Overview

The codes in this replication package construct the data and run the analyses for "Improving Household and Community Disaster Recovery: Evidence on the Role of Insurance," by Xuesong You and Carolyn Kousky. Because some of the data sources used are proprietary, the full set of code will only run when the appropriate datasets are available.

Data Availability and Provenance Statements

Statement about Rights

- I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

- Some data **cannot be made** publicly available.

Summary of Data Sources

- Survey Data

  Please contact the authors of You and Kousky (2024) for data access.

- Foot traffic data

  We obtained access to monthly foot traffic data from SafeGraph. The sample used in this paper includes visitation patterns in FL from April 2018 to October 2019. While we are not able to directly share the raw data files following the terms of the agreement, the data is available for academic, noncommercial use by submitting a data access request to SafeGraph. The name of the dataset is "SafeGraph Patterns Data."

- Flood depth data

  We obtained the Hurricane Michael flood depth shapefile from First Street Foundation (FSF), which has restrictions on the dissemination of the raw data file for free but is available for purchase at FSF. We use the shapefile to calculate the weighted average flood depth (in feet) across developed land area within a census tract and to extract the level of flooding (in feet) at a business' location.

- FEMA flood zone data

  We used FL's FEMA flood map (National Flood Hazard Layer with an effective date of 08/31/2016) to identify the share of developed area in the floodplain by census tract. We downloaded the dataset on September 6, 2016 from the following URL: https://www.fema.gov/flood-maps/national-flood-hazard-layer.

- Land cover data

  We used the 2016 National Land Cover Database from the Multi-Resolution Land Characteristics Consortium to identify the developed area of a census tract. We downloaded the file on March 13, 2022 from the following URL: https://www.mrlc.gov/data. The name of the dataset is "National Land Cover Database 2016."

- NFIP flood insurance policies data

  We downloaded NFIP flood insurance policies data on August 5, 2022 from the following URL: https://www.fema.gov/about/openfema/data-sets. The name of the dataset was "OpenFEMA Dataset: FIMA NFIP Redacted Policies - v1," which has since been deprecated and removed by FEMA in September 2023. The current version, "OpenFEMA Dataset: FIMA NFIP Redacted Policies – v2," is available to download. Our analysis is not affected by such updates.

- NFIP flood insurance claims data

  We obtained NFIP flood insurance claims data through a Freedom of Information Act (FOIA) request to FEMA. The data discloses the census tract of the flooded property, as well as the timing and amount of the claim payment.

- Facebook Social Connectedness Index

  We downloaded Facebook Social Connectedness Index data on March 13, 2022 from the following URL: https://data.humdata.org/dataset/social-connectedness-index. The name of the dataset is "US zip code-US zip code."

- Census tract-level socio-economic demographics

  We obtained socio-economic demographic variables by census tract from the 2017 American Community Survey 5-year dataset. The R script.

- Census tract-level household savings status

  We obtained tract-level household savings status as of 2017 (i.e., share of households with a traditional savings account) from the Claritas Financial CLOUT dataset accessed via S&P Market Intelligence database. We downloaded the data on June 7, 2022.

- Local business categories

  Local business categories are defined by Delgado et al. (2016) and the file was downloaded from the following URL: https://www.clustermapping.us/content/cluster-mapping-methodology. The file name is "US Cluster Definitions 2017 v2020-0323.csv"

- Disaster Declarations data

  We used FEMA's disaster declarations data to identify the area impacted by Hurricane Michael and areas that had flood-related disaster declarations from 2016-2017. We downloaded the data on October 29, 2023 from the following URL: https://www.fema.gov/about/openfema/data-sets. The name of the dataset is "OpenFEMA Dataset: Disaster Declarations Summaries - v2."

- Tract-to-ZIP and ZIP-to-County crosswalk files

  We downloaded the 2017-Q4 tract-to-ZIP and ZIP-to-County crosswalk file on September 16, 2022 from the following URL: https://www.huduser.gov/portal/datasets/usps_crosswalk.html.

- U.S. county and tract boundary file

We downloaded the 2017 U.S. county boundary shapefile and the 2017 FL tract boundary shapefile on October 29, 2023 from the following URL: https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.2017.html#list-tab-1556094155. The names of the shapefiles are "cb_2017_us_county_500k.zip [11 MB]" and "cb_2017_12_tract_500k."

Instructions to replicators

- "01_Code/build" includes scripts that produce all the variables used in You and Kousky (2024) for analysis. See Table 1 for detail.
- "01_Code/analysis" includes scripts that produce all the figures and tables in the main paper. See Table 2 for detail.
- Without access to the proprietary data used in the paper, much of the code will not be runnable by replicators.

