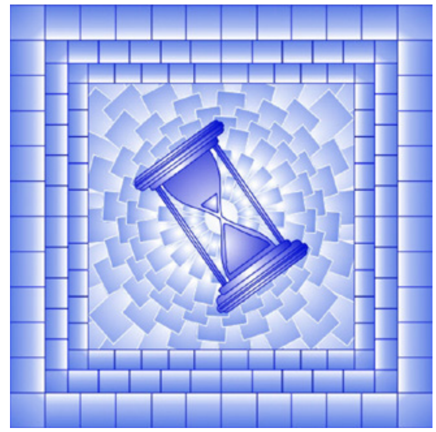
Catalogue no. 62F0014F2018003
ISBN 978-0-660-27244-3

Internet Access Services Index Methodology in the Consumer Price Index



Release date: August 7, 2018

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

email at STATCAN.infostats-infostats.STATCAN@canada.ca

telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

• Statistical Information Service	1-800-263-1136
• National telecommunications device for the hearing impaired	1-800-363-7629
• Fax line	1-514-283-9350

Depository Services Program

• Inquiries line	1-800-635-7943
• Fax line	1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Internet Access Services Index Methodology in the Consumer Price Index

1- Introduction

Residential Internet access is a high-technology service. As such, its calculation requires methods that are not typically applied to other components of the CPI. This paper describes the methodology that Statistics Canada has adopted to measure the price change of residential Internet access services.

2 - Background

The Internet access services aggregate is a component of the Communications aggregate. It accounts for 28% of the Communications aggregate in the 2015 basket. Conceptually, the Internet access services index covers all Internet access services to a residence. However, the prices of wireless plans to the home are not collected. With 4% of Canadian residential subscriptions in 2016 (Canadian Radio-television and Telecommunications Commission [CRTC] 2017a), they do not represent an important part of the market.

3 - Data Collection

Data for the Internet access services index is collected by price evaluators in Statistics Canada's Head Office from the websites of the Internet Service Providers (ISP)s. Residential internet access is marketed and sold in a similar manner across Canada. A monthly subscription to a plan from an ISP allows a consumer to access the Internet.

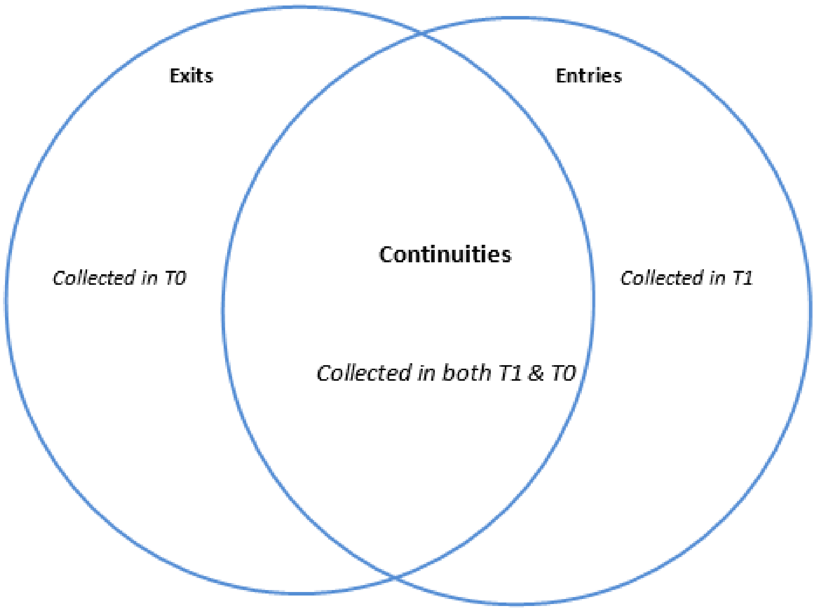
Prices are collected on a provincial basis from all ISPs with at least a 10% market share by revenue. This ensures a coverage of 75% of the market in each province. Market shares are sourced from the Annual Survey of Telecommunications, conducted jointly by Statistics Canada and the CRTC, and are updated annually.

