Business Closures and (Re)Openings in Real-Time Using Google Places: Proof of Concept

Daniel E. Rigobon 1, Thibaut Duprey 2,*, Artur Kotlicki 3, Philip Schnattinger 3, Soheil Baharian 2 and Thomas R. Hurd 4

1 Operations Research & Financial Engineering (ORFE), Princeton University, Princeton, NJ 08540, USA; drigobon@princeton.edu
2 Bank of Canada, Ottawa, ON K1A 0G9, Canada; sbaharian@bank-banque-canada.ca
3 Bank of England, London EC2R 8AH, UK; artur.kotlicki@bankofengland.co.uk (A.K.); philip.schnattinger@gmail.com (P.S.)
4 Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S 4K1, Canada; hurdt@mcmaster.ca
* Correspondence: tduprey@bankofcanada.ca

Abstract: We present a new estimation of business opening and closure rates using data from Google Places—the data set behind the Google Maps service. Our algorithm, through a bisection routine, counts the appearance and disappearance of “pins” that represent unique businesses. As a proof of concept, we compute business opening and closure rates for the city of Ottawa during the reopening phase of the COVID-19 pandemic in mid-2021. The lifting of restrictions coincides with a wave of re-entry of temporarily closed businesses, suggesting that government support may have facilitated the survival of hibernating businesses. Our entry estimates are validated by a survey of new businesses. This methodology allows policymakers to monitor business dynamics in quasi-real-time during rapidly unfolding crises.

Keywords: business entry rate; business exit rate; Google Places; real-time; small and medium enterprises

JEL Classification: D22; E32; C55; C81

1. Introduction

The COVID-19 pandemic is the largest public health crisis in recent history. To contain the spread of the virus and protect lives, countries around the world instituted lockdowns or other restrictions to both individual and business activities. While these policy measures were effective in reducing infection rates, they also placed businesses under substantial financial stress, especially those relying on in-person interactions. Governments’ financial support for businesses was a balancing act between doing too-little-too-late that leads to a permanent loss of productive firms, versus doing too-much-too-broadly that facilitates the survival of non-productive “zombie” firms (Favara et al. 2021). To help policymakers calibrate business support policies during fast-paced crises, how can we measure business dynamics in a timely manner?

National statistical offices often report rates of business entry and exit. For example, Statistics Canada (2020b) recently began computing business entry and exit rates using tax records. However, the pandemic highlighted two limitations. First, traditional estimates are produced at a given frequency that cannot be adjusted to match the speed of the crisis. Second, estimates are released with a lag of several months, which limits the real-time value to policymakers.

In this paper, we use Google Places, the database behind Google Maps, to derive a proof of concept on how to compute rates of entry, reopening, temporary closure, and exit
for labor-intensive and customer-facing business establishments. Specifically, we ask the following questions: Do we observe a surge of business entry after the end of the COVID-19 lockdown using Google Places data? If so, is the recovery driven by new businesses or hibernating businesses that reopened once allowed to do so? If hibernating businesses reopened, this would suggest that some viable businesses managed to adapt and did not close permanently, possibly as a result of government support, thereby limiting the long run productivity cost of the lockdown and speeding up the recovery. Our approach approximates business dynamics by tracking changes in the unique “pins” visible on Google Maps. Essentially, we develop an algorithm to identify and count all businesses of given types, on a given date and within a given geographical area. A comparison between two dates allows us to compute the desired rates. Our methodology provides several advantages. First, data can be collected at any frequency and without delay. Second, our scraping procedure allows for any choice of granularity. Third, Google Places distinguishes between temporary and permanent closures. Finally, it is possible to include information about customer reviews, location, and cost range, which provides additional insights into a business’ activity.

We apply our method to the retail and food service (e.g., restaurant and cafe) sectors of Ottawa in the spring and summer of 2021—during the first reopening phase of the pandemic. We observed a substantial wave of business entries immediately after restrictions were lifted in Ontario. Sectoral heterogeneity in the phases of the lifting of restrictions is associated with different peaks in the entry rates. Entries peak in June for the food and retail sectors, after phase one of the reopening on June 11, while they peak in July for night clubs that are allowed to reopen only after phase three on July 16. To provide external validity to our estimates, we conduct a survey of recently (re)opened business establishments and validate at least 71% of the business entries in the food sector, and 44% in the retail sector. These results indicate that Google Places data is capable of capturing the response of business establishments to policy interventions, and that the data quality is best for the food service sector. This may reflect the importance of an online presence for the food sector, perhaps for takeouts, deliveries, and reviews.

