

Supply chain diversity buffers cities against food shocks

Michael Gomez¹, Alfonso Mejia^{1*}, Benjamin L. Ruddell², and Richard R. Rushforth²

¹Department of Civil and Environmental Engineering, The Pennsylvania State University, University Park, Pennsylvania, USA.

²School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, Arizona, USA.

*Corresponding author. E-mail: aim127@psu.edu

Food supply shocks are increasing worldwide^{1,2}, particularly the type of shock wherein food production or distribution loss in one location propagates through the food supply chain to other locations^{3,4}. Analogous to biodiversity buffering ecosystems against external shocks^{5,6}, ecological theory suggests that food supply chain diversity is crucial for managing the risk of food shock to human populations^{7,8}. Here we show that boosting a city's food supply chain diversity increases the resistance of a city to food shocks of mild to moderate severity up to 15%. We develop an intensity-duration-frequency model linking food shock risk to supply chain diversity. The empirical-statistical model is based on annual food inflow observations from all metropolitan areas in the United States during the years 2012-2015, years when most of the country experienced moderate to severe droughts. The model explains a city's resistance to food shocks of a given frequency, intensity, and duration as a monotonically declining function of the city's food inflow supply chain's Shannon diversity. This model is simple, operationally useful, and hazard-agnostic. Using this method cities can improve their resistance to food supply shocks with policies that increase the food supply chain's diversity.

Food supply shock is a pressing issue that may be increasing worldwide^{1,2}. Extreme-weather events, possibly exacerbated by climate change^{9,10}, are a main driver of food supply shocks¹¹. Furthermore, the risk of simultaneous global breadbasket failure is also likely rising¹², posing a threat for systemic and catastrophic food production losses¹³. Global and national food supply chains increase exposure to shocks compared with local food supply chains^{4,7,14,15}, but also add diversity and resilience^{14,16}.

Network topological diversity and connectivity are key attributes of resilient social-ecological^{7,17} systems. Food supply chains, along with other material inflows such as water and energy, are a close analogy to an ecological food web^{18,19}. Resilience to shock has “three R’s”²⁰; resistance to changing food inflows, robustness to a wide range of hazards, and recovery time after a shock. Therefore, from the ecological and resilience theory^{21,22}, a food shock resilience model should relate the diversity and/or connectivity of the food supply chain network to explain a city's resistance to food shocks. Supply chain diversity provides adaptive options for the city to exploit when some of its supply chains suffer shock^{23,24}, thus boosting resilience to shock. The

methods in this paper measure resistance to food supply shock, which is a specific subtype of resilience as explained above.

If cities, companies, and nations had access to a model estimating their ability to buffer food supply chain shocks, this model could be used in policy and management to optimize supply chains and control the risk of shocks²⁴. The ideal model would be simple, quantitative, accurate, operationally useful, applicable at the scale of cities which consume and process most food²⁵, hazard-agnostic for all causes of shocks, and would explain shock risk as a function of factors that a city, company, or nation can control. We are skilled at modeling hazard-specific risk management, but general hazard-agnostic models of resilience are more ideal because they can potentially handle unexpected extreme events that continue to threaten our human systems. Here, we propose a statistical-empirical model meeting these ideal criteria, explaining the resistance of United States cities to food supply shocks as a function of the topological diversity of the city's food supplier network.

Results

Using annual timescale food inflow supply networks (crops, live animals, feed, and meat) for the cities of the United States, we extract the annual domestic food inflow subgraph of each metropolitan area for the period 2012-2015, which is the period with available data²⁶ and when food production systems were substantially impacted by drought and agricultural production shocks on the Great Plains and western United States^{27,28}. In this model a shock occurs when food inflows drop by more than the intensity threshold (ranging from 1% to 15%) for a duration of a year, with the drop measured against the average inflows of the four-year time period.

Using observations of thousands of food inflows to hundreds of United States cities across four years and four types of food, we calculate the probability (or frequency) of exceeding a given shock intensity for an annual duration for each United States city. The resistance of a city to food shocks is the complement to the probability of the shock. We find that resistance relates positively to the Shannon diversity of a city's food inflows. This assumes that all cities' supply chains were exposed to shocks of many intensities during the study period. This is a justified assumption in the United States for the study period of 2012-2015, because in 2012 the United States Great Plains experienced an exceptionally severe drought²⁹ and in 2012-2013 the Western

United States experienced a severe drought⁹, both with widespread losses reported in crops and livestock^{29,30}.

The resulting model takes the operationally useful form of an intensity-duration-frequency (IDF) relationship that is widely used for risk and hazard engineering and as a basis for design codes and policies of risk³¹. This model provides an all-hazards or “hazard agnostic” approach, because- although the empirically observed shocks underlying our analysis are mostly due to drought affecting food suppliers- the model is valid in principle for all kinds of shocks to the city’s food supply network.

Relationship between food supply shocks and supply chain diversity

For cities in the United States, the probability of an annual food supply shock S being greater than a shock intensity s , $P(S > s)$ (methods), declines as the diversity D of a city’s food inflows supply chain increases (Fig. 1a). Our measure of supply chain diversity is calculated using the Shannon diversity of a city’s supply chain network of domestic trading partners based on a combination of five different indicators (Methods). Using data for 284 cities and 4 food sectors, the annual probabilities of food supply shocks are calculated by measuring for each city and food sector the maximum food supply departure from the annual average during 2012-2015 (Methods). We utilize a total of 4,544 buyer-supplier subgraphs and 1,136 time series to calculate D and $P(S > s)$.