Table 1: Variable List

Variable Name	Definition	Variation	Data Sources & Notes	Scripts for Construction	Data Location
visitor_count_home_tract	# of visitors from a census tract to a business in a given month	tract-business-time	SafeGraph Monthly Patterns Dataset (*Proprietary*) *Note*: The original data is constructed at business-time level. The script on the right describes how we convert it to tract-business-time level.	*01_Code/build/01-FootTraffic.R*	*02_Data/FL_visitor_1804_1910_tract.Rdata* *Note*: As per the terms of agreement, we are unable to publicly share this data. Values of all variables are replaced with *
NAICS_2digit_name	a business' industry (based on NAICS 2 digits)	business			
dummy_brands	dummy variable equal to one if a business is a chain of a larger brand	business			
visitor_home_tract	census tract in which visitors' homes are located	tract			
placekey	unique ID for a business	business			
visitor_home_tract_placekey	visitor_home_tract × placekey	tract-business			
month	month of visitation	time			
Post	dummy variable equal to one for months after Hurricane Michael (month ≥ October 2018)	time			
FloodDepth_ft_location	flood depth at a business' location (in feet)	business	1) Business lat/lon from SafeGraph Patterns Dataset (*Proprietary*) 2) First Street Foundation Hurricane Michael flood depth shapefile (*Proprietary*)	Extracted with ArcGIS Pro	
FloodDepth_ft_tract	weighted average flood depth across developed land area within a census tract (in feet)	tract	1) First Street Foundation Hurricane Michael flood depth shapefile (*Proprietary*) 2) 2017 FL tract boundary shapefile from Census Bureau (public) 3) 2016 National Land Cover Database from the Multi-Resolution Land Characteristics Consortium (public)	Extracted with ArcGIS Pro	*02_Data/tract_flood_FL.Rdata* *Note*: As per the terms of agreement, we are unable to publicly share this data. Values of all variables are replaced with *
FloodInsTakeup	a tract's pre-event flood insurance take-up rate (as of the end of 2017)	tract	1) NFIP flood insurance policies data (public) 2) 2017 American Community Survey 5-year data from Census Bureau (public)	*01_Code/build/02-NFIPTakeup.R*	*02_Data/NFIPTakeup.Rdata*

FloodZone_Share_Dev_Area_tract	a tract's share of developed area in the FEMA-designated 100-year floodplain	tract	1) FL's National Flood Hazard Layer with an effective date of 08/31/2016 (public) 2) 2017 FL tract boundary shapefile from Census Bureau (public) 3) 2016 National Land Cover Database from the Multi-Resolution Land Characteristics Consortium (public)	Extracted with ArcGIS Pro	*02_Data/FloodZone.Rdata*
log_medincome2017	a tract's median household income in 2017 (logged value)	tract	2017 American Community Survey 5-year data from Census Bureau (public)	*01_Code/build/03-ACS.R*	*02_Data/tract2017.Rdata*
log_population2017	a tract's total population in 2017 (logged value)	tract			
ptg_Unemployed2017	a tract's unemployment rate in 2017	tract			
ptg_white2017	a tract's population share identifying as white in 2017	tract			
ptg_OwnerOccupied2017	a tract's share of homes that are owner occupied in 2017	tract			
ptg_Mortgage2017	a tract's share of households with a home mortgage in 2017	tract			
ptg_Bachelor2017	a tract's share of population with a Bachelor's degree or above in 2017	tract			
Gini2017	a tract's Gini Index of income inequality in 2017	tract			
ptg_SavingsAccount2017	a tract's share of households with a traditional savings account in 2017	tract	Claritas Financial CLOUT dataset accessed via S&P Market Intelligence database	Require an S&P Market Intelligence account	*02_Data/SavingsAccount.Rdata*
ClaimAmount	total dollar volume of NFIP claims that had been paid by the end of each month in a census tract	tract-time	Hurricane Michael NFIP flood insurance claims data from FEMA *Note*: Claim data (obtained via an FOIA request to FEMA) is at	*01_Code/build/04-NFIPClaim.R*	*02_Data/NFIP_Michael_Claim.Rdata*

WAvgPaidWeek	among all NFIP claims that had been paid by the end of each month in a census tract, the average number of weeks that these claims had been paid out by the end of each month, weighted by dollar values	tract-time	individual claim level. The script on the right describes how to construct the aggregated variables at tract-month level for analysis.		
SCI_Flooded_Distant	a tract's social connectedness to geographically distant areas (at least 750 miles away) that had flood-related disaster declarations from 2016-2017	tract	1) Facebook Social Connectedness Index data (public) 2) Tract-to-ZIP and ZIP-to-County crosswalk files from HUD (public) 3) FEMA's disaster declarations summaries dataset (public) 4) 2017 U.S. county boundary shapefile from Census Bureau (public)	*01_Code/build/05-SCI.R*	*02_Data/SCI_v3.Rdata*
SCI_MichaelArea	a tract's social connectedness to Hurricane Michael affected areas	tract	1) Facebook Social Connectedness Index data (public) 2) Tract-to-ZIP and ZIP-to-County crosswalk files from HUD (public) 3) FEMA's disaster declarations summaries dataset (public)		
SCI_HighIncome	a tract's social connectedness to high-income areas	tract	1) Facebook Social Connectedness Index data (public) 2) Tract-to-ZIP and ZIP-to-County crosswalk files from HUD (public) 3) 2017 American Community Survey 5-year data from Census Bureau (public)		
SCI_LowIncome	a tract's social connectedness to low-income areas	tract			
All survey variables	See You and Kousky (2024) for details	individual	Disaster survey collected by You and Kousky (2024)	*01_Code/build/06-Survey.R*	Please contact the authors of You and Kousky (2024) for access.

Table 2: List of Tables/Figures and Programs

Figure / Table #	Program
Figure 1	01_Code/analysis/07-Survey-Figure1
Figure 2	01_Code/analysis/03-FootTraffic-Figure23
Figure 3	01_Code/analysis/03-FootTraffic-Figure23
Table 1	01_Code/analysis/08-Survey-Table1
Table 2	01_Code/analysis/01-FootTraffic-Table2
Table 3	01_Code/analysis/02-FootTraffic-Table3
Table 4	01_Code/analysis/04-FootTraffic-Table4
Table 5	01_Code/analysis/05-FootTraffic-Table5
Table 6	01_Code/analysis/06-FootTraffic-Table6
Table 7	01_Code/analysis/09-Survey-Table7
Table 8	01_Code/analysis/10-Survey-Table8910
Table 9	01_Code/analysis/10-Survey-Table8910
Table 10	01_Code/analysis/10-Survey-Table8910