Article

Supply chain diversity buffers cities against food shocks

https://doi.org/10.1038/s41586-021-03621-0

Received: 12 October 2020

Accepted: 7 May 2021

Published online: 7 July 2021

Check for updates

Michael Gomez[1], Alfonso Mejia[1]✉, Benjamin L. Ruddell[2] & Richard R. Rushforth[2]

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 by up to 15 per cent. 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 USA during the years 2012 to 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 addresses any kind of hazard. 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 shocks[11]. The risk of simultaneous global breadbasket failure is also probably rising[12], posing a threat to systemic and catastrophic food production losses[13]. Geopolitical crises and policy changes are also responsible for a large proportion of shocks to different food systems[1]. Moreover, threats to public health can disrupt food supply chains, as shown by the coronavirus disease 2019 pandemic in several national economies around the world[14]. Global and national food supply chains increase exposure to shocks compared with local food supply chains[4,7,15,16], but also add diversity and resilience[15,17].

Network topological diversity and connectivity are key attributes of resilient social–ecological[7,18] systems. Food supply chains, along with other material inflows such as water and energy, are a close analogy to an ecological food web[19,20]. Resilience to shock has "three R's"[21,22]: resistance to changing food inflows, robustness to a wide range of hazards and recovery or reorganization time after a shock. Therefore, from ecological and resilience theory[23,24], 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[25–27], thus boosting resilience to shock. The methods in this paper measure resistance to food supply shock, which is a specific subtype of resilience[21].

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 shocks[26,28]. The ideal model should be simple, quantitative, accurate, operationally useful, applicable at the scale of cities (which consume and process most food)[29], '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 modelling hazard-specific risk management, but general hazard-agnostic models of resilience are better 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 USA 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 USA, we extracted the annual intranational food inflow subgraph of each metropolitan area for the period 2012–2015, which is the period with available data[30] and when food production systems were substantially affected by drought and agricultural production shocks on the Great Plains and in the western USA[31,32]. In this model a shock occurs when food inflows drop by more than the intensity threshold (ranging from 3% 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 USA 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 USA city. The resistance of a city to food shocks is the

[1]Department of Civil and Environmental Engineering, The Pennsylvania State University, University Park, PA, USA. [2]School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA. ✉e-mail: aim127@psu.edu

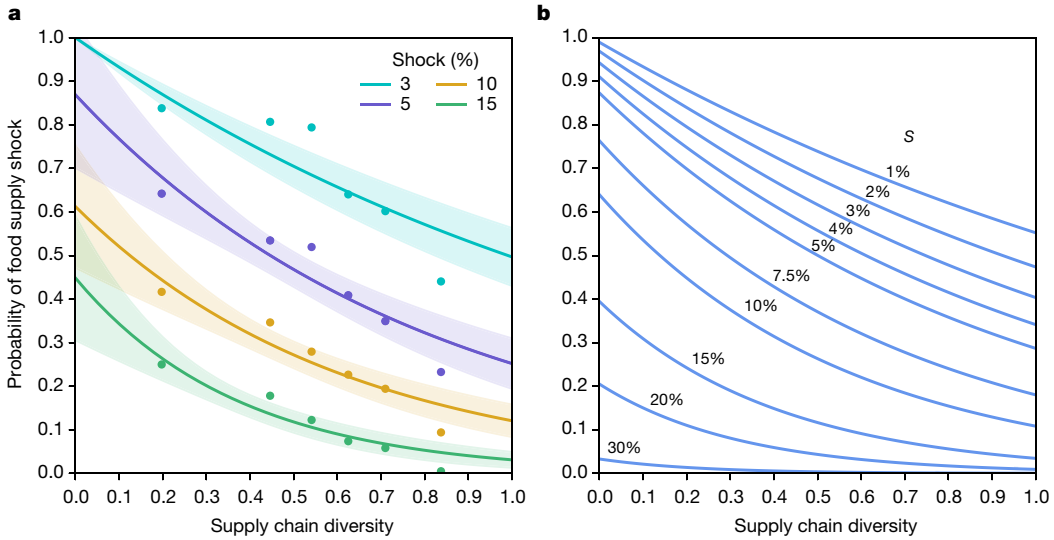


Fig. 1 | Relationship between the probability of shock and supply chain diversity. a, **b**, Empirical (**a**) and modelled (**b**) relationship between the probability of food supply chain shock $P(S > s)$ and supply chain diversity D for different shock intensities s. Food systems—each food system consists of the supply chain for one of the cities' four food sectors—are classified into 6 bins of supply chain diversity (limits 0, 0.395, 0.497, 0.585, 0.665, 0.755 and 0.92) with 204, 202, 204, 203, 206 and 202 observations in each bin. These 6 bins each have an average (with standard error given in parentheses) supply chain diversity of 0.28 (0.007), 0.45 (0.002), 0.54 (0.002), 0.62 (0.002), 0.71 (0.002) and 0.81 (0.002), respectively. The probability of food supply chain shock in each bin is calculated as the number of food systems with shocks larger than a shock intensity s divided by the total number of food systems in the bin (Methods). For clarity, panel **a** shows the probabilities of food supply chain shock for $s \in \{3, 5, 10, 15\}$ (for other s values see Supplementary Fig. 1). For each s value, the empirical value of $P(S > s)$ is fitted against D using a constrained exponential function (curves in **a**). The confidence bounds represent ±1 s.d. of the fitted curves (degrees of freedom = 4). The parameter values and goodness-of-fit results for these exponential fits are shown in Extended Data Table 1. With the exponential fit of the empirical relationship, the curves in panel **b** are obtained by relating the parameters from the exponential fits to the shock intensity s using linear regression (Extended Data Fig. 1a, b). The parameter values and goodness-of-fit results for these linear fits are shown in Extended Data Table 2.

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 USA for the study period of 2012–2015, because in 2012 the USA Great Plains experienced an exceptionally severe drought[33] and in 2012–2013 the western USA experienced a severe drought[9], both with widespread losses reported in crops and livestock[33,34].

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 risk[35]. 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.

Food shocks and supply chain diversity

For cities in the USA, the probability of an annual food supply shock S being greater than a shock intensity s, $P(S > s)$ (see 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 intranational 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 4,884 buyer–supplier subgraphs and 1,221 time series to calculate $P(S > s)$ and D.

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 values 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). The relationship in Fig. 1a holds when controlling for changes in demand using population or food inflows as proxy for demand (Methods). It also holds when the analysis is performed using the original food flow data with 69 cities rather than the 284 cities in Fig. 1a (Methods). In addition, although the five indicators used to calculate diversity have a varied influence on the relationship in Fig. 1a, the inclusion of all indicators in the supply chain diversity measure increases the Pearson correlation between the data in Fig. 1a (Methods).

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 ($R^2 \geq 0.73$; 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 scale with the city size at the metropolitan level[36,37], where population is typically used to quantify city size. Hence, we evaluate whether D varies in a systematic

