Manuscript

Urban Analytics and City Science

EPB: Urban Analytics and City Science
2022, Vol. 0(0) 1–14

DOI: 10.1177/23998083221090246
journals.sagepub.com/home/epb
SAGE

Unveiling spatial patterns of disaster impacts and recovery using credit card transaction fluctuations

Faxi Yuan, Amir Esmalian, Bora Oztekin and Ali Mostafavi
Texas A&M University, College Station, TX, USA

Abstract

The objective of this study is to examine spatial patterns of disaster impacts and recovery of communities based on fluctuations in credit card transactions (CCTs). Such fluctuations could capture the collective effects of household impacts, disrupted accesses, and business closures and thus provide an integrative measure for examining disaster impacts and community recovery. Existing studies depend mainly on survey and sociodemographic data for disaster impacts and recovery effort evaluations, although such data has limitations, including large data collection efforts and delayed timeliness results. Also, there are very few studies have concentrated on spatial patterns of disaster impacts and short-term recovery of communities, although such investigation can enhance situational awareness during disasters and support the identification of disparate spatial patterns of disaster impacts and recovery in the impacted regions. This study examines CCTs data Harris County (Texas, USA) during Hurricane Harvey in 2017 to explore spatial patterns of disaster impacts and recovery duration from the perspective of community residents and businesses at ZIP-code and county scales, respectively, and to further investigate their spatial disparities across ZIP codes. The results indicate that individuals in ZIP codes with populations of higher income experienced more severe disaster impact and recovered more quickly than those located in lower income ZIP codes for most business sectors. Our findings not only enhance the understanding of spatial patterns and disparities in disaster impacts and recovery for better community resilience assessment but also could benefit emergency managers, city planners, and public officials in enhanced situational awareness and resource allocation.

Keywords

Smart resilience, community-scale big data, disaster impact, recovery duration, data analytics

Corresponding author:
Faxi Yuan, Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, TX 77843, USA;
Email: faxi.yuan@tamu.edu

Introduction

Examining community resilience, including disaster impacts and recovery duration (Vugrin et al., 2010), from a systems perspective calls for the inclusion of interactions among system components (Mostafavi et al., 2012). Community systems include businesses which provide the products and services, residents who use the products and services, and infrastructure (e.g., roads) that provides access to businesses (Dong et al., 2020). Perturbations caused by natural hazards (e.g., hurricanes and flooding) could impact residents, businesses, and infrastructure systems. The collective effects of these impacts could be captured using credit card transactions (CCTs) as a proxy for population activity data.

Understanding the state of the community, including immediate disaster impacts and short-term recovery duration, provides the basis for resource allocation and prioritization of recovery strategies (Yuan and Liu. 2020; Zhai et al., 2020). Short-term recovery lasts from days to weeks and includes assessment of the damages and needs, restoration of basic infrastructure, and mobilization of recovery resources (FEMA 2016). Current approaches such as survey (Maquiling et al., 2021), for assessing disaster impacts and recovery of communities, however, do not consider the dynamic interactions among the systems and their components.

Understanding spatial patterns and disparities related to impacts and short-term recovery across various regions not only enhances situational awareness but also supports identifications of regions with more severe disaster impacts and slower recovery (Chen et al., 2020; Yuan et al., 2021b). Different regions in a community may experience varying disturbance impacts (Esmalian et al., 2020). The disparate impacts and recovery patterns from disasters are not limited to unequal exposure of the areas to hazards (Coleman et al., 2019; Peacock and Ragsdale 1997). The infrastructure condition, business state, and demographic characteristics play a role in the extent to which certain areas are impacted and recover from disturbances. With the growing availability of population activity data, it is now possible to explore such spatiotemporal aspects and further resolve the previous knowledge gap in quantifying community resilience in terms of disaster impacts and short-term recovery duration (Knüsel et al., 2019; Yabe et al., 2020a). Population activity data (e.g., mobility) provide researchers with fundamental human activity information (Hasan et al., 2013). Current applications of population activity data in disaster studies include examining evacuation patterns (Han et al., 2019), conducting rapid damage assessment (Yuan and Liu 2018; Yuan and Liu 2020), understanding social impacts (Dong et al., 2018; Juhász and Hochmair 2020), evaluating economic resilience (Alatrista-Salas et al., 2020; Martinez et al., 2016), and assessing business recovery (Yabe et al., 2020c). Population activity data could effectively capture interactions of residents with physical systems, such as businesses and infrastructures. Data representativeness issue within social media data (Yuan et al., 2021a) and the concentration on migration patterns with mobility data during disasters (Lu et al., 2016; Yabe et al., 2020b), however, can only reflect limited awareness of community states, disaster impacts and recovery durations of essential business sectors for community's preparation, response, and recovery. Nevertheless, limited studies using population activity data to comprehensively explore disaster impacts and short-term recovery across various business and essential sectors (e.g., drugstores) exist.

In this study, we use Safegraph's CCTs data as an aggregated measure to examine the community in terms of disaster impacts and recovery duration (Supp Figure S3). The CCTs data include information about ZIP codes where individual residents live and their CCTs history, which provides information regarding business categories and transaction dates to enable determination of fluctuations in CCTs prior to, during, and in the aftermath of disasters. This data is captured and evaluated to find the spatiotemporal patterns. Perturbations caused by disasters could affect households, infrastructure systems that provide access to business, as well as businesses themselves. The collective effects of these disruptions could be captured in the CCT fluctuations. For example, if households are economically impacted by disasters, or if they could not access businesses because of road inundation/closure, or if businesses are closed due to damage, effects of these perturbations are reflected in the CCTs. Accordingly, we employ the CCT data to unveil the spatial patterns and disparities of disaster impacts and recovery duration across the ZIP codes in the context of Hurricane Harvey which made landfall in August 2017 in Harris County, Texas.

