

Catalogue no. 71-607-X
ISBN 978-0-660-39495-4

Real-time Local Business Conditions Index: Concepts, data sources, and methodology

Release date: July 7, 2021
Updated: October 28, 2022



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

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

© His Majesty the King in Right of Canada as represented by the Minister of Industry, 2022

All rights reserved. Use of this publication is governed by the Statistics Canada [Open Licence Agreement](#).

An [HTML version](#) is also available.

Cette publication est aussi disponible en français.

Acknowledgments

The development of this index benefited from the collaboration and knowledge sharing with Marco Marini, Alberto Sanchez Rodelgo, and Jim Tebrake of the Statistics Department at the International Monetary Fund. Michael Stepner and colleagues of the Economic Analysis Division of Statistics Canada provided relevant feedback and review. These colleagues are acknowledged for their valuable inputs, any errors and omissions remain the sole responsibility of the authors.

Table of contents

Overview	5
Features of the RT-LBCI	5
Real time: weekly frequency or higher	5
Geographic granularity: neighbourhood and city-level data	6
Minimum processing time: hours between data extraction and release.....	6
Integration of alternative data with Statistics Canada data holdings.....	6
Conceptual structure	6
Business size – static component.....	6
Business operating conditions – real-time component.....	7
Business level of activity – real-time component	7
Data sources	7
Business Register.....	7
Commercial APIs.....	7
Road traffic.....	8
Computation methods	8
Computation of the VRR	8
Extraction and computation of PBC.....	9
Extraction and computation of DIT	9
Interpretability and open methodological issues	9
Current coverage and update schedule	10
Ongoing developments	10
Appendix	11
References	12

This document is part of the Real-Time Local Business Conditions Index data visualization product. You can view the visual at www.150.statcan.gc.ca/n1/pub/71-607-x/71-607-x2021017-eng.htm.

Overview

Improving the timeliness of statistics has been a long-term objective of Statistics Canada. Since the beginning of the COVID-19 pandemic, near real-time statistics have gained prominence as a way to monitor rapidly changing social and economic activities, as well as the impact of containment measures and the subsequent recovery.

Statistics Canada has launched several initiatives to this effect, including flash estimates for several macroeconomic statistics in addition to a new quarterly survey on business conditions. The Real-Time Local Business Conditions Index (RT-LBCI) is implemented as a part of this effort. This document accompanies the inaugural experimental release of the RT-LBCI and presents the conceptual structure of the index, the data and computation methods, and current limitations.

The RT-LBCI is released as an experimental statistic. It is intended to provide a real-time signal on business activities following the disruptions brought about by the pandemic and through the recovery phase. It is acknowledged that this signal does not capture all dimensions of business conditions. Moreover, the methodological developments of this index are ongoing. New data sources may be incorporated into the index computation over time and, consequently, the computation methods may too be revised over time.

Features of the RT-LBCI

In the development of the RT-LBCI, the existing literature is of limited guidance. Existing research generally focuses on macroeconomic or financial indicators, usually available at a monthly frequency and at the national level (or State-level in the USA).

Much of this research originated from analysis and forecasting of business cycles, with the use of leading, coincident and lagging macroeconomic indicators (Lahiri and Moore 1991; Zarnowitz 2007). Since the economic downturn of 2008 and the increased economic volatility that followed, there has been a renewed attention devoted to real-time economic indicators, supported also by increasing availability of real-time data. For instance, the Aruoba-Diebold-Scotti Business Conditions Index¹ (Aruoba *et al.* 2009) tracks real business conditions at a high – generally monthly – observation frequency. Lewis *et al.* (2020) developed a Weekly Economic Index (WEI) using ten indicators of real economic activity which cover consumer behavior, the labour market, and production. The index is scaled to align with four-quarter Gross Domestic Product growth rates. Moreover, indices that utilise non-standard sources such as a weekly GDP tracker based on Google Trends (Woloszko, 2020) have more recently emerged. Indeed, these examples put an emphasis on monthly macroeconomic and financial indicators; however, the local dimension is rarely captured in such analyses.

Against this backdrop, four desirable features guided the development of the RT-LBCI. These features make the specification of the index unique, and are expected to remain at the core of further development and exploratory work of this measure.