Our results indicate that with greater supply chain diversity D cities are more likely to avoid or resist shocks of increasing intensity (3, 5, 10, and 15%; Fig. 1a). The shock intensity is quantified as the occurrence of a food supply loss greater than a specified percentage threshold s . On average, 1 in 4 cities (probability of shock 0.25; Fig. 1a) with D of 0.20 experience a supply shock greater than 15% in any of their food sectors, while for the same threshold cities with D of 0.45, 0.54, 0.63, 0.71, and 0.83 experience food supply shocks 1 out of 6, 8, 14, 17, and 202 times, respectively (probabilities of 0.18, 0.12, 0.07, 0.06 and 0.004, respectively; Fig. 1a). The same trend is observed for shock intensities greater than 3, 5, and 10% (Fig. 1a).

We use the observed empirical relationship between $P(S > s)$ and D to build a statistical IDF model of food supply shocks. Letting $F_s(D) = \ln(P(S > s))$, the model takes the following form

$$F_s(D) = -k_s(D - D_{0,s}) \propto D - D_{0,s}, \quad (1)$$

where k_s and $D_{0,s}$ are fitting parameters that depend on the shock intensity s . The parameters are estimated using nonlinear least squares. The exponential function in equation (1) provides a good fit to the data ($R^2 > 0.74$; Extended Data Table 1) and it has the desirable characteristics of being simple to implement and bounded at $P(S > s) \in [0,1]$ and $D \in [0,1]$.

Many different (social, economic, infrastructural, and environmental) urban indicators of city functioning have been shown to depend on the city size at the metropolitan level³², where population is typically used to quantify city size. Hence, we test whether D varies in a systematic way with the city size. For all food sectors, D only shows a weak positive trend with population (Extended Data Fig. 1). Therefore, the supply chains of larger cities are only slightly and weakly more diverse than those of small cities, on average.

A general IDF model of food supply shocks to cities

We use the model parameters fitted using equation (1) to derive and extrapolate a family of IDF curves covering a wide range of shock intensities. The model parameters k_s and $D_{0,s}$ are each linearly regressed against the shock intensity s to obtain parameters values for different intensities (Extended Data Fig. 2a-b). The fitted linear regressions show good performance ($R^2 > 0.83$; Extended Data Table 2). Using these regressed parameter values in equation (1), we obtain the IDF curves in Fig. 1b. The probability of shock equations for different shock intensity values are included in Extended Data Table 3. For example, assuming a design frequency of $P = 0.25$ per year (a “four-year” shock), a city with $D = 0.2$ expects to experience at least one 15% or greater food shock. If the same city increased its food supply chain diversity to $D = 0.8$, the same 15% shock occurs less often with frequency $P = 0.05$ (a “twenty-year” shock). The curves in Fig. 1b are valid for food supply chain shock to United States cities and regions at roughly the metropolitan area scale during the period 2012-2015, but may possibly be valid for other regions, time periods, and supply chain types as well.

In the standard engineering design application³¹, design codes or standards will set the maximum tolerable frequency and shock intensity (e.g. annual $P < 0.01$ and $s < 5\%$), then the city would enact policies and investments to modify D to bring the expected frequency and intensity below that level of risk. Insurance and emergency management strategies would then be developed to cover the remaining risk. These IDF curves can be directly employed by engineers,

insurers, policymakers, and planners to measure and control the risk of shock to the food supply chain and to design solutions that reduce the risk of shock.

Using the IDF curves to create a map of the so-called “100-year” shock with a 1% annual exceedance probability (frequency F is $P(S > s) = 0.01$, duration D is 1 year), we find that the expected shock intensity s varies from 21% to 31% in cities and rural regions across the United States (Fig. 2). The narrow 10% range of variability is due to the relative similarity of food supply chain structure and diversity across United States communities. The cities with the highest expected “100-year” shock intensities are Grand Junction, Colorado; Corpus Christi, Texas; Beaumont, Texas; and Steamboat Springs, Colorado (Fig. 2); and the cities with the lowest expected intensities are Florence, South Carolina; Cleveland, Ohio; Roanoke, Virginia; and Columbus, Ohio. The shock intensities are on average greater in the western United States (west of the 100th meridian) than in the East, with an average shock intensity of $25.4 \pm 1.9\%$ (average ± 1 standard deviation) versus $23.9 \pm 1.5\%$, respectively. In addition, some cities show a lower shock intensity compared to their neighboring cities. For example, Los Angeles has a lower shock intensity compared to other cities in the state of California, with a shock intensity of 22.6 versus $26.2 \pm 0.9\%$, respectively. The lower expected shock intensity of Los Angeles is due to its unusually diversified, $D=0.87$, food supply chain. Spatial patterns for shocks with other frequencies are similar.