Only wired, broadband (non-dialup) plans are priced. For the purposes of this index, a plan is defined by its ISP, download speed, upload speed and usage cap. These characteristics are recorded for each selected plan when it is priced and are also used to identify a plan across time.

Each pricing month, the price evaluator selects and records the price and characteristics of a plan at every available download speed for each ISP. The collected price corresponds to the regular, unbundled, price of the plan. If there are several plans with the same download speed, only one of them is selected, based on the price evaluator's judgment of which one is the most representative plan. As long as a plan remains available it will continue to be priced.

When a new download speed is offered, the price evaluator immediately selects and prices a plan corresponding to this new download speed. This means that new plans can "enter" the sample even without a corresponding "exit". In the same way, a plan can exit the sample without a corresponding entry. However, in the case of an exit, the price evaluator will select and price a new plan with the same download speed from the same ISP, if possible. The dynamic nature of entries and exits means that the sample size is not fixed and that it will adjust automatically in response to the introduction of new plans and the discontinuation of older plans. This is necessary due to the high churn rate; at the ISP level, on average 15.1% of plans exited the sample on a quarterly basis between the first quarters of 2016 and 2018. Figure 1 graphically represents the difference between entries and exits.

Figure 1:
Enries, Exits and Continuities



4 - Weighting by Download Speed

As described in Section 3, prices corresponding to the full range of available download speeds are collected every period from each ISP. This is done to ensure that new plans enter the sample immediately and to combat sample deterioration. It would not be reasonable to accord the same weight to both a newly introduced, high-speed, plan and a plan with a more representative download speed. It is, therefore, important to weight plans below the ISP level to help ensure that the index reflects the price changes experienced by consumers. Ideally plans would by weighted by their revenue or expenditure share but at this time expenditure data at the individual plan level is unavailable. In absence of external quantity or expenditure data at individual plan level, plans are weighted based on their observed download speed. This is done based on the distribution of the available download speeds within an ISP.

The plans of a given ISP, *l*, in a given period, t, are assigned a weight vector, q_l^{t*} , such that the weighted sum of their download speeds equals their median download speed. The weight vector q_l^{t*} can be regarded as a vector of quantity shares. This approach assumes that the distribution of available download speeds is reflective of consumer demand. There is a wide body of research that supports the conclusion that the purchasing decisions of consumers can be strongly influenced by "anchors". While some studies have shown that completely arbitrary anchors such as Social Security Numbers can be effective (Ariely, Loewenstein & Prelec 2003), there is evidence that the presentation of similar products with extreme yet plausible prices is a more effective anchoring mechanism (Kishna, Wagner & Yoon 2006) (Sugden, Zheng & Zizzo 2013).

In addition to the constraint that the weighted sum of download speeds is equal to the median download speed, weights are assigned such that their variance is minimized. This ensures that weights are spread out as evenly as possible, rather than simply assigning all weights to the fastest and slowest plans. Combined with the constraint that the weights sum to one, this minimization of variance is equivalent to minimizing the squared difference between the generated weights and a vector of equal weights, given the constraint that the weighted average download speed equals the median download speed.

The generation of weight vector q_l^{t*} can be formulated as the solution to the following minimization problem. Defining a_l^t as the median download speed of ISP *l* in period t, s_{il} as the download speed of plan i from ISP *l* and n as the number of available plans:

$$(1)\quad q_l^{t*} = \underset{q_{il}}{argmin} \sum_{i=1}^{n} \left(q_{il} - \frac{1}{n}\right)^2$$

