

What Yelp Data Can Tell Us About The Minimum Wage (and other policies)

By Michael Luca

In April 1992, New Jersey increased the state's hourly minimum wage from \$4.25 to \$5.05. The change was controversial, as some policymakers raised concerns that higher minimum wages might have the unintended consequence of increasing unemployment rates. Business leaders also expressed mixed opinions, with some worrying that higher minimum wages might be bad for business.

Princeton economists David Card and Alan Krueger were watching the discussion unfold, and set out to understand exactly how the wage hike would affect New Jersey jobs.¹ They made a plan to survey fast food

restaurants before and after the new wage went into effect to see if it had an impact on employment in an industry concentrated with minimum wage workers. In need of a control group where there was no increase in the minimum wage, the researchers chose the neighboring state of Pennsylvania.

Flipping through telephone books in the months before the new wage became effective, the authors identified 473 stores in New Jersey and Pennsylvania to be in their sample. A team of interviewers sat down and called each of the stores, sometimes up to nine times before someone answered, and asked questions about employment, starting wages, prices, and other store characteristics. With the interviewers' persistent efforts, they

achieved relatively high response rates. Ultimately, they were able to complete 410 phone interviews, for a response rate of 87 percent.

Several months after the minimum wage increase, the interviewers called the same stores again for a follow-up survey. While the majority of stores picked up the phone, 39 did not, so the research team drove to all 39 holdouts and asked them to complete the survey in person. This resulted in a 99.8 percent response rate for the follow-up survey.

The labor-intensive survey yielded important results. Despite the large hike in the minimum wage, the authors found that employment at these establishments did not seem to suffer. Now, 25 years later, the findings are still regularly cited in discussions of the minimum wage. Moreover, the study is still taught for its methodological approach of comparing changes in New Jersey with changes in the control state of Pennsylvania — a canonical example of what economists call a difference-in-differences analysis.

In the decades since the Card and Krueger paper, there has been a

¹ Card, David, and Alan B. Krueger. 1994. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review*, 84(4): 772-793.

About the Author



Michael Luca is a Visiting Scholar at SIEPR and an assistant professor at Harvard Business School. His research and teaching focus on the economics of digitization and on using data to improve policy and managerial decisions.

large body of research exploring the impact of minimum wage changes — ranging from employment and wage effects to business outcomes, including prices of products and whether (and which) businesses go out of business after an increase in the minimum wage.

Recent years have witnessed the rise of review platforms, including Yelp, TripAdvisor, and others, allowing for online reviews of everything from restaurants to doctors and providing customers with unprecedented amounts of information about the quality of goods and services. While these digital platforms are designed with consumers in mind, it turns out that they also have the potential to transform the way that researchers and policymakers do their work.

A New Approach

Two types of datasets have formed the backbone of analyses studying minimum wage hikes. First, researcher-administered surveys, such as ones used by Katz and Krueger (1992) and Card and Krueger (1994), have been an important data source.² Second, government datasets from

agencies such as the U.S. Census Bureau and the Bureau of Labor Statistics (BLS) have been frequently used to estimate the impact of the minimum wage in the U.S. (Dube, Lester and Reich, 2010; Aaronson et al., forthcoming).^{3,4}

In a recent paper, “Survival of the Fittest? The Impact of the Minimum Wage on Firm Exit,”⁵ Dara Lee Luca and I set out to understand whether the minimum wage was causing restaurants to close — and if so, which ones. Our interest in the minimum wage grew with the renewed interest among policymakers in increasing the minimum wage. Recent proposals from policymakers have ranged from eliminating the federal minimum wage entirely to increasing it to \$15 per hour. Beyond the \$7.25 federal minimum wage, states — and more recently, cities — have increasingly set their own higher local minimum wages.

3 Dube, Arindrajit, T. William Lester, and Michael Reich. 2010. “Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties.” *The Review of Economics and Statistics*, 92(4): 945-964.

4 Aaronson, Daniel, Eric French, Isaac Sorkin, and Ted To. (forthcoming). Industry Dynamics and the Minimum Wage: A Putty-Clay Approach. *International Economic Review*.

