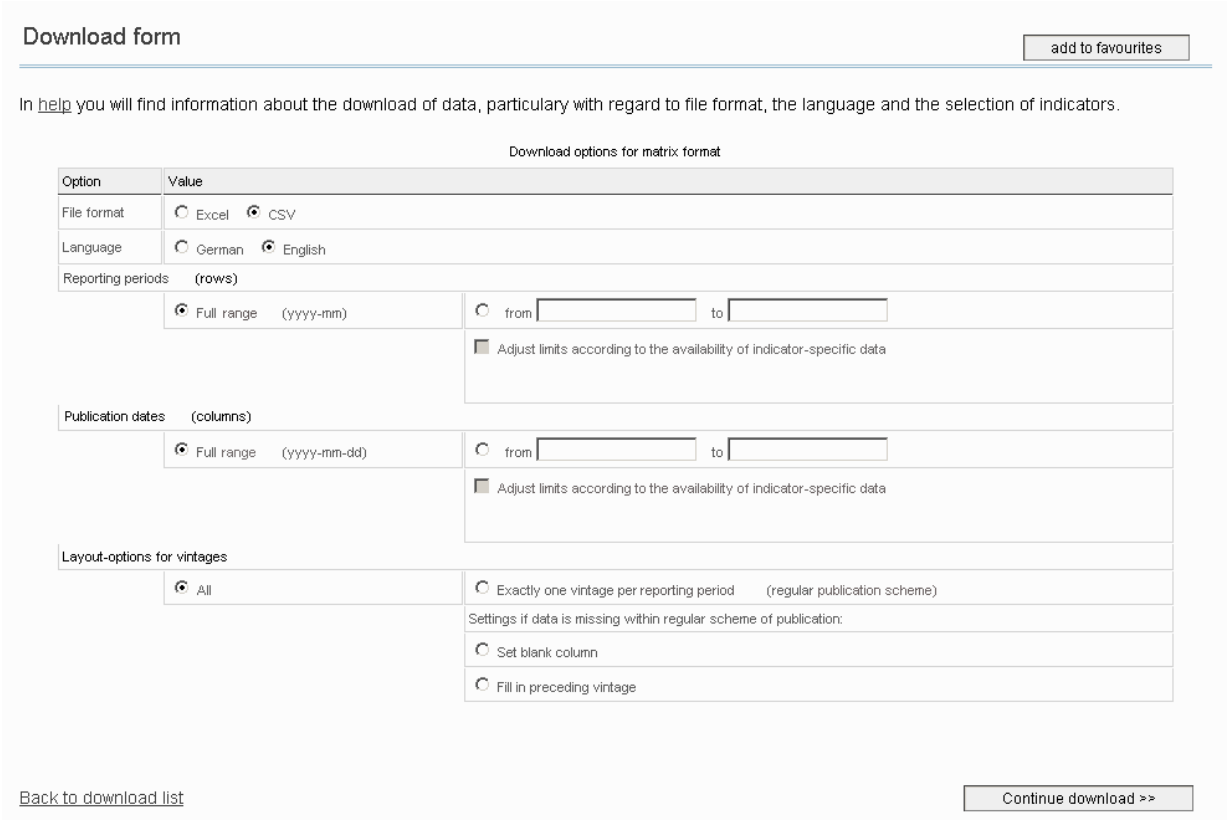
In general, extensive download lists may be established and maintained. However, effective downloads should be as concise as possible. The reason for this is that real-time data is generally vast, which makes simultaneous downloads of a large number of indicators

time-consuming and, potentially at times, less stable. Hence, users are asked to select, out of the download list, the set of indicators for which real-time data should be effectively downloaded. This is done by marking the squares which are placed to the left of the listed indicators. Then, either "Download in matrix format" or "Download in snapshot format" must be pressed to proceed.