Materials and methods

This study started with collection, categorization, and processing of CCTs data. Then, we quantified two indices of community state—disaster impact and recovery effort—as suggested by Vugrin et al. (2010) across ZIP codes in Harris County. Since this research relied on the number of days to reveal the recovery effort, we employed recovery duration to represent the recovery effort. We conducted spatial autocorrelation analyses to reveal spatial patterns of disaster impacts and recovery duration of communities. Subsequently, we utilized the stepwise regression model to examine spatial disparities in disaster impacts and recovery duration with respect to demographic and flood claim data across different ZIP code regions. Finally, we quantified the disaster impacts and recovery duration of various business sectors at the county level in Harris County. Details for methods are available at Supp Methods.

Case study

Harris County, home to Houston, Texas, the fourth largest city in the United States, has experienced a rapid population growth over the past decades (Qian 2010). Harris County is among the most flood-prone counties in the United States due to its location in a coastal area, burgeoning urban development, and the lack of flood control infrastructure development in parallel with the development and population growth. Accordingly, the county has experienced several severe flood events such as Hurricane Harvey in 2017. Each event caused extensive flooding with losses ranging from hundreds of millions to hundreds of billion dollars. Therefore, Harris County was selected as the testbed for this study.

Data and categorization

Credit card transactions data was provided by Safegraph company. Each CCT record contains the transaction date, cardholder's residential ZIP code, the daily number of unique cards from the ZIP code involved in transactions, the daily number of unique transactions on cards from the ZIP code, and the daily total amount spent on cards from the ZIP code. In this study, we employed the fluctuations of total spent for our analysis. Referring to Visa Merchant Data Standards Manual (Visa, 2021), we categorized 396 MCCs (merchant category codes) into 21 business sectors such as drugstore sector with MCCs 5122 and 5912 (Supp Table S2).

Determining impact and recovery

Using CCT transactions from 01/08/2017 to 21/08/2017, we computed seven transaction baselines from Monday to Sunday and further calculated fluctuations of daily transactions from 01/08/2017 to 30/09/2017. We plotted the curve of daily fluctuations of CCT transactions based on the approach proposed by Nan and Sansavini (2017) to obtain the resilience curve (please see Supp Figure S2). Using the resilience curve, we determined disaster impact and recovery duration (Vugrin et al., 2010). Disaster impact was determined based on the maximum drop of CCT fluctuations by each sector post-Hurricane Harvey landfall (Supp equations (1)–(3)). Recovery duration was computed as number of days taken for the daily CCT fluctuations to return to zero post disruption (Supp equation (4)).

Spatial patterns and disparities

With the daily CCT fluctuations of sectors and adjacency of ZIP codes, we conducted global and local spatial autocorrelation analyses with global (Supp equations (5)–(7)) and local Moran's I (Supp equations (5), (6), and (8)), respectively. Then, we used stepwise regression method to explore the spatial disparities of disaster impact and recovery where flood damage and socio-demographic characteristics were used as potential factors.

Results

Spatial pattern analysis

Geographic distribution of disaster impact and recovery duration. We computed disaster impacts and recovery duration for business sectors for 142 ZIP codes in Harris County. Using ArcGIS, maps of these two indices for the grocery sector as an example are presented in Figure 1. Figure 1(a) shows disaster impacts; Figure 1(b) shows recovery duration. In both figures, oval 1 indicates clusters of similar disaster impact and recovery duration levels; oval 2 indicates clusters of dissimilar disaster impact and recovery duration levels. The clusters of either similar or dissimilar disaster impact levels and recovery durations reveal their spatial patterns.

Global spatial pattern. For investigating spatial patterns across 142 ZIP codes, we included 12 business sectors (Supp Table S3), as their CCT fluctuation curves present continuity and are consistent match with the transitions in the resilience curve (Supp Figure S2). Using the disaster impacts and recovery duration for the 12 business sectors and their corresponding spatial lag values at each ZIP code, we computed their global Moran's I and p-values as summarized in Supp Table S3. In Supp Table S3, we can see that global Moran's I related the disaster impacts in eight business sectors are significant (p-value < .1). Global Moran's I values for disaster impacts in drugstore, home supply, recreation, and restaurant are not significant. Global Moran's I values for recovery duration are significant (p-value < .1) for four business sectors: clothing, drugstore, internet and telecommunication, and restaurant. For the eight remaining business sectors, global Moran's I values are not significant.

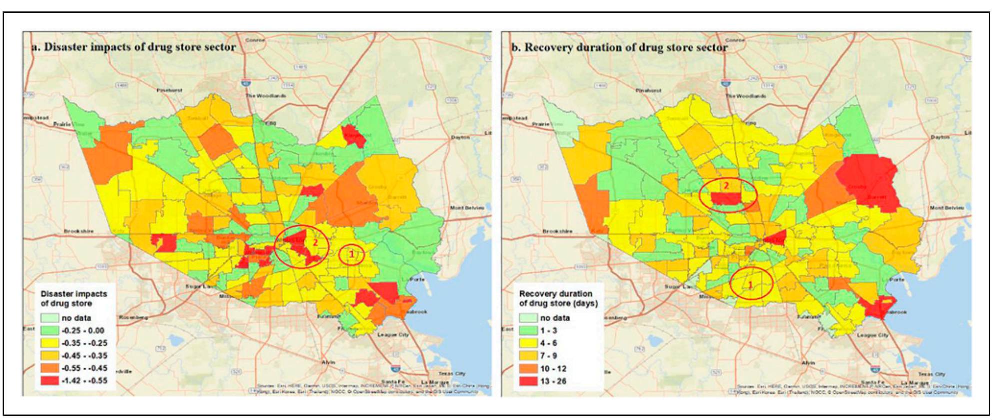

Figure 1. Geographic spatial distribution for the grocery sector: (a), disaster impact; (b), recovery duration.

