What Yelp Reviews Can Tell Us About the State of the Economy

Data from online sources like Yelp can help you take the pulse of local business activity in the absence of updated official statistics.

Businesses generate a tremendous amount of digital exhaust – that virtual trail of **data they collect as part of their operations**. When you buy a shirt online, for example, the type of credit card you use, the time of your purchase and other such information not core to the transaction is captured and stored. The question is, how useful could it be exactly?

Using data from the online reviews platform Yelp, my co-authors* and I sought to find out. In a **paper** forthcoming in *Big Data for Twenty-First Century Economic Statistics* published by the US National Bureau of Economic Research, we show that Yelp's crowdsourced data can help measure economic activity at the local level in close to real time, in contrast with official data that is often published years after. Changes in the number of businesses and restaurants reviewed on Yelp can help "**nowcast**" changes in the corresponding official statistics before they are released. In short, sources like Yelp could complement official data in business and policymaking.

Real-time 'nowcasting'

Take the US Census Bureau's **County Business Patterns**. CBP publishes annual statistics on the number of businesses, employees and payroll by state, county, metropolitan area, ZIP code and congressional district levels – but with a significant time lag. As of January 2021, the latest available CBP data was from 2018, aggregated to the ZIP code level.

Data collected by online platforms such as Google, LinkedIn and Yelp may fill this gap. For our paper, my co-authors and I zoomed in on Yelp, which by the end of 2016 had listed over 3.7 million businesses with 65.4 million "recommended reviews", or reviews that are deemed to be authentic or helpful.

We began our analysis with a comparison of Yelp and CBP data between 2009 and 2015. We only counted businesses as open if they had received at least one Yelp recommended review. We limited our analysis to ZIP codes with at least one business in CBP and Yelp in 2009.

In 2015 (the last year of CBP data available), in the restaurant sector, CBP listed 542,029 businesses in 24,790 ZIP codes, and Yelp listed 576,233 in 22,719 ZIP codes. The Yelp-to-CBP coverage ratio was, therefore, 106 percent, meaning Yelp might have captured more restaurants than it missed. This pattern was most pronounced in densely populated, wealthier areas, pointing to a combined effect of Yelp covering smaller joints with no employees – such businesses are excluded from CBP – and the

penchant of people living in urban, more affluent regions to eat out and share their experience online.

We then explored whether Yelp data can predict changes in the overall number of companies as well as that of restaurants in CBP before the official statistics are released. We found that, after accounting for what historical CBP data could project, Yelp data could account for 26 percent of the remaining change in restaurant openings, and 29 percent of the remaining business openings in CBP.

Limits of data exhaust

Further analysis showed that Yelp is more predictive in richer, more densely populated and more educated ZIP codes, likely for reasons mentioned earlier. Each new Yelp business is associated with 0.75 extra CBP establishments in an area like New York City's Upper East Side (population: 60,453; area: 1.22 square kilometres), compared to 0.5 in a less affluent region and 0.2 in places that are less educated, poorer and sparsely populated.

Assessing Yelp's predictive power by industry, we found that Yelp data could predict 8.5 to 10.2 percent of CBP changes for retail, leisure, hospitality, as well as professional and business services, compared to 0.9 to 8.2 percent in public services, goods manufacturing or transportation and wholesale trade. This finding further highlights the limitations of using Yelp data for nowcasting for mayors and business owners alike.

Our study demonstrates how digital exhaust could be repurposed for uses quite different from what generated the data in the first place. Increasingly, companies are beginning to engage in selling data as one of their revenue streams, arousing concerns over privacy. For businesses and managers seeking to leverage such incidental data, our research suggests they would benefit from a careful evaluation of the potential pay-offs as this is by no means an all-purpose tool.

***Edward Glaeser**, the Fred and Eleanor Glimp Professor of Economics at Harvard University, and **Michael Luca**, the Lee J. Styslinger III Associate Professor of Business Administration at Harvard Business School.

Hyunjin Kim is an Assistant Professor of Strategy at INSEAD. She studies how firms can manage data and algorithms to improve their strategic decision making, and how these technologies change how firms compete and build competitive advantage.

*INSEAD Knowledge is now on **LinkedIn**. Join the conversation today.*

Visit INSEAD Knowledge
http://knowledge.insead.edu

*Follow INSEAD Knowledge on **Twitter** and **Facebook**.*

Find article at
https://knowledge.insead.edu/strategy/what-yelp-reviews-can-tell-us-about-the-state-of-the-economy-16266

Download the Knowledge app for free