## 5.1 Matrix format

The download in matrix format sorts the time-series vintages of an economic indicator next to each other in chronological order where columns are headed by the publication date. The export may be either in EXCEL or CSV format. In the former case, an EXCEL file is created containing as many spreadsheets as selected indicators. As the download in CSV format is possible only for each indicator separately, further selection is necessary. In any format, the output is organized such that the first row is reserved for publication dates, while reporting periods are written in the first column. The download can be either in German or English formatting conventions.

Figure 9: Download in matrix format



Apart from the file format and language, the download form asks users to determine the dimensions of the real-time data matrix and to select output options for vintages (see Figure 9). By default, available real-time data is exported completely with respect to reporting periods and vintages, i.e. the option "Full range" is activated in both cases. However, the ranges in both matrix dimensions can be specified alternatively. In the standard option, their limits are seen as fixed, implying that the output sheets contain totally congruent data matrices. Consequently, rows and columns are left blank at the beginning and at the end of each matrix dimension whenever available data do not reach the pre-specified limits. However, if ranges are to be determined only as a maximum, the box "Adjust limits according to the availability of indicator-specific data" can be marked.[20]

As regards reporting periods (i.e. rows in the export matrix), range limits have to be declared in the format "yyyy-mm", where the first part indicates the year and the second part the month to be addressed. Note that, even in the context of quarterly or annual data, ranges have to be specified in months, although it is irrelevant which month of the limiting period is actually chosen (see Example 5.1). As regards publication dates (i.e. columns in the export matrix), the format takes the structure "yyyy-mm-dd" where the additional third component indicates the day within a month, i.e. 01, 02, ..., 31. While the system expects the limits to be fully specified in the nine-digits format, users need not know the exact publication date. When no vintage is found at the set date, the system automatically chooses the vintage which is closest.

**Example 5.1** The aim is to export the real-time data matrix of GDP which comprises only vintages starting in the first quarter of 2000 and ending in the fourth quarter of 2006. To determine the range, either 2000-01, 2000-02 or 2000-03 may be written in the field after "from" and either 2006-10, 2006-11 or 2006-12 in the field after "to".

In the case of irregularities in the pattern and timing of publications, users may additionally control the output of vintages. The standard is to output all vintages whose publication dates fall into the pre-specified range. In many applications, however, it may be advantageous for the real-time data matrices to be delivered in a format which satisfies a triangular structure at the bottom. Such a matrix will emerge if the publication schedule is regular, i.e. there is precisely one issue of data per reporting period, which implies that the observations in both matrix dimensions possess the same fixed time interval. The merit of this structure is that diagonals bear a special meaning. In particular, the diagonals counted from the bottom stack the releases in ascending order, i.e. the time series of indicator values which are published first, second, third and so on. Differences between the diagonals can thus be directly interpreted as revisions.

The overwhelming majority of economic indicators, by their nature, have a regular publication scheme. However, when real-time data matrices are to be trimmed to a triangular structure in the case of irregularities, users should mark the option "Exactly one vintage per reporting period (regular publication scheme)". In this context, however, the system

[20]The output of this option will not differ from the one obtained under the choice "Full range" if the export list comprises a single member.

has to know how (potential) irregularities have to be dealt with. Technically speaking, deviations from the regular scheme will generally emerge if either no or more than one vintage is issued within a time interval defined by the length of the reporting period.

First, indicators may be published regularly or irregularly more often than is suggested by the frequency of the underlying data. In this case, neighboring vintages possess the same number of reporting periods, which implies the real-time data matrix is (partly) block-triangular, rather than triangular (see Example 5.2). The triangular structure is generally ensured by opting for a regular publication scheme. Which vintage among those of the same column length is systematically chosen depends on the date of the first vintage to be output. In this case, it is recommended that the range of vintages be specified, at least to its lower limit, instead of simply using the option "Full range".

