Detailed data description

Transport flows

Our data on shipment of live swine was an estimate of the total number of swine moved between all ordered pairs of states over the course of a year. We refer to these estimates as transport flows. They were generated by the USDA Economic Research Service [1] and are available on the web.[1] Although these flows are based on data from 2001 and likely somewhat outdated, Ref. [2] found that for cattle the 2001 data seemed broadly to agree with estimated flows for 2009.

Cases and litter rates

We obtained the case data from the January 8 version of a publicly available report of laboratory testing activity [3] via the "Number of New Cases Reported" link on the webpage of the American Association of Swine Veterinarians (AASV). We used the data in Tables 2 and 4 in this report. Table 2 contained the number of positive diagnostic case submissions stratified by the state of origin and the week of submission. Table 4 contained the number of positive diagnostic case submissions stratified by the state of origin and the age class of the swine from which the tissue sample was taken.

The precise meaning of the values in the time series changed in the week of June 16. Prior to that week, the values are the number of farms testing positive for PEDV. Beginning with that week, the values are the number of diagnostic cases submissions, or accessions, and there may be many of these accessions for each infected farm. However, in the 9 weeks prior to June 16, the difference between numbers of positive case submissions and numbers of positive farms was typically less than 5. Thus it seems that the number of case submissions approximates the number of infected farms reasonably well. As mentioned in the main text, positive accessions are correlated with the number of positive farms available in more recent reports.

The litter estimates are generally considered accurate but have limitations as indicators of PEDV burdens. For example, the 2013–2014 winter was unusually harsh and may explain some fraction of the losses. Also, the survey was not designed to estimate PEDV losses by state and may not have obtained a representative sample for each state.

Predictors of cumulative burden

In the fitted regression models for cumulative burdens, many states in the Northeast had close centroids and very similar residuals, suggesting a lack of independence at this spatial scale. Using single linkage clustering, we found two groups of states that formed chains of states with centroids less than 175 km apart: (1) MD, DE, and NJ; and (2) VT, NH, CT, MA, and RI. Because of the lack of independence among data from these states suggested by our initial fits, we created a reduced data set where the values of both predictive variables and response variables for these two groups of states were averaged to form single observations. The results presented are based on that data, for

[1] http://webarchives.cdlib.org/sw1rf5mh0k/http://ers.usda.gov/Data/InterstateLivestockMovements/StateShipments.xls

which no spatial autocorrelation was indicated by maximum likelihood fits of a model of exponentially decaying covariance in the residuals.

Counts of farms of different sizes were obtained from a database application available from the USDA [4]. This application contains data from the 2007 Census of Agriculture.

The balance sheet variables for each state came from estimates for the period of December 2011 to December 2012 in Ref. [5]. These variables were swine inventory, which is the total number of swine; pig crop, which is the number of pigs born that survive the first few weeks of life; inshipments, which is the number of swine imported to the state; and marketings, which is the number of swine either exported from the state or slaughtered at a commercial facility.

Farm resource regions classify counties into one of 9 general groups based on a wide variety of criteria including farm characteristics and crops and livestock produced [6]. A list mapping counties to these regions was obtained from a USDA spreadsheet.[2] We included a variable for each region and each state was assigned a value of in [0,1] for each such variable that was equal to the proportion of farms with 25 or more swine in the counties of that region.

Nearby cases were calculated in various ways based on different possible models of spread. To represent cumulative exposure from a spatial model, we calculated for each state the average number of cases in other states with shared borders and used the logarithm of this average as a potential predictor. To represent cumulative exposure via shipment of pigs, we calculated for each state a weighted sum of cases in other states, where the weights were given by the flows from that state. Because we found with our simulations that the dependence of cases in one area can be expected to increase sublinearly with the flows from that area, we used various power transforms of the flows as weights. Specifically, we used the flows raised to the powers of 1, 1/2, 1/4, 1/8, and 1/16. These weighted sums were then log transformed to create a series of potential predictors.

The summary statistics for farm density were average number of farms per county, average number of farms in counties with at least one farm, median number of farms per county, maximum number of farms per county. Some states had a median of zero farms per county. When log transforming, one half of the smallest positive median was added to values of all states before transformation. For this analysis, we defined farms as operations with 25 or more swine.

References

1. USDA ERS (2003). Interstate livestock movements. By D. Shields and K. Mathews. Available: `http://www.ers.usda.gov/publications/ldpm-livestock,-dairy,-and-poultry-outlook/lpdm10801.aspx#.U26fN1Qt5Mk`. Accessed 14 November 2013.

2. Buhnerkempe MG, Tildesley MJ, Lindström T, Grear DA, Portacci K, et al. (2014) The impact of movements and animal density on continental scale cattle disease outbreaks in the United States. PLoS One 9: e91724.

3. USDA APHIS VS (2014). Porcine epidemic diarrhea virus (PEDv) testing data from NAHLN laboratories. Available: `http://www.aasv.org/pedv/PEDV_weekly_report_140108.pdf`. Accessed 14 January 2014.

[2] `http://www.ers.usda.gov/Briefing/ARMS/resourceregions/reglink.xls`

4. USDA NASS (2009). 2007 Desktop Data Query Tool 1.02. Available: http://www.agcensus.usda.gov/Publications/2007/Online_Highlights/Desktop_Application/dataquery.zip. Accessed 14 January 2013.

5. USDA NASS (2013). Meat Animals Production, Disposition, and Income 2012 Summary. Available: http://usda01.library.cornell.edu/usda/current/MeatAnimPr/MeatAnimPr-04-25-2013.zip. Accessed 31 July 2013.

6. USDA ERS (2000). Farm Resource Regions. Available: http://www.ers.usda.gov/ersDownloadHandler.ashx?file=/media/926929/aib-760_002.pdf. Accessed 3 April 2014.