5 Luca, Dara Lee, and Michael Luca. 2017. “Survival of the Fittest: The Impact of the Minimum Wage on Firm Exit.” Harvard Business School NOM Unit Working Paper No. 17-088.

We noticed that there had been 21 local minimum wage changes in the San Francisco Bay Area in a few short years, creating an ideal test bed for understanding the impact of these changes. Similar to Card and Krueger, we focused on the restaurant industry. However, instead of conducting a survey or flipping through phonebooks, we turned to Yelp — the online review giant headquartered in San Francisco. Through a data-sharing agreement, we were able to construct a rich dataset on restaurant characteristics and outcomes, including each business’s Yelp rating, approximate price range, and an indicator capturing whether and when restaurants had gone out of business.

Data in hand, our analyses revealed two main findings. First, at any wage level, lower-rated restaurants are more likely to go out of business than higher-rated ones. This relationship persists after controlling for price and other business characteristics. So, for example, a lower-rated pizza place is more likely to go out of business than a higher-rated one. This suggests that lower-quality restaurants generally struggle more and may be more sensitive to increased costs.

Second, the impact of minimum wage increases is greater for lower-rated businesses than for higher-rated

2 Katz, Lawrence F., and Alan B. Krueger. 1992. “The Effect of the Minimum Wage on the Fast-Food Industry.” *Industrial and Labor Relations Review*, 46(1): 6-21.

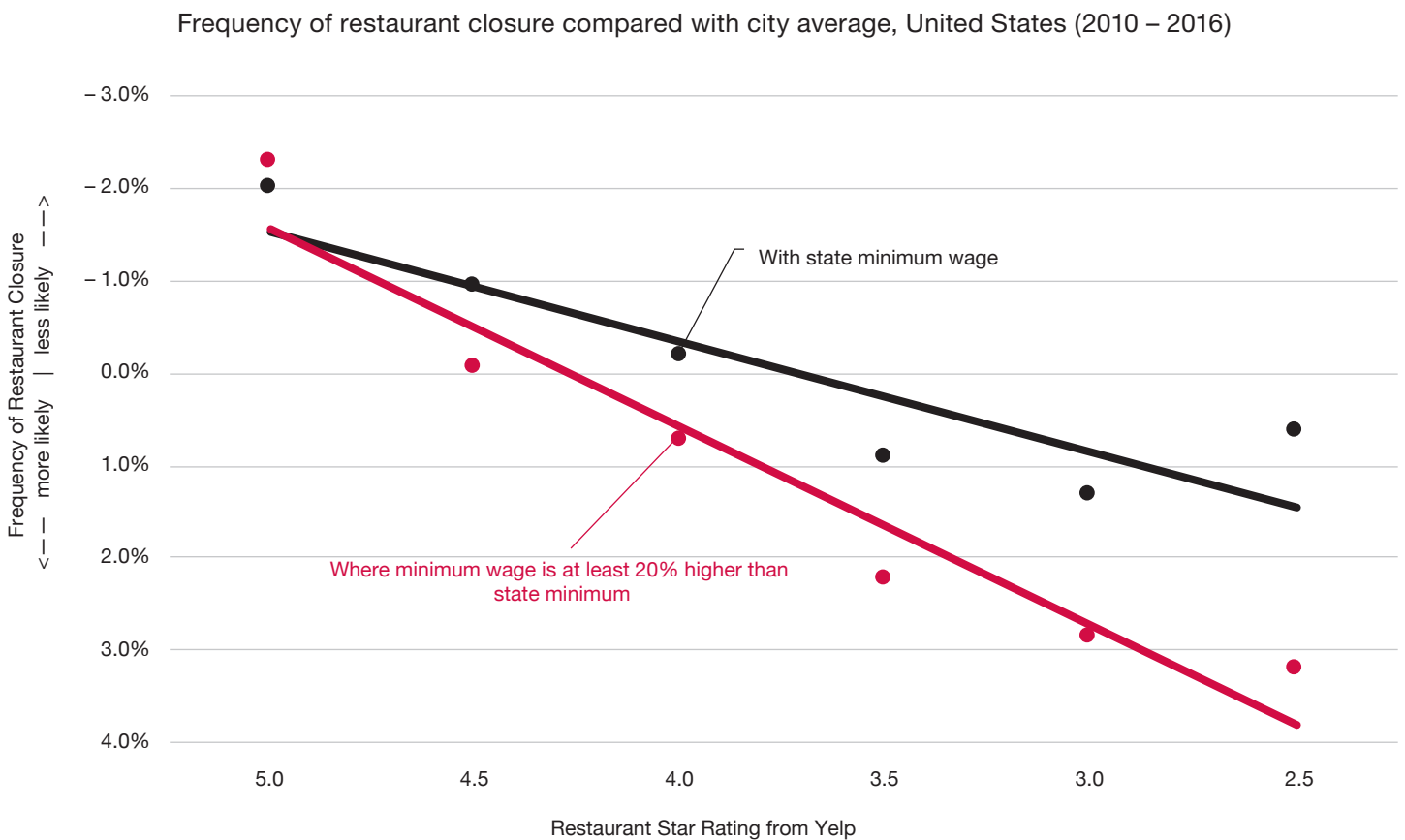
ones, as shown in Figure 1. For example, our estimates suggest that a \$1 increase in the minimum wage would lead to a 14 percent increase in the likelihood of going out of business for a 3.5-star rated business but has no effect on 4.5- or 5-star rated restaurants. (Five stars is Yelp’s highest rating.) So, increases to the minimum wage will increase closures among lower-rated pizza places, while higher-rated ones