**Example 5.2** In German short-term business statistics, it is common practice to report the monthly index of retail turnover twice a month. The preliminary announcement of retail turnover in period $t$ is published regularly at the end of the following month, i.e. in period $t+1$. In the middle of period $t+2$, and before the retail turnover of period $t+1$ is first published, there is an immediate revision of the preliminary announcement for the retail turnover of period $t$. If the aim is to extract a triangular matrix containing only the first announcements, the regular publication scheme option has to be chosen when determining the lower limit of the vintage range in the sense that the date should be, or at least close to, the first of the month.

Second, an irregularity occurs when statistical agencies cannot meet the delivery dates of the regular publication scheme. In particular, it sometimes happens that the publication of a new issue of data is postponed for more than the length of a reporting period (see Example 5.3). Regular publications are then said to be "missing". For the real-time data matrix, such an event also implies a deviation from the triangular structure because the column lengths of consecutive vintages may differ for more than one entry at the bottom. With the option "Exactly one vintage per reporting period (regular publication scheme)", a quasi-triangular structure is produced by either leaving missing vintages blank or charging the columns, which would actually be empty, with the latest data issued at the relevant times (option "Fill in preceding vintage"). The attribute "quasi" is used because the extracted matrix in fact shows irregular steps at the bottom. However, as the property of equal and fixed time intervals in the matrix dimensions is preserved, diagonals (containing missing values in this case) keep their conceptual meaning.[21]

**Example 5.3** In spring 2006, some regional statistical agencies in Germany were affected by strikes in the public sector. As a result, the monthly indices of the orders received were not able to be published in time. In particular, the indicator was not published at all in April and July while earlier missing reports were submitted later in May and June.

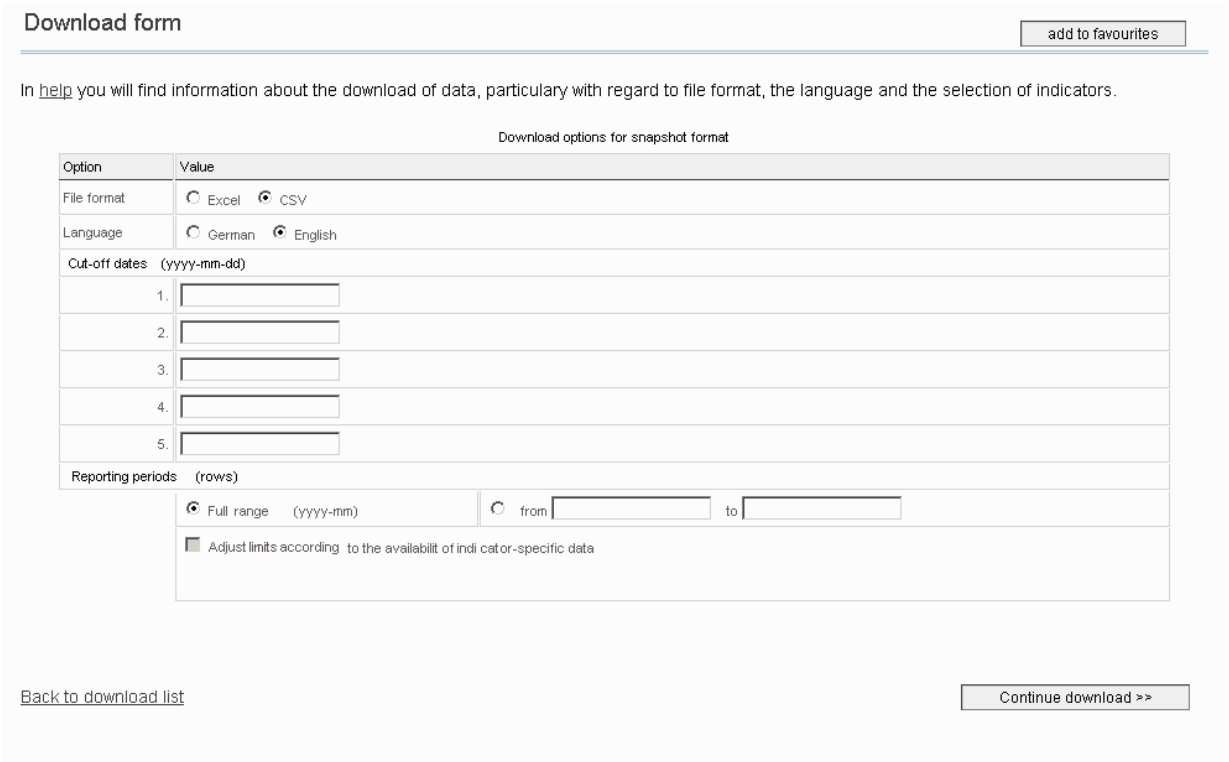
[21] It is worth stressing that releases are understood in terms of the regular publication scheme and the publication schedule. This means that the issue of a new figure is said to be omitted if it is not issued around the regular publication date. In this sense, there is no first release for this reporting period, as is indicated by a missing value in the lowermost diagonal. However, when moving along this reporting period to the right, one will find a first entry which is a first announcement in broader terms.