Specifically, we can see the disaster impacts for two business sectors—home supply and restaurant—show negative values of global Moran's I. Their global Moran's I values, however, are not significant, meaning these two sectors show no spatial cluster pattern for the disaster impacts. For the other 10 business sectors, disaster impacts show positive values of global Moran's I. The global Moran's I values of the disaster impacts for drugstore and recreation sectors are not significant (p-value > .1). Positive values of global Moran's I mean that similar values are likely to be geographically clustered with each other. In our case, positive values of global Moran's I for disaster impacts of the remaining eight business sectors reveal that the overall trend is that neighboring ZIP codes are more likely to have similar levels of disaster impact from Hurricane Harvey.

For recovery duration, five business sectors presenting negative global Moran's I values (home supply, market, recreation, retail and transportation) are not at significant levels (p-value > .1). For the seven remaining business sectors, the global Moran's I values for only four of them are significant and positive. This reveals that for these four business sectors, including clothing, drugstore, internet and telecommunication, and restaurant, ZIP codes with long (or short) recovery duration are more likely to have neighboring ZIP codes with long (or short) recovery duration. There is no significant spatial cluster trend for recovery durations of most business sectors such as auto and grocery.

Local spatial patterns. We computed the local Moran's I for disaster impacts and recovery duration for the 12 business sectors. Grocery sector plays a critical role for the communities to prepare and respond to disasters so that situational awareness of disaster impacts and recovery duration of grocery are necessary for emergency management agencies to design and implement response strategies (Spiegler et al., 2016). Using grocery sector as an example, we present the geographic distributions of disaster impacts and recovery duration across the 142 ZIP codes in Figure 2(a) and Figure 3(a). Considering statistical significances, we also distinguished ZIP codes with local Moran's I at high significance level (p-value < .1) in Figures 2(b) and 3(b). Local Moran's I is significant in the black-shaded ZIP codes areas; areas shaded gray have local Moran's I of low significance levels. To identify the locations of clusters with similar and dissimilar values at high significance level (p-value < .1), we plotted ZIP codes with clusters of similar values, including high-high and low-low for disaster impacts and recovery duration, and those with dissimilar values (high-low and low-high) in Figures 2(c) and 3(c). The ZIP codes shaded blue and red are clusters of similar values; those shaded pink and light blue are clusters with dissimilar values. In Figures 2(c) and 3(c), we can examine the blue clusters to identify ZIP codes with severe disaster impact, and red clusters to pinpoint regions with long recovery durations, respectively. These insights can inform disaster managers and public officials in their resource allocation and prioritization during response and recovery efforts.

For the 11 remaining business sectors, we present the Local Indicators of Spatial Association (LISA) maps in Supp Figs 24 to 45. Supp Tables S4 and S5 summarize counts of clusters of ZIP codes of similar and dissimilar values with high significance levels (p-value < .1) for disaster impact and recovery duration for the 12 business sectors. For instance, for the grocery sector, among the 23

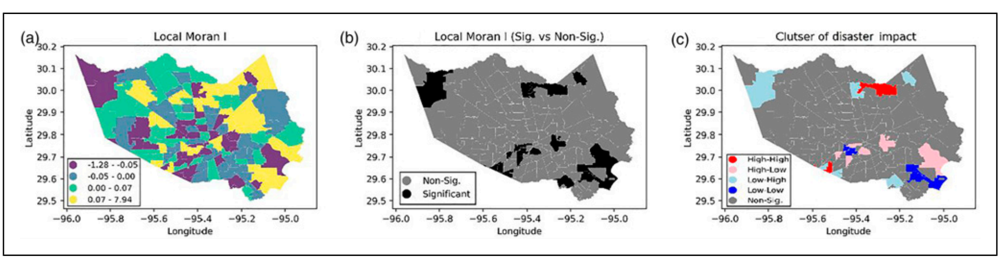

Figure 2. Local indicators of spatial association (LISA) cluster map for disaster impact of grocery sector.

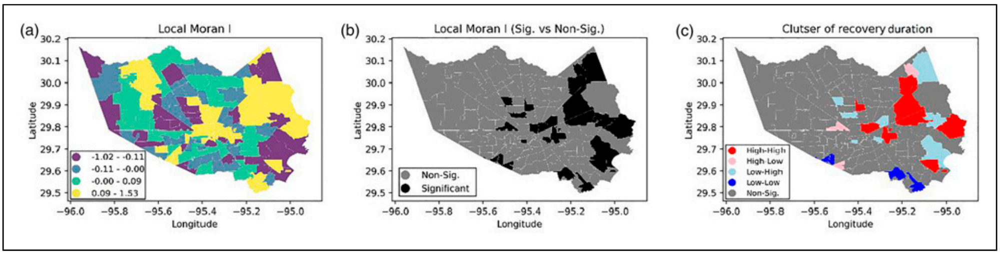

Figure 3. Local indicators of spatial association cluster map for recovery duration of grocery sector.

ZIP codes with local Moran's I at significant level (p-value < 0.1) for disaster impacts, 11 ZIP codes are clusters of similarity. Seven of these 11 ZIP code regions are clusters of low disaster impact values (blue-shaded areas, Figure 2(c)), and four are clusters with high disaster impact values (red-shaded areas, Figure 2(c)). Since we used CCT fluctuations based on changes in total transactions to quantify disaster impacts, their values are negative (indicating decline in transactions due to the disaster). Low values of disaster impacts refer to severe hurricane impact; high values represent slight hurricane impact (Figure 1(a)). From the clusters of similarly low (or high) values for disaster impacts in the grocery sector, we can see that ZIP codes with severe (or slight) hurricane impacts are more likely to have neighboring ZIP code regions having suffered similar hurricane impacts. As the similarity cluster of low values accounts for 63.64%, we can say the dominant similarity cluster type of disaster impacts in grocery sector is low-low. Hence, we can conclude that the dominant local correlation trend in disaster impacts of grocery sector is that ZIP codes with severe disaster impacts are more likely physically close to others with severe disaster impacts. Considering flood propagation characteristics, this trend can be expected as flooded regions are more likely to be located close to other flooded regions.