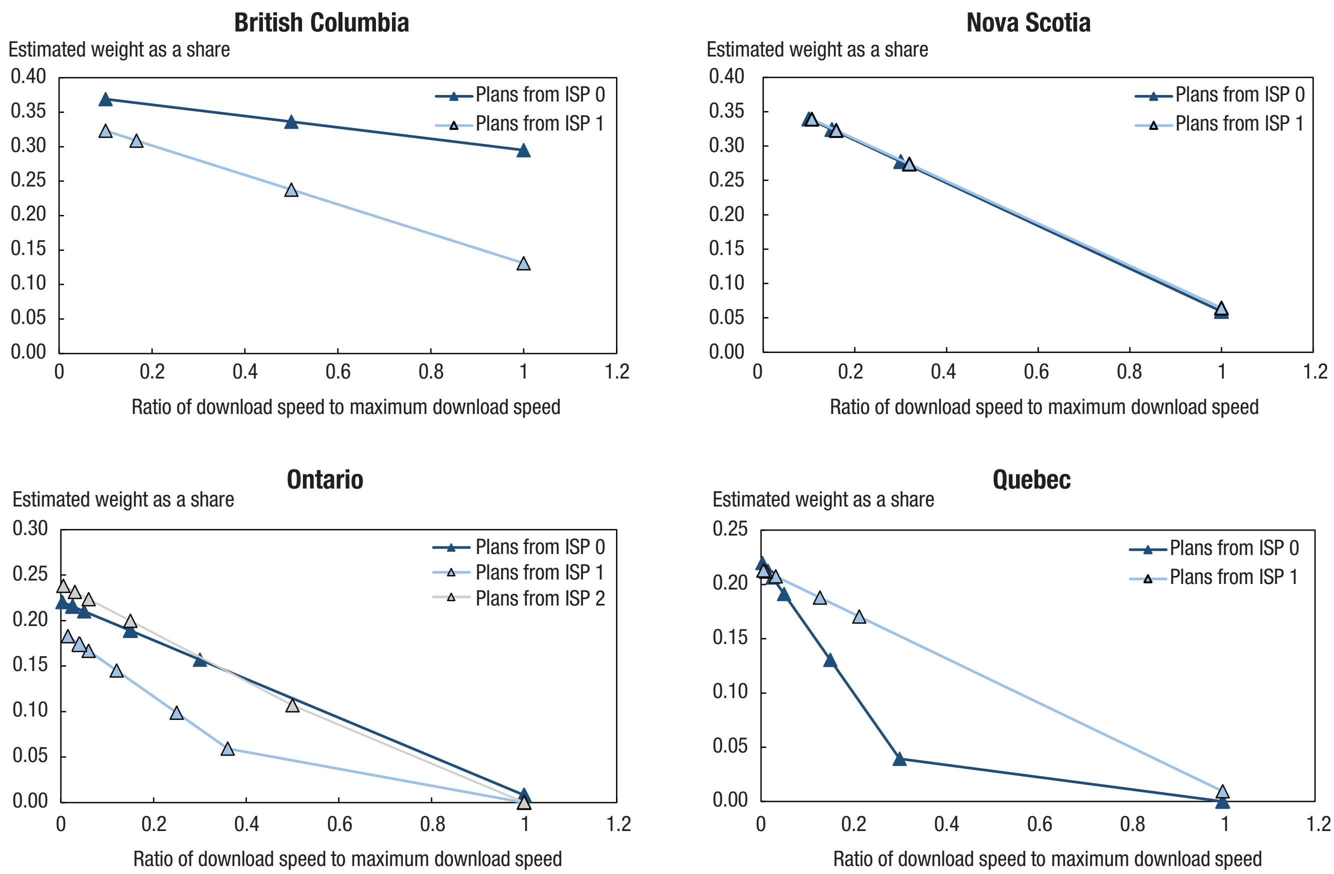
subject to:

$$\sum_{i=1}^{n} q_{il} * s_{il} = a_l^t;$$
$$\sum_{i=1}^{n} q_{il} = 1;$$
$$\text{and } q_{il} \geq 0$$