Our approach uniquely allows us to distinguish between temporarily and permanently closed business establishments. The wave of business establishment entry after the lifting of restrictions was largely driven by reopenings. Specifically, during the stay-at-home order of April 2021, about 30% of night clubs and bars, 15% of cafes and restaurants, and 10% of retail stores were temporarily closed. About half of these businesses reopened by the end of September. Our results are indicative of hibernation during the pandemic, which may have been supported by government subsidies.

The paper is organized as follows. Section 2 reviews the most relevant literature. Section 3 outlines the data, our scraping methodology, and computation of entry and exit rates. Section 4 presents our main results and figures for businesses in Ottawa. Next, Section 5 discusses the results with relevant limitations, and Section 6 concludes.

2. Literature Review

The use of alternative data for economic measurement is not new. For instance, Woloszko (2020) uses keywords from Google Trends combined via a machine learning neural network approach to construct real-time proxy of GDP, and Cavallo and Rigobon (2016) collect millions of online retailers’ prices to create new micro-based consumer price indices that proxy for inflation.

However, the concern for timely measures of business resilience was largely spurred by the pandemic. To our knowledge, only four pandemic-related analyses closely relate to ours.

Yelp (2020) used its platform’s business reviews by customers to compute the relative importance of temporary and permanent closures during the early phase of the COVID-19 crisis. They reported the fraction of businesses on their platform that were neither active nor formally closed and showed that many businesses were merely hibernating.
Crane et al. (2020) highlighted the shortcomings of lagging official statistics on business dynamics and reviewed what alternative data sources could be used to measure business exits in quasi-real-time during the COVID-19 pandemic (e.g., Google keyword searches, paycheck issuance, and phone-tracking data). Their measures indicate that, on the one hand, exits were elevated in certain sectors during the first year of the pandemic, but, on the other hand, many industries experienced lower-than-usual exit rates, possibly due to government support and optimism about their survival prospects.

Kurmann et al. (2021) matched small US business establishment records from Homebase (a software provider) with information on business activity as recovered from Google, Facebook, and SafeGraph. Focusing primarily on the first wave of COVID-19 lockdowns, they found that small business employment contracted most in the hardest-hit service sectors but also rebounded more strongly, with many closed businesses subsequently reopening. Although business closures accounted for 70% of the initial decline in small business employment, two thirds of closed businesses have reopened. They also show that businesses that received earlier government support were less likely to close, suggesting that the government support to small US businesses helped mitigate some effects of the pandemic. We instead comprehensively estimate opening and closure rates in a given geographical area across the most severely hit sectors and assess the quality of the estimates. Similar to them, we find that about 50% of the Canadian businesses that closed during the later wave of lockdown (in early 2021) eventually reopened.

In contemporaneous work, Statistics Canada (2021) construct a real-time local business conditions index to provide a real-time signal on business activities following the disruptions brought about by the pandemic and through the recovery phase. They combined two key sources of data. On the one hand, they used a sampling strategy to collect data from Google Places to approximate the fraction of new or closed businesses, thereby capturing the extensive margin of business dynamics. On the other hand, they used road traffic data to approximate for the extent of the economic activity of opened businesses, thereby capturing the intensive margin of business dynamics. We instead focus on the extensive margin only by suggesting an alternative bisection method to comprehensively compute rates of opening and closure by sectors while distinguishing reopenings from new entries. In addition, we assess the quality of our data with a survey of recently opened businesses.

3. Materials and Methods

3.1. Data

We use Google Places to identify and count businesses in a geographic area. The information is likely to be comprehensive and timely, but it relies on updates by Google coming from: business owners who have a business account, customers who provide reviews (about 80% of businesses in our data set have at least one review), users who report inaccurate listings, or other publicly available information.

We focus on establishments in the retail and food service (e.g., restaurants, bars, cafes, night-clubs) sectors in the downtown core of Ottawa/Gatineau. We do so for the following reasons. First, sectors with face-to-face consumer interactions likely have a better reporting status, given their incentives to be on Google Maps. Second, areas with the most foot traffic may have better data quality due to reviews by Google users. Third, these sectors were the most affected by lockdowns during the COVID-19 crisis and thus most relevant to track in a timely manner.