For LISA clusters of disaster impacts for other business sectors, we can use the same analysis process to derive their cluster characters. According to Supp Table S4, the dominant cluster patterns for the disaster impacts of various business sectors are as follows: (1) auto: low-low (ZIP codes with severe hurricane impacts are more likely to have neighboring ZIP codes with severe hurricane impacts); (2) clothing: low-low; (3) drugstore: low-low; (4) home supply: high-low (ZIP codes with slight hurricane impacts are more likely to have neighboring ZIP codes with severe hurricane impacts); (5) internet and telecommunication: low-low; (6) market: low-low; (7) recreation: low-low; (8) restaurant: low-low; (9) retail: low-low; (10) service: high-high (ZIP codes with slight hurricane impacts are more likely to have neighboring ZIP codes with slight hurricane impacts); and (11) transportation: low-low. As a result, the dominant LISA cluster pattern is low-low for the disaster impacts for 10 out of 12 business sectors.

In terms of the LISA clusters of recovery duration, Supp Table S5 illustrates the dominant cluster patterns: (1) auto: none; (2) clothing: none; (3) drugstore: low-high (ZIP codes with short recovery duration are more likely to have neighboring ZIP codes with long recovery duration); (4) grocery: high-high (ZIP codes with long recovery duration are more likely to have neighboring ZIP codes with long recovery duration); (5) home supply: low-high; (6) internet and telecommunication: low-high; (7) market: low-high; (8) recreation: none; (9) restaurant: low-high; (10) retail: low-high; (11) service: low-high; and (12) transportation: low-high. Therefore, the dominant LISA cluster pattern is low-high for the recovery duration of eight out of 12 business sectors. The results indicate that ZIP codes with sufficient recovery resources (i.e., short recovery duration) are surrounded by ZIP codes with insufficient recovery resources (i.e., long recovery duration). This result indicates the spatial heterogeneity of short-term recovery durations.

Supp Tables S4 and S5 illustrate that LISA cluster counts for disaster impact range from 21 to 44, and for recovery duration, 18 to 33, at significant level (p-value < .1). The spatial patterns of disaster impacts and recovery obtained from recorded CCTs were verified in two ways. First, we compared the LISA maps with flood maps of Hurricane Harvey provided by FEMA (Supp Figure. S4). Based on the flood depth grid map during Hurricane Harvey from Federal Emergency Management Administration (FEMA 2018), we identified the 46 flooded ZIP codes (flooded areas are larger than 10% of the ZIP area). From the comparison, we found that most of the significant LISA clusters (Figures 2(b) and 3(b)) were exactly located in the flooded ZIP-code regions. This is consistent with the expectation that ZIP codes outside flooded areas do not show significant LISA clusters. Second, spatial clusters of disaster impact and recovery durations based on CCTs were compared with household survey data collected from households affected by Hurricane Harvey in Harris County. The survey asked households to rate the impact of Hurricane Harvey on their lives and respondents reflected their household's recovery duration (Esmalian et al., 2020). Patterns of household impact and recovery from the survey agreed with spatial clusters based on CCTs. For example, regarding the spatial patterns of drugstores, the high-high cluster in the northeast and north of the downtown area also have a high reported recovery duration in the survey; a similar pattern is observed for the low-low clusters. Similarly, the low-low cluster of the disaster impact for the grocery sector is consistent with the reflected impact of Hurricane Harvey on households' lives. The findings could also shed light on identifying hotspots of impacts and recovery delays during floods by identifying significant LISA clusters. Disaster management agencies can then conduct detailed investigations within these regions and make their plans to allocate resource in a more efficient and data-driven manner.

Spatial disparities analysis results

Statistics of disaster impact and recovery duration. We first describe the spatial disparities of disaster impact and recovery duration for the defined business sectors across the 142 ZIP codes. The summary of statistical analysis—mean, median, and standard deviation—of disaster impact and recovery duration is presented in Supp Table S6. According to the results related to disaster impact in Supp Table S6, the home supply sector has the largest mean, median, and standard deviation (std = 0.53), while the restaurant sector has the lowest standard deviation value (std = 0.12). This result means that the home supply sector has seen the most significant spatial variations among the ZIP codes, and the restaurant sector has shown the least significant spatial variations. For recovery duration, the internet and telecommunication sector shows the largest mean and median values. In addition, we can see clothing sector possesses the greatest standard deviation (std = 6.82), while recovery duration for the grocery sector is the lowest (std = 3.49). This result reveals that the recovery duration of clothing sector varies most significantly across ZIP codes, while the recovery duration of grocery sector varies the least among the 12 business sectors in Harris County. As business sectors such as clothing are not as essential as drugstore and grocery sectors, we do not discuss their spatial patterns and disparities in details in the body of the manuscript. Related figures and results are provided in the supplementary information.

Spatial disparities. For spatial disparities of disaster impact and recovery duration, we conducted a stepwise regression model selection. The modeling approach was implemented in our regression analysis for disaster impact (absolute value) and recovery duration of the 12 business sectors. As drugstore sector plays an essential role for the communities from the preparedness, response, and recovery stages, understanding of which factors result in the geographic fluctuations in disaster impacts and recovery duration for drugstore sector across various ZIPs can support the future design of urban plan strategy to resolve the disparities and further enhance community resilience (Beatty et al., 2019). Hence, the stepwise regression results for disaster impact and recovery duration of drugstores are presented in Table 1 and Table 2, respectively. The regression results for disaster impact and recovery duration of other business sectors are presented in Supp Tables S7–28.