Effect of shock co-occurrence on the IDF curves

The model in equation (1) is valid if only one food sector is shocked. However, food supply shocks to multiple sectors (e.g. crops, live animals) can and do co-occur in time¹, meaning that shocks from different sectors are simultaneously experienced by a city or region. Shock co-occurrence is due to multiple factors, such as the tendency of droughts to impact large areas and of industries to collocate in space and to form strong interdependencies^{27,28}. Even though drought and extreme heat may have a more immediate impact on crop and pasture losses, those losses can quickly propagate to other food sectors, for example, by reducing the crop inputs required to produce animal feed and by inducing livestock culling^{27,28}. The contextual details of a cascading shock are hard to predict, making it important that the model consider both single and also co-occurring shocks.

In our dataset shock co-occurrence is widespread, with most cities in the study period 2012–2015 experiencing at least one shock co-occurrence of 2 or more food sectors (Extended Data Fig. 3). Accounting for shock co-occurrence, we find that the probability $P'(S > s)$ of an annual food supply shock S exceeding intensity threshold s declines with increasing D (Fig. 3a). For example, on average 1 in 4 cities ($P'(S > s) = 0.25$), with D of 0.20 experience food supply shocks of $s > 10\%$ in more than one food sector simultaneously. Our empirical results suggest that supply chain diversity can also boost the resistance of cities to rarer but more dangerous co-occurring shocks.

To consider the effect of shock co-occurrence on the IDF curves, we modify the model in equation (1) as follows

$$F'_s(D) = -k'_s(D - D'_{0,s}), \quad (2)$$

where the fitting parameters k'_s and $D'_{0,s}$ now account for the effect of co-occurrence on the probabilities of food supply shock to a city. The parameters are estimated using nonlinear least squares (Extended Data Table 4). Each of the estimated parameters is related to the shock intensity s using linear regression (Extended Data Fig. 2c-d and Extended Data Table 5), which are in turn used, together with equation (2), to create a family of IDF curves that account for co-occurrence effects (Fig. 3b).

For a given intensity, the model expects co-occurring food supply shocks to be less frequent than single food supply shocks (Figs. 1b and 3b). For example, when neglecting co-occurrence, for an annual-duration shock intensity of 10%, a city with $D = 0.2$ has a frequency of shock of $P = 0.45$, whereas for co-occurring shocks the expected frequency decreases to $P = 0.29$. For example, the frequency with which San Diego ($D = 0.35$), California, expects a single shock of 10% intensity is $P = 0.36$, but for two or more co-occurring shocks, $P = 0.22$.

Discussion

Analogous to biodiversity buffering ecosystems against external shocks^{5,6}, our results show that cities with greater diversity of food suppliers have a lower probability of suffering a food supply shock for any reason. This method aims at operationalizing ecological theory and network theory to form a valid engineering operations risk management framework for supply chains, grounded in empirical science. This study is a step towards quantifying, explaining, and

managing the risk faced by cities and regions. The all-hazard nature of the method is important, because it promises the ability to build resilience in the face of unpredictable events that increasingly characterize hazards in our fast-paced and highly connected world. This method also holds promise for managing supply chain shock risk and resilience at the scales of companies, neighborhoods, and nations, as well as for other kinds of supply chains beyond food.

Approaches to reduce or avoid food supply shocks have been extensively explored at the national level^{14,15,33}, and the method presented in this paper corroborates- partially- prior national scale findings, while presenting a simpler answer that applies at finer and more manageable scales. Moreover, food supply shocks are most harmful to vulnerable populations within cities, so boosting food supply resilience in cities- and especially in vulnerable neighborhoods of cities- is an important policy goal. While food insecurity is on average low in developed countries, approximately 10-12% in the United States³⁴, it can be high for marginalized groups in cities and in rural areas; 20-40% in the United States³⁴. Food insecurity is positively related to food price variability across American cities³⁵. Food supply shocks may increase food prices and price variability³⁶. Therefore it seems plausible that higher food supply diversity can reduce food insecurity. Because of the varied socio-economic geography of the United States, our dataset includes cities and regions with markedly different population, density, wealth, race, culture, and climate characteristics. Therefore, our results may be directly relevant for a wide range of cities, cultures, timeframes, and nations.

Pimm et al.³⁷ emphasize the need to test and utilize empirical measures to operationalize resilience concepts. Our empirical model linking supply chain diversity to the probability of shock contributes to that operational goal. Cities need a multifaceted and adaptive approach to manage risk from multiple forms of shocks and to build resilience. In an era of expanding urbanization and connectivity, cities are posed to play a key role in global sustainability³⁸. Cities will need to actively manage their supply chain connections to deal with the causes and consequences of shocks to their critical lifelines such as food. Using this method to guide the policy objective of diversification of supply chains, cities and communities can engage in demand-side policies that scientifically manage risk and build resilience into their supply chains. The IDF framing of this method makes it directly applicable using the design-code framework that is already broadly employed by policymakers to manage risk by design.

A city's food supply chains- like most supply chains- represent the agglomeration of efforts by many independent companies and logistics operations that grow, manufacture, ship, distribute, store, and retail food products. It is a complex supply chain with many parts. Translating the high-level design framework inherent to our proposed IDF model into actionable steps and effective regulations for individual operators remains a challenge for the future. Diversity by design, at the level of communities, will take coordination, and possibly regulation, among many parties.