## 5.2 Snapshot format

The download in snapshot format comes out as a spreadsheet containing the data status of a specified set of indicators as it existed at a given point in time. The choice is not restricted to a single cut-off date. In EXCEL format, the download creates a file consisting of several spreadsheets, each of them designated by the respective cut-off date. Downloads in CSV format, however, have to be initiated separately for each cut-off date. In snapshots, rows comprise the reporting periods while columns sort indicators according to the download list. The full identifier forms the headline of each time series. It includes the indicator code and the exact publication dates, which might be useful information in this context, too. The download can be in either German or English formatting conventions.

In the download dialogue, users are further asked to specify the range of reporting periods and the cut-off dates (see Figure 10). As regards reporting periods, users may either select the option "Full range" or determine the desired sample by filling in the fields referring to range limits in the format "yyyy-mm". As in the matrix download, a month has to be specified even if quarterly or annual data are to be exported (recall Example 5.1). Cut-off dates have to be specified in the format "yyyy-mm-dd".

Figure 10: Download in snapshot format



Note that whenever the download list comprises indicators measured at different frequencies, the highest frequency is chosen as reference. The time series of the indicators, which are measured at a longer periodicity, possess missing values by construction (see Example 5.4).

**Example 5.4** The aim is to reconstruct the data of GDP and industrial production as they existed at 15 July 2007. While industrial production is available at monthly frequency, GDP is measured only quarterly. As the periodicity of the exported EXCEL sheet is monthly, the GDP series contain missing values in the first two months of each quarter, while the quarterly value is written into the cell referring to the third month.

# References


**Bald-Herbel, Christiane (2009)**, *Umstellung der Konjunkturstatistiken im Produzierenden Gewerbe auf Basis 2005*, Wirtschaft und Statistik 3/2009: 223-231.

**Braakmann, Albert, Norbert Hartmann, Norbert Räth and Wolfgang Strohm (2005)**, *Revision der Volkswirtschaftlichen Gesamtrechnungen 2005 für den Zeitraum 1991 bis 2004*, Wirtschaft und Statistik 5/2005: 425-462.

**Croushore, Dean and Tom Stark (2001)**, *A Real-Time Data Set for Macroeconomists*, Journal of Econometrics 105: 111-130.

**Deutsche Bundesbank (1991)**, *Data, adjusted for seasonal and working-day variations, on the expenditure components of GNP*, Monthly Report April 1991: 35-40.

**Deutsche Bundesbank (1992)**, *The economic scene in Germany in spring 1992*, Monthly Report June 1992.

**Deutsche Bundesbank (1995)**, *Results of the national accounts for Germany as a whole*, Monthly Report October 1995: 45-58.

**Deutsche Bundesbank (2001)**, *New concepts in short-term statistics*, Monthly Report September 2001: 7.

**Deutsche Bundesbank (2005)**, *Revision of Germany's national accounts*, Monthly Report May 2005: 36-37.

**Deutsche Bundesbank (2007)**, *Extension of the reporting categories in the Deutsche Bundesbank's negotiated pay rate statistics*, Monthly Report August 2007: 50.