Regression results for disaster impact of the drugstores sector (Table 1) include reported loss of the residents (Mobley et al., 2021), total population of the ZIP code, median age, and income level of the residents. The model results suggest that a larger amount of reported loss is positively related to the disaster impact (based on change in transactions in drugstores). This result means that ZIP codes with larger amounts of total loss sustained a greater disaster impact for the drugstore sector. The reported loss is an indicator of the extent of direct physical damage in the affected area. Since most drugstores are visited by local residents, this result could be due to drugstores being flooded, residents losing access to drugstores, or residents relocating to temporary housing. Total population and median age within the ZIP codes are negatively associated with the disaster impact. Those ZIP codes with a higher population and areas with a higher median age experienced less disaster impact related to drugstores. The results show that areas with a higher population experienced less impact during Hurricane Harvey. With a higher population, more drugstores may exist within the ZIP-code area, providing redundancy and lower impact. Furthermore, ZIP Code areas with a higher percentage of the elderly experienced less disaster impact. Elderly populations may have a higher need for the drugstores and make more drugstore-related purchases in other areas; therefore, the greater transactions for this category could contribute to the less disaster impact upon the drugstore sector. Finally, areas with higher income experienced a higher level of impact for the transactions related to drugstores, as suggested by the positive relationship between income and disaster impact. This result implies that wealthier residents could continue fulfilling their drugstore-related needs even if they live in a flood-affected area.

Table 2 shows the Poisson regression results for the recovery duration related to drugstores. The results suggest that those areas with a larger number of flood insurance claims required greater recovery duration. Similar to the relationship between loss and disaster impact, this association could be explained by the higher physical damage in areas with a higher reported claim in comparison with other regions. In addition, Table 2 reveals that percentage of white population and

Table 1. Gaussian regression analysis for disaster impact on the drugstore sector.

Variable	Coefficient	Std. error	*p*-values	VIF
Intercept	0.89	0.24	<.001	—
Reported loss	2.63 e-05	1.87 e-05	.16	1.13
Total population	−3.32 e-06	1.16 e-06	.004	1.04
Age	−1.51 e-02	7.71 e-0–	.05	1.84
Income	5.09 e-06	1.35 e-06	<.001	1.67

VIF: variance inflation factor.

Table 2. Poisson regression analysis for recovery duration of the drugstore sector.

Variable	Coefficient	Std. error	*p*-values	VIF
Intercept	1.82	0.12	<.001	—
Number of claims	9.03	2.21	<.001	1.12
White (%)	0.28	0.19	.136	1.10
Income	−5.45 e-06	1.83 e-06	.003	1.21

VIF: variance inflation factor.

income are main factors resulting in the varying recovery durations of ZIP codes with similar disaster impact levels. When controlling the number of claims in the regression model, we can see these two factors have either positive (percentage of white population) or negative (income) relationship with recovery durations. Specifically, higher income residents recovered quicker as indicated by their drugstore-related transactions. Considering the significant and positive relationship between the income and the disaster impact, we can see that although the communities with higher incomes experienced a greater impact, they recovered more quickly than others. Such a pattern is also observed in other business sectors, which highlights the influences of socioeconomic factors in the resilience of residents in response to the similar-level disaster-induced perturbations. The percentage of the white residents has a positive relationship with recovery duration level. This indicates that ZIP codes with a higher white population percentage took longer to recover based on drugstore transaction changes. ZIP codes with a higher percentage of white population (such as neighborhoods near the reservoirs) were the most severely affected by flooding during Hurricane Harvey. The more severe physical damage to the drugstore sector caused by flooding could explain the slower recovery in ZIP codes with large white populations. The VIF (variance inflation factor) values for all variables included in the model are less than 2 which suggests that the models are not affected by the multicollinearity issue.

Disaster impact and recovery duration for various business sectors

Drugstore and health care. The drugstore sector refers mainly to drugstores and pharmacies, while the health care sector represents hospitals, medical service centers, and doctors and physicians. Both sectors provide community residents with medical resources to cope with disasters. CCT fluctuations based on the total transaction values for drugstores and health care are illustrated in Figures 4(a) and (b). Figure 4(a) documents a significant increase in the total transactions in the drugstore sector before Hurricane Harvey made landfall in Harris County (red vertical line). Immediately after landfall, a sharp decrease occurs. This phenomenon may reflect the possibility that community residents in Harris County used their credit cards to purchase necessary medicines from the drugstores for preparing to Hurricane Harvey. Transactions for the purchase of medications show an immediate drop after the landfall; recovery took about one and a half weeks to the new steady level. Also, the new steady level is higher than the previous baseline level. This phenomenon indicates that community residents made more transactions in the drugstore sector in the recovery period, starting 9 September 2017. Compared with the drugstore sector, no discernible increase of health care transactions occurred before Hurricane Harvey's landfall (Figure 4(b)). This situation is reasonable as medical services from either hospitals or medical centers cannot be stockpiled to prepare for a hurricane. Figure 4(b) indicates sharp decreases in total transactions after landfall of the hurricane, with overall recovery to baseline level occurring over the course of 2 weeks.

In addition, the disruption (orange line segment) and recovery (green line segment) periods of drugstore and health care sectors are marked in Figures 4(a) and (b). Using the maximum drop of total transactions from their baseline levels illustrated in their 7-days moving average curves, we quantify the disaster impacts on drugstore and health care sectors. We added the periods of the orange and green line segments to calculate the recovery duration (Supp Methods) for total transactions to return to their baseline levels.