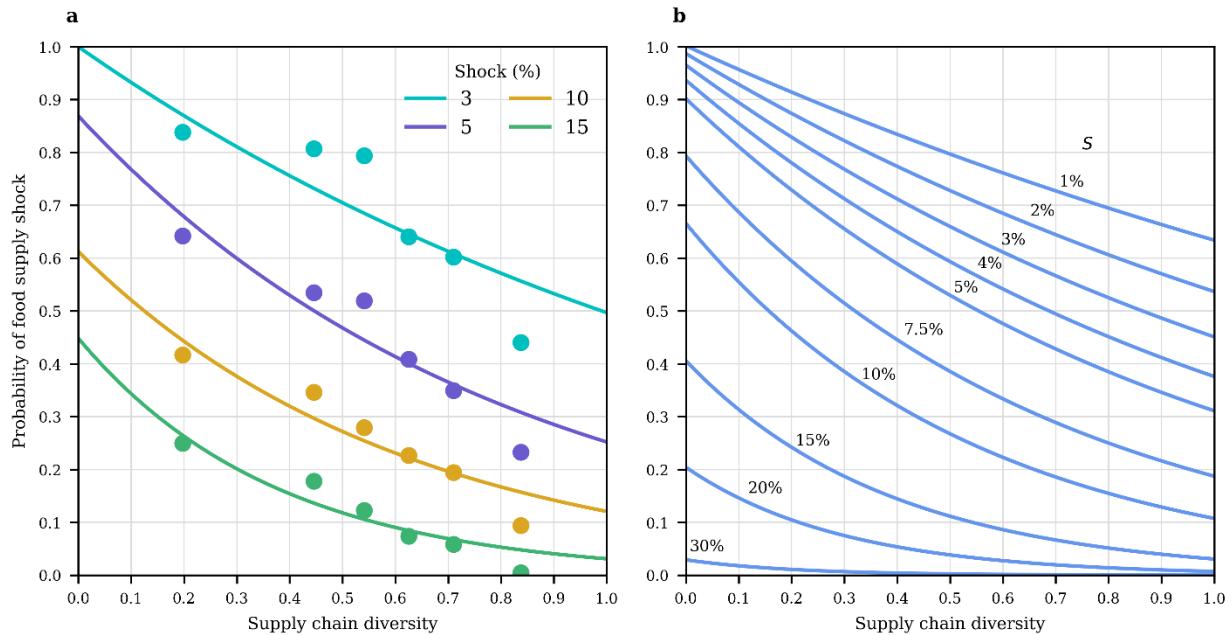


Fig. 1: Empirical (a) and modeled (b) relationship between the probability of food supply chain shock $P(S > s)$ and supply chain diversity D for different shock intensities s . For each shock intensity s , the empirical value of $P(S > s)$ is fitted against D using a constrained exponential function. With the exponential fit of the empirical relationship, IDF curves are obtained by relating the parameters from the exponential fits to the shock intensity s .

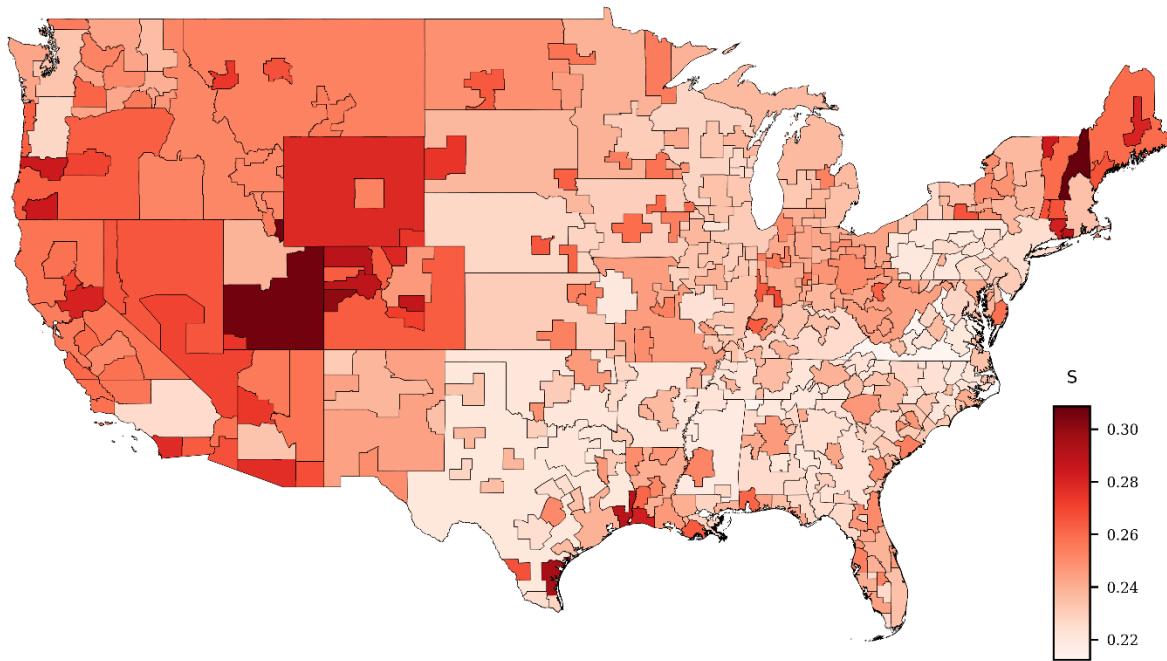


Fig. 2: Map of expected shock intensities of an annual duration food supply chain shock with annual occurrence probability of 1%. The shock intensities are represented as a fraction of the average annual inflows. The annual occurrence probability of 1% is also known as a 100-year food supply shock.