We use the functionality of “Nearby Search” in queries to Google Places API. Instead of searching for a specific business as in Kurmann et al. (2021), this query returns all businesses of a given type within a bounding circle—defined by a point (in latitude and longitude) and a radius (in meters). Out of 96 possible business types to query, we use “store”, “gas_station”, “restaurant”, “bar”, “cafe”, and “night_club”. Those keywords allow us to match North American Industry Classification System (NAICS) codes 44/45 and 722 for the retail and food sectors, respectively.
3.2. Cross-Section of Businesses

Each query returns at most 20 places, with a flag indicating whether more than 20 places are available, but not returned by the query. We use this flag to design a simple bisection algorithm to find a set of queries such that (1) each query returns no greater than 20 results and (2) a desired area is fully covered.

We begin with a single large square and query the circle that circumscribes it. Whenever the query flags that there are more than 20 results, we sub-divide the square into four smaller squares and re-run the query on each. This terminates when there are no more than 20 results per query. Figure 1 shows that, the higher the density of businesses (the dots), the finer the search grid needs to be (the squares).

Figure 1. Illustration of the bisection algorithm for keyword “restaurant” in downtown Ottawa. Note: The blue shape is the bounding box of the set of Forward Sortation Area with postal codes K1A, K1N, K1R, K1S, K1P, K2P. The vertical and horizontal axes represent latitude-longitude coordinates. The red squares are those inscribed in the coverage disks of each query, and the green points indicate the businesses found. This data was collected from Google Places in May 2021.

3.3. Time Series of Openings and Closures

Since the Google Places API returns only the most recent information and no historical data, we repeatedly scrape the same area monthly to build a time series. To save on costs, instead of beginning each data collection with an uninformative grid of a single square, we initialize the algorithm by using the grid of squares resulting from the previous month’s scrape.

We identify an exit if a business’ unique identifier `place_id` is removed from the data set. We further identify temporary closures by using the feature `business_status`, which indicates whether a business is currently operational or temporarily closed. The closure rate is computed as the fraction of exiting or temporarily closed businesses compared to the operational businesses in the previous month (see Table 1 for details).

Table 1. Business openings (entries and reopenings) and closures (exits and temporary closures).

<table>
<thead>
<tr>
<th>Open in t with:</th>
<th>Temporarily</th>
<th>Non-Existen</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤10 reviews</td>
<td>&lt;10 reviews</td>
<td>closed in t</td>
</tr>
<tr>
<td>&gt;10 reviews</td>
<td></td>
<td>in t</td>
</tr>
<tr>
<td>continuing</td>
<td></td>
<td>Temporary closure Exit</td>
</tr>
<tr>
<td>temporarily closed in t – 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-existent in t – 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Likewise, we identify an entry when a new unique `place_id` appears in the data set. A reopening corresponds to a business that was previously temporarily closed but is now operational again. The opening rate is computed as the fraction of entrant or reopening businesses compared to the previous month.

If a business establishment is not immediately included in Google Places’ registry upon opening, it may enter the data set later when reporting is improved. By then, it is likely to have accumulated customer reviews already. Conversely, a business opening in a given month is unlikely to have many reviews yet. Therefore, we require new openings in a month to have at most 10 reviews. The choice of the 10-review cutoff is informed by a survey we discuss below.

4. Results

We reviewed a total of about 3000 unique businesses\(^3\) observed from April to September 2021. About 55% are retail and 45% are food service businesses.

Figure 2 shows the monthly opening and closure rates. The opening rate peaks in June, consistent with the timing of the lifting of the lockdown and stay-at-home orders on June 2nd for Ontario. Restrictions were lifted in June for retail businesses, bars and restaurants, but only in July for night clubs, which explains the delayed peak for night clubs. Note, however, that some establishments can be classified as night clubs, bars and restaurants at the same time, such that some night clubs re-started some of their operations at the same time as other food businesses. As expected, business openings were larger than business closures from the end of spring onward, but both have balanced out by the end of the summer of 2021. The spike of business openings in late spring is largely driven by reopenings rather than new entries.

Our data is suggestive of the existence of hibernating businesses during the pandemic. Figure 3 shows the fraction of businesses identified as temporarily closed. The share of temporary closures was highest for night clubs and bars at around 30% in April 2021, prior to the lifting of COVID restrictions. Night clubs were not all temporarily closed, because many are also restaurants, allowing for takeout. If one considers only those businesses recorded as temporarily closed in April or May, about 50% reopened by the end of September, with the remainder still classified as temporarily closed or definitively disappearing from the data set.