In both figures, the cyan curve represents the daily CCT fluctuation from rough data, while the blue curve shows the CCTV by the 7-day moving average. Specifically, the golden part of the blue curve reveals the perturbed stage (i.e., disruption stage), while the green part reflects the recovery stage. The red vertical line marks the date when Hurricane Harvey made landfall in Harris County. The gray horizontal line highlights where the total transactions equal to their corresponding baseline values. These descriptions apply to Supp Figure 5–23. Using the CCT curves for all defined business

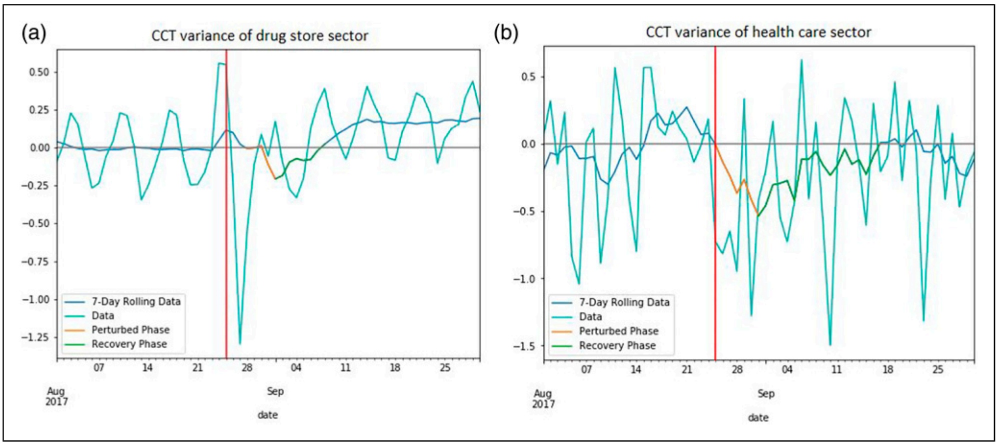

Figure 4. Credit card transactions fluctuation curves for drug store (4a) and health care (4b) sectors.

sectors, we computed their disaster impacts and recovery duration and summary is available in Supp Figure S46.

Discussions and conclusion

Credit card transaction fluctuation data can reveal the spatial patterns of disaster impacts on community residents, business, and infrastructure systems. With the adapted resilience curves (daily fluctuation curve of CCTs) for our defined business sectors, we examined changes in community state in terms of disaster impacts and recovery duration with CCT fluctuations from both community residents' perspective at ZIP-code scale and businesses' perspective at county scale. At the ZIP-code scale, we also inspected the spatial patterns and disparities of these community state indices. The results show the value of using human activity data (such as CCTs) for capturing the combined effects of household impacts, infrastructure disruptions, and business impacts as part of community resilience evaluations.

Results of spatial analytics reveal that the overall pattern of disaster impacts for the auto, clothing, grocery, internet and telecommunication, market, retail, service, and transportation sectors is a cluster of similar disaster impact levels, while no significant spatial cluster trend was found for recovery durations of most business sectors. With LISA cluster maps, our results identify clusters with severe disaster impacts or long recovery duration. For instance, the blue-shaded areas representing ZIP code areas in Figure 2(c) refer to the locations of clusters with severe disaster impact of the grocery sector, while red-shaded areas in Figure 3(c) mark locations of clusters with long recovery duration. With maps illustrated in Figure 2(c) and 3(c), disaster managers can monitor and evaluate the spatial and temporal patterns of disaster impact and recovery across different areas, and further prioritize resource allocation in more impacted areas in the blue (severe disaster impact) and red (long recovery duration) cluster regions. The insights obtained from CCT fluctuation analysis can complement the other on the ground data, as well as crowdsourced data to inform disaster response and recovery decisions.

In addition, our findings in spatial disparities reflect that factors of reported loss, total population, age, and household income may explain the spatial disparities of disaster impact for the drugstore sector. Meanwhile, variables such as number of insurance claims, percentage of the white population, and household income can explain the spatial disparities of the recovery duration for drugstore sector. Integrating regression analysis results of spatial disparities of disaster impact and recovery durations, we found that socioeconomic factors (income) play an essential role in residents' resilience. Furthermore, when controlling the number of claims, we find that the increase of percentage of white can increase the recovery durations (positive relationship) while the increase of income will decrease the recovery durations of drugstores (negative relationship).

The findings of this study highlight the value of the proposed approach in analyzing fluctuations in credit card transactions to inform disaster response and recovery. Identification of hotspots with severe disaster impacts can help rapid evaluation of disaster impact on households and businesses and their recovery to inform resource allocation and prioritization decisions. For instance, Table 1 shows that disaster impact on drugstores were more severe in areas with older adults and lower income households. This finding can inform public health agencies in their disaster response efforts. In addition, investigation of patterns of recovery durations in different regions can help disaster managers and public officials to design strategies for better recovery in future hazards.

The study and findings contribute to the emerging field of smart flood resilience focusing on harnessing community-scale big data to enhance disaster resilience capabilities including predictive flood risk monitoring, rapid impact assessment, and situational awareness (Yuan et al., 2021c; Yuan et al., 2021d, Yuan et al., 2021e), as smart situational awareness in response and recovery stages is one of its essential components (Podesta et al., 2020). This research demonstrates that using CCTs can enhance situational awareness of spatial patterns of disaster impacts and recovery of communities across flood stages. Although the findings of spatial patterns and disparities cannot be directly applied to regions beyond Harris County, the analytical framework for harnessing CCTs to enhance situational awareness of disaster impacts and recovery can be generally applied to other regions.