will be largely insulated from these changes. For policymakers, these results provide additional nuance to our understanding of how businesses are affected by the minimum wage. The minimum wage does seem to cause some businesses to close, but there is substantial and predictable heterogeneity in the effect.

Aside from the direct insight into the impact of the minimum wage, these analyses highlight the way in which

nontraditional data sources — in this case, Yelp reviews — can help to complement the types of data economists and policy wonks have traditionally relied on. The next sections will discuss the value of Yelp data in this context, provide other examples of how Yelp data can shed light on policy questions, and conclude with ways that policymakers could further integrate new sources of data to guide their decisions.

Figure 1. The Impact of the Minimum Wage by Yelp Rating



Why Yelp Data?

While government datasets have been critical to our understanding of the minimum wage and the economy more generally, the effects we identify in this paper would have been difficult to observe using standard datasets. The growth of online review platforms such as Yelp, which has generated 127 million reviews for millions of businesses, allows for unique insights into the economy. First, we can use each restaurant's rating as a proxy for its reputation, a measure that is not captured by conventional datasets. This lets us evaluate whether the minimum wage differentially impacts lower quality businesses.

Second, we are able to use the data in close to real time, whereas BLS and census data become publicly available only after a lag. Working closer to real time allows researchers and policymakers to more quickly understand the impacts of different economic policies.

Third, we are able to observe granular data on businesses, whereas the public versions of the U.S. Census and BLS data are aggregated to coarser geographic levels, such as by county or zip code. In principle, researchers can request restricted access to business-level data via an extensive application process, but the

current waiting period for access to the government data, even among approved applications, is estimated to be two years. For example, a researcher trying to understand the impact of a policy change in 2017 would not be able to examine firm-level microdata from the census until at least 2020. Yelp data allowed us to circumvent these challenges.

Similarly, researcher-implemented surveys, like the one conducted by Card and Krueger, have yielded important findings. These surveys have the advantage of allowing researchers to control the timing of data collection and the questions that are being asked. However, surveys are also expensive, time consuming, and can suffer from low response rates and selection bias. Card and Krueger's survey of 410 restaurants was not an easy feat.

