The Consumer Price Index: changes for 2020

The Consumer Price Index (CPI) follows each month the prices of a fixed basket of products. This basket is updated each year in order to remain representative of the consumption and in order to take into account the evolution in the consumer habits (Laspeyres-type index). Each year, a chain-linking of the index is performed in January and enables to update the basket of products and their weights and possibly permits to introduce some methodological improvements.

The update of the products in the CPI basket

Each year, the product sample is updated in order to take into account changes in the household consumption. The products that are no more representative in 2019 are dropped from the CPI basket meanwhile new products (new services in particular), which represent a huge or an increasing market share, are introduced. It's also an opportunity to take into account new habits of consumption (internet consumption in particular) and sometimes to adapt the collection protocol to track prices at best (durable goods, convergence of characteristics of the product sample in Metropolitan area and in oversea territories, revision of the sample of outlets in Reunion Island to be more representative of the whole Reunion Island, taking into account of the Rents and Charges survey and the survey on rents in the social sector for the price changes in rents in Martinique, Guyana and Reunion Island). The prices of these new products are collected from December 2019 in order to measure their price evolution between December 2019 and January 2020 and so that they contribute to the CPI from January 2020.

In addition to these updates, the basket has been adapted to take into account the integration of scanner data from January 2020 (see below).

Integration of scanner data in the calculation of the consumer price index

From January 2020 on, the French CPI uses scanner data on the scope of manufactured food products, of cleaning products and of the products for personal care sold in supermarket and hypermarket in metropolitan France. On this scope, prices were previously collected by price collectors on the field, in the outlets (about 32 000 prices collected each month). Conversely, scanner data are gathered by retailers when consumers go to pay for their goods in store. Scanner data enable to follow prices and also purchased quantities for each day, in each outlet, and for each product (identified thanks to its bar-codes). They are daily received by Insee and their transmission to Insee has been made compulsory by an implemented order of 13 April 2017.

The implemented methodology for processing scanner data is consistent with the current CPI concepts: a fixed basket of products (77 millions of products) is tracked month after month; in case one product of this basket disappears, it is replaced and a quality adjustment is performed. Because of the huge number of data, some treatments that were previously performed by the price collectors are now automated as the sorting of the products in the classification -about 600 consumption segments were created-, the identification of relaunches, the choice of the replacing product and the computation of quality adjustments.

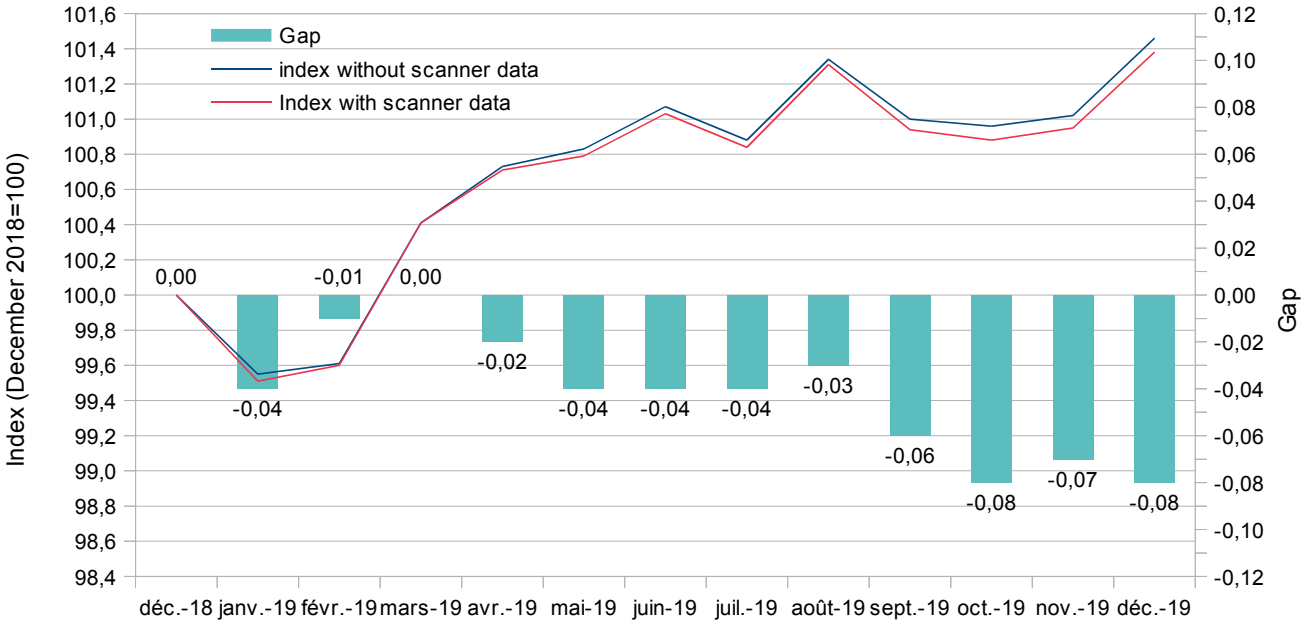
The huge number of data, their comprehensiveness on the considered scope, the detailed knowledge of purchased quantities, the daily tracking of prices as well as the more effective tracking of the prices paid by consumers (including personal rebate) rather than the prices displayed in the stores improve the quality of the CPI.