Real time: weekly frequency or higher

The concept of timely data has evolved with changing economic context and new information technology. Today, the term “real-time” generally implies data of a weekly frequency or greater. To the extent possible, this is the type of data used in the development of this experimental index. The implication of this is a significant reliance of alternative data sources, as many traditional economic indicators have a considerable lag between their release

1. [Aruoba-Diebold-Scotti Business Conditions Index](#).

date and their reference period. As such, a large part of this real-time information is generated through online digital platforms.

Geographic granularity: neighbourhood and city-level data

The impact of the pandemic and its various waves was spatially uneven. Regions and metropolitan areas experience different measures and public health restrictions. It is expected that the recovery process will vary across space. Hence, the need for a high degree of geographic granularity was a guiding principle in the development of the index. To the extent possible, data are processed at location-level before being aggregated to the neighbourhood and municipality levels.

Minimum processing time: hours between data extraction and release

The use of weekly data imposes similarly short processing timelines. In the development of the index, specific attention was paid to the automation of data extraction processing and computation. A seamless processing pipeline was implemented in the cloud and is written in Python to reduce the extraction time and generate the index in a matter of hours.

Integration of alternative data with Statistics Canada data holdings

Finally, while alternative data sources are utilised for their high frequency and spatial granularity, the index is supplemented with Statistics Canada data holdings. These holdings provide robust and comprehensive business data which is infeasible, if not impossible, to obtain from alternative data sources. In the current version of the index, this was achieved at a geographic unit level, since microdata linkage was not attainable given compressed production timelines. Further integration with Statistics Canada's monthly survey results will be included alongside ongoing development of the index.

Conceptual structure

The current version of the RT-LBCI has three components. The first is a static component intended to capture the economic size of a business district at the local level; the second captures the operating conditions of the businesses in a given area; and, the third dimension captures the level of business activity in the area. These three components are further explained below.

Business size – static component

The static component captures where businesses are located, using the highest possible level of geographic granularity, and the economic size of these businesses. This dimension is used to define “business dense areas” within each population centre.

The business density measure is generated from Statistics Canada's business data. Data on the economic size of businesses at the establishment level is lacking from alternative data sources and thus this Statistics Canada data fills a data gap. Hence, this component is what currently provides the nexus between Statistics Canada's microdata holdings and the alternative (and real-time) data sources used for the other components of the index. The business size is computed using the total revenue of the businesses.

Business operating conditions – real-time component

The business operating conditions component is generated from a number of data sources acquired from private sector APIs that measure business closures (sources described in the next section). The measure that is generated based on these data is the percentage of businesses reporting closure (temporary or permanent) in a given local area.

The information is collected for businesses that can be generally defined as ‘Main Street’ and ‘non-essential’ businesses (more details on business types are provided in the following section). The sample of businesses used for this computation is drawn primarily from business dense areas. Closures are weighted by the economic size of the neighbourhood.

Business level of activity – real-time component

The business level of activity component is intended to capture mobility flows in the business district. In the current version of the index, this is estimated using road traffic flows. The existing literature suggests that traffic flows are a major determinant of the level of economic activity in the surrounding areas. This correlation was observed for foot traffic and retail sales (Perdikaki et al. 2011; Dong et al. 2017) as well as for road traffic and economic activity of the area (van Ruth 2014). While it is likely that the rapid expansion on e-commerce and online sales brought about by the COVID-19 pandemic has weakened this correlation, traffic flows are expected to remain a key determinant of economic performances of commercial districts as well as key determinants of broader economic vitality of a neighbourhood.

The measure that is currently generated from this data is the change from pre-pandemic (2019) to mid-pandemic levels for a given week, and is weighted by the size of the area, as measured by aggregate revenue. The methodological section provides more details on the specification of this component.

Data sources

There are three main sources of data that are used in the current version of the RT-LBCI, each used for a single component of the index: Business Register microdata, commercial APIs on business operation status, and road traffic data.

Business Register

The Business Register (BR) is a continuously maintained central repository of businesses and institutions operating in Canada. It provides national coverage of baseline information on total business revenue and employment. Microdata for businesses from 27 industries (see Appendix) were extracted and aggregated to the Dissemination Block² (DB) level for the year 2019 to capture pre-pandemic levels.

Commercial APIs

Data on business operating conditions are generated from three commercial APIs sources: Google Places, Yelp Fusion, and Zomato.