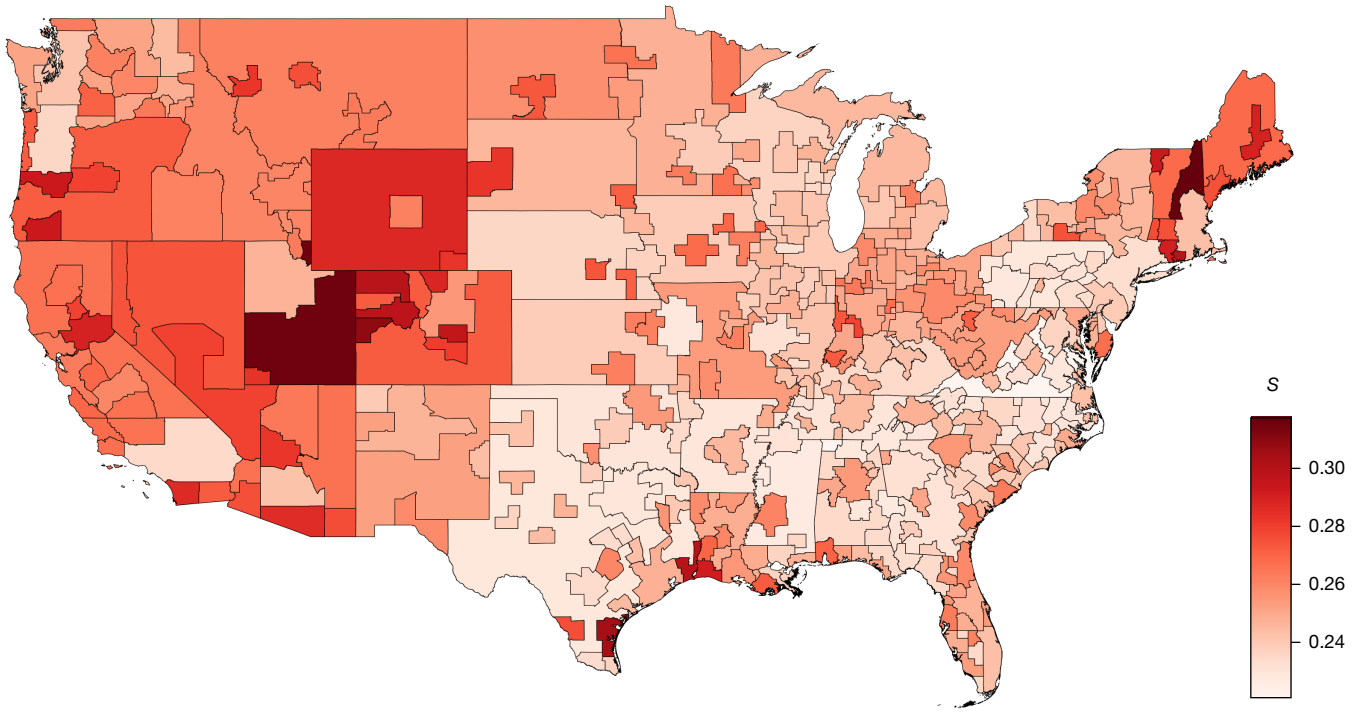


Fig. 2 | Map of expected shock intensities of an annual duration food supply chain shock with annual occurrence probability of 1%. The shock intensities are represented as a fraction of the average annual inflows. The annual occurrence probability of 1% is also known as a 100-year food supply shock.

way with the city size. For all food sectors, D shows a very weak positive trend with population (Supplementary Fig. 2). Since larger cities tend to have slightly more diverse supply chains (Supplementary Fig. 2), they are slightly less likely to experience shocks of increasing intensity.

A general model of food 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 parameter values for different intensities (Extended Data Fig. 1a, b). The fitted linear regressions show good performance ($R^2 \geq 0.80$; 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 $D = 0.2$ can expect to experience at least one food shock of 15% or greater. 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 USA 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 application[35], design codes or standards will set the maximum tolerable frequency and shock intensity (for example, annual $P < 0.01$ and $s < 5\%$), and 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 22% to 32% in cities and rural regions across the USA (Fig. 2). The narrow 10% range of variability is due to the relative similarity of food supply chain structure and diversity across USA 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 USA (west of the 100th meridian) than in the eastern USA (Fig. 2), with an average shock intensity of 26.3 ± 1.9% (average ± 1 standard deviation) versus 24.8 ± 1.5%, respectively. In addition, some cities show a lower shock intensity than their neighbouring cities. For example, Los Angeles has a lower shock intensity compared to other cities in the state of California, with a shock intensity of 23.5% versus 27.2 ± 0.9%, respectively. The lower expected shock intensity of Los Angeles is due to its unusually diversified (with $D = 0.69$) food supply chain. Spatial patterns for shocks with other frequencies are similar.

Co-occurrence of food shocks

The model in equation (1) is valid if only one food sector is shocked. However, food supply shocks to multiple sectors (for example, crops, live animals) can and do co-occur[1], meaning that shocks from different sectors can be simultaneously experienced by a city or region. Shock co-occurrence is due to multiple factors, such as the tendency of droughts to affect large areas and of industries to collocate and to form strong interdependencies[31,32]. 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

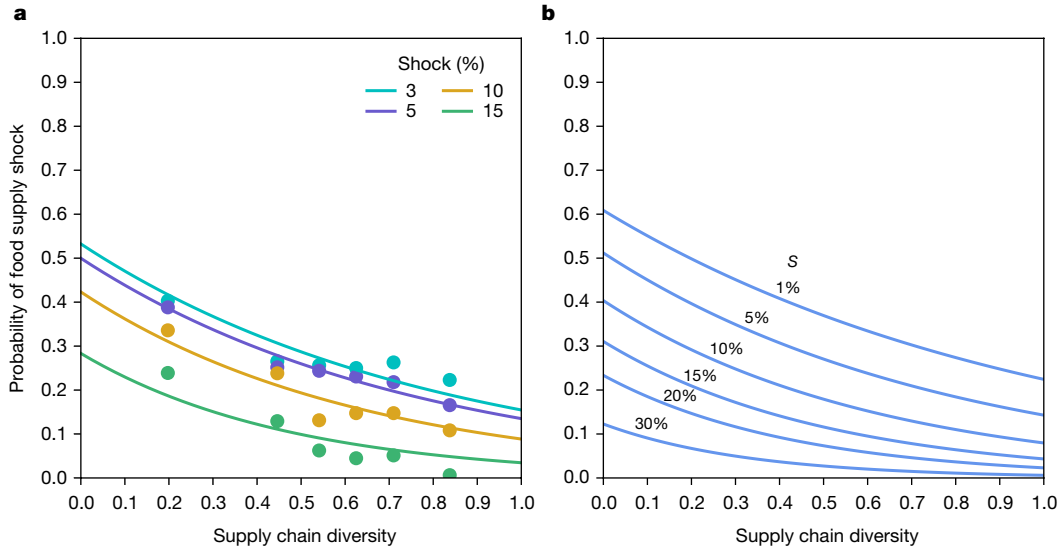


Fig. 3 | Relationship between the probability of co-occurring shocks and supply chain diversity. a, **b**, Empirical (**a**) and modelled (**b**) relationship between the probability of co-occurring food supply chain shock $P'(S>s)$ and supply chain diversity D for different shock intensities s. Food systems are classified into 6 bins of supply chain diversity (limits 0, 0.395, 0.497, 0.585, 0.665, 0.755 and 0.92) with 204, 202, 204, 203, 206 and 202 observations in each bin. These 6 bins each have an average (with standard error given in parentheses) supply chain diversity of 0.28 (0.007), 0.45 (0.002), 0.54 (0.002), 0.62 (0.002), 0.71 (0.002) and 0.81 (0.002), respectively. The probability of co-occurring food supply chain shocks in each bin is calculated as the number of cities with shocks larger than a shock intensity s to 2 or more food sectors divided by the total number of cities in the bin (Methods). For each s, the empirical value of $P'(S>s)$ is fitted against D using a constrained exponential function (curves in panel **a**). The parameter values and goodness-of-fit results for these exponential fits are shown in Extended Data Table 4. With the exponential fit of the empirical relationship, the curves in panel **b** are obtained by relating the parameters from the exponential fits to the shock intensity s using linear regression (Extended Data Fig. 1c, d). The parameter values and goodness-of-fit results for these linear fits are shown in Extended Data Table 5. Panel **b** shows the curves for $s \in \{3, 5, 10, 15\}$; for other s values see Supplementary Fig. 3.

inducing livestock culling[31,32]. The contextual details of a cascading shock are difficult 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. 2). 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 (that is, $P'(S>s)=0.25$), 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. 1c, 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-occurrent 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-occurrent shocks, $P=0.22$.

Discussion

Analogous to biodiversity buffering ecosystems against external shocks[5,6], our results show that cities with a 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 the 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, neighbourhoods and nations, as well as for other kinds of supply chains beyond food.