**Deutsche Bundesbank (2008)**, *Impact of the 2008 index recalculation on the national price index and the Harmonised Index of Consumer Prices*, Monthly Report April 2008: 32-33.

**Deutsche Bundesbank (2009)**, *Publication of a comprehensive real-time database for the German economy*, Monthly Report August 2009: 52-53.

**Elbel, Günther and Ute Egner (2008)**, *Verbraucherpreisstatistik auf neuer Basis 2005*, Wirtschaft und Statistik 4/2008: 339-350.

**Fritsch, Sigrid (2006)**, *Erwerbstätigkeit in Deutschland – Beschleunigte Veröffentlichung gesamtwirtschaftlicher Erwerbstätigenzahlen*, Wirtschaft und Statistik 9/2006: 936-946.

**Garratt, Anthony and Shaun P. Vahey (2006)**, *UK real-time macro data characteristics*, Economic Journal 116: F119-F135.

**Gerberding, Christina; Miriam Kaatz, Franz Seitz and Andreas Worms (2005)**, *A real-time data set for German macroeconomic variables*, Schmollers Jahrbuch 125: 337-346.

**Herbel, Norbert and Joachim Weisbrod (1999)**, *Auswirkungen des neuen Konzepts der Produktionserhebungen auf die Berechnung der Produktionsindizes ab 1999*, Wirtschaft und Statistik 4/1999: 293-298.

**Jacobs, Jan and Jan-Egbert Sturm (2005)**, *Do Ifo indicators help explain revisions in German industrial production*, in: Sturm, Jan-Egbert and Timo Wollmershäuser (eds.), *Ifo survey data in business cycle and monetary policy analysis*, Heidelberg and New York: Physica: 93-114.

**Knetsch, Thomas A. and Hans-Eggert Reimers (2006)**, *How to treat benchmark revisions? The case of German production and orders statistics*, Economic Research Centre of the Deutsche Bundesbank: Discussion Paper 38/2006.

**Knetsch, Thomas A. and Hans-Eggert Reimers (2009)**, *Dealing with benchmark revisions: The case of German production and orders Statistics*, Oxford Bulletin of Economics and Statistics 71, 2: 209-235.

**Lorenz, Andreas (2008)**, *Revision analysis and the role of metadata*, Contribution to the OECD/Eurostat Task Force on "Performing Revision Analysis for Sub-Annual Economic Statistics".

**Lützel, Heinrich (1991)**, *Revision der Volkswirtschaftlichen Gesamtrechnungen 1970 bis 1991*, Wirtschaft und Statistik 4/1991: 227-247.

**Nowack, Marlene and Joachim Weisbrod (1995)**, *Auswirkungen der NACE-Verordnung und der PRODOM-Verordnung auf die kurzfristigen Statistiken im Bergbau und im Verarbeitenden Gewerbe*, Wirtschaft und Statistik 3/1995: 192-200.

**Peter, Florian (2009)**, *Die Indizes der Außenhandelspreise auf Basis 2005*, Wirtschaft und Statistik 7/2009: 711-723.

**Rheinisch-Westfälisches Institut für Wirtschaftsforschung (2007)**, *Ein Echtzeit-Datensatz für Daten der VGR für Deutschland*, RWI: Datenzeche #3, November 2007.

**Statistisches Bundesamt (2005)**, *Methodische Änderungen bei den Konjunkturindizes*, Wirtschaft und Statistik 3/2005: 179.

**Strohm, Wolfgang, Norbert Hartmann, Hartmut Essig and Peter Bleses (1999)**, *Revision der Volkswirtschaftlichen Gesamtrechnungen 1999 – Anlaß, Konzeptänderungen und neue Begriffe*, Wirtschaft und Statistik 4/1999: 257-281.

**Vorholt, Hubert (2008)**, *Neuberechnung der Baupreisindizes auf Basis 2005*, Wirtschaft und Statistik 9/2008: 808-822.