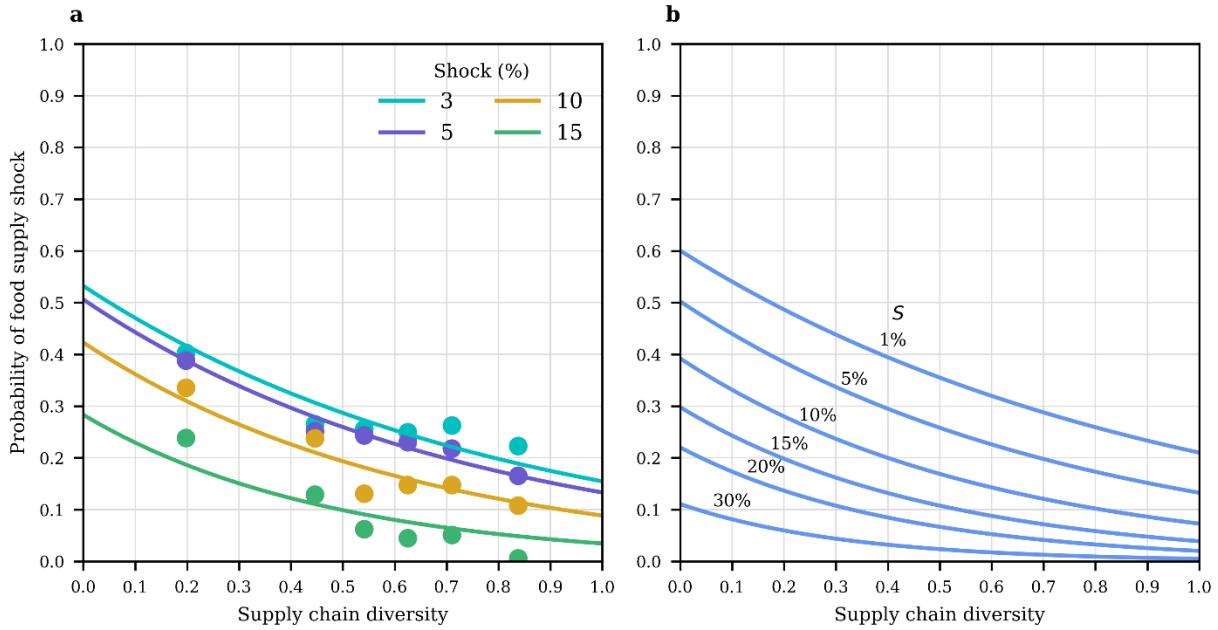


Fig. 3: Empirical (a) and modeled (b) relationship between the probability of co-occurring food supply chain shock $P'(S > s)$ versus supply chain diversity D for different shock intensities s . For each shock intensity s , the empirical value of $P'(S > s)$ is fitted against D using a constrained exponential function. With the exponential fit of the empirical relationship, IDF curves are obtained by relating the parameters from the exponential fits to the shock intensity s .