Approaches aimed at reducing or avoiding food supply shocks have been extensively explored at the national level[15,16,38], 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. Food supply shocks are most harmful to vulnerable populations within cities, so boosting food supply resilience in cities—and especially in vulnerable neighbourhoods of cities—is an important policy goal. Although food insecurity is on average low in developed countries, approximately 10–12% in the USA[39], it can be high for marginalized groups in cities and in rural areas; 20–40% in the USA[39]. Food insecurity is positively related to food price variability across American cities[40]. Food supply shocks may increase food prices and price variability[41]. Therefore, it seems plausible that higher food supply diversity could reduce food insecurity. Because of the varied

socioeconomic geography of the USA, 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.[42] 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 have a key role in global sustainability[43]. 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[44]. 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.

Food businesses (for example, grocers, restaurants and distributors) and—to a lesser extent, individuals—can voluntarily contribute to a resilient food policy by intentionally favouring a diverse supply chain where possible, and by maintaining slightly larger food inventories in locales that are known to be at higher risk of food shock (for example, southern Utah). Local, state, and federal governments and mission agencies can collect data on the diversity of their businesses' suppliers, can set policy targets to achieve minimum supply chain diversity, and can create regulations or incentives to achieve those minimums. In the USA, this could be relevant to several government-sponsored programmes aimed at reducing food insecurity, such as the Supplementary Nutrition Assistance Program[45] and the National School Lunch Program[46].

Insurers can price in shock risk where appropriate to incentivize diversity and resilience, and a niche business sector could emerge to mitigate that risk by managing supply chains or building buffers. Large agribusiness can reduce the supply-side risk by diversifying its supply chains. Governments can scientifically size and locate food buffers (stockpiles) to fill the gap between the IDF food security risk metrics presented in this paper and the desired level of food security—or can prioritize emergency recovery assets to higher-risk locations. The precise and actionable statistics in this model form a basis for scientifically designed food security and risk management systems that fit neatly within the existing risk management and design paradigm already used by insurers, engineers, emergency managers and policymakers.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03621-0.

1. Cottrell, R. S. et al. Food production shocks across land and sea. *Nat. Sustainability* **2**, 130–137 (2019).
2. Wenz, L. & Levermann, A. Enhanced economic connectivity to foster heat stress–related losses. *Sci. Adv.* **2**, e1501026 (2016).
3. Bren d'Amour, C., Wenz, L., Kalkuhl, M., Christoph Steckel, J. & Creutzig, F. Teleconnected food supply shocks. *Environ. Res. Lett.* **11**, 035007 (2016).
4. Tu, C., Suweis, S. & D'Odorico, P. Impact of globalization on the resilience and sustainability of natural resources. *Nat. Sustainability* **2**, 283–289 (2019).
5. Isbell, F. et al. Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature* **526**, 574–577 (2015).
6. Bennett, J. A. et al. Resistance of soil biota and plant growth to disturbance increases with plant diversity. *Ecol. Lett.* **23**, 119–128 (2020).
7. Nyström, M. et al. Anatomy and resilience of the global production ecosystem. *Nature* **575**, 98–108 (2019).
8. Renard, D. & Tilman, D. National food production stabilized by crop diversity. *Nature* **571**, 257–260 (2019).
9. Diffenbaugh, N. S., Swain, D. L. & Touma, D. Anthropogenic warming has increased drought risk in California. *Proc. Natl Acad. Sci. USA* **112**, 3931–3936 (2015).
10. Naumann, G. et al. Global changes in drought conditions under different levels of warming. *Geophys. Res. Lett.* **45**, 3285–3296 (2018).
11. Lesk, C., Rowhani, P. & Ramankutty, N. Influence of extreme weather disasters on global crop production. *Nature* **529**, 84–87 (2016).
12. Gaupp, F., Hall, J., Hochrainer-Stigler, S. & Dadson, S. Changing risks of simultaneous global breadbasket failure. *Nat. Clim. Change* **10**, 54–57 (2020).
13. Mehrabi, Z. Food system collapse. *Nat. Clim. Change* **10**, 16–17 (2020).
14. Scheelbeek, P. F. D. et al. United Kingdom's fruit and vegetable supply is increasingly dependent on imports from climate-vulnerable producing countries. *Nat. Food* **1**, 705–712 (2020).
15. Puma, M. J., Bose, S., Chon, S. Y. & Cook, B. I. Assessing the evolving fragility of the global food system. *Environ. Res. Lett.* **10**, 024007 (2015).
16. Suweis, S., Carr, J. A., Maritan, A., Rinaldo, A. & D'Odorico, P. Resilience and reactivity of global food security. *Proc. Natl Acad. Sci. USA* **112**, 6902–6907 (2015).
17. Marchand, P. et al. Reserves and trade jointly determine exposure to food supply shocks. *Environ. Res. Lett.* **11**, 095009 (2016).
18. Folke, C. et al. Regime shifts, resilience, and biodiversity in ecosystem management. *Annu. Rev. Ecol. Evol. Syst.* **35**, 557–581 (2004).
19. Rushforth, R. & Ruddell, B. The vulnerability and resilience of a city's water footprint: the case of Flagstaff, Arizona, USA. *Wat. Resour. Res.* **52**, 2698–2714 (2016).
20. Vallina, S. M. & Le Quéré, C. Stability of complex food webs: resilience, resistance and the average interaction strength. *J. Theor. Biol.* **272**, 160–173 (2011).
21. Allen, C. R., Angeler, D. G., Chaffin, B. C., Twidwell, D. & Garmestani, A. Resilience reconciled. *Nat. Sustainability* **2**, 898–900 (2019).
22. Helfgott, A. Operationalising systemic resilience. *Eur. J. Oper. Res.* **268**, 852–864 (2018).
23. Arese Lucini, F., Morone, F., Tomassone, M. S. & Makse, H. A. Diversity increases the stability of ecosystems. *PLoS ONE* **15**, e0228692 (2020).
24. Gao, J., Barzel, B. & Barabási, A.-L. Universal resilience patterns in complex networks. *Nature* **530**, 307–312 (2016).
25. Bai, X., Nagendra, H., Shi, P. & Liu, H. Cities: build networks and share plans to emerge stronger from COVID-19. *Nature* **584**, 517–520 (2020).
26. Garnett, P., Doherty, B. & Heron, T. Vulnerability of the United Kingdom's food supply chains exposed by COVID-19. *Nat. Food* **1**, 315–318 (2020).
27. Rotz, S. & Fraser, E. D. G. Resilience and the industrial food system: analyzing the impacts of agricultural industrialization on food system vulnerability. *J. Environ. Stud. Sci.* **5**, 459–473 (2015).
28. Gomez, M., Garcia, S., Rajtmajer, S., Grady, C. & Mejia, A. Fragility of a multilayer network of intranational supply chains. *Appl. Netw. Sci.* **5**, 71 (2020).
29. Davies, E. G. R. Cities drive food and water security. *Nat. Sustainability* **1**, 120–121 (2018).
30. Hwang, H. et al. *The Freight Analysis Framework Version 4 (FAF4)* ORNL/TM-2016/489 (Oak Ridge National Laboratory, 2016).
31. Countryman, A. M., Paarlberg, P. L. & Lee, J. G. Dynamic effects of drought on the U.S. beef supply chain. *Agric. Resour. Econ. Rev.* **45**, 459–484 (2016).
32. Leister, A. M., Paarlberg, P. L. & Lee, J. G. Dynamic effects of drought on U.S. crop and livestock sectors. *J. Agric. Appl. Econ.* **47**, 261–284 (2015).
33. Rippey, B. R. The U.S. drought of 2012. *Weather Clim. Extrem.* **10**, 57–64 (2015).
34. Schnoor, J. L. The U.S. drought of 2012. *Environ. Sci. Technol.* **46**, 10480 (2012).
35. Mays, L. *Water Resources Engineering* 3rd edn (Wiley, 2019).
36. Bettencourt, L. M. A. The origins of scaling in cities. *Science* **340**, 1438–1441 (2013).
37. Mahjabin, T., Garcia, S., Grady, C. & Mejia, A. Large cities get more for less: water footprint efficiency across the US. *PLoS ONE* **13**, e0202301 (2018).
38. Seekell, D. et al. Resilience in the global food system. *Environ. Res. Lett.* **12**, 025010 (2017).
39. Coleman-Jensen, A., Rabbitt, M. P., Gregory, C. A. & Singh, A. *Household Food Security in the United States in 2015*. https://www.ers.usda.gov/webdocs/publications/79761/err-215.pdf?v=9349.1 (Economic Research Service, U.S. Department of Agriculture, 2016).
40. Gregory, C. A. & Coleman-Jensen, A. Do high food prices increase food insecurity in the United States? *Appl. Econ. Perspect. Policy* **35**, 679–707 (2013).
41. Gilbert, C. L. & Morgan, C. W. Food price volatility. *Phil. Trans. R. Soc. Lond. B* **365**, 3023–3034 (2010).
42. Pimm, S. L., Donohue, I., Montoya, J. M. & Loreau, M. Measuring resilience is essential to understand it. *Nat. Sustainability* **2**, 895–897 (2019).
43. Elmqvist, T. et al. Sustainability and resilience for transformation in the urban century. *Nat. Sustainability* **2**, 267–273 (2019).
44. Davis, K. F., Downs, S. & Gephart, J. A. Towards food supply chain resilience to environmental shocks. *Nat. Food* **2**, 54–65 (2021).
45. United States Department of Agriculture, Food and Nutrition Service. *Supplemental Nutrition Assistance Program (SNAP)* https://www.fns.usda.gov/snap/supplemental-nutrition-assistance-program (USDA).
46. United States Department of Agriculture, Food and Nutrition Service. *National School Lunch Program* https://www.fns.usda.gov/nslp (USDA).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Methods