Since weight vector q_l^{t*} is treated as a vector of quantity shares, the corresponding vector of expenditure shares, v_l^{t*}, is obtained by multiplying these quantity shares by their respective prices and subsequently dividing by their sum so that:

$$(2)\quad v_{il}^{t*} = \frac{q_{il}^{t*} * p_{il}^{t}}{\sum_{i=1}^{n} q_{il}^{t*} * p_{il}^{t}}$$

Figure 2
Estimated Distribution of Weights by Normalized Download Speed for Selected Provinces in February 2018



The weight generation procedure generally results in a reduced weight being assigned to the faster plans. This is because ISPs usually offer a range of download speeds that increase in an exponential fashion[1]. However, average download is constrained to equal median speed in a linear fashion. This means that any weight that is assigned to a faster plan results in a disproportionate shift of the average download speed, a shift that has to be compensated by increasing the weight of a slower plan. Figure 2 provides a graphical representation of the results of this estimation procedure for selected provinces in February 2018. Estimated weights are plotted against download speeds; download speeds have been normalized by dividing by the maximum download speed from the same ISP. One can readily observe a negative relationship between download speed and weight. This relationship also tends to be linear, except for two kinks, one in Ontario and one in Quebec. In both cases, the fastest plan has received a minimal weight that is very close to zero. Overall, the results of weight generation procedure are more plausible than the alternative of assigning all plans an equal weight.

5 - Hedonic Adjustment of Entries and Exits

For products that undergo rapid technological change and an associated churn in models, the matched model breaks down and can lead to substantial bias (Silver & Heravi, 2005). This problem applies to Internet access plans since an average of 15.1% of plans exit the sample each quarter. This high churn rate can be explained by both improved technology such as increasing download speeds as well as marketing tactics of the ISPs. Under such conditions, hedonic methods can be applied to account for price change that would be missed under the matched model approach (Triplett, 2004).

A benefit of this application of hedonics is that it is easier for CPI analysts to interpret and explain due to its similarity to other methods of quality adjustment employed in the CPI.

5.1 - Regression Functional Form

A regression is estimated every collection period on the same sample from which the index is calculated. This means that the plans from the two periods are pooled in the same regression. As a result, the estimated coefficients of the characteristics are constrained to be the same between the two periods. Given that all plans were observed in two adjacent quarter, this would seem to be a reasonable restriction.

Least squares are used to solve for the B vector of parameters in the following formulation:

$$(3) \quad \ln p_i^t = B \cdot X_i^t + \in_i$$

where p_i^t is the price of plan i from period t, $\in_i$ is a random error term with an expected value of zero and X_i is plan i's vector of characteristics.

There are a limited number of ISPs providing service to a given residence, and it is assumed that a consumer will not change residences simply to change ISPs. It is therefore readily apparent that there is no Canada-wide market for Internet access services. Ideally separate regressions would be run or dummies included for each ISP. However, in this case there would be too few observations to estimate the relationship between prices and characteristics with much confidence. Instead, all plans from current and previous period are pooled together for one regression.

The regression is weighted in order to ensure that all combinations of characteristics and prices are accorded their proper importance. The results of the optimization model, as specified in equations (1) and (2), are used to assign a regression weight to each plan. The model is constrained so that, for both periods, all providers within a given province are equally weighted and each of the ten provinces receives an equal weight. If a particular plan is not available in a given period, this is equivalent to it having no weight in the regression for that period. Plans from the three territories are excluded from the regression because they are not representative of the overall Canadian market and should not influence the adjustment of plans in the provinces. However, plans from the territories are still subject to same adjustment procedure as the plans from the provinces.