- 1 Cottrell, R. S. *et al.* Food production shocks across land and sea. *Nature Sustainability* **2**, 130-137, doi:10.1038/s41893-018-0210-1 (2019).
- 2 Wenz, L. & Levermann, A. Enhanced economic connectivity to foster heat stress-related losses. *Science Advances* **2**, e1501026, doi:10.1126/sciadv.1501026 (2016).
- 3 Bren d'Amour, C., Wenz, L., Kalkuhl, M., Christoph Steckel, J. & Creutzig, F. Teleconnected food supply shocks. *Environmental Research Letters* **11**, 035007, doi:10.1088/1748-9326/11/3/035007 (2016).
- 4 Tu, C., Suweis, S. & D'Odorico, P. Impact of globalization on the resilience and sustainability of natural resources. *Nature Sustainability* **2**, 283-289, doi:10.1038/s41893-019-0260-z (2019).
- 5 Isbell, F. *et al.* Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature* **526**, 574-577, doi:10.1038/nature15374 (2015).
- 6 Bennett, J. A. *et al.* Resistance of soil biota and plant growth to disturbance increases with plant diversity. *Ecology Letters* **23**, 119-128, doi:10.1111/ele.13408 (2020).
- 7 Nyström, M. *et al.* Anatomy and resilience of the global production ecosystem. *Nature* **575**, 98-108, doi:10.1038/s41586-019-1712-3 (2019).
- 8 Renard, D. & Tilman, D. National food production stabilized by crop diversity. *Nature* **571**, 257-260, doi:10.1038/s41586-019-1316-y (2019).
- 9 Diffenbaugh, N. S., Swain, D. L. & Touma, D. Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences* **112**, 3931, doi:10.1073/pnas.1422385112 (2015).
- 10 Naumann, G. *et al.* Global Changes in Drought Conditions Under Different Levels of Warming. *Geophysical Research Letters* **45**, 3285-3296, doi:10.1002/2017GL076521 (2018).
- 11 Lesk, C., Rowhani, P. & Ramankutty, N. Influence of extreme weather disasters on global crop production. *Nature* **529**, 84-87, doi:10.1038/nature16467 (2016).
- 12 Gaupp, F., Hall, J., Hochrainer-Stigler, S. & Dadson, S. Changing risks of simultaneous global breadbasket failure. *Nature Climate Change* **10**, 54-57, doi:10.1038/s41558-019-0600-z (2020).
- 13 Mehrabi, Z. Food system collapse. *Nature Climate Change* **10**, 16-17, doi:10.1038/s41558-019-0643-1 (2020).
- 14 Puma, M. J., Bose, S., Chon, S. Y. & Cook, B. I. Assessing the evolving fragility of the global food system. *Environmental Research Letters* **10**, 024007, doi:10.1088/1748-9326/10/2/024007 (2015).
- 15 Suweis, S., Carr, J. A., Maritan, A., Rinaldo, A. & D'Odorico, P. Resilience and reactivity of global food security. *Proceedings of the National Academy of Sciences* **112**, 6902, doi:10.1073/pnas.1507366112 (2015).
- 16 Marchand, P. *et al.* Reserves and trade jointly determine exposure to food supply shocks. *Environmental Research Letters* **11**, 095009, doi:10.1088/1748-9326/11/9/095009 (2016).
- 17 Folke, C. *et al.* Regime Shifts, Resilience, and Biodiversity in Ecosystem Management. *Annual Review of Ecology, Evolution, and Systematics* **35**, 557-581, doi:doi:10.1146/annurev.ecolsys.35.021103.105711 (2004).
- 18 Rushforth, R. & Ruddell, B. The vulnerability and resilience of a city's water footprint: The case of Flagstaff, Arizona, USA. *Water Resources Research* **52**, 2698-2714, doi:10.1002/2015WR018006 (2016).

- 19 Vallina, S. M. & Le Quéré, C. Stability of complex food webs: Resilience, resistance and the average interaction strength. *Journal of Theoretical Biology* **272**, 160-173, doi:<https://doi.org/10.1016/j.jtbi.2010.11.043> (2011).
- 20 Allen, C. R., Angeler, D. G., Chaffin, B. C., Twidwell, D. & Garmestani, A. Resilience reconciled. *Nature Sustainability* **2**, 898-900, doi:10.1038/s41893-019-0401-4 (2019).
- 21 Arese Lucini, F., Morone, F., Tomassone, M. S. & Makse, H. A. Diversity increases the stability of ecosystems. *PLOS ONE* **15**, e0228692, doi:10.1371/journal.pone.0228692 (2020).
- 22 Gao, J., Barzel, B. & Barabási, A.-L. Universal resilience patterns in complex networks. *Nature* **530**, 307-312, doi:10.1038/nature16948 (2016).
- 23 Bai, X., Nagendra, H., Shi, P. & Liu, H. Cities: build networks and share plans to emerge stronger from COVID-19. *Nature* **584**, 517-520, doi:10.1038/d41586-020-02459-2 (2020).
- 24 Garnett, P., Doherty, B. & Heron, T. Vulnerability of the United Kingdom's food supply chains exposed by COVID-19. *Nature Food* **1**, 315-318, doi:10.1038/s43016-020-0097-7 (2020).
- 25 Davies, E. G. R. Cities drive food and water security. *Nature Sustainability* **1**, 120-121, doi:10.1038/s41893-018-0038-8 (2018).
- 26 Hwang, H. *et al.* The Freight Analysis Framework Version 4 (FAF4), ORNL/TM-2016/489. *Oak Ridge National Laboratory* (2016).
- 27 Countryman, A. M., Paarlberg, P. L. & Lee, J. G. Dynamic Effects of Drought on the U.S. Beef Supply Chain. *Agricultural and Resource Economics Review* **45**, 459-484, doi:10.1017/age.2016.4 (2016).
- 28 Leister, A. M., Paarlberg, P. L. & Lee, J. G. Dynamic effects of drought on U.S. crop and livestock sectors. *Journal of Agricultural and Applied Economics* **47**, 261-284, doi:10.1017/aae.2015.8 (2015).
- 29 Rippey, B. R. The U.S. drought of 2012. *Weather and Climate Extremes* **10**, 57-64, doi:<https://doi.org/10.1016/j.wace.2015.10.004> (2015).
- 30 Schnoor, J. L. The U.S. Drought of 2012. *Environmental Science & Technology* **46**, 10480-10480, doi:10.1021/es303416z (2012).
- 31 Mays, L. *Water Resources Engineering 3rd Edition*. (Wiley, 2019).
- 32 Bettencourt, L. M. A. The Origins of Scaling in Cities. *Science* **340**, 1438-1441, doi:10.1126/science.1235823 (2013).
- 33 Seekell, D. *et al.* Resilience in the global food system. *Environmental Research Letters* **12**, 025010, doi:10.1088/1748-9326/aa5730 (2017).
- 34 Coleman-Jensen, A., Rabbitt, M. P., Gregory, C. A. & Singh, A. (ed Economic Research Service U.S. Department of Agriculture) (2016).
- 35 Gregory, C. A. & Coleman-Jensen, A. Do High Food Prices Increase Food Insecurity in the United States? *Applied Economic Perspectives and Policy* **35**, 679-707, doi:10.1093/aep/ppt024 (2013).
- 36 Gilbert, C. L. & Morgan, C. W. Food price volatility. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 3023-3034, doi:10.1098/rstb.2010.0139 (2010).
- 37 Pimm, S. L., Donohue, I., Montoya, J. M. & Loreau, M. Measuring resilience is essential to understand it. *Nature Sustainability* **2**, 895-897, doi:10.1038/s41893-019-0399-7 (2019).

