Exploring the Spatiotemporal Dynamics of Google Trends Data: An Application in Estimating Childcare Demand

Sean C. Reid & Jacob A. Cooper
International Population Conference 2025

September 2024

Introduction

Extensive interdisciplinary research underscores the critical importance of infancy and early life experiences in shaping long-term academic, social, and financial outcomes [1]. Child care, or early care and education (ECE), programs that serve young children during their critical period of brain development are vital in fulfilling both the quality care and educational needs of many families, particularly those with working parents. The ECE industry in the United States is decentralized with programs often being privately-operated. State and federal funding for the industry comes in the form of government-funded systems (such as Head Start and public pre-kindergarten) or, more often, through block grants and subsidies. This patchwork approach to such a critical industry has long been insufficiently serving working families living in the US, and this was greatly exacerbated by the COVID-19 pandemic. The pandemic "brought to light the lack of safe, reliable, and affordable child care options" nationwide as ECE providers closed their doors in 2020 [2].

Understanding accessibility of ECE resources is influenced by many factors like socioeconomic status, parental preferences, and residential locations. Researchers and practitioners alike have developed methods to quantify supply and demand for ECE resources to better help policymakers enhance ECE access [3, 4, 5]. The majority of methodological ECE access literature have been focused on the supply side of the industry due to more readily available data. Validation data for ECE demand is lacking across time and space and is often quantified using proxy data from the American Community Survey (ACS). Using ACS data is a viable proxy but has several key limitations: census data census data may inaccurately quantify the number of children in under-served communities [3], they are only annually updated, and they represent only the *potential* demand for child care instead of an actual number of families actively seeking or needing such services. Therefore, to reduce its reliance upon static, point-in-time data, the ECE field and industry would benefit from a proxy demand dataset that would point towards trends in ECE demand across time and space at a regularly updated interval. In this study we propose utilizing Google Trends data as an alternative proxy for estimating ECE demand across diverse temporal and spatial contexts.

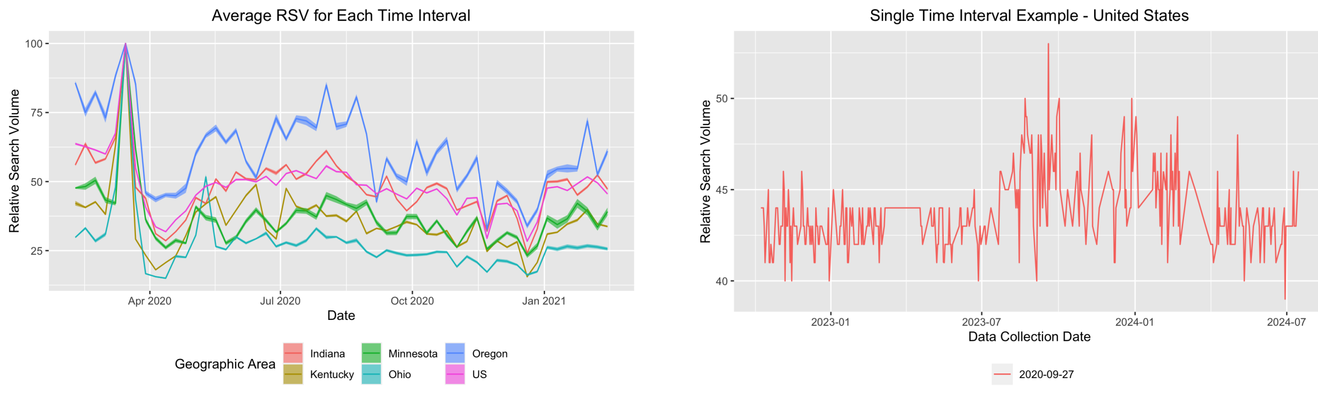
The rise of the internet to ubiquitous use around the world has provided a great opportunity for researchers to leverage publicly available demographic and geographic data for research endeavors. Platforms like Facebook, Airbnb, Twitter, and Google Trends have been popular sources of data because of their large numbers of users and data dissemination infrastructure. Each of the data sources listed above are versatile and wide ranging but have unique intricacies, nuances, biases, and limitations. These considerations require researchers utilizing digital data in their demographic and spatial studies to acknowledge and attempt to account for the impact on their research findings. These limitations and biases are often not reasons to abandon data sources in geographic research; indeed, leveraging accessible data to make the best possible estimates in light of limited alternatives is a critical first step in evaluating data-poor topics despite the wide range of uncertainty they may produce. Instead, scholars and practitioners must be transparent about these limitations throughout the life cycle of their research.

