

endogeneity issues, our econometric model is able to quantify the economic and social value of the extracted features on local demand from a causal perspective.

On a broader note, the objective of this paper is to illustrate how multiple and diverse sources of publicly available crowdsourced data can be mined and incorporated into the prediction of local demand to enhance the understanding of users' economic behavior through its interactions with local businesses. Our study demonstrates the potential of how we can best make use of the large volumes of user-generated content and geotagged social media data to create matrices that capture multidimensional characteristics in a manner that is fast, cheap, accurate, and meaningful. Local businesses can use this information to proactively design their business strategies (e.g., advertising and promotions) when facing a potential change of its neighborhood city services. Furthermore, it can help government decision makers to understand local economic trends. For example, it is useful for urban planners to be able to quantify the opportunity cost, and moreover, the overall expected economic outcome of an urban project or event in a location, under various urban and economic conditions. Since our data come from publicly available channels, we can easily apply our methodology to other categories of local businesses in various locations. Such analyses can help small businesses gain insights into their local urban systems and economies, which, in turn, increases their success and the sustainability of urban neighborhoods.

Our research also has implications for location-based services, such as Google Maps, by making it possible to incorporate data into understanding local neighborhoods. Specifically, they can use the model we propose to specify the location efficiency scores in predicting the economic potential for a new market. For example, one possibility would be to provide an "economic index" of each neighborhood for new businesses to predict their demand in different locations and, thus, optimize their location selection.

Our work has several limitations, some of which can serve as fruitful areas for future research. Our analysis is based on a randomly selected subset of Twitter and Foursquare data. It can be improved by leveraging more data from other crowdsourced channels to gain a more comprehensive understanding of traffic and human mobility conditions. Also, in order to better predict the local demand, future work can look into not only the geographic and socioeconomic perspectives of cities, but also other natural and environmental aspects, such as climate and pollution factors, healthcare, etc. Such research would help us draw a comprehensive picture of the overall urban system and to study the economic dynamics and social interactions more precisely.

7. REFERENCES

- [1] M. Anderson and J. Magruder. Learning from the crowd: Regression discontinuity estimates of the effects of an online review database*. *The Economic Journal*, 122(563):957–989, 2012.
- [2] G. C. Blomquist, M. C. Berger, and J. P. Hoehn. New estimates of quality of life in urban areas. *The American Economic Review*, pages 89–107, 1988.
- [3] BusinessWeek. It costs \$333 million to shut down boston for a day. 2013a. www.businessweek.com/articles/2013-04-19/it-costs-333-million-to-shut-down-boston-for-a-day.
- [4] P. Calthorpe. *The next American metropolis: Ecology, community, and the American dream*. Princeton Architectural Press, 1993.
- [5] Z. Cheng, J. Caverlee, K. Lee, and D. Z. Sui. Exploring millions of footprints in location sharing services. *ICWSM*, 2011:81–88, 2011.
- [6] P. C. Cheshire and S. Magrini. Population growth in european cities: weather matters—but only nationally. *Regional studies*, 40(1):23–37, 2006.
- [7] J. A. Chevalier and D. Mayzlin. The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3):345–354, 2006.
- [8] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD*, pages 1082–1090. ACM, 2011.
- [9] R. Ewing and S. Handy. Measuring the unmeasurable: urban design qualities related to walkability. *Journal of Urban design*, 14(1):65–84, 2009.
- [10] R. Florida. The economic geography of talent. *Annals of the Association of American geographers*, 92(4):743–755, 2002.
- [11] C. Forman, A. Goldfarb, and S. Greenstein. How did location affect adoption of the commercial internet? global village vs. urban leadership. *Journal of urban Economics*, 58(3):389–420, 2005.
- [12] A. Ghose, P. G. Ipeirotis, and B. Li. Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content. *Marketing Science*, 31(3):493–520, 2012.
- [13] J. A. Hausman. Valuation of new goods under perfect and imperfect competition. In *The economics of new goods*, pages 207–248. 1996.
- [14] D. Karamshuk, A. Noulas, S. Scellato, V. Nicosia, and C. Mascolo. Geo-spotting: Mining online location-based services for optimal retail store placement. In *Proceedings of the 19th ACM SIGKDD*, pages 793–801. ACM, 2013.
- [15] D. Lambiri, B. Biagi, and V. Royuela. Quality of life in the economic and urban economic literature. *Social Indicators Research*, 84(1):1–25, 2007.
- [16] J. Lindqvist, J. Cranshaw, J. Wiese, J. Hong, and J. Zimmerman. I'm the mayor of my house: examining why people use foursquare—a social-driven location sharing application. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2409–2418. ACM, 2011.
- [17] T. Litman. Economic value of walkability. *Transportation Research Record: Journal of the Transportation Research Board*, (1828):3–11, 2003.
- [18] X. Lu, S. Ba, L. Huang, and Y. Feng. Promotional marketing or word-of-mouth? evidence from online restaurant reviews. *Information Systems Research*, 24(3):596–612, 2013.
- [19] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web*, pages 851–860, 2010.
- [20] J. M. Villas-Boas and R. S. Winer. Endogeneity in brand choice models. *Management Science*, 45(10):1324–1338, 1999.