Dataset of food flow networks

We derive annual, intranational food flow networks for the USA using the Freight Analysis Framework version 4 (FAF4) database[30]. The derived networks are for different food sectors and include all metropolitan areas in the USA. The FAF4 database consists of annual commodity flows during 2012–2015 for 115 geographic areas in the USA and 43 different sectors. We focus on the following four food sectors in the FAF4 database: crops, live animals, animal feed and meat. The 115 geographic areas in the FAF4 database cover the entire contiguous USA, including 69 metropolitan statistical areas and 46 remainders of states (the remainder is the area of a state that is not part of a FAF4 metropolitan area).

To obtain food flows for all metropolitan areas in the USA, we disaggregate the FAF4 database from 115 to 329 areas (Supplementary Fig. 4), out of which 284 are metropolitan or combined statistical areas (120 metropolitan and 164 combined statistical areas). The disaggregation is performed using different socioeconomic 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 remainder of state as follows:

$$U^c_{o'd} = \frac{U^c_{od}}{a_o} \times a_{o'}, \tag{3}$$

where $U^c_{o'd}$ (in tons per year) is the disaggregated supply for food sector c in origin o' that satisfies demand at the FAF4 destination d, U^c_{od} (in tons per year) 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, $U^c_{o'd}$ is further disaggregated into demand E using:

$$E^c_{o'd'} = \frac{U^c_{o'd}}{a_d} \times a_{d'}, \tag{4}$$

where $E^c_{o'd'}$ (in tons per year) 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 food flow data at both the state level (48 states) and metropolitan level (115 areas including 69 metropolitan areas). Prior to performing the disaggregation, we jointly use the FAF4 state data and the FAF4 metropolitan data to select the best performing attractor variables. That is, we first use equations (3) and (4) to disaggregate the FAF4 state-level data to the metropolitan-level for the metropolitan and remainder-of-state areas in FAF4. By comparing the performance of our disaggregated flow data against the empirical FAF4 metropolitan-level data, we select the best attractor variable for each food sector. The following attractor variables are considered: population[47], employment[48], wages[48], number of establishments[48] and cropland area[49]. These variables are selected on the basis of previous analysis and data availability[50].

To assess the performance of the attractor variables, we use the Pearson correlation coefficient between the empirical FAF4 flows and the disaggregated flows for the metropolitan areas and remainder-of-state areas in FAF4 (Extended Data Fig. 3). The performance is high with correlation values greater than 0.87 and an average of 0.95. Using the best-performing disaggregation variables, we build the food flow networks employed in this study. The nodes in the networks represent metropolitan and remainder-of-state areas, and the weighted links represent annual food flows during 2012–2015 for crops, live animals, feed and meat (Supplementary Fig. 5).

The FAF4 metropolitan and remainder-of-state areas we used to select the attractor variables span a wide range of populations, cropland areas, and number of establishments, since these FAF4 areas include the largest cities in the USA and a broad range of medium-size cities. The values of the attractor variables used in the disaggregation are within the ranges implied by the FAF4 metropolitan data (Supplementary Fig. 6), indicating that the variables are reliable. The exception to this is population, which is only used to disaggregate meat demand. Population, however, has a high disaggregation performance with a correlation coefficient of 0.97 (Extended Data Fig. 3). In addition, the use of population to disaggregate meat demand is consistent with previous scaling results for metropolitan areas in the USA[51] that have shown that metropolitan-level variables that are related to resource consumption scale approximately linearly with population.

Food inflows supply chain diversity

To determine annual supply chain diversity, we extract the annual food buyer–supplier subgraph of each city and food sector from the food flow networks[19]. We refer to each of these subgraphs as a food system. The food buyer–supplier subgraph of a city i consists of all the supply chain interactions with its trading partners or neighbours j for a specific food sector. Our measure of supply chain diversity is based on the notion of functional distance[52]. 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. The indicators are described below in the 'Functional distance indicators' section of the Methods. We also perform statistical analyses to evaluate the influence of the attractor variables on these indicators (Methods). The indicators represent stable characteristics of cities and therefore tend to remain fairly constant during our study period.

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

$$d^r_{ij} = N^{-1}|r_k - r_i|, \tag{5}$$

where the normalization constant N is determined as the maximum value of $|r_k - r_i|$ between any node k in the network and i. In equation (5), $d^r_{ij} = 0$ for functionally similar nodes and $d^r_{ij} = 1$ for dissimilar nodes.

For each city's buyer–supplier subgraph and food sector, any pair of connected nodes has 5 different functional distance indicators associated with it. To combine these distance indicators into a single measure, we calculate the average functional distance indicator $\langle d^r_{ij} \rangle$ as the arithmetic average of the 5 functional distance indicators for any pair (i,j) of connected nodes. We use the discrete probability distribution of food inflows binned by $\langle d^r_{ij} \rangle$ categories, together with Shannon entropy[53], to calculate the supply chain diversity $D^t_{i,c}$ of node i and sector c at year t:

$$D^t_{i,c} = \frac{-\sum_{k=1}^{K} Y^t_{i,c}(k) \ln Y^t_{i,c}(k)}{\log K}. \tag{6}$$

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

$D^t_{i,c}$ is sensitive to the total number of bins K. Thus, for each node in our food flow networks, we tested the sensitivity of $D^t_{i,c}$ to the total number of bins K. For $K = 15$, D values stabilize (Supplementary Fig. 7); therefore, we used 15 bins when performing all calculations of functional diversity.

Functional distance indicators

The average functional distance between a city and its trading partners is based on the following five indicators:

(1) Physical distance indicator (PDI). The PDI is obtained by calculating the Euclidean distance from the centroid of each geographic area to the centroid of all other areas. The geometric centroids of all geographical areas are calculated using the GIS software ArcMap (https://desktop.arcgis.com/en/arcmap/).

(2) Climate indicator (CI). To account for different climates in cities across the USA, the Palmer Drought Severity Index (PDSI) is used[54]. The monthly PDSI is obtained from the National Oceanic and Atmospheric Administration for the years 1895–2015 at the climate division geographic level. An area-weighted average is performed to aggregate the PDSI data to the metropolitan level. The CI is obtained by calculating the monthly correlation between an area and all other areas.

(3) Urban classification indicator (UCI). To identify the urbanization level of a geographical area, the Urban-Rural Classification indicator of the National Center for Health Statistics is employed[55]. This indicator classifies counties using a scale from 1 to 6, where a value of 1 indicates the county is highly rural and a value of 6 means highly urban. The UCI is obtained at the metropolitan level using an area-weighted average of the county-level values.