This project presents a novel methodology for harnessing Google Trends data to approximate demand for childcare services across time and space. We utilize Google Trends based on findings that internet search engines are the second most common method of finding childcare behind personal referrals [6]. We build on a growing body of literature examining best use practices of Google Trends for research purposes by providing a fresh understanding of the dynamics of Google's reporting of search activity on its platform [cite some here]. This methodology is one of the first to critically evaluate the spatial aspects of Google's internet search trends data. Taking a GIScience approach allows greater understanding and versatility in how scholars and practitioners can use Google Trends data to evaluate demographic and geographic phenomena alike. For industries like Early Childhood Education (ECE), where these data are scarce or come with questionable reliability, the approach presented here can provide relative search intensity data with a measurable quantification of error as a valuable surrogate.

Data and Methodology

Google Trends data is a free tool developed by Google that allows for users to sample their search engine history. This data source is commonly used by academics and professionals spanning disciplines from economics to tourism to public health. There are several key features of the data researchers should be familiar with. The first is the search term or topic of interest. Users can utilize Google generated topics that include related sub searches and similar items related to your search topic. Alternatively users can define their own search query using boolean type logic to chain together several search terms to return a more limited and specific subset of the search history. Second is the temporal resolution and period of the data set which can span the entire life of the Google search engine. The temporal range will directly influence the time unit in which the data is reported. For example, if the temporal range is one day then the temporal unit of search interest will be in hours. If the temporal range is set between six months and five years then the temporal unit of search interest will be in weeks. Perhaps the most important feature of the data for this project is the geographic resolution. Users can refine their to specific countries, states/provinces, and metropolitan areas. Additionally, each of these the aspects of the data can be further stratified by the search method like web search, YouTube, images, etc.

The next key aspect of the data to consider is how Google samples their database and how they report search interest values. The immense amount of data included in the search history is far too large for users to query over a web interface. Therefore, Google samples their data twice per day as the dataset users can query as described above. It's important to note that Google Trends does not provide actual search volume but rather a relative volume. Based on the restricted sample, Google then reports user defined search terms as a relative search volume (RSV) compared to all other search terms in the sample, in the temporal range, and in the specified geographic area. The time and geography that have the highest relative search volume is rescaled to be 100 and the rest of the time points and geographies are scaled to be proportional to the reference. Each of these transformations introduces bias into the search results and research has shown that executing the exact same query on different days can result in different data. This presents a consistency problem with the data that can be detrimental to using Google Trends as a viable dataset. There are recent studies that explore methodological techniques to account for the consistency issue by querying the data multiple times across different days and aggregating the relative search volumes across the data collection period [7, 8, 9, 10, 11]. The number of replicate queries needed for consistent data is inconclusive but recommendations have been made based on search popularity but we feel these recommendations are not appropriate across all spatial scales. We expect there are additional temporal and spatial intricacies that must be considered when calibrating the number of data replicates in a deterministic way.



(a) (b)