In contrast, Yelp data allowed us to observe outcomes and quality for roughly 35,000 restaurants across the Bay Area over an eight-year time span. To get a rough sense of how much it would cost to obtain similar data through a survey, consider recent estimates we obtained for conducting a single survey of about 1,000 restaurants in the Bay Area, which ranged from \$50,000 to \$100,000, even for a basic survey instrument with relatively little

follow-up with restaurants, and for a single year. If we wanted to conduct this survey for eight years to replicate the full dataset, we would have needed to anticipate the project eight years ago, and the costs would quickly multiply.

These comparisons give a sense of the relative advantages of different datasets and how they might complement each other in a policy analyst's toolkit.

How Policy Analysts Can Leverage New Data Sources

This paper is part of a larger endeavor in collaboration with Yelp, policymakers, and other academics, in which we are exploring ways to improve policy and policy research, using Yelp data.

In a 2013 paper⁶ with Jun Seok Kang, Polina Kuznetsova, and Yejin Choi, we developed an algorithm using Yelp data to predict which restaurants are most likely to have health code violations. Building

⁶ Kang, Jun Seok, Polina Kuznetsova, Michael Luca, and Yejin Choi. 2013. "Where Not to Eat? Improving Public Policy by Predicting Hygiene Inspections Using Online Reviews." In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1443-1448.

Continued on Back Flap

on this, Ed Glaeser, Andrew Hillis, Scott Kominers, and I partnered with the City of Boston⁷ to develop, implement, and test the potential for this type of algorithm in practice. Boston is now using an algorithm to target its inspections, and my collaborators and I are in the process of analyzing the results. In ongoing work, Ed Glaeser, Hyunjin Kim, and I have combined Yelp data with BLS data and are exploring the potential to measure and forecast local economic activity across the United States well in advance of public reports. Taken together, these projects highlight the ways in which new data sources are poised to complement more traditional ones to provide a richer understanding of policy and to allow policymakers to better allocate scarce resources.

7 Glaeser, Edward L., Andrew Hillis, Scott Duke Kominers, and Michael Luca. 2016. "Crowdsourcing City Government: Using Tournaments to Improve Inspection Accuracy." *American Economic Review*, 106(5): 114-18.

In "Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life,"⁸ my collaborators Ed Glaeser, Scott Kominers, Nikhil Naik, and I expand on this and explore various ways that new data sources can improve policy.

One theme that emerges is that, relative to more traditional datasets used in policy analyses, new data sources can at times be more up-to-date, more granular, and can have variables that are not contained elsewhere. This has the potential to improve at least two types of policy analyses: It can improve policy evaluations focused on causal effects, as in the case of the minimum wage, and it can improve predictions and forecasts that policymakers use to inform resource allocation, such as

8 Glaeser, Edward L., Scott Duke Kominers, Michael Luca, and Nikhil Naik. 2016. "Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life." *Economic Inquiry*. DOI: 10.1111/ecin.12364.

the case of hygiene inspections. Of course, no single dataset is perfect, and one practical goal of research in this area is to better understand how to work with different types of data, how datasets fit together, and what the limitations of emerging and more traditional datasets are.

Occasionally, I am asked whether new data sources will make the U.S. Census and other large-scale government data collection efforts irrelevant. The answer, at this point, is a resounding no. Traditional datasets will continue to be important. However, the agencies that collect and manage these datasets should continue to explore ways to complement them with new data, by adjusting existing questions, adding new ones, and merging datasets. And researchers and policy analysts should look broadly for the right combination of datasets to answer the question at hand.

**About the Stanford Institute for
Economic Policy Research**

We support research that informs economic policymaking while engaging future leaders and scholars. We share knowledge and build relationships among academics, government officials, the business community and the public.

Policy Briefs

SIEPR Policy Briefs summarize research by our affiliated faculty and researchers. They reflect the views and ideas of the author only. SIEPR is a nonpartisan research institute.

For Additional Copies

Please visit SIEPR.stanford.edu

Location

John A. and Cynthia Fry Gunn Building
366 Galvez Street
Stanford, CA 94305-6015

Online

siepr.stanford.edu

 [@siepr](https://twitter.com/siepr)

 facebook.com/SIEPR/