In 2019, the CPI was computed twice, once for the dissemination and a second time using scanner data, in order to assess the impact of scanner data; the impact on the all-item index is quite low (chart 1) : no more than 0.08 points on the index level and -0.1 point some months on the month-on-month and the year-on-year changes. The measured inflation is slightly lower with the scanner data. However, scanner data impacts are

higher on classes for which the use of scanner data is more important. For food and non-alcoholic beverage, the gap between the two indexes (with and without scanner data) is about 0.4 points in December 2019; for furnishings, household equipment and routine household maintenance, about 0.2 points and for miscellaneous goods and services about 0.1 points.

A detailed analysis of discrepancies shows that they are essentially explained by three factors: (i) the huge amount of data improves the index precision (with the field collection and the limited number of prices collected, sometimes, prices that were tracked could have a non usual dynamic); (ii) the comprehensiveness of scanner data and the information on purchased quantities improve the product coverage (previously, without any information about quantities, consumption segments of the CPI basket might have been less representative of the household consumption), (iii) information about paid prices (rather than displayed prices) as well as knowledge of the daily purchases enable to take into account special offers more adequately.

Chart 1: All-item CPI (December 2018=100) computed with and without scanner data



New methodology for calculating price indices for rail transport

Until 2019, the data used to calculate the mainline passenger transport index within “the passenger transport by railway” item (item 07.3.1.1.1) were list prices. These list prices were updated every six months and no longer reflected the prices actually paid by consumers. Indeed, yield management is used by railway companies: this consists of optimizing the turnover of a service by adapting, in real time, the price of the service according to the occupancy rate and the consumer. These methods are also used by airline companies, or in the tourism sector. For these services with very volatile prices, Eurostat recommends the constitution of a sample of prices representative of the behavior of the consumers, by including in particular the anticipation of the purchase and the period during which the service is consumed.

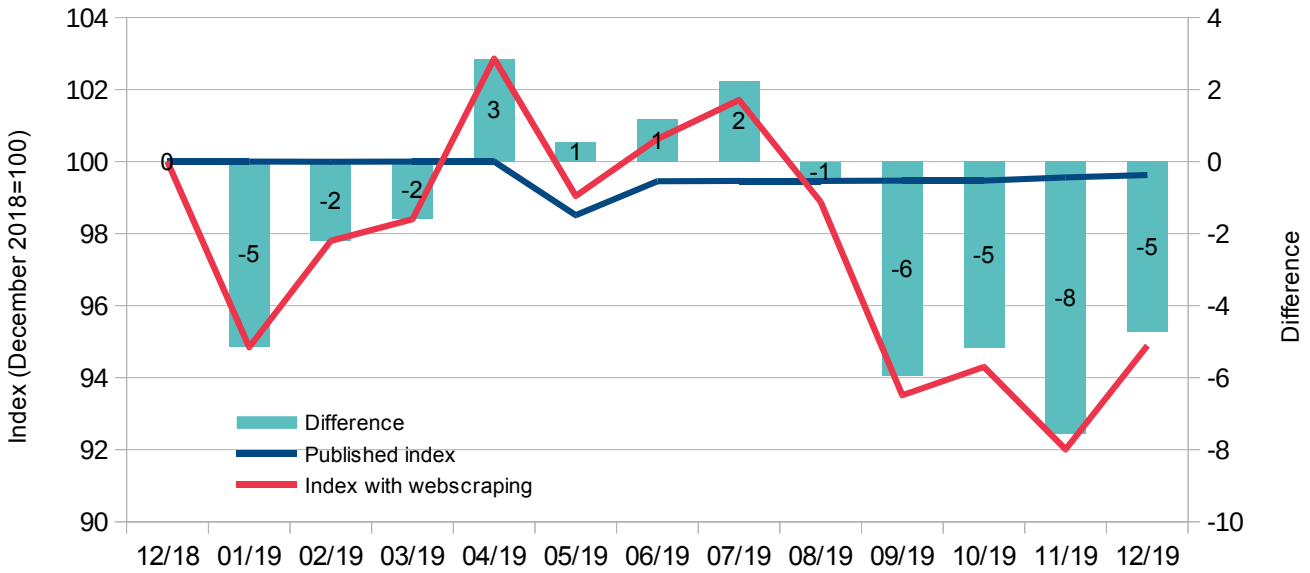
Thus, an automated internet data collection (webscraping) has been implemented on the train ticket websites. An experiment was carried out for high-speed trains (classic or low-cost) in 2018 and 2019. Daily, a robot collects ticket prices with four anticipation dates (2 days, 10 days, 30 days and 60 days before the train departs), according to two consumer profiles (with or without a reduction card) for a sample of 250 journeys