A limitation in this study falls in the lack of proper explanations for the observed CCT fluctuations across the hurricane period, as CCTs could be impacted by various additional factors, such as major community events or holidays (e.g., Labor Day on 4 September 2017). Future research will investigate the local news and holiday/festival calendars across communities and make the corresponding adjustments on the CCT to normalize the impacts from the hurricane on the state of community. Also, future research can explore the CCT in the same period of 2016 to use as a baseline to compute the CCT fluctuations. Another limitation in this study is the representativeness of the CCT data for the general population. Existing data shows that seven in 10 Americans possess at least one credit card which indicates a good representativeness of CCTs data (Gonzalez-Garcia and Johnson, 2020). SafeGraph collected the CCT data from a panel of ~40 MM mobile devices in the United States and Canada and this process can result in biases and outliers in the CCT data (Utrera 2021). Our future research will design and conduct a survey among the affected residents in Harris County during Hurricane Harvey to learn about their purchase activities through credit cards and which business sectors they have spent their money. The survey results will be used to build the CCT fluctuation curve and further evaluate our curve based on the SafeGraph CCT data.

Despite limitations, this research contributes to the body of knowledge related to community resilience evaluation through two main aspects: first, the study and results demonstrate the potential for utilizing human activity data for more integrative evaluation of the spatial and temporal patterns of impacts and recovery in community resilience assessments. Second, the approach could be utilized for enhancing situational awareness to better inform resource allocation and recovery strategy prioritization efforts by disaster managers and public agencies.

Acknowledgments

The authors would like to acknowledge the funding support from the SafeGraph for providing the credit card transaction data. Any opinions, findings, conclusions, or recommendations expressed in this research are those of the authors and do not necessarily reflect the view of the funding agencies.

Author Contributions

F.Y., **A.E.**, and **A.M.** conceived and designed the research. **F.Y.**, **A.E.**, and **B.O.** performed the numerical calculations. **F.Y.**, **A.E.**, and **A.M.** analyzed the results, and wrote and edited the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Science Foundation CAREER Award under grant number 1846069.

ORCID iD

Faxi Yuan https://orcid.org/0000-0003-1275-8233

Supplemental Material

Supplemental material for this article is available online.

References

Alatrista-Salas H, Gauthier V, Nunez-del-Prado M, et al (2020) Impact of natural disasters on consumer behavior: case of the 2017 el nino phenomenon in peru. *PLoS One* 16(1): 1–28.

Beatty TKM, Shimshack JP and Volpe RJ (2019) Disaster preparedness and disaster response: Evidence from sales of emergency supplies before and after hurricanes. *Journal of the Association of Environmental and Resource Economists* 6(4): 633–668.

Chen Z, Gong Z, Yang S, et al. (2020) Impact of extreme weather events on urban human flow: a perspective from location-based service data. *Computers, Environment and Urban Systems* 83: 101520.

Coleman N, Esmalian A and Mostafavi A (2019) Equitable resilience in infrastructure systems: empirical assessment of disparities in hardship experiences of vulnerable populations during service disruptions. *Natural Hazards Review* 21(4): 1–13.

Dong S, Esmalian A, Farahmand H, et al. (2020) An integrated physical-social analysis of disrupted access to critical facilities and community service-loss tolerance in urban flooding. *Computers, Environment and Urban Systems* 80: 101443.

Dong X, Meyer J, Shmueli E, et al. (2018) Methods for quantifying effects of social unrest using credit card transaction data. *EPJ Data Science* 7(1): 1–25.

Esmalian A, Dong S and Mostafavi A (2020) Susceptibility curves for humans: empirical survival models for determining household-level disturbances from hazards-induced infrastructure service disruptions. *Sustainable Cities and Society* 66: 102694.

FEMA (2016) *National Mitigation Framework*. 2nd Edition. fema.gov. Available at: https://www.fema.gov/sites/default/files/2020-04/National_Mitigation_Framework2nd_june2016.pdf.

FEMA (2018) FEMA–harvey flood depths grid. Available at: https://www.hydroshare.org/resource/165e2c3e335d40949dbf501c97827837. (accessed 25 Febuary 2021).

Han SY, Tsou M, Knaap E, et al. (2019) How do cities flow in an emergency? Tracing human mobility patterns during a natural disaster with big data and geospatial data science. *Urban Science* 3(2): 51.

Gonzalez-Garcia J and Johnson A (2020) Credit card ownership statistics. Available at: https://www.creditcards.com/credit-card-news/ownership-statistics/ (accessed 23 January 2022).

Hasan S, Zhan X and Ukkusuri SV (2013) Understanding urban human activity and mobility patterns using large-scale location-based data from online social media. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Chicago, Illinois, USA, August 11–14, 2013, ACM.

Juhász L and Hochmair H (2020) Studying spatial and temporal visitation patterns of points of interest using safegraph data in Florida. *GI Forum* 1: 119–136.

Knüsel B, Zumwald M, Baumberger C, et al. (2019) Applying big data beyond small problems in climate research. *Nature Climate Change* 9(3): 196–202.

Lu X, Wrathall DJ, Sundsøy PR, et al. (2016) Unveiling hidden migration and mobility patterns in climate stressed regions: A longitudinal study of six million anonymous mobile phone users in Bangladesh. *Global Environmental Change* 38: 1–7.

Maquiling KSM, Sala SDLa and Rabé P (2021) Urban resilience in the aftermath of tropical storm washi in the philippines: the role of autonomous household responses. *Environment and Planning B: Urban Analytics and City Science* 48: 1025–1041.

Martinez EA, Rubio MH, Martinez RM, et al. (2016) *Measuring Economic Resilience to Natural Disasters with Big Economic Transaction Data*. arXiv.

Mobley W, Sebastian A, Blessing R, et al. (2021) Quantification of continuous flood hazard using random forest classification and flood insurance claims at large spatial scales: a pilot study in southeast Texas. *Natural Hazards and Earth System Sciences* 21(2): 807–822.