(4) Network modularity indicator (NMI). This indicator identifies geographical areas (network nodes) that belong to the same community. A community is a group of nodes whose strength interactions are stronger than with the rest of the network. To identify the network's communities, we aggregate the flows from the four food sectors (crops, live animals, feed and meat) into a single-layer network. The communities are identified by maximizing the modularity measure of Newman[56,57] using the greedy optimization algorithm of Blondel et al.[58,59]. Network nodes that lie in the same community are assigned a NMI of 1 and 0 otherwise.

(5) Economic specialization indicator (ESI). Each geographical area is assigned a score based on its dominant economic supply sector. Supply is quantified using the FAF4 intranational commodity flows[30]. Areas with a dominant meat sector are assigned an ESI of 1, crops an ESI of 2, fruits and vegetables an ESI of 3, animal feed an ESI of 4, live animals an ESI of 5, milled grains an ESI of 6, and industrial products an ESI of 7.

Probabilities of food supply chain shock

The annual probability of food supply chain shock is calculated as the probability that food inflows to a city fall below a percentage of the average inflows for that city during 2012–2015[8]. To compute this probability, we group all nodes from the 4 food flow networks (1,221 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 chain shock S_i^c as

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

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 $s \in \{3, 4, 5, \ldots, 15\}$, 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:

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

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 chain shock,

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

Furthermore, we adapt equations (8) and (9) to calculate the probability of a food supply chain 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 fashion similar to that described above. We bin the network's nodes into 6 groups from lowest to highest diversity and determine the percentage of food supply chain shock with equation (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 chain shock S being greater than the shock intensity s in bin b is now

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

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

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

Statistical analyses

We use the disaggregated food flow data to calculate both the probability of food supply chain shock and supply chain diversity. Therefore, we perform two complementary analyses to test whether the attractor variables used in the disaggregation are causing a circularity issue in the empirical relationship between the probability of food supply chain shock and supply chain diversity. For the first analysis, we determine the Pearson correlation between the functional distance indicators (PDI, CI, UCI, NMI and ESI) and the attractor variables (Supplementary Fig. 8). We find that the attractor variables are weakly correlated with the functional distance indicators (Supplementary Table 1). For the second analysis, we determine the Pearson correlation between the food supply chain shock intensities and attractor variables for the 4 food sectors (Supplementary Fig. 9). The attractor variables are also weakly correlated with the food supply chain shock intensities (Supplementary Table 2). Thus, circularity is not unduly influencing the empirical relationship between the probability of shock and supply chain diversity.

We also evaluate whether the empirical relationship between the probability of food supply chain shock and supply chain diversity is driven by the disaggregation of the original FAF4 data. For this, we recalculate the probability of shock and supply chain diversity using the FAF4 data. For the FAF4 data, the probability of shock also declines with rising supply chain diversity (Supplementary Fig. 10), similar to the reduction observed using the disaggregated food flow data (Fig. 1a), suggesting that the latter data are not driving the relationship.

Furthermore, we test whether the relationship between the probability of food supply chain shock and supply chain diversity holds for different demand levels. To control for demand, we stratify all the data into low, medium and high demand levels using population or food inflows as proxies for demand. For both stratifications, the bounds are chosen so that each level has approximately the same number of data points. Using the stratified data, we recalculate the Pearson correlation between the probability of shock at 3%, 5%, 10% and 15% shock intensities and supply chain diversity for each level of population or food inflows (Supplementary Table 3). We find that the relationship between the probability of food supply chain shock and supply chain diversity holds for these different demand levels (Supplementary Table 3). Using the same exponential model in equation (1) to fit the relationship for the

stratified data (Supplementary Fig. 11), we determine the exponential model parameters k_s and $D_{0,s}$ for each demand level (Extended Data Table 6). These parameters fall within the 95% confidence interval of the parameters of the exponential model in the main text (Extended Data Table 1), indicating that the model is robust.

We perform two different sensitivity analyses to assess the influence of the five distance indicators on the empirical relationship between the probability of food supply chain shock and supply chain diversity. The first analysis compares single-indicator diversity measures against the multi-indicator diversity measure calculated using all 5 indicators. Five different single-indicator diversity measures are compared, one measure for each of the 5 indicators: PDI, CI, UCI, NMI and ESI. For the second sensitivity analysis, we leave out one indicator at a time to calculate diversity using the 4 remaining indicators, which results in another 5 different diversity measures. The diversity measures for the sensitivity analyses are all calculated following the approach in the 'Food inflows supply chain diversity' section of the Methods. To perform the sensitivity analyses, we plot the empirical relationship between the probability of food supply chain shock and each diversity measure (Supplementary Figs. 12 and 13), and calculate the Pearson correlation of the data (Extended Data Table 7). The correlation coefficients are used to quantify the influence of the distance indicators on the relationship between the probability of food supply chain shock and supply chain diversity (Extended Data Table 7). The probabilities of food supply chain shock are calculated following the approach in the 'Probabilities of food supply chain shock' section of the Methods. We find that the five indicators have a varied influence on the relationship between the probability of food supply chain shock and supply chain diversity (Extended Data Table 7). The inclusion of all 5 indicators, however, in the supply chain diversity measure increases the Pearson correlation between the probability of food supply chain shock and supply chain diversity (Extended Data Table 7).

Data availability

The commodity flows data are from the Freight Analysis Framework version 4 dataset, prepared by the Federal Highway Administration of the United States Department of Transportation (https://ops.fhwa.dot.gov/freight/freight_analysis/faf/). The socioeconomic and agricultural data used are all publicly available at https://www.census.gov/programs-surveys/popest.html (population), https://www.bls.gov/cew/downloadable-data-files.htm (employment and wages), and https://www.fsa.usda.gov/news-room/efoia/electronic-reading-room/frequently-requested-information/crop-acreage-data/index (cropland area). Additional data used for the functional distance indicators are available at https://www.ncdc.noaa.gov/temp-and-precip/drought/historical-palmers/ (Palmer drought severity index) and https://www.cdc.gov/nchs/data_access/urban_rural.htm (urban–rural classification). Calculated metropolitan level food flows and functional diversity are published in the HydroShare data repository at https://doi.org/10.4211/hs.caebbbb68cce49f2bbf351f28d865794. The base map of the USA is from the United States Census Bureau TIGER/Line database (https://www.census.gov/cgi-bin/geo/shapefiles/index.php).

Code availability

Requests for code related to the analyses performed can be directed to M.G. (michael_gomezs@hotmail.com).

