Supplemental methods

For each CBSA and season combination, we computed a family of 11 correlation matrices where the time series of this CBSA lagged 0, ± 1, ± 2, ± 3, ± 4, and ± 5 weeks relative to the time series of the other CBSAs. We limited the analysis to ± 5 weeks for clarity and because we considered longer than a month one category. We computed the first singular value (SV) for each of the 11 correlation matrices in this family. Among these 11 SVs, we chose the time lag corresponding to the largest SV as a measure of relative timing of that CBSA relative to the United States for that season. This time lag also corresponded to the strongest collective correlation between CBSAs compared to the other time lags considered.

Next, we provide the definitions neccesary for this analysis. Let $\{1, ..., N\}$ index the CBSAs included for analysis, and let $n \in \{1, ..., N\}$. Let $\{1, ..., T\}$ index the weeks of the 2010–2011 through 2015–2016 seasons. Denote the remainder of the seasonal-trend decomposition using LOESS of the weekly case counts as $X = \{X_{i,t} \mid i = 1,...,N \text{ and } t = 1,...,T\}$ (Cleveland et al., 1990). Let $l \in \{-5,...,5\}$ be a time lag in weeks. Let $s \in$ {2010–2011,...,2015–2016} be an influenza season. Let $S \subset \{1,...,T\}$ be those weeks in season s, and let $(S + l) = \{t + l \mid t \in S\}$. Define $\langle X_{i,S} \rangle$ as the average of $X_{i,t}$ over $t \in S$ for a fixed $i \in \{1,...,N\}$. Define $\langle X_{i,S+l} \rangle$ as the average of $X_{i,t}$ over $t \in S \cap (S + l)$ with a fixed $i \in \{1,...,N\}$, where the intersection limits the data to season s. Define $\langle X_{i,S+l}\, X_{j,S} \rangle$ as the average of $(X_{i,t+l}\ X_{j,t})$ over $t \in (S - l) \cap S$, with fixed $i, j \in \{1,...,N\}$ such that $i \neq j$. Define $\langle X_{i,S}\ X_{j,S+l} \rangle$ similarly. Denote the standard deviation of $X_{i,t}$ over $t \in S$ as $\sigma_{i,s}$ with a fixed $i \in \{1,...,N\}$. Because we assume the time series is stationary, we use this value of the standard deviation for all values l. Define $\mathscr{M}_{n,s,l} = [m_{i,j}]$ as the $N \times N$ correlation matrix where the n-th component of X during season s lags by l weeks:

- if $i = j$, then $m_{i,j} = 1$,
- if $i = n$ and $i \neq j$, then $m_{i,j} = (\langle X_{i,S+l} X_{j,S}\rangle - \langle X_{i,S+l}\rangle\langle X_{j,S}\rangle)/\sigma_{i,s}\sigma_{j,s}$,
- if $j = n$ and $i \neq j$, then $m_{i,j} = (\langle X_{i,S} X_{j,S+l}\rangle - \langle X_{i,S}\rangle\langle X_{j,S+l}\rangle)/\sigma_{i,s}\sigma_{j,s}$, and
- if $i \neq n$ and $j \neq n$, then $m_{i,j} = (\langle X_{i,S}\ X_{j,S}\rangle - \langle X_{i,S}\rangle\langle X_{j,S}\rangle)/\sigma_{i,s}\sigma_{j,s}$.

Write the singular values of $\mathscr{M}_{n,s,l}$ in decreasing order: $\lambda_1(\mathscr{M}_{n,s,l}) \geq \lambda_2(\mathscr{M}_{n,s,l}) \geq \ldots \geq \lambda_N(\mathscr{M}_{n,s,l})$. Then, the spectral norm of $\mathscr{M}_{n,s,l}$ is $|||\mathscr{M}_{n,s,l}|||_2 = \lambda_1(\mathscr{M}_{n,s,l})$, the matrix norm induced by the usual l_2-norm (Horn and Johnson, 2013). Take $\tau \in$ {-5, ..., 5} such that $|||\mathscr{M}_{n,s,\tau}|||_2 \geq |||\mathscr{M}_{n,s,l}|||_2$ for all $l \in$ {-5, ..., 5}. Finally, define τ as the time lag corresponding to the n-th CBSA during season s. In summary, the pairwise correlations of the N CBSAs during season s are collectively larger with respect to the spectral norm when the n-th component lags by τ weeks compared to when the n-th component lags by any other time between -5 weeks and 5 weeks. We use this time lag as the estimate of "how many weeks influenza activity in this CBSA precedes or follows influenza activity in the other CBSAs?"

The mathematics underpinning our use of the the spectral norm for investigating collective phenomena are straightforward. Let M be a correlation matrix of order n. By definition, $M = X^T X$ for some matrix X, so M is positive semidefinite; and, every eigenvalue of M is non-negative. As M is real symmetric, we may find a real orthogonal matrix Q such that $M = QDQ^T$, where D is a diagonal matrix such that the diagonal elements of D are in non-increasing order. As QDQ^T is also a singular value decomposition of M, the eigenvalues of M coincide with the singular values of M. So, the first singular value of M, the spectral radius of M, and the spectral norm of M are all identical. As the diagonal of M is all ones, the sum of the eigenvalues of M is n. If the columns of X are perfectly uncorrelated, then the off diagonal elements of M are all zeros, and every eigenvalue of M is one. In this case, the largest singular value of M is one. By the pigeonhole principle, one is the greatest lower bound on the first singular value; otherwise, the trace of M is strictly less than n. If the columns of M are perfectly correlated, then the off diagonal elemests of M are all ones, the rank of M is one, and the row sums of M are all equal to n. Further, the column vector of all ones is an eigenvector of M corresponding to the eigenvalue n. In this case, the largest singular value of M is n. By Geršgorin's Theorem, n is the least upper bound on the first singular value (Horn and Johnson, 2013). To summarize, the spectral norm ranges from one to n, representing the range of collective behaviour.

References for supplemental methods

Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I., 1990. STL: a seasonal-trend decomposition procedure based on loess. J. Off. Stat. 6, 3-33.

Horn, R.A., Johnson, C.R., 2013. Matrix Analysis, 2 ed. Cambridge University Press, New York, NY.