(one way), which corresponds to more than 10,000 requests. In 2020, this collection method was also extended to other mainline trains and regional trains.

The prices obtained are then aggregated by a geometric average at the elementary level by priority, profile, route and period (week or weekend). They are finally aggregated by a Laspeyres formula. The weights for the different journeys were calculated from aggregated rail traffic data and data collected by webscraping, which allow to know the frequency and prices of trains on each line.

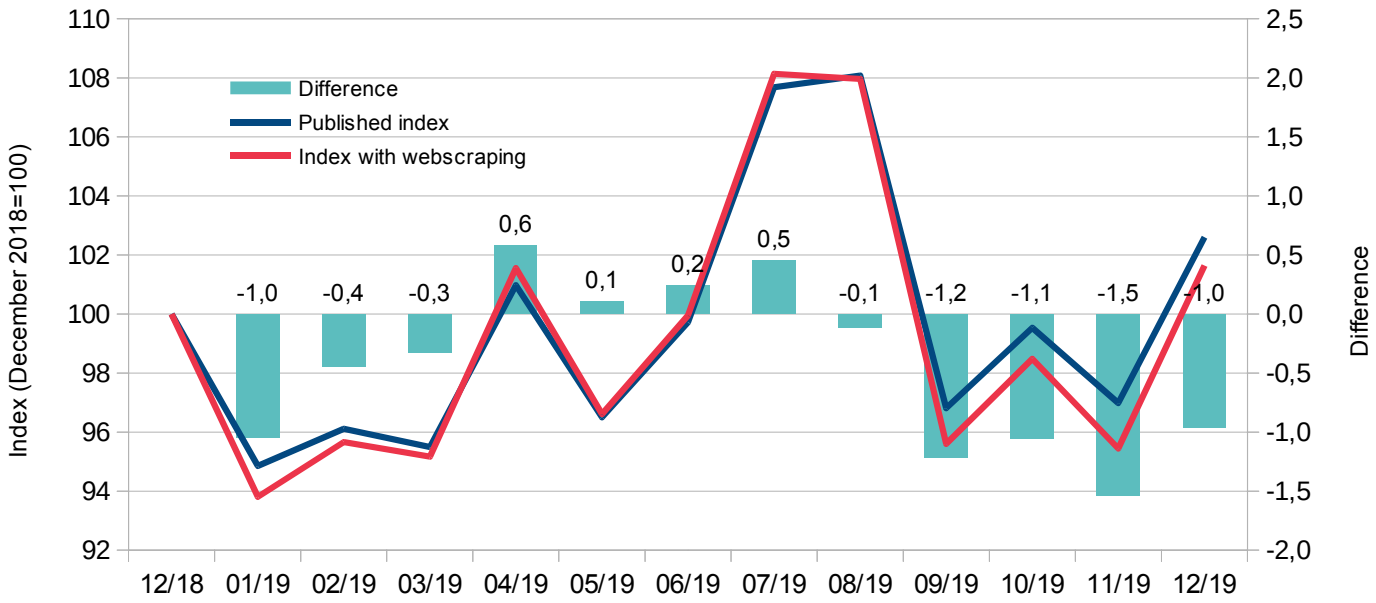
The new index obtained from data collected by webscraping therefore exhibits volatility - prices increasing sharply during the school holiday periods -, unlike the index calculated with the current method. The maximum difference on the passenger transport by railway item was 7.6 points in November (see Chart 2). The difference over one year between the two indices is -4.7 points. In May 2019, a major change in train ticket prices resulted in an observable decline from August (the first month in which all tickets were purchased with the new pricing). The impact of the methodological change was 1.5 points in November on the "Transport services" group (07.3) and 0.04 points on the overall index .