Google’s Places API provides access to one of the most comprehensive spatial databases on establishments and points of interest. This information is similar to what is publicly available on Google Maps. The primary variables of interest from the Places API for this index are the business status, location, and type of establishment. Google Places data are collected for the following types of establishments: restaurant, meal delivery, meal takeaway, café, bar, night club, home goods store, furniture store, hardware store, gym, beauty salon, hair care, supermarket, grocery, liquor store, department store, clothing store, jewelry store, shoe store, florist, casino, movie theatre, art gallery, museum, bowling alley, aquarium, zoo, amusement park, electronics store, bicycle store, book store, and bakery. The Yelp Fusion API provides similar information to what is available through Google Places, but is limited to businesses with operational and permanently closed statuses. Thus, it is used to supplement data obtained

2. A Dissemination Block is defined as an area bounded on all sides by roads and/or boundaries of standard geographic areas; generally, in urban areas DBs correspond to a city block. For detailed information see [Illustrated Glossary — Dissemination block \(DB\)](#).

from Google Places. Additionally, the Zomato API is used as a supplemental source for the food industry as it contains information on dine-ins, takeout, and similar features which are much less commonly found in other sources.

Due to API designs, returning a small number of businesses and limiting the number of queries, it is infeasible for them to be exhaustively queried in order to obtain information for the population of establishments. To mitigate this issue, a sampling approach was required. A variety of algorithms were developed and tested to maximize the coverage of the sample while minimizing the number of queries. The strategy used is referred to as the 'density-based cursory search'. This search strategy utilises BR data to identify the densest points of in-scope establishments in population centres by calculating the inverse distance between establishments. Searching these areas provides the most cost-effective results. The Places API text search functionality is used with the location, radius and keyword parameters.

Road traffic

TomTom's Historical Traffic Stats provides access to worldwide, anonymized GPS location data since 2008. It can provide insights on travel speed, travel time, and sample size on the road network. As with other data sources, there are challenges and constraints to using this data. Firstly, it is infeasible to create a measure of traffic that exhaustively covers all streets in a given population centre. Therefore, it is necessary to sample specific routes that are within business dense areas in order to get the most accurate measure from a limited amount of road. The second constraint is that the sample of road traffic is comprised of TomTom users exclusively. As such, there is an implicit assumption that TomTom users are representative of other users of the road network.

Additionally, there exists the overall challenge that only a proportion of traffic on the road will be consumers looking to engage (shop, dine, etc.) with establishments. To control for this, only routes within business dense areas were chosen, as road users in these areas are more likely to be consumers of establishments' products and services. Furthermore, as the pandemic has persisted, many establishments have moved to digital-based solutions to help weather the adverse economic impacts. For example, restaurants have moved to takeout-based solutions, and stores have moved toward web-based solutions, ultimately changing the relationship between physical traffic and revenue. Lastly, because the analysis is limited to a sample of routes, there is a risk that some routes may change significantly over time. For example, a road may be shut down for construction. Because of this, care must be taken when selecting routes to sample.

Computation methods

The index combines the static component on business size, derived from the BR, with the real-time data from commercial APIs. For each geographic unit, the general specification is

$$\text{RT-LBCI} = \text{VRR} \times (1 - \text{PBC}) \times (1 + \text{DIT})$$

where VRR is the value of retail revenue (business density measure), PBC is the percentage of business reporting closure, and DIT is the difference in traffic between the pre-pandemic and post-pandemic periods. The computation of this component is explained below.

Computation of the VRR

The static component on business size is implemented using a modified version of the gravity model presented in Alasia *et al.* (2020). The measure is based on a gravity model that accounts for the distance between a reference dissemination block (DB) and all the DBs within a given distance within which the service is available. The density measures also account for the businesses located in the reference DB.

Distances are based on driving network distances between the centroids of dissemination blocks (as opposed to straight-line distances). This revenue density measure is calculated with a 1 km buffer, meaning that all DBs within 1 km on the road network are included in the calculation for a given DB. This buffer distance was chosen as it is small enough to capture neighbourhood-level effects, but not so large as to capture the effects of many

other surrounding neighbourhoods. These DB level results are then aggregated to the Census Tract³ (CT) so to be combined with the percentage of business closures (PBC) component, which is also at the CT level.