38 Elmqvist, T. *et al.* Sustainability and resilience for transformation in the urban century.
Nature Sustainability **2**, 267-273, doi:10.1038/s41893-019-0250-1 (2019).

Methods

Dataset of food flow networks. We derive annual, subnational food flow networks for the United States using the Freight Analysis Framework version 4 (FAF4) database²⁶. The networks are for different food sectors and include all metropolitan areas in the United States. The FAF4 database consists of annual commodity flows during 2012-2015 for 132 geographic areas in the United States and 43 different sectors. We focus on the following four food sectors in the FAF4 database: crops, live animals, animal feed, and meat. The 132 geographic areas in the FAF4 database cover the entire contiguous United States, including 84 metropolitan statistical areas and 48 remainders of states (area of a state that is not part of a FAF4 metropolitan area).

To obtain food flows for all metropolitan areas in the United States, we disaggregate the FAF4 database from 132 to 329 areas, out of which 284 are metropolitan or combined statistical areas (120 metropolitan and 164 combined statistical areas). The disaggregation is performed using different economic and agricultural-related variables as attractors of supply and demand. For each food sector, a flow with origin o and destination d in the FAF4 database is disaggregated to a metropolitan-level flow with origin o' and destination d' using a disaggregation variable a as the best attractor of supply or demand.

The disaggregation is performed in two stages. In the first stage, the supply U of each FAF4 remainder of state is disaggregated to include all the metropolitan areas in that state as follows

$$U_{o'd}^c = \frac{U_{od}^c}{a_o} \times a_{o'}, \quad (3)$$

where $U_{o'd}^c$ [tons/yr] is the disaggregated supply for food sector c in origin o' that satisfies demand at the FAF4 destination d , U_{od}^c [tons/yr] is the FAF4 food flow for sector c between areas o and d , and $a_{o'}$ and a_o are the attractor variables for the new origin o' and FAF4 origin o , respectively. In the second stage, $U_{o'd}^c$ is further disaggregated into demand E using

$$E_{o'd'}^c = \frac{U_{o'd}^c}{a_d} \times a_{d'}, \quad (4)$$

where $E_{o'd'}^c$ [tons/yr] is the demand at destination d' for food sector c supplied by origin o' , while $a_{d'}$ and a_d are the attractor variables at the disaggregated destination d' and FAF4 destination d , respectively.

The FAF4 database includes both state-level and regional-level (132 areas) commodity flows. Prior to performing the metropolitan-level disaggregation, we use the state-level flows to evaluate the performance of different attractor variables. That is, we first use equations (3) and (4) to disaggregate the FAF4 flows from state-level to regional-level. Then, by comparing the performance of our state-level disaggregated flows against the empirical regional-level FAF4 flows, we select the best attractor variable for each food sector. The following attractor variables are considered: population³⁹, employment⁴⁰, wages⁴⁰, number of establishments⁴⁰, and available cropland⁴¹. These variables are selected based on previous analysis and data availability⁴².

We use the Pearson's correlation coefficient between the empirical FAF4 flows and the flows disaggregated from state-level to regional-level using equations (3) and (4) to choose the best attractors of supply and demand for each food sector (Extended Data Fig. 4). Using the best-performing disaggregation variables, we build the metropolitan-level food flow networks employed in this study. The nodes in the networks represent metropolitan areas and remainders

of states, and the weighted links represent annual food flows during 2012-2015 for crops, live animals, feed, and meat.

Food inflows supply chain diversity of a city. To determine annual supply chain diversity, we extract the annual food buyer-supplier subgraph of each city and food sector from the food flow networks¹⁸. The food buyer-supplier subgraph of a city i consists of all the supply chain interactions with its trading partners or neighbors j for a specific food sector. Our measure of supply chain diversity is based on the notion of functional distance⁴³. We compute the functional distance d between i and any of its trading partners j by combining five different indicators: physical distance, climate correlation, urban classification, economic specialization, and network modularity. These indicators are described in the supplementary information. The indicators represent stable characteristics of cities and therefore tend to remain fairly constant during our study period.

The functional distance d_{ij}^r for an indicator r between any pair of connected nodes (i,j) is calculated as

$$d_{ij}^r = N^{-1} |r_j - r_i|, \quad (5)$$

where the normalization constant N is determined as the maximum value of $|r_j - r_i|$ among all the neighbors j of i . In equation (5), $d_{ij}^r = 0$ for functionally similar nodes and $d_{ij}^r = 1$ for dissimilar nodes.