To validate our method, we conducted a survey of business openings in Ottawa/Gatineau from May to September 2021 by either calls or in-person checks. We confirmed 71% of the new openings in the food sector. However, if we also treat as entries the businesses that first appear in the data set with more than 10 customer reviews, then the ratio of confirmed openings would be reduced to 62%. This is in line with our identification assumption that true business entries are those associated with few reviews upon entry. In our survey, we also confirmed 44% of the new openings in the retail sector. Survey results are consistent with the view that data quality is superior for the food sector in busy areas, likely because of active reviews by customers and incentives to maintain an online presence for takeout. Anecdotally, well-known chains that opened or closed locations during our study were systematically correctly reported. The percentages reported by this survey should be understood as a lower-bound. First, the full address or the phone number of a business may not be available. Second, some businesses may operate from an office building that is not accessible given COVID-19 restrictions. Finally, a few businesses are located in residential buildings, mostly related to online retailers that operate from the owner’s residence.
Figure 2. Business opening and closure rates in 2021. Note: The monthly opening and closure rates for the core downtown of Ottawa/Gatineau are derived from Google Places. The food sector is the aggregation of the results by the keywords bar, cafe, restaurant and night_club.
Figure 3. Evolution of the rate of businesses temporarily closed in 2021. Note: The monthly rate of temporarily closed businesses for the core downtown of Ottawa/Gatineau is derived from Google Places. The food sector is the aggregation of the results by the keywords bar, cafe, restaurant and night_club. Note that the point in April for the retail sector is an estimate based on a smaller sample. The dark grey shaded area corresponds to the lockdown and stay-at-home policies, while the light grey area corresponds to the extended period of restrictions only for night clubs.

5. Discussion

We note a few limitations of our method. First, pre-pandemic data cannot be collected. The absence of long-term data currently prevents us from benchmarking our opening and closure rates with normal times. Second, the reopening of a business (extensive margin) does not imply it returns to the level of economic activity prior to its closure, such that other data would be needed to measure the intensive margin. For example, Statistics Canada (2021) use traffic data around businesses to proxy for the intensive margin. Third, business closures are harder to assess because a business no longer exists to confirm the timing of its closure. Still, our estimates correlate well with data from Statistics Canada (2020a) for a range of Canadian cities (Duprey et al. 2022). Fourth, the quality of our estimates is partly outside our control as it is inherently tied to the quality of Google Places’ data itself. This may further complicate the measurement of business closures. For instance, a closed business not reported as such—by the owner or customers—may remain in the data set until a new business opens at the same location. Lastly, our method cannot infer the rationale behind business closures, but it captures a broad definition of business closures that includes voluntary closures that are usually harder to track. Indeed there were fewer court-mediated bankruptcies and more voluntary closures during the pandemic, because many businesses voluntarily closed due to restrictions or low expected profits.

Still, our method can be scaled and applied to other cities and sectors. The wide availability of Google Places’ data allows for a straightforward transfer to other areas lacking high frequency data on business establishment dynamics. Given the cost constraint (which in our case amounts to about CAD 1700 for a year’s worth of data, collected monthly, on downtown Ottawa’s retail and food sectors) to download a comprehensive data coverage across time, sectors and geographical areas, further optimization of the scraping algorithm or sampling methods as in Statistics Canada (2021) might be needed. As the digitalization of economies continues, the coverage and reliability of a data set such
as Google Places will continue to improve, and Google itself might consider compiling and monetizing dedicated business health statistics for research and policy-making.

Our new method also successfully captures business establishment dynamics at a micro level. As such, it fills a gap in the understanding of small firms, for which micro data are usually not publicly available. Typically the literature on small- and medium-sized firms relies either on confidential micro-data (e.g., Horvath and Lang (2021) studied how subsidized loans affect firms’ investment and employment) or survey data (e.g., Barikowska et al. (2020) studied business turnover and profits during the COVID-19 pandemic). A combination of our data set with tax and banking records could allow for a deeper look at the extensive margin of employment in labor-intensive small businesses (see, e.g., Neumark et al. 2016; Kurmann et al. 2021) or the role of banking intermediation and relationship lending on the probability of business creation and survival (see, e.g., Beck et al. 2018). Following Bahaj et al. (2022), we could also link our data with other novel high-frequency data, for example to include job vacancies data in real-time.