1. For example, an ISP might offer plans at the following speeds: 20 Mbps, 40 Mbps, 80 Mbps and 160 Mbps.

5.2 – Adjustment Procedure

Once the regression model has been estimated and the coefficients estimated, the missing prices of entering and exiting plans are calculated. This is done by applying an adjustment factor to the predicted price from the regression. The adjustment is necessary since the hedonic regression is constrained to hold characteristics coefficients constant across the two-time periods. Thus, the missing of price of plan i in period t from ISP l is calculated as:

$$(4) \quad p_{il}^{t} = e^{\hat{\beta} \cdot X_{il}^{t}} \times A_{il}^{t}$$

Here $\hat{\beta}$ is the vector of characteristic coefficients estimated from the regression; note the lack of time subscript, t, on the β vector of coefficients. X_{il}^{t} is the missing plan's vector of characteristic while A_{il}^{t} is the adjustment factor calculated from the plans available in period t from ISP l that are most similar to the missing plan. Having noted that the adjustment procedure only considers plans from the same ISP, henceforth the ISP subscript l will be dropped for legibility reasons.

The selection of the most similar plans is fully automated. The download speed of plan i, s_i, compared to the download speeds of all plans available from ISP l in period t, the set S_l^t If plan i's download speed is either higher or lower than any other plan, i.e.:

$$(5) \quad s_i > \underset{s \in S_l^t}{\operatorname{argmax}}(s) \quad \text{or } s_i > \underset{s \in S_l^t}{\operatorname{argmin}}(s)$$

then the following equation holds:

$$(6) \quad A_i^t = \frac{p_j^t}{e^{\hat{\beta} \cdot X_j}}$$

Here p_j^t is the price of the plan from ISP l available in period t with the closest speed to plan i's. This corresponds to Cases 1 and 2 in Table 1. The above calculation also applies if $s_{jl} = s_{il}$, i.e. when there is a plan from ISP l in period t with a download speed exactly equal to the download speed of the missing plan i. This corresponds to Case 3 in Table 1. If none of these conditions are satisfied, then A_i^t is calculated as:

$$(7) \quad A_i^t = \left(\frac{p_k^t}{e^{\hat{\beta} \cdot X_j}}\right)^{0.5} \times \left(\frac{p_j^t}{e^{\hat{\beta} \cdot X_j}}\right)^{0.5}$$

In this case, the adjustment factor is based on the price of two different plans, j and k, both available from ISP l in period t. Their download speeds, s_j and s_k, bound the download speed of plan i, such that $s_k > s_i > s_j$. In addition, s_j and s_k satisfy the following conditions:

$$(8) \quad s_k = \underset{s \in S_l^t;\, s > s_i}{\operatorname{argmin}} (s - s_i) \qquad s_j = \underset{s \in S_l^t;\, s_i > s}{\operatorname{argmax}} (s_i - s)$$

In other words, plan k is the plan offered by ISP l in period t with a download speed that is faster than s_j, but by a minimal amount. Plan j is the plan offered by ISP l in period t plan that is slower than s_j, but by minimal amount. The different methods of calculating the adjustment factor A_i^t are summarized in Table 1, while Tables 2-5 provide examples (entries and exits are shaded).

Table 1

Case	Condition	Calculation of A_{il}^t	Where
I	$s_i \geq \underset{s \in S_l^t}{\operatorname{argmax}}(s)$ *The missing plan has a download speed that is higher than or equal to the download speed of any plan offered in period t, the period for which the plan is missing*	$A_i^t = \dfrac{p_j^t}{e^{\hat{\beta} \cdot X_j}} = p_j^t e^{-\hat{\beta} \cdot X_j}$	$s_j = \underset{s \in S_l^t}{\operatorname{argmax}}(s)$
2	$s_i \geq \underset{s \in S_l^t}{\operatorname{argmax}}(s)$ *The missing plan has a download speed that is lower than or equal to the download speed of any plan offered in period t, the period for which the plan is missing*	$A_i^t = \dfrac{p_j^t}{e^{\hat{\beta} \cdot X_j}} = p_j^t e^{-\hat{\beta} \cdot X_j}$	$s_j = \underset{s \in S_l^t}{\operatorname{argmax}}(s)$
3	$s_i = s_j$ *The missing plan has a download speed that is exactly equal to the download speed of a plan offered in period t, the period for which the plan is missing*	$A_i^t = \dfrac{p_j^t}{e^{\hat{\beta} \cdot X_j}} = p_j^t e^{-\hat{\beta} \cdot X_j}$	$s_j \in S_l^t$
4	Else *There are two plans offered in period t with download speeds that bound the download speed of the missing plan*	$A_i^t = \left(\dfrac{p_k^t}{e^{\hat{\beta} \cdot X_j}}\right)^{0.5} \times \left(\dfrac{p_j^t}{e^{\hat{\beta} \cdot X_j}}\right)^{0.5}$	$s_k > s_i > s_j$ $s_k = \underset{s \in S_l^t}{\operatorname{argmin}}(s - s_i)$ $s_j = \underset{s \in S_l^t}{\operatorname{argmin}}(s_i - s)$

