Supporting Material (SM): Contact, travel, & transmission

S1 Medical claims data

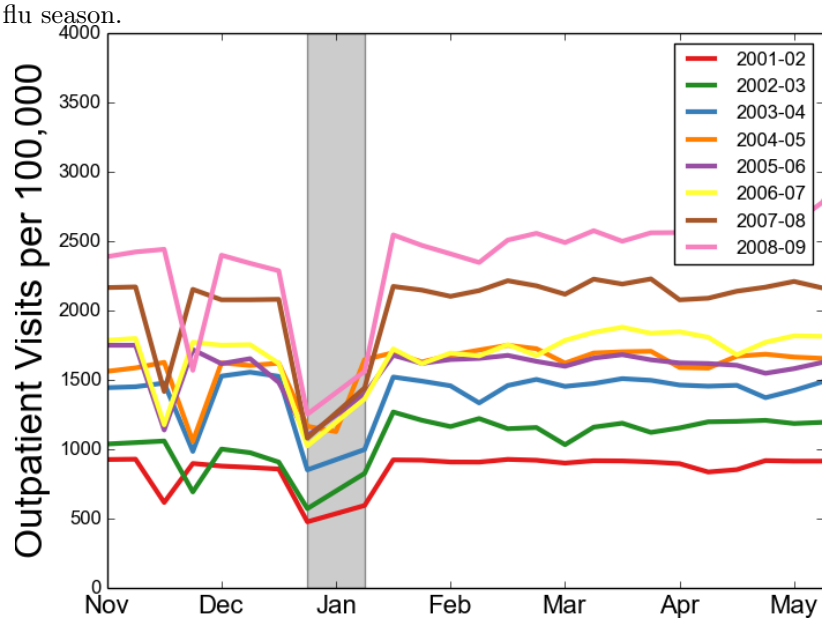
S1.1 Definition of ILI incidence ratio

We use the ILI incidence ratio as defined in [25]. Parameters referenced in the equation are described in Table S1.

$$r_{w,s} = d_{w,s}/v_{w,s} \times (p_s/100,000) \tag{1}$$

The incorporation of visits into this metric helps to account for artificialities in the medical claims data related to to physician office closures and changes to care-seeking behavior during the holidays, and increasing database coverage over time (Figure S1).

Figure S1: **Total visits to physicians reported by the medical claims dataset drops during holiday periods.** The grey bar highlights the typical school holiday for winter break, and a short dip (not highlighted) in late November demonstrates a similar pattern during the Thanksgiving holiday. Coverage in the medical claims dataset increases over time, as witnessed by the rising visit rates in each flu season.



S1.2 Estimation of the effective reproductive number

To account for changes in health care seeking behavior during holiday periods (Figure S1) and changes to medical claims coverage over time, we adjusted the raw ILI medical claims data for input into the EpiEstim program for estimation of the effective reproductive number R_t. Parameters referenced in the equation are described in Table S1.

$$d^*_{t,s} = (d_{w,s}/C) \times (v_{w_s}/v_{w,s}) \times (1/\phi) \tag{2}$$

Table S1: **Parameters in adjustments to ILI medical claims data** The same notation was used to describe the ILI incidence ratio and the adjustments for the effective reproductive number analysis.

Common terms	
$d_{w,s}$	raw ILI cases at week w in season s
$v_{w,s}$	total visits to health care facilities for any diagnosis at week w in season s
Specific to equation 1	
$r_{w,s}$	ILI incidence ratio (IR) at week w in season s
p_s	population size in season s
Specific to equation 2	
$d^*_{t,s}$	adjusted ILI cases at day t in season s, input for effective reproductive number analysis
v_{w_s}	total visits to health care facilities for any diagnosis at winter reference week w_s
ϕ	estimate of the proportion of the total population that seeks care for symptomatic influenza-like illness (0.45)
Indicator variables	
w_s	winter reference week (chosen as the week of November 1 in a given season s)
t	indicator for time in days
s	indicator for flu season
w	indicator for time in weeks
C	number of days in a week (7 days)

S2 Metapopulation model

S2.1 Comparison of baseline and holiday air travel networks

We compared the baseline and holiday travel networks across common network measures for unweighted and weighted networks (Table S2). Weights represented the average number of monthly passengers between two metro areas, and weights were normalized by the maximum of the average number of monthly passengers between two destinations for a given network (i.e., the largest weight is one) for weighted network measures. The baseline network had a greater number of edges, perhaps because it represented data across three months, while the holiday network had a greater maximum average number of passengers and a lower unweighted mean degree. This indicates that during the holidays, there is simultaneously a greater volume of travelers and a lower average connectivity between airports, perhaps suggesting that this increased volume of holiday travelers seeks out locations with fewer connections (e.g., smaller cities). Comparisons of the degree distributions between the baseline and holiday networks reveals that the holiday network is less right-skewed than the baseline network, corroborating the smaller mean degree and mean weighted degree for the holiday network (Figure S2).

Table S2: **Comparison of baseline and holiday air travel networks across network measures.**

	Baseline	Holiday
Descriptive characteristics		
Number of nodes	236	235
Number of edges	5,071	3,794
Maximum of the average number of monthly passengers between two destinations	7,276,232	7,586,460
Unweighted measures		
Mean degree	42.97	32.29
Transitivity	0.55	0.51
Average clustering coefficient	0.66	0.63
Average shortest path length	1.97	2.09
Weighted measures		
Mean degree		
Average clustering coefficient	0.0028	0.0045
Average shortest path length	2,047.27	6,337.30

Figure S2: **The holiday air travel network was more right-skewed than the baseline network across the unweighted and weighted degree distributions.** A) The unweighted degree distribution has an inset that represents each air traffic network visually. B) The weighted degree distributions has a lower range than the unweighted degree distribution, as expected, since weights were normalized to a maximum of one.