More broadly, firm dynamics play an important role to understand aggregate business cycle fluctuations (Clementi and Palazzo 2016). The properties of entering and exiting firms are crucial for the welfare consequences of monetary policy (Hamano and Zanetti 2022) and supply chain disruptions (Bilbiie and Melitz 2020). Google Places information on firms sector, price level, popularity, location and hours of operation could serve as a proxy for the demand and competitiveness. These properties and their distribution among the firm population are ultimately determining the allocative efficiency and aggregate welfare in standard macroeconomic models (Dhingra and Morrow 2019).

Eventually, the extension of our data set over both time and space could be used to study several other topics. First, our data set could shed some light on new trends among small businesses. Anecdotal evidence from our survey suggests the existence of self-employed retail businesses that operate only online from the owners’ residence, e.g., “Instagram sellers”. We can also use the high-frequency data to capture the effect of natural catastrophes or other sudden supply chain disruptions on business dynamics. With greater geographic coverage, one could investigate the impact of input or output price shocks—such as an increase to one region’s minimum wage or large oil price swings in oil-intensive regions—on business closure rates. With a longer time series, another possible avenue of work would be to compare the dynamics in downtown core areas to that of commercial areas outside the main city centers, for instance after the re-zoning of a commercial area or the creation of a new public transportation infrastructure.

6. Conclusions

We introduce a new method to track business dynamics in a timely manner using Google Places, the data behind the Google Maps service. By uniquely identifying and counting the “pins” on the map when business establishments are added or removed from Google Maps, we provide policymakers with a useful tool to monitor business dynamics in quasi-real-time. This is especially useful during rapidly unfolding crises, i.e., the COVID-19 pandemic. We successfully identify waves of business entry, which are largely driven by reopenings. During the COVID-19 pandemic, we show that a significant fraction of businesses were only temporarily closed and reopened shortly after regional restrictions were lifted. The reopening of viable businesses suggests that they managed to adapt, thereby limiting the long run productivity cost of the lockdown. This also suggests a meaningful impact of government policies that supported hibernating businesses, thereby speeding up the recovery once the economy reopened.

Funding: Funding to graduate students at the time was provided as part of the Extended Problem Solving Workshop jointly organized by the The Fields Institute for Research in Mathematical Sciences and the Organisation for Economic Co-operation and Development. Data acquisition was funded by the Bank of Canada.

Data Availability Statement: The data was downloaded from the API of Google Places https://developers.google.com/maps/documentation/places/web-service (accessed on 12 April 2022), the data behind the Google Maps service, in accordance with their terms of service. Data was collected approximately monthly, between April and September of 2021.

Acknowledgments: This short paper builds on the outcome of the 2021 OECD-Fields Extended Problem Solving Workshop. We are grateful to Matheus Grasselli for making this opportunity happen. We would like to thank Angus Armstrong, Gino Cateau, William Hynes, Mikael Kahn, Alan Kirman, and Sharon Kozicki for their encouragements, as well as André Kurmann, Danny Leung and Nicholas Sander for their comments on earlier drafts.

Conflicts of Interest: The authors declare no conflict of interest. The views expressed in the paper are those of the author(s), and neither represent those of the Bank of Canada nor of the Bank of England or its committees.

Notes
1 See https://support.google.com/local-listings/answer/98510997?authuser=0&hl=en&visit_id=63753932185538809-1024761200&rd=1 (accessed on 12 April 2022) for more details on Google’s data collection.
3 Google Places has very good coverage, but benchmarking is difficult. For instance, statistics on businesses by Statistics Canada are available at the enterprise level, whereas it is available at the establishment level for Google Places. In addition, publicly available statistics cover the whole metropolitan area, not just the downtown core, and the breakdown by sector is not always available.
4 In April, we did not collect comprehensive data for the retail sector, such that some statistics comparing April to May cannot be computed.
5 This cost estimate does not include potential free credits offered by Google Places.

References
Cavallo, Alberto, and Roberto Rigobon. 2016. The billion prices project: Using online prices for measurement and research. Journal of Economic Perspectives 30: 151–78. [CrossRef]
Hamano, Masashige, and Francesco Zanetti. 2022. Monetary policy, firm heterogeneity, and product variety. European Economic Review 144: 104089. [CrossRef]