Table 2
Case #1 - Entry

Period	Download Speed	Upload Speed	Cap	Actual Price	$e^{\widehat{\ln p}}$	A^t	Calculation	Adjusted Price	Missing Period	Relative
T	40	4	250	$99.95	$76.11	99.95/76.11 = 1.3132	1.3132 X 86.46	$113.54	T	107.95/113.54 = 0.9508
T+1 Entry	80	8	250	$107.95	$86.46					

Table 3
Case #2 - Exit

Period	Download Speed	Upload Speed	Cap	Actual Price	$e^{\widehat{\ln p}}$	A^t	Calculation	Adjusted Price	Missing Period	Relative
T Exit	6	1	100	$55.00	$48.84	60.00/59.24 = 1.0128	48.84 X 1.0128	$49.47	T+1	49.47/55.00 = 0.8995
T+1	15	1	150	$60.00	$59.25					

Table 4
Case #3 - Entry

Period	Download Speed	Upload Speed	Cap	Actual Price	$e^{\widehat{\ln p}}$	A^t	Calculation	Adjusted Price	Missing Period	Relative
T	5	1	20	$46.95	$39.92	46.95/39.92 = 1.1761	1.1761 X 43.23	$49.67	T	46.95/49.67 = 0.9452
T+1 Entry	5	1	40	$46.95	$43.23					

Table 5
Case #4 - Exit

Period	Download Speed	Upload Speed	Cap	Actual Price	$e^{\widehat{\ln p}}$	A^t	Calculation	Adjusted Price	Missing Period	Relative
T Exit	35	3	120	$67.99	$67.78	$(69.99/75.65)^{0.5}$ x $(61.99/62.91)^{0.5}$ = 0.9548	67.78 x 0.9548	$64.72	T+1	64.72/67.99 = 0.9519
T+1	60	10	120	$69.99	$75.65					
T+1	30	5	70	$61.99	$62.91					

6 - Aggregation

Provincial and territorial elementary price indexes are calculated in a two-stage fashion. In the first stage, plan level price relatives are aggregated to produce ISP level price movements. The price movements of all plans, whether continuing, entering or exiting, are included in the calculation, after applying the hedonic adjustment procedure described in Section 5. These price movements are weighted with an average of the estimated weight vectors v^t and v^{t-1}, from equation (2) in Section 4. If a plan is not available in certain period due to entering or exiting the sample, it receives a weight of zero in that period. For example, if plan i has entered the sample in period t, then v_i^{t-1} will equal zero. This aggregation to the ISP level movement resembles a chained Tornquist-Thiel index. The following equations show the calculation of price change for ISP l:

$$(9) \quad I_l^{t\text{-}1:t} = \prod_i \left(\frac{p_{il}^t}{p_{il}^{t-1}}\right)^{\frac{v_{il}^{t-1} + v_{il}^t}{2}}$$

Where:

$$\sum_i v_{il}^{t-1} = 1; \ \sum_i v_{il}^{t-1} = 1$$

In the second stage, provincial and territorial elementary price indexes are calculated as the weighted geometric mean of the ISP specific price movements. The weight vector, w, contains the per ISP shares of residential internet access revenues from the most recent Annual Survey of Telecommunications. As described in Section 2, these weights are updated on an annual basis. Aggregation of price movements using a geometric mean with fixed weights is generally known as the geometric Young (ILO et al., 2004). The following equations show the calculation of price change for a given province.