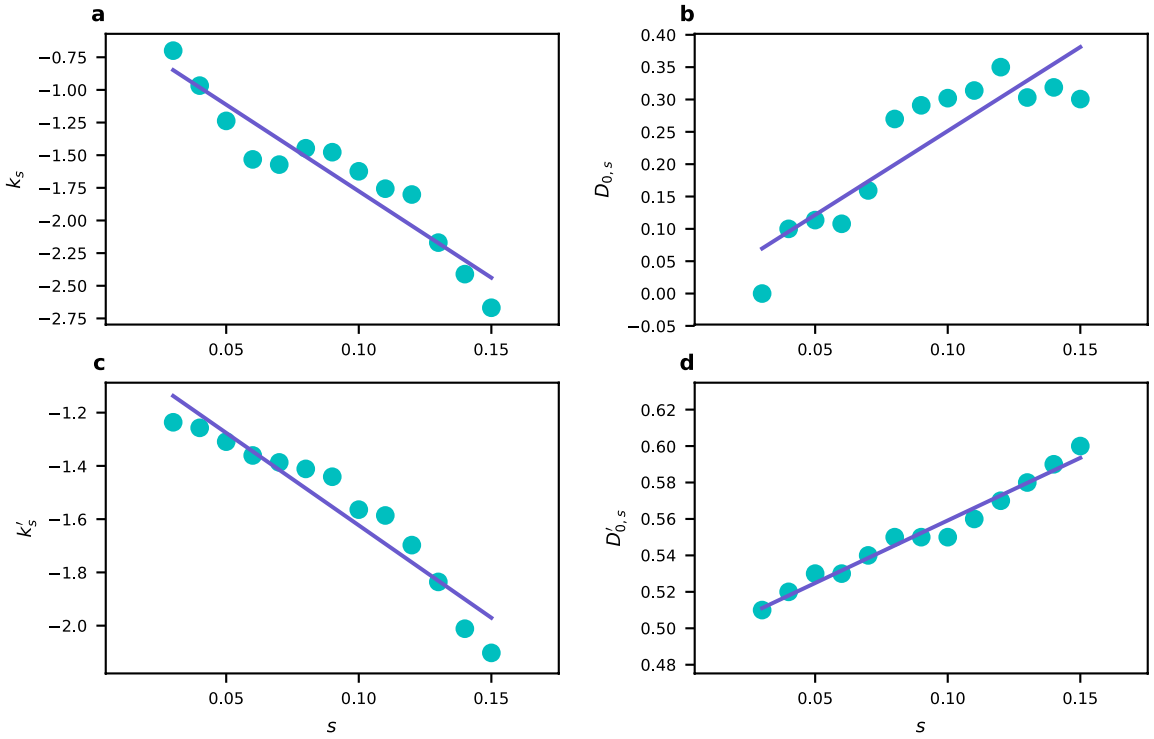
47. *Annual Resident Population Estimates 2010–2016.* https://www.census.gov/programs-surveys/popest.html (United States Census Bureau, Population Division, 2017).
48. *Quarterly Census Of Employment And Wages 2012.* https://www.bls.gov/cew/downloadable-data-files.htm (United States Bureau of Labor Statistics, 2015).
49. *Crop Acreage Data Reported To The Farm Service Agency 2012.* https://www.fsa.usda.gov/news-room/efoia/electronic-reading-room/frequently-requested-information/crop-acreage-data/index (United States Department of Agriculture, Farm Service Agency, 2013).
50. 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).
51. Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C. & West, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl Acad. Sci. USA* **104**, 7301–7306 (2007).
52. Walker, B., Kinzig, A. & Langridge, J. Plant attribute diversity, resilience, and ecosystem function: the nature and significance of dominant and minor species. *Ecosystems* **2**, 95–113 (1999).
53. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
54. United States National Oceanic and Atmospheric Administration *Historical Palmer Drought Indices.* https://www.ncdc.noaa.gov/temp-and-precip/drought/historical-palmers/ (NOAA).
55. Ingram, D. & Franco, S. *NCHS Urban–Rural Classification Scheme for Counties*. https://www.cdc.gov/nchs/data_access/urban_rural.htm (National Center for Health Statistics, 2013).
56. Newman, M. E. J. Modularity and community structure in networks. *Proc. Natl Acad. Sci. USA* **103**, 8577 (2006).
57. Girvan, M. & Newman, M. E. J. Community structure in social and biological networks. *Proc. Natl Acad. Sci. USA* **99**, 7821–7826 (2002).
58. Blondel, V. D., Guillaume, J. P., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008).
59. Garcia, S. & Mejia, A. Characterizing and modeling subnational virtual water networks of US agricultural and industrial commodity flows. *Adv. Water Resour.* **130**, 314–324 (2019).

Acknowledgements We thank S. Spencer for advising on the form of exponent to use in the model. Computations for this research were performed on the Pennsylvania State University's Institute for Computational and Data Sciences' Roar supercomputer. 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.

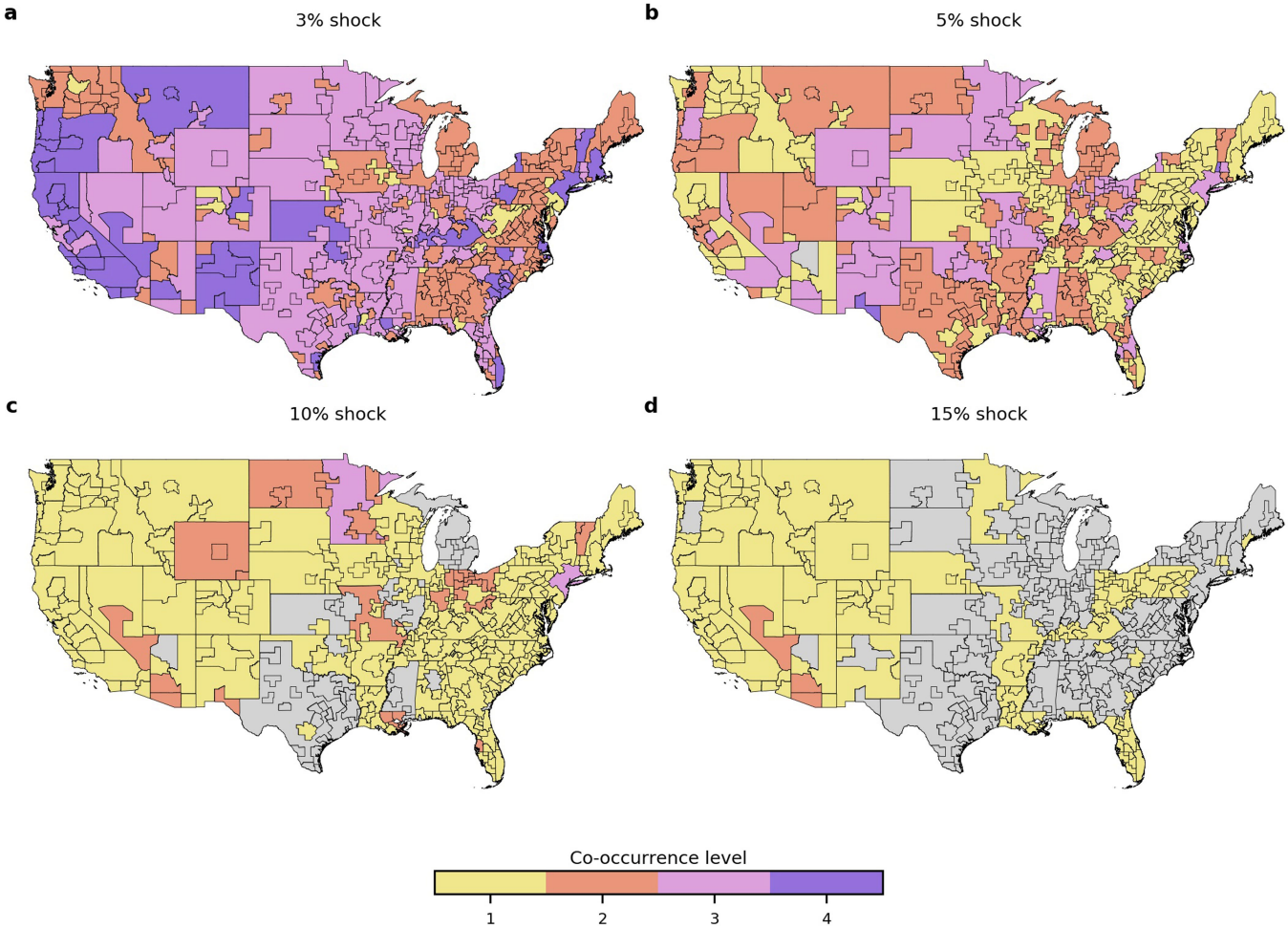
Author contributions M.G. carried out the data analysis and generated the results with input from A.M. and B.L.R. All authors devised the research, contributed to the analysis of results, and wrote the manuscript.

Competing interests The authors declare no competing interests.