For each city's buyer-supplier subgraph and food sector, any pair of connected nodes has five different functional distance indicators associated with it. To combine these distance indicators into a single measure, we calculate the average functional distance indicator $\langle d_{ij}^r \rangle$ as the arithmetic average of the five functional distance indicators for any pair (i,j) of connected nodes. We use the discrete probability distribution of food inflows binned by $\langle d_{ij}^r \rangle$ categories, together with Shannon entropy⁴⁴, to calculate the supply chain diversity $D_{i,c}^t$ of node i and sector c at year t ,

$$D_{i,c}^t = \frac{-\sum_{k=1}^K Y_{i,c}^t(k) \ln Y_{i,c}^t(k)}{\log K}. \quad (6)$$

For sector c and year t , $Y_{i,c}^t(k)$ is the proportion of food inflows to node i within bin k . The k bin is obtained by binning all the $\langle d_{ij}^r \rangle$ values for node i into a total number of K bins.

$D_{i,c}^t$ is sensitive to the total number of bins K . Thus, for each node in our food flow networks, we tested the sensitivity of $D_{i,c}^t$ to the total number of bins K . For $K=15$, D values stabilize (Extended Data Fig. 5); therefore, we use 15 bins when performing all calculations of functional diversity.

Probabilities of food supply shock. The annual probability of food supply shock is calculated as the probability that food inflows to a city fall below a percentage of the average inflows for that city during 2012-2015⁸. To compute this probability, we group all nodes from the 4 food flow networks (1221 observations) into 6 diversity bins ordered from lowest to highest functional diversity D . The bin size is selected to obtain bins with similar number of

observations, approximately 204 observations in each bin. For each city i and food sector c in a bin, we calculate the food supply shock S_i^c as

$$S_i^c = \left[1 - \frac{\min(I_i^c)}{\langle I_i^c \rangle} \right] \times 100, \quad (7)$$

where I_i^c is the time series of total food inflows to node i for sector c during 2012-2015, and $\min(I_i^c)$ and $\langle I_i^c \rangle$ are the minimum and average values of the time series I_i^c , respectively.

For each diversity bin b , we count the number of observations n_b that meet the criteria $S_i^c > s$ for $s \in \{3, 5, 10, 15\}$, with s being the shock intensity threshold. The probability of a food supply shock S being greater than s in bin b is calculated as

$$P_b(S > s) = \frac{n_b}{N_b}, \quad (8)$$

where N_b is the total number of observations in bin b . Thus, for each shock intensity s , we obtain a set of probabilities of food supply shock,

$$P(S > s) = P_b(S > s) \text{ for } b = \{1, \dots, 6\}. \quad (9)$$

Furthermore, we adapt equations (8) and (9) to calculate the probability of a food supply shock S being greater than s , $P'(S > s)$, under co-occurrence conditions. We define co-occurrence as any city that experiences a shock to 2 or more food sectors during 2012-2015. With this definition, $P'(S > s)$ is calculated in a similar fashion as described above. We bin the network's nodes into 6 groups from lowest to highest diversity and determine the percentage of food supply shock with Eq. (7). Letting n'_b be the total number of cities for which $S_i^c > s$ holds for 2 or more food sectors, the probability of a food supply shock S being greater than the shock intensity s in bin b is now

$$P'_b(S > s) = \frac{n'_b}{N'_b}, \quad (10)$$

where N'_b is the total number of cities in bin b . Thus, under co-occurrence conditions, the new set of probabilities for each shock intensity s is

$$P'(S > s) = P'_b(S > s) \text{ for } b = \{1, \dots, 6\}. \quad (11)$$

- 39 *Annual Resident Population Estimates 2010-2016* (United States Census Bureau, Population Division, 2017).
- 40 *Quarterly census of employment and wages 2012* (United States Bureau of Labor Statistics, 2015).
- 41 *Crop acreage data reported to the Farm Service Agency 2012* (United States Department of Agriculture, Farm Service Agency, 2013).
- 42 Rushforth, R. R. & Ruddell, B. L. The Hydro-Economic Interdependency of Cities: Virtual Water Connections of the Phoenix, Arizona Metropolitan Area. *Sustainability* **7**, 8522-8547 (2015).
- 43 Walker, B., Kinzig, A. & Langridge, J. Original Articles: Plant Attribute Diversity, Resilience, and Ecosystem Function: The Nature and Significance of Dominant and Minor Species. *Ecosystems* **2**, 95-113, doi:10.1007/s100219900062 (1999).

- 44 Shannon, C. E. A mathematical theory of communication. *Bell System Technol. J.* **27**,
379-423 (1948).

Data availability

All data used in the analysis are publicly available and accordingly cited in the paper, Extended Data and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Acknowledgments

This material is based upon work supported by the National Science Foundation (NSF) under grant ACI-1639529. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the NSF.

Author contributions

M.G. carried the data analysis and generated the results with input from A.M. All authors devised the research, contributed to the analysis of results and wrote the manuscript.

Competing interests

The authors declare that they have no competing interests.