Chart 2: Passenger transport by railway item (07.3.1.1.1) (December 2018=100)



Reading note: In November 2019, the index obtained with the old method is 99.6 and the index obtained with webscraping is 92.0; the difference is therefore 7.6 points.

Chart 3: Transport services index (073) (base 100 in December 2018)



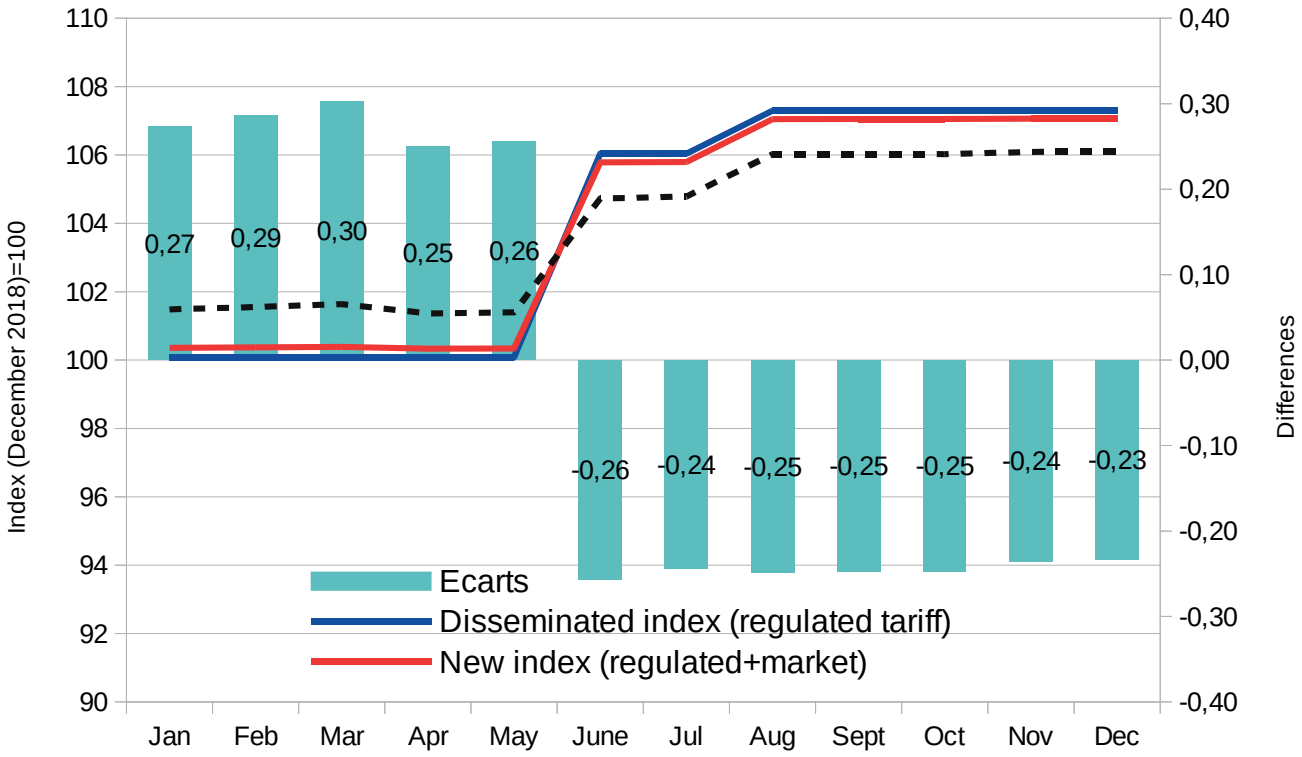
Introduction of market tariffs in the electricity consumption index

Before the opening of the French electricity market to competition for residential customers, the only offer available to households was a regulated tariff set by the public authorities. Since the opening to competition, all consumers can subscribe to a market offer proposed by the historical supplier of electricity or by an alternative supplier. From now on, 22% of the contracts correspond to a market offer.