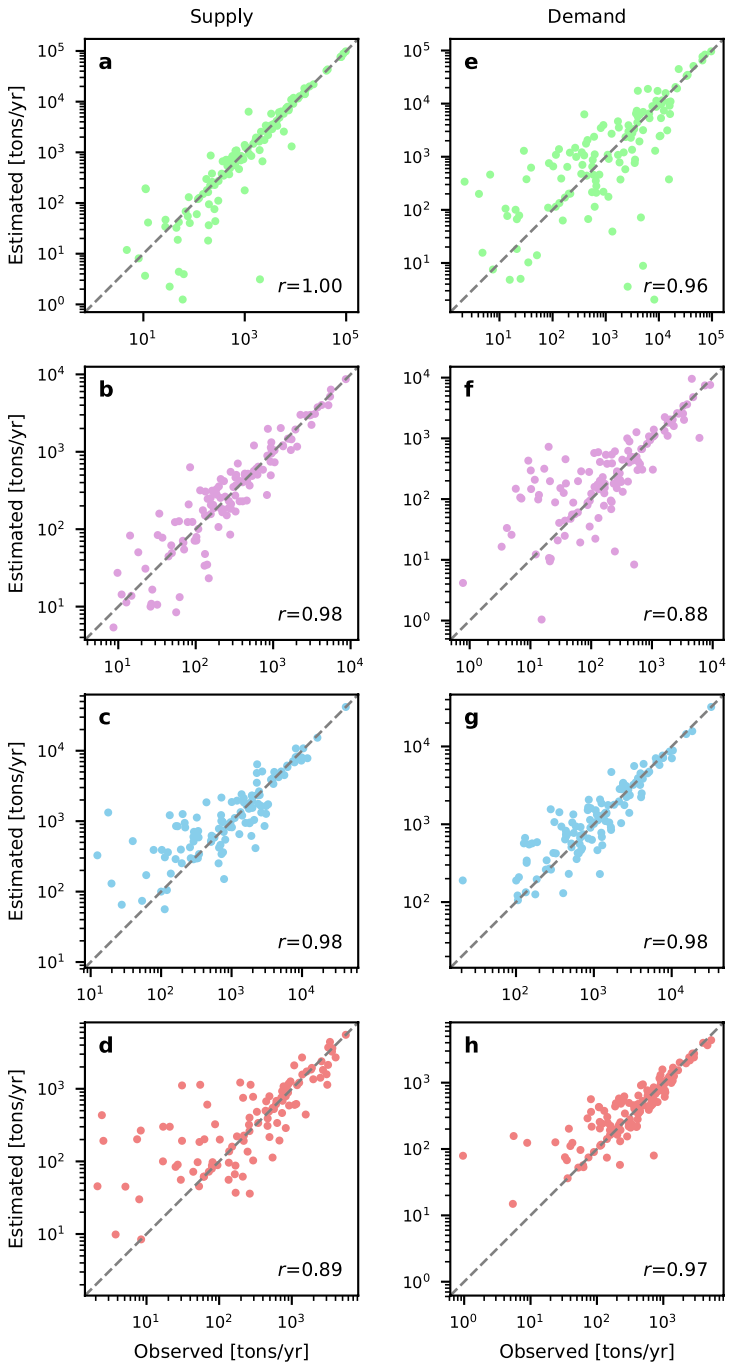
Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03621-0.
Correspondence and requests for materials should be addressed to A.M.
Peer review information *Nature* thanks Zia Mehrabi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.
Reprints and permissions information is available at http://www.nature.com/reprints.



Extended Data Fig. 1 | Shock intensity versus exponential regression parameters. Relationship between the supply chain shock intensity s and the exponential regression parameters k_s and $D_{0,s}$ for single (**a**, **b**) and k'_s and $D'_{0,s}$ for co-occurring shocks (**c**, **d**). The lines are linear regression fits. The goodness-of-fit results for single and co-occurring shocks are summarized in Extended Data Tables 1 and 4, respectively.



Extended Data Fig. 2 | Level of co-occurrence of food supply chain shocks across the USA. For a given shock intensity (3%, 5%, 10% and 15%), the co-occurrence level is measured as the number of food supply chain shocks from different food sectors to a geographic location during 2012–2015, out of a possible maximum of 4 shocks. The maps show the level of co-occurrence for different shock intensities. The overall co-occurrence level across the USA tends to decline as the threshold increases, meaning that co-occurrence becomes less likely for higher shock intensities.



Extended Data Fig. 3 | Performance of food flow disaggregation approach. Comparison between observed and disaggregated food supply (**a**–**d**) and demand (**e**–**h**) flows for crops (**a**, **e**), live animals (**b**, **f**), feed (**c**, **g**), and meat (**d**, **h**) based on the FAF4 metropolitan and remainder-of-state areas. Different socioeconomic and agricultural variables (Methods) were tested and those that resulted in the highest values of the Pearson correlation coefficient *r* were used in the disaggregation. For the crop sector, disaggregated flows using cropland area as an attractor yielded the highest *r* between observed and disaggregated flows for both supply and demand. For the live animals and feed sector, the flows disaggregated by number of establishments resulted in the highest *r* for both supply and demand. For the meat sector, the number of establishments resulted in the highest *r* value for supply and population for demand.

Extended Data Table 1 | Goodness-of-fit results for the exponential regression between the probability of food supply chain shock and supply chain diversity for shock intensities *s* ranging from 3 to 15%

s	k_s	$D_{0,s}$	R^2	Adjusted R^2	RMSE	SSE
3%	-0.6994 (-0.9242, -0.4746)	0 (fixed at bound)	0.7295	0.7295	0.08018	0.03214
4%	-0.9672 (-1.17, -0.7649)	0.1000 (fixed at bound)	0.8432	0.8432	0.06499	0.02112
5%	-1.237 (-1.916, -0.559)	0.1135 (-0.2164, 0.4434)	0.8691	0.8691	0.05944	0.01413
6%	-1.532 (-2.538, -0.5262)	0.1078 (-0.2669, 0.4826)	0.8283	0.7854	0.07468	0.02231
7%	-1.572 (-2.599, -0.5456)	0.1594 (-0.2425, 0.5612)	0.831	0.7887	0.06863	0.01884
8%	-1.447 (-2.585, -0.3087)	0.2699 (-0.3055, 0.8453)	0.7741	0.7176	0.07027	0.01975
9%	-1.477 (-2.498, -0.456)	0.2909 (-0.2268, 0.8087)	0.8141	0.7677	0.05974	0.01428
10%	-1.623 (-2.487, -0.7595)	0.3021 (-0.09523, 0.6994)	0.8798	0.8497	0.04433	0.007861
11%	-1.756 (-2.427, -1.085)	0.3139 (0.02878, 0.5991)	0.9329	0.9162	0.03037	0.00369
12%	-1.801 (-2.755, -0.847)	0.35 (-0.06165, 0.7616)	0.9032	0.8791	0.03904	0.006096
13%	-2.169 (-3.141, -1.197)	0.3032 (-0.01076, 0.6171)	0.9153	0.8941	0.03227	0.004166
14%	-2.411 (-3.558, -1.263)	0.3187 (-0.01263, 0.6501)	0.9103	0.8878	0.03023	0.003655
15%	-2.668 (-4.132, -1.203)	0.3008 (-0.06109, 0.6628)	0.8909	0.8636	0.03274	0.004287

The regression has the form $P_s = e^{k_s(D+D_{0,s})}$ and is bounded at $P(S > s) \in [0, 1]$ and $D \in [0, 1]$. The parameters k_s and $D_{0,s}$ are shown with the 95% confidence intervals in parentheses (degrees of freedom = 4). RMSE, root-mean-square error; SSE, sum of squared errors.

Extended Data Table 2 | Goodness-of-fit results for the linear regression between the parameters k_s and $D_{0,s}$ and the supply chain shock intensity s

Regression Parameters	k_s	$D_{0,s}$
Intercept	-0.451	-0.008
Slope	-13.249	2.595
p-value	<0.001	<0.001
R^2	0.900	0.800
95% confidence interval	-16.19, -10.31	1.74, 3.46