Mostafavi A, Abraham DM and Lee J (2012) System-of-systems approach for assessment of financial innovations in infrastructure. *Built Environment Project and Asset Management* 2(2): 250–265.

Nan C and Sansavini G (2017) A quantitative method for assessing resilience of interdependent infrastructures. *Reliability Engineering & System Safety* 157: 35–53.

Peacock WG and Ragsdale AK (1997) Social systems, ecological networks and disasters: Toward a socio-political ecology of disasters. In: Peacock WG, Morrow BH and Gladwin H. *Hurricane Andrew: Ethnicity, Gender, and the Sociology of Disasters*. NY: Routledge, 20–35.

Podesta C, Coleman N, Esmalian A, et al. (2020) Quantifying community resilience based on fluctuations in visits to point-of-interest from digital trace data. arXiv, 1–23.

Qian Z (2010) Without zoning: Urban development and land use controls in Houston. *Cities* 27(1): 31–41.

Spiegler VLM, Potter AT, Naim MM, et al. (2016) The value of nonlinear control theory in investigating the underlying dynamics and resilience of a grocery supply chain. *International Journal of Production Research* 7543: 1–22.

Utrera F (2021) Tips for normalizing foot traffic data. Available at: https://www.safegraph.com/blog/tips-for-normalizing-foot-traffic-data. (accessed 23 January 2022).

Visa (2021) Visa merchant data standards manual. Available at: https://usa.visa.com/content/dam/VCOM/download/merchants/visa-merchant-data-standards-manual.pdf (accessed 21 April 2022).

Vugrin ED, Warren DE, Ehlen MA, et al. (2010) A framework for assessing the resilience of infrastructure and economic systems. *Sustainable and resilient critical infrastructure systems*. Berlin, Heidelberg: Springer, 77–116.

Yabe T, Rao PSC and Ukkusuri SV (2020a) Regional differences in resilience of social and physical systems: case study of Puerto Rico after Hurricane Maria. *Environment and Planning B: Urban Analytics and City Science*. 48: 1042–1057.

Yabe T, Tsubouchi K, Fujiwara N, et al. (2020b) Understanding post-disaster population recovery patterns. *Journal of the Royal Society, Interface* 17(163): 20190532.

Yabe T, Zhang Y and Ukkusuri S (2020c) Quantifying the economic impact of extreme shocks on businesses using human mobility data: a bayesian causal inference approach. arXiv:2004.11121.

Yuan F and Liu R (2018) Feasibility study of using crowdsourcing to identify critical affected areas for rapid damage assessment: Hurricane Matthew case study. *International Journal of Disaster Risk Reduction* 28: 758–767.

Yuan F and Liu R (2020) Mining social media data for rapid damage assessment during hurricane matthew. *Feasibility Study* 34(3): 1–14.

Yuan F, Li M, Liu R, et al. (2021a) Social media for enhanced understanding of disaster resilience during Hurricane Florence. *International Journal of Information Management* 57: 102289.

Yuan F, Liu R, Mao L, et al. (2021b) Internet of people enabled framework for evaluating performance loss and resilience of urban critical infrastructures. *Safety Science* 134: 105079.

Yuan F, Xu Y, Li Q, et al. (2021c) Spatio-temporal graph convolutional networks for road network inundation status prediction during Urban flooding. arXiv preprint arXiv:2104.02276.

Yuan F, Yang Y, Li Q, et al. (2021d) Unraveling the temporal importance of community-scale human activity features for rapid assessment of flood impacts. arXiv preprint arXiv:2106.08370.

Yuan F, Fan C, Farahmand H, et al. (2021e) Smart flood resilience: harnessing community-scale big data for predictive flood risk monitoringrapid impact assessment, and situational awareness. arXiv preprint arXiv: 2111.06461.

Zhai W, Peng Z-R and Yuan F (2020) Examine the effects of neighborhood equity on disaster situational awareness: Harness machine learning and geotagged Twitter data. *International Journal of Disaster Risk Reduction* 48: 101611.

Author Biographies

Faxi Yuan is a postdoctoral researcher in the UrbanResilience.AI Lab. He received his Ph.D. degree in Construction Management from the University of Florida. Prior to the doctoral study, he obtained his master's degree in Civil Engineering for Risk Mitigation from Politecnico di Milano in Italy, and bachelor's degree in Construction Management from Southeast University in China. His interdisciplinary research is at the nexus of network sciences, infrastructure engineering, and social science. His research topics include disaster informatics, community resilience, infrastructure network, and IoT enabled intelligent infrastructure.

Amir Esmalian is a Ph.D. candidate in construction engineering management at Texas A&M University. He joined the UrbanResilience.AI Lab in Fall 2017 as a research assistant. He received his bachelor's degree in Civil Engineering from Sharif University of Technology. His research interests lie in community resilience and infrastructure systems. He hopes his research work could serve people and be a part of the movement toward creating a better world.

Bora Oztekin is a junior student in the Department of Computer Science and Engineering, Texas A&M University. As he learned more about natural disasters and discovered that relief organizations took great interest in the work of the research group, he was inspired to join the UrbanResilience.AI Lab in Fall 2019. His research focuses on creating versatile models from our data using data science tools. He is also interested in machine learning and natural language processing and looks forward to being a part of its application in the lab.

Ali Mostafavi is the Director of the UrbanResilience.AI Lab. At Texas A&M University, he is a fellow of the Institute for Sustainable Communities, the Institute for Science, Technology, and Public Policy, as well as the Hazard Reduction and Recovery Center. His research focuses on analyzing, modeling, and improving network dynamics in the nexus of humans, disasters, and the built environment to foster convergence knowledge of resilient communities. He received several awards and honors such as an Early-career Research Fellowship from the National Academies' Gulf Research Program and a NSF CAREER Award.