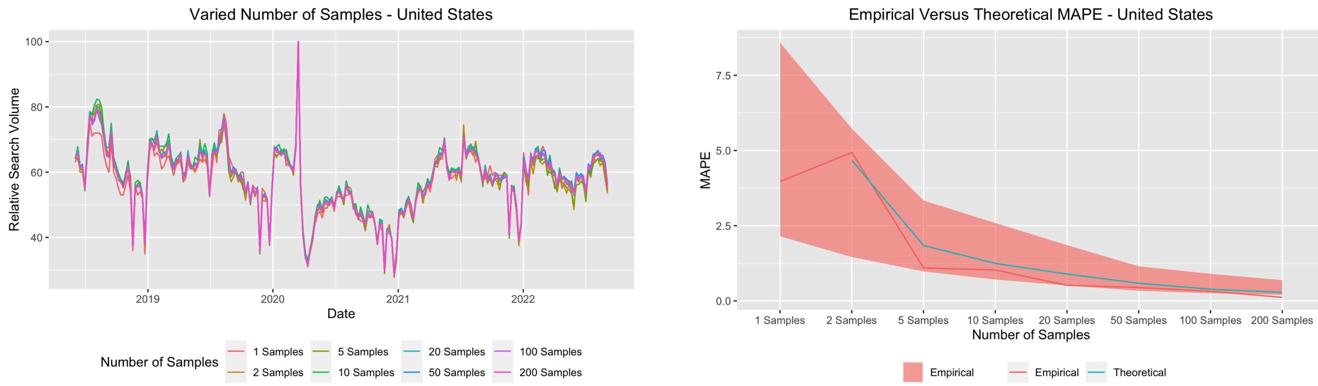
Figure 1

To examine this idea we collected over 400 replicate data extracts from Google Trends with the following query specifications. The search topic of "Child Care" was used to account for varying child care searches. The temporal range for our query included pre and post pandemic periods of February 16th, 2020 to February 20th, 2021 at a weekly temporal unit. We pulled several geographic specifications of this dataset for the United States, states of Oregon, Indiana, Minnesota, Kentucky, Ohio, and metropolitan area around Eugene, Oregon. Figure 1a shows the average RSV value for each day in our temporal range for all geographies. Figure 1b illustrates what data that comprises the single data point for the week of September 27th, 2020 for the United States. You can see that each day we extracted the data the RSV value varied to some degree. We then generate descriptive statistics for each time interval across all geographies to better understand the variation in our data. We then followed a recommended methodology account for inconsistency in the data due to search volume [9]. We then randomly sample a progressively larger number of extractions from our dataset for testing. We then replicate methods from Cebrian 2024 [9] to examine how accounting for only search term popularity at various geographic levels influences consistency. We first calculate search term popularity in each geographic unit. Second we calculate the Mean Absolute Percentage Error (MAPE) by comparing the average RSV for the progressively larger number of samples against the average RSV for the whole dataset. We present a theoretical or expected relationship between MAPE and number of extractions to an empirical bootstrap of the data. In the preliminary results below we illustrate early findings from our analysis.

Preliminary Results

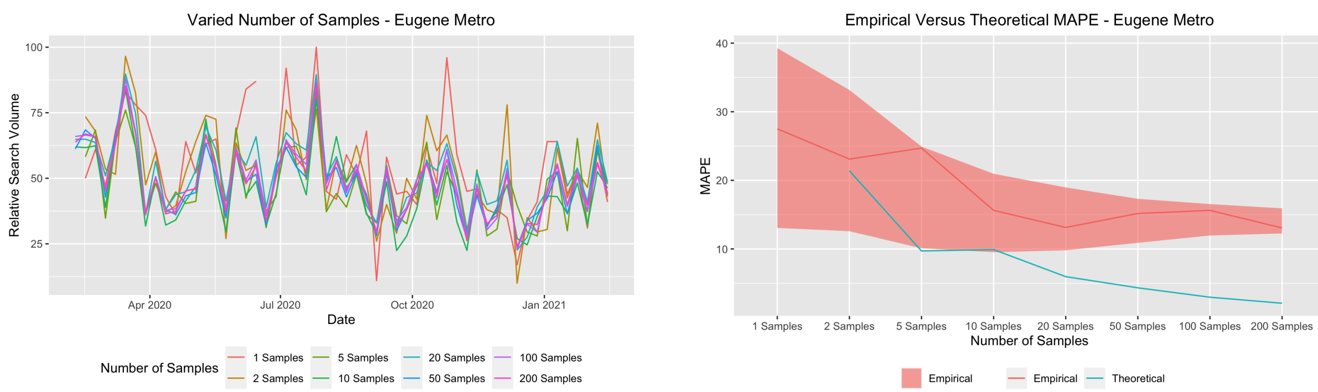
Based on the methods outlined in previous research [9], we expected the theoretical number of extractions to match the empirical bootstrapped data. Larger geographic areas fail to represent heterogeneity at smaller spatial scales, an idea related to the Modifiable Areal Unit Problem (MAUP), and therefore will not display geographic influence on Google Trends consistency. At smaller geographic scales, where heterogeneity in the sample population and smaller sample size are more apparent, we expect to observe a deviation between the theoretical number of extractions and empirical bootstrapped data. These assumptions were verified and illustrated graphically in for the United States (Figure 2) and Eugene Metro (Figure 3).