Extended Data Table 3 | Probability of shock equations for different shock intensities *s*

s [%]	Equation	*P*(*D* = 0)	*P*(*D* = 1)
1%	$P = e^{-0.583(D+0.018)}$	0.990	0.552
2%	$P = e^{-0.716(D+0.044)}$	0.969	0.474
3%	$P = e^{-0.848(D+0.070)}$	0.942	0.403
4%	$P = e^{-0.981(D+0.096)}$	0.910	0.341
5%	$P = e^{-1.113(D+0.122)}$	0.873	0.287
6%	$P = e^{-1.246(D+0.148)}$	0.832	0.239
7%	$P = e^{-1.378(D+0.174)}$	0.787	0.198
8%	$P = e^{-1.511(D+0.200)}$	0.740	0.163
9%	$P = e^{-1.643(D+0.226)}$	0.690	0.133
10%	$P = e^{-1.776(D+0.252)}$	0.640	0.108
11%	$P = e^{-1.908(D+0.277)}$	0.589	0.087
12%	$P = e^{-2.041(D+0.303)}$	0.538	0.070
13%	$P = e^{-2.173(D+0.329)}$	0.489	0.056
14%	$P = e^{-2.306(D+0.355)}$	0.441	0.044
15%	$P = e^{-2.438(D+0.381)}$	0.395	0.034
16%	$P = e^{-2.571(D+0.407)}$	0.351	0.027
17%	$P = e^{-2.703(D+0.433)}$	0.310	0.021
18%	$P = e^{-2.836(D+0.459)}$	0.272	0.016
19%	$P = e^{-2.968(D+0.485)}$	0.237	0.012
20%	$P = e^{-3.101(D+0.511)}$	0.205	0.009
21%	$P = e^{-3.233(D+0.537)}$	0.176	0.007
22%	$P = e^{-3.366(D+0.563)}$	0.150	0.005
23%	$P = e^{-3.498(D+0.589)}$	0.127	0.004
24%	$P = e^{-3.631(D+0.615)}$	0.107	0.003
25%	$P = e^{-3.763(D+0.641)}$	0.090	0.002
26%	$P = e^{-3.896(D+0.667)}$	0.074	0.002
27%	$P = e^{-4.028(D+0.693)}$	0.061	0.001
28%	$P = e^{-4.161(D+0.719)}$	0.050	0.001
29%	$P = e^{-4.293(D+0.745)}$	0.041	0.001
30%	$P = e^{-4.426(D+0.771)}$	0.033	0.000

The minimum, *P*(*D* = 1), and maximum, *P*(*D* = 0), probability values are also shown.

Extended Data Table 4 | Goodness of fit results for the exponential regression between the probability of co-occurring food supply chain shocks and supply chain diversity

s	k'_s	$D'_{0,s}$	R^2	Adjusted R^2	RMSE	SSE
3%	-1.2362 (-1.899, -0.5731)	0.5100 (-0.01714, 1.037)	0.7500	0.6875	0.0356	0.0051
4%	-1.2571 (-1.759, -0.7554)	0.5200 (0.1246, 0.9154)	0.8915	0.8644	0.0261	0.0027
5%	-1.3090 (-1.676, -0.9424)	0.5300 (0.2511, 0.8089)	0.9538	0.9423	0.0179	0.0013
6%	-1.3608 (-1.422, -1.299)	0.5300 (0.2778, 0.7822)	0.9710	0.9710	0.0143	0.0010
7%	-1.3869 (-1.458, -1.315)	0.5400 (0.2484, 0.8316)	0.9648	0.9468	0.0161	0.0013
8%	-1.4113 (-1.485, -1.338)	0.5500 (0.2498, 0.8502)	0.9631	0.9631	0.0161	0.0013
9%	-1.4407 (-1.82, -1.061)	0.5500 (0.2857, 0.8143)	0.9710	0.9637	0.0157	0.0010
10%	-1.5643 (-2.448, -0.6806)	0.5500 (-0.01055, 1.111)	0.8882	0.8602	0.0322	0.0042
11%	-1.5862 (-2.763, -0.4092)	0.5600 (-0.1821, 1.302)	0.8335	0.7918	0.0413	0.0068
12%	-1.6975 (-2.989, -0.4064)	0.5700 (-0.1906, 1.331)	0.8523	0.8153	0.0397	0.0063
13%	-1.8357 (-3.423, -0.2483)	0.5800 (-0.2827, 1.443)	0.8251	0.7813	0.0414	0.0068
14%	-2.0114 (-3.71, -0.3126)	0.5900 (-0.2481, 1.428)	0.8412	0.8015	0.0359	0.0052
15%	-2.1022 (-2.467, -1.738)	0.6000 (-0.4348, 1.635)	0.8071	0.8071	0.0367	0.0067

The regression has the form $P'_s = e^{k'_s(D+D'_{0,s})}$ bounded at $P'_s(S > s) \in [0, 1]$ and $D \in [0, 1]$. The parameters and $D'_{0,s}$ are shown with the 95% confidence intervals in parentheses (degrees of freedom = 4).

Extended Data Table 5 | Goodness-of-fit results for the linear regression between the parameters k'_s and $D'_{0,s}$ and the supply chain shock intensity s

Regression Parameters	k'_s	$D'_{0,s}$
Intercept	-0.929	0.490
Slope	-6.943	0.687
p-value	<0.001	<0.001
R2	0.915	0.972
95% confidence Interval	-8.34, -5.54	0.61, 0.76

Extended Data Table 6 | Exponential fit parameters k_s and $D_{0,s}$ for the relationship between the probability of food supply chain shock and supply chain diversity at different demand levels and shock intensities

Shock intensity *s*	Low		Mid		High	
	k_s	$D_{0,s}$	k_s	$D_{0,s}$	k_s	$D_{0,s}$
Population						
3%	-0.81	0	-0.75	0	-0.58	0
5%	-1.5	0.02	-1.26	0.17	-1.3	0
10%	-2.19	0.08	-1.35	0.52	-1.57	0.3
15%	-3.03	0.19	-2.62	0.35	-2.11	0.49
Food inflows						
3%	-0.75	0	-0.72	0	-0.65	0
5%	-1.63	0.03	-1.43	0	-1.13	0.14
10%	-1.92	0.18	-2.59	0.02	-0.98	0.82
15%	-2.76	0.18	-2.95	0.31	-1.72	0.8

We stratify all data into low, medium and high demand levels using population or food inflows as proxy for demand. For each demand stratum, we recalculate the probability of food supply chain shock for different shock intensities: 3%, 5%, 10% and 15%.

Extended Data Table 7 | Sensitivity of the supply chain diversity results to different configurations of the basket of indicators

	s = 3%			*s* = 5%			*s* = 10%			*s* = 15%		
	r	**p-value**	***δ***	***r***	**p-value**	***δ***	***r***	**p-value**	***δ***	***r***	**p-value**	***δ***
Average of indicators												
d	-0.900	0.014	-	-0.969	0.001	-	-0.982	<0.001	-	-0.992	<0.001	-
Single indicator												
CI	-0.847	0.034	-0.05	-0.941	0.005	-0.03	-0.968	0.002	-0.01	-0.994	<0.001	0.00
PDI	-0.905	0.013	0.00	-0.871	0.024	-0.10	-0.858	0.029	-0.12	-0.933	0.007	-0.06
UCI	-0.317	0.540	-0.58	-0.877	0.022	-0.09	-0.740	0.092	-0.24	-0.730	0.100	-0.26
ESI	-0.112	0.833	-0.79	0.105	0.843	-1.07	0.257	0.623	-1.24	-0.101	0.848	-0.89
NMI	-0.796	0.058	-0.10	-0.869	0.025	-0.10	-0.846	0.034	-0.14	-0.788	0.063	-0.20
One indicator removed												
without CI	-0.930	0.007	0.03	-0.964	0.002	0.00	-0.954	0.003	-0.03	-0.969	0.001	-0.02
without PDI	-0.835	0.038	-0.06	-0.948	0.004	-0.02	-0.977	<0.001	0.00	-0.996	<0.001	0.00
without UCI	-0.887	0.018	-0.01	-0.908	0.012	-0.06	-0.881	0.020	-0.10	-0.957	0.003	-0.03
without ESI	-0.887	0.018	-0.01	-0.948	0.004	-0.02	-0.985	<0.001	0.00	-0.972	0.001	-0.02
without NMI	-0.851	0.032	-0.05	-0.951	0.004	-0.02	-0.972	0.001	-0.01	-0.994	<0.001	0.00

The 5 indicators are: climate indicator (CI), physical distance indicator (PDI), urban classification indicator (UCI), economic specialization indicator (ESI), and network modularity indicator (NMI). We performed two complementary sensitivity analyses to evaluate how supply chain diversity relates to the probability of food supply chain shock for different configurations of the basket of indicators (Methods). For each configuration of the basket, the Pearson correlation coefficient *r* for the relationship between the probability of food supply chain shock and diversity was calculated. The difference between correlation values is denoted by δ.