$$(10) \quad I^{t\text{-}1:t} = \prod_l \left(I^{t\text{-}1:t}\right)^{w_l}$$

Where:

$$\sum_l w_l = 1$$

7 - Conclusion

This paper has covered the methods adopted in the calculation of the Internet access services component of the Canadian CPI. Hedonic adjustment of entering and exiting plans as well as estimation of plan level weights are key features of the adopted methodology. Statistics Canada is continually engaged in efforts to improve the Consumer Price Index program. Access to timely administrative data directly from telecommunications companies could greatly benefit the Internet access services index.

8 - Additional Information

More information about the concepts and use of the Consumer Price Index (CPI) is available in *The Canadian Consumer Price Index Reference Paper* (Catalogue number 62-553-X https://www150.statcan.gc.ca/cgi-bin/IPS/display?cat_num=62-553-X).

Two videos, "An Overview of Canada's Consumer Price Index (CPI)" (https://www.youtube.com/watch?v=qfKmJe3CK6E) and "The Consumer Price Index and Your Experience of Price Change," (https://www.youtube.com/watch?v=U0xDcqE-zNs) are available on Statistics Canada's YouTube channel.

Contact information

For more information, or to enquire about the concepts or methods, contact us (toll-free 1-800-263-1136; 514-283-8300; STATCAN.infostats-infostats.STATCAN@canada.ca)

References

Analoui, F., & Kharmai, A. (2003). *Strategic management in small and medium enterprises*. Cengage Learning EMEA.

Ariely, D., Loewenstein, G., & Prelec, D. (2003). "Coherent arbitrariness": Stable demand curves without stable preferences. *The Quarterly Journal of Economics*, 118(1), 73-106.

Canadian Radio-Television and Telecommunications Commission. (n.d.). *Internet - Our Role*: Retrieved from https://crtc.gc.ca/eng/internet/role.htm.

Canadian Radio-Television and Telecommunications Commission. (2017a). *2017 Communications Monitoring Report*: Retrieved from https://crtc.gc.ca/eng/publications/reports/policymonitoring/2017/index.htm.

Canadian Radio-Television and Telecommunications Commission. (2017b). *Telecom Regulatory Policy CRTC 2017-104*: Retrieved from https://crtc.gc.ca/eng/archive/2017/2017-104.htm.

Krishna, A., Wagner, M., Yoon, C., & Adaval, R. (2006). Effects of Extreme-Priced Products on Consumer Reservation Prices. *Journal of Consumer Psychology*, 16(2), 176-190.

OECD/Eurostat. (2014). *Eurostat-OECD Methodological Guide for Developing Producer Price Indices for Services: Second Edition*: OECD Publishing.

Silver, M., & Heravi, S. (2002). Why the CPI matched models method may fail us: results from an hedonic and matched experiment using scanner data." *European Central Bank (ECB) Working Paper* 144.

Silver, M., & Heravi, S. (2005). A failure in the measurement of inflation: Results from a hedonic and matched experiment using scanner data. *Journal of Business & Economic Statistics*, 23(3), 269-281.

Statistics Canada. (n.d). *A Revision of the Methodology of the Internet Access Services Component of the Consumer Price Index beginning with the March 2008 CPI*. Retrieved from http://www23.statcan.gc.ca/imdb-bmdi/document/2301_D40_T9_V1-eng.pdf

Statistics Canada. (2015a). *The Canadian Consumer Price Index Reference Paper*: Catalogue No. 62-553-X.

Statistics Canada. (2015b, November 18). *Consumer Price Index, November 2015*. Retrieved from https://www.statcan.gc.ca/daily-quotidien/151218/dq151218b-eng.htm

Sugden, R., Zheng, J., & Zizzo, D.J. (2013). Not all anchors are created equal. *Journal of Economic Psychology,* 39, 21-31.

Triplett, J. (2004). *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*: OECD Publishing.