Extraction and computation of PBC

The second component gives the percentage of business closures and is derived from commercial APIs providing information on operational businesses status. The proportion of closures is calculated for each census tract and then multiplied by the volume of retail revenue for the corresponding census tract. In effect, the proportion of business closures in each census tract are weighted by the volume of retail revenue of that same census tract.

Extraction of business conditions information from commercial APIs is implemented with a density-based cursory search. At its core, the idea of the algorithm is to identify the most establishment and revenue dense areas while ensuring these points are far enough away so that results from the API are unlikely to be duplicated. The first step is to identify and record the densest area. Next, the densest area and all other areas within a specified distance from said area are removed from the data set. The next densest area from the remaining areas is then located, and the process continues iteratively until the maximum number of areas are obtained.

The selected areas are then used to narrow down the search of establishments in the APIs. The establishments retrieved from the APIs are geocoded to Census geographies and aggregated to the CT level, where the proportion of closed businesses is derived. The percentage of business closures is then integrated with the volume of retail revenue component. These results are aggregated to the population centre⁴ level to be integrated with the difference in traffic component.

Extraction and computation of DIT

The third component is the difference in road traffic from pre-pandemic levels. This component is designed with the assumption that traffic flows in commercial areas are correlated with business activity. The measure is derived from TomTom's historical Traffic Stats data and is calculated from the percent change in the current week from the corresponding week in 2019 (pre-pandemic). A three-week moving average is used.

Because there are a small number of routes that are used in each population centre, the traffic component is aggregated to the population centre level by taking the mean of the traffic sample size for each route. The percent change is then calculated from this aggregate measure.

Interpretability and open methodological issues

The current version of the RT-LBCI is released as normalized index values with a baseline value set equal to 100 at the starting period of the series. An increase in the index value implies an improvement of business conditions in the geography of reference, while a decline of the index value implies a deterioration of business conditions. There are implications and limitations to this specification which are further explained below.

A normalized version is preferred for ease of comparability across jurisdictions. The absolute value of the index varies considerably between population centres, since the VRR component varies significantly among major metropolitan areas (specifically, Toronto, Montreal and Vancouver to name a few). In the current version, interpretation focuses on the direction and relative changes of trends, both within and across jurisdictions.

A second methodological decision is the choice of the baseline value. Data collection for the index started in the summer of 2020, when the first wave of the pandemic was recessing but a variety of containment measures were still in place in different regions and municipalities. Moreover, not all the data sources used for the index provided access to historical data. Hence, the starting period may reflect business conditions in different phases of the containment measures adopted at the local level. It should be kept in mind that the choice of the baseline suffers from this limitation.

3. Census tracts (CT) are small geographic units created in census metropolitan areas (CMAs) and larger census agglomerations (CAs). They provide a level of geography between the CMA or CA and the dissemination block allowing for further economic and social analysis. For detailed information see [Illustrated Glossary — Census tract \(CT\)](#).

4. Population centres are defined by Statistics Canada as having a population of at least 1,000 and a population density of 400 persons or more per square kilometer. This analysis focuses on some of the large urban population centres, with a population of 100,000 or more. For detailed information see [Illustrated Glossary — Population centre \(POPCTR\)](#).

In addition to these considerations on the interpretability of the index, there are other open methodological issues as well as assumptions that users should keep in mind. First, sampling methods are not standard statistical methods, because there are cost and technical considerations that limit the sampling strategy. Similarly, the choice of road network used to monitor traffic is guided by an informed decision rather than a form of statistical sampling. Finally, in an ideal scenario, microdata on operating status would be matched with corresponding business records in the BR; however, this process is not feasible in a compressed production process timeline like the one employed here. Although this analysis relies on a highly granular level of geography, the integration of BR data and commercial API data at the geographic unit level may introduce some form of aggregation bias, which at this point remains unassessed.