The average RSV did not vary a great deal based on the number of samples which can be seen in Figure 2a. In Figure 2b, the relationship between search term popularity and number of extractions was consistent with previous findings. The MAPE value declined as the number of extractions increased until leveling off around 50 extractions. In contrast, we see a noisy graph for average RSV in Figure 3a for the Eugene Metro as the number of extractions varied. In Figure 3b, we see a departure in the theoretical relationship between MAPE and the number of extractions compared with the empirical bootstrapped data. A similar leveling off occurs around 50 extractions but there is a disparity between the two lines. We interpret this as the deterministic process for consistent Google Trends data developed in other studies potentially inappropriate for data at fine geographic scale. We seek to expand upon our findings and offer a standardized methodology to ensure consistent Google Trends data across geographic scales. Our case study is focused on estimating child care demand but we anticipate our findings can be broadly applied to any search topic.



(a) (b)

Figure 2



(a) (b)

Figure 3

References

[1] Dolores Acevedo-Garcia, Clemens Noelke, Robert W. Ressler, and Leah Shafer. Improving The Infrastructure For Neighborhood Indices To Advance Health Equity. *Health Affairs Forefront.*

[2] Nikita Kalluri, Colleen Kelly, and Arvin Garg. Child Care During the COVID-19 Pandemic: A Bad Situation Made Worse. *Pediatrics*, 147(3):e2020041525, March 2021.

[3] John Fantuzzo, Katherine Barghaus, Kristen Coe, Whitney LeBoeuf, Cassandra Henderson, and Caroline C. DeWitt. Expansion of quality preschool in Philadelphia: Leveraging an evidence-based, integrated data system to provide actionable intelligence for policy and program planning. *Children and Youth Services Review*, 127:106093, August 2021.

[4] Rasheed Malik, Katie Hamm, Leila Schochet, Cristina Novoa, Simon Workman, and Steven Jessen-Howard. America's Child Care Deserts in 2018.

[5] Christian Sandström, Karl Wennberg, Martin W. Wallin, and Yulia Zherlygina. Public policy for academic entrepreneurship initiatives: A review and critical discussion. *The Journal of Technology Transfer*, 43(5):1232–1256, October 2018.

[6] Heather Sandstrom, Catherine Kuhns, Sarah Prendergast, Teresa Derrick-Mills, and Laura Wagner. Parental Search and Selection of Child Care and Early Education: A Literature Review. OPRE Report 2024-082. Technical report, Office of Planning, Research and Evaluation, April 2024.

[7] Alessandro Rovetta. Reliability of Google Trends: Analysis of the Limits and Potential of Web Infoveillance During COVID-19 Pandemic and for Future Research. *Frontiers in Research Metrics and Analytics*, 6, May 2021.

[8] Eduardo Cebrián and Josep Domenech. Is Google Trends a quality data source? *Applied Economics Letters*, 30(6):811–815, March 2023.

[9] Eduardo Cebrián and Josep Domenech. Addressing Google Trends inconsistencies. *Technological Forecasting and Social Change*, 202:123318, May 2024.

[10] Amaryllis Mavragani, Gabriela Ochoa, and Konstantinos P. Tsagarakis. Assessing the Methods, Tools, and Statistical Approaches in Google Trends Research: Systematic Review. *Journal of Medical Internet Research*, 20(11):e9366, November 2018.

[11] Joshua Wilde, Wei Chen, Sophie Lohmann, and Jasmin Abdel Ghany. Digital Trace Data and Demographic Forecasting: How Well Did Google Predict the US COVID-19 Baby Bust? *Population and Development Review*, n/a(n/a).