Thus, from January 2020, the electricity price index will retrace the evolution of regulated tariffs as well as market offers from different suppliers (historical or not). If the regulated tariffs apply to all households, whether former or new subscribers, the collection of market offers will only concern the new contracts (the prices for the clients having signed a contract in the past being unavailable).

Taking unregulated tariffs into account would have increased the price index observed in the first semester of 2019 by 0.3 points, and would have reduced it in the second half by 0.2 to 0.3 points (see Chart 4). At the level « Housing, water, electricity, gas and others fuels », the difference between the published index and that including market offers would have been at most +0.05point in the 1st semester and -0.04 point in the 2nd. The difference in the global Consumer Prices index between the new index and the published one would have been 0.007 points in absolute terms.

Chart 4 : Comparison between electricity price indexes (December 2018=100)



The weight update

Each year in January CPI and HICP weights are updated for the current year. These weights are used in order to aggregate elementary indices for different consumption segments. These weights represent the share of each consumption segment in the total household consumption (in the scope of the CPI). They are mainly obtained from national accounts (in basis 2014) and their annual estimates of household consumption but also from different detailed specialist sources. In addition, the turnover from scanner data was used at the finest level and compared with the amounts of consumption expenditure from other sources; this confrontation led to some reevaluations, in particular the weight of supermarket and hypermarket in the CPI basket.

The update of seasonal adjustments and seasonal adjusted indices

As each year; seasonal adjustments for the all-item index (France, all household) and for four indexes of core inflation have been revised over the period January 2000- December 2019 taking into account the 2019 data.

Collection schedule

The consumer price index (CPI) is based on scanner data, price collected by collectors on the field or recorded centrally. The centralized collection is made throughout the calendar month.

The field collection is carried out according to a specific calendar, fixed a year in advance. This collection calendar differs from the civil calendar. Every month, prices for CPI are collected during 20 days, throughout the working days of four consecutive weeks.

Each product in the sample is tracked on a specific day among the 20 days of filed collection (numbered from 1 to 20) and the price collector responsible for this collection returns every month to observe in the same outlet, the same product, the same day among the collection month: this organization ensures that we

measure changes on average over a month, guarantees the opening of the outlet and neutralizes possible « day of the week » effects on the prices.

A calendar month consists of 28 to 31 days and doesn't correspond to a whole number of weeks. So, every year, Insee adapts the field collection calendar so that the 48 weeks of collection coincide at best with the civil calendar. This adaptation consists of fixing weeks without collection, four on average in a year.

Regardless of this exercise, the month-on-month changes in the CPI integrate calendar effects which may affect the year-on-year comparability of monthly changes. In general, the calendar effects in the index level disappear after one or two months and are limited to specific consumption segments. For example, every year, the school holidays or some days off are not located on the same month. These generate variations in the index evolution for the accommodation and passenger transports. If such effects are seeing, they are commented in the publication *Informations Rapides* published at the same time as the CPI. It is the same for the sale calendar when it changes.

The Insee's field collection calendar aims as much as possible to reproduce the calendar effects : a shift in the sales observed in the civil calendar will have to be found, as far as possible, in Insee's collection calendar.

In 2020, the field collection calendar is a little different from that in 2019. Indeed, in 2020 (as previously in 2013), it has been necessary to introduce a fifth week in the year without collection in order to insure that the beginning of the following collection year coincides with the beginning of the civil year. If the CPI calendar will reflect the decrease in the number of sales days in February and August (due to a sale period reduced to 4 weeks from 2020 (compared to 6 weeks before)), the decrease in this number in July will be hidden.

Table : Number of sale days in the civil calendar and the CPI calendar

	Calendar	2013	2014	2015	2016	2017	2018	2019	2020
Winter sales									
January	**CPI**	13	18	18	18	13	13	13	13
	civil	23	24	25	26	21	22	23	24
February	**CPI**	12	12	12	12	17	17	17	7
	civil	12	11	17	16	21	20	19	4
Summer sales									
June	**CPI**	0	3	3	3	0	0	0	0
	civil	5	6	7	9	3	4	5	7
July	**CPI**	20	20	19	19	17	18	18	18
	civil	30	29	31	31	31	31	31	21
August	**CPI**	0	0	2	2	7	7	7	0
	civil	0	0	4	2	8	7	6	0