Current coverage and update schedule

The RT-LBCI covers 30 population centres which collectively house approximately 73% of the Canadian population⁵ and about 69% of employer businesses⁶. These population centres are:

- | | |
|---------------|------------------------------------|
| 1. Abbotsford | 16. Oshawa |
| 2. Barrie | 17. Ottawa–Gatineau |
| 3. Calgary | 18. Québec |
| 4. Edmonton | 19. Regina |
| 5. Guelph | 20. Saskatoon |
| 6. Halifax | 21. Sherbrooke |
| 7. Hamilton | 22. St. Catharines – Niagara Falls |
| 8. Kanata | 23. St. John's |
| 9. Kelowna | 24. Toronto |
| 10. Kingston | 25. Trois-Rivières |
| 11. Kitchener | 26. Vancouver |
| 12. London | 27. Victoria |
| 13. Milton | 28. White Rock |
| 14. Moncton | 29. Windsor |
| 15. Montréal | 30. Winnipeg |

The index is released weekly, every Friday at 8:30 a.m. (EST) for the preceding reference week. The release schedule may be subject to change, for instance, when statutory holidays coincide with release dates.

Ongoing developments

This is an experimental product which is expected to be further developed, revised, and expanded.

Geographic coverage is expected to expand over time to other major urban areas across Canada and include additional near-real-time and granular information. This is likely to result in changes to the methodology and computational processes. Finally, further integration with Statistics Canada data holdings, in particular monthly surveys will be further explored.

5. [Municipalities in Canada with the largest and fastest-growing populations between 2011 and 2016.](#)

6. [Canadian Business Counts, with employees, census metropolitan areas and census subdivisions, June 2020.](#)

Appendix

Table 1
NAICS table

NAICS Code	Industry group / Industry
311811	Retail bakeries
442	Furniture and home furnishings stores
44314	Electronics and appliance stores
444	Building material and garden equipment and supplies dealers
445	Food and beverage stores
44711	Gasoline stations with convenience stores
448	Clothing and clothing accessories stores
451113	Cycling equipment and supplies specialty stores
451310	Book stores and news dealers
452	General merchandise stores
4531	Florists
51213	Motion picture and video exhibition
6212	Offices of dentists
71211	Museums
71213	Zoos and botanical gardens
713110	Amusement and theme parks
71321	Casinos (except casino hotels)
713940	Fitness and recreational sports centres
71395	Bowling centres
72241	Drinking places (alcoholic beverages)
72251	Full-service restaurants and limited-service eating places
8121	Personal care services

References

- Alasia, Alessandro, Nick Newstead, Joseph Kuchar, and Marian Radulescu (2021). [Measuring proximity to services and amenities: An experimental set of indicators for neighbourhoods and localities](#). *Reports on Special Business Projects*. Catalogue no. 18-001-X. Statistics Canada.
<https://www150.statcan.gc.ca/n1/pub/18-001-x/18-001-x2020001-eng.htm>
- Aruoba, S.B., F.X. Diebold, and C. Scotti (2009). “[Real-Time Measurement of Business Conditions](#),” *Journal of Business and Economic Statistics* vol. 27, no. 4 (October): 417-27.
https://www.jstor.org/stable/27799095?seq=1#metadata_info_tab_contents
- Dong Lei, Sicong Chen, Yunsheng Cheng, Zhengwei Wu, Chao Li, and Haishan Wu (2017). Measuring economic activity in China with mobile big data. *EPJ Data Science* 6:29. <https://doi.org/10.1140/epjds/s13688-017-0125-5>
- Lahiri, Kajal, and Geoffrey H. Moore, Edited by (1991). *Leading economic indicators. New approaches and forecasting records*. Cambridge University Press.
- Lewis, D., K. Mertens, and J. Stock (2020). “Monitoring Real Activity in Real Time: The Weekly Economic Index,” *Federal Reserve Bank of New York Liberty Street Economics*, March 30.
- Perdikaki, Olga, Saravanan Kesavan, and Jayashankar M. Swaminathan (2011). Effect of Traffic on Sales and Conversion Rates of Retail Stores. *Manufacturing & Service Operations Management* 14 (1) 145-162
<https://doi.org/10.1287/msom.1110.0356>
- van Ruth, Floris (2014). [Traffic intensity as indicator of regional economic activity](#). *Discussion Paper 21*.
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.671.5907&rep=rep1&type=pdf>
- Woloszko, N. (2020). “[Tracking activity in real time with Google Trends](#)”. *OECD Economics Department Working Papers*, No. 1634, OECD Publishing.
https://www.oecd-ilibrary.org/economics/tracking-activity-in-real-time-with-google-trends_6b9c7518-en
- Zarnowitz, Victor (2007). “Composite Indexes of Leading, Coincident, and Lagging Indicators”. *Business Cycles*, Chicago: University of Chicago Press, pp. 316-356.