1 **Title:**

2 Network analysis of cattle movements in Uruguay: quantifying heterogeneity for risk-based
3 disease surveillance and control

4

5 **Authors**

6 Kimberly L. VanderWaal (kvw@umn.edu)^{a*}

7 Catalina Picasso (picas001@umn.edu)^{a,b}

8 Eva. A Enns (eenns@umn.edu)^c

9 Meggan E. Craft (craft@umn.edu)^a

10 Julio Alvarez (jalvarez@umn.edu)^a

11 Federico Fernandez (FFernandez@mgap.gub.uy)^b

12 Andres Gil (adgr@adinet.com.uy)^d

13 Andres Perez (aperez@umn.edu)^a

14 Scott Wells (wells023@umn.edu)^a

15

16 ^a Department of Veterinary Population Medicine, University of Minnesota, 1365 Gortner
17 Avenue, St. Paul, MN, 55108

18 ^b Animal Health Bureau, Ministry of Livestock, Agriculture, and Fisheries, 1476
19 Constituyente, Montevideo, 11200 Uruguay, 11200

20 ^c Division of Health Policy and Management, School of Public Health, University of
21 Minnesota, 420 Delaware Street SE, MMC 729, Minneapolis, MN, 55455

22 ^d Facultad de Veterinaria, Universidad de la Republica, 1550 Alberto Lasplaces, Montevideo,
23 Uruguay, 11100

24 * Corresponding author: Kimberly VanderWaal, kvw@umn.edu (612-624-7179)

25

26 **Abstract**

27   Movement of livestock between premises is one of the foremost factors contributing
28 to the spread of infectious diseases of livestock. In part to address this issue, the origin and
29 destination for all cattle movements in Uruguay are registered by law. This information has
30 great potential to be used in assessing the risk of disease spread in the Uruguayan cattle
31 population. Here, we analyze cattle movements from 2008-2013 using network analysis in
32 order to understand the flows of animals in the Uruguayan cattle industry and to identify
33 targets for surveillance and control measures. Cattle movements were represented as
34 seasonal and annual networks in which farms represented nodes and nodes were linked
35 based on the frequency and quantity of cattle moved. At the farm level, the distribution of
36 the number of unique farms each farm is connected to through outgoing and incoming
37 movements, as well as the number of animals moved, was highly right-skewed; the majority
38 of farms had few to no contacts, whereas the 10% most highly connected farms accounted
39 for 72-83% of animals moved annually. This extreme level of heterogeneity in movement
40 patterns indicates that some farms may be disproportionately important for pathogen
41 spread. Different production types exhibited characteristic patterns of farm-level
42 connectivity, with some types, such as dairies, showing consistently higher levels of
43 centrality. In addition, the observed networks were characterized by lower levels of
44 connectivity and higher levels of heterogeneity than random networks of the same size and
45 density, both of which have major implications for disease dynamics and control strategies.
46 This represents the first in-depth analysis of farm-level livestock movements within South

America, and highlights the importance of collecting livestock movement data in order to understand the vulnerability of livestock trade networks to invasion by infectious diseases.

Keywords

Disease risk, Livestock movement, Pathogen transmission, Social network analysis, Surveillance

INTRODUCTION

Between-farm livestock movement is among the most important risk factors predicting infection at the farm-level as well as the vulnerability of livestock trade networks to large-scale disease outbreaks (Gilbert et al., 2005; Ortiz-Pelaez et al., 2006; Humblet et al., 2009; Natale et al., 2009; Lal Dutta et al., 2014). Notably, epidemics of high economic consequence, such as foot-and-mouth disease (FMD), have been facilitated by high rates of livestock movement (Kao, 2002; Kao et al., 2006). Animal traceability is key to understanding and limiting infectious disease outbreaks, and many countries now require all movements of cattle between premises to be registered (Caporale et al., 2001; Scudamore and Harris, 2002). Knowledge of animal movement patterns enhances our understanding of the epidemiology of endemic and epidemic diseases, thus leading to more accurate assessments of risk and better informed control measures, including regionalization, compartmentalization, and targeted surveillance (Martínez-López et al., 2009a; Bajardi et al., 2012; OIE, 2014).

Network analysis has been used extensively to analyze livestock movements (Green et al., 2006; Kiss et al., 2006; Webb, 2006; Kao et al., 2007; Robinson et al., 2007; Natale et

al., 2009; Aznar et al., 2011; Nöremark et al., 2011; Mweu et al., 2013; Brooks-Pollock et al., 2014). Specifically, analysis of network structure allows for identification of targets for surveillance, intervention, and control (Christley et al., 2005; Kiss et al., 2006; Natale et al., 2009; Bajardi et al., 2012). In a network framework, livestock farms are represented as nodes and linked according to patterns of movement between farms. Metrics of a farm's connectivity or centrality within the network, such as the number of farms to which it is connected, correlate with the probability of becoming infected and time to infection (Christley et al., 2005; Natale et al., 2009). These metrics typically exhibit extreme levels of heterogeneity across farms, creating complex network structures, with some farms functioning as movement hubs. Such farms are potential super-spreaders and may be disproportionately important to pathogen spread (Lloyd-Smith et al., 2005). Complex structure and population-level connectivity can be captured by network-level metrics, which in turn predict the speed and extent to which a pathogen can spread (Keeling and Eames, 2005; Salathé and Jones, 2010). Network analysis has been used to explore risk-based surveillance and control strategies based on identifying farms that either lie on bottlenecks within the network or are highly connected (Kao et al., 2006; Ortiz-Pelaez et al., 2006; Kao et al., 2007; Natale et al., 2009). Such targeted measures are much more effective at limiting disease spread than non-targeted efforts (Kiss et al., 2006; Natale et al., 2009).

In this analysis, we describe patterns of cattle movement within the Uruguayan livestock industry. In response to a major epidemic of FMD in 2001, all movements of cattle in Uruguay are traced to assist in outbreak investigations. A comprehensive analysis of these data has yet to be undertaken. Although a small number of studies have described patterns of cattle movement between geographic regions in South America (Aznar et al.,

2011), this is the first analysis of a farm-level cattle movement network in South America. In contrast to the movement networks observed in European and North American livestock industries, the Uruguay cattle industry is unique in that cattle management tends to be more extensive rather than intensive. For example, dairy and fattening farms are largely pasture-based. Feedlots account for only 5-15% of total slaughter (Joseph, 2013). Thus, characterizing of movements in this system broadens our understanding of livestock networks in other management contexts. Here, we applied network analysis techniques to describe cattle movement networks in Uruguay over multiple years. Specifically, we analyzed major flows of animals within and between production types, quantified heterogeneity in the centrality of individual farms with respect to production type and herd size, and described the global structure and connectivity of the cattle industry.

MATERIALS AND METHODS

Data source

Details of between-farm cattle movement from the study period of July 2008 to May 2013 were obtained from the Uruguay's Ministry of Livestock, Agriculture, and Fisheries. Data provided by the government were stored in four databases: demographic attributes, milk producer registry, geolocation data, and movement records. These datasets were related to one another using a premise ID number that is unique to each farm. Demographic attributes included the number of animals of each age-class reported on a given farm each year. Classes included calves, heifers age 1-2 years, heifers >2 years, steers age 1-2 years, steers 2-3 years, steers >3 years, bulls, and breeding cows. We defined herd

size as the total number of cattle of any age-class reported on the farm. Geolocation data listed the UTM (Universal Transverse Mercator) coordinates of each farm. Movement records consisted of the date of each movement, total number of animals of each age-class moved, the premise ID of the source location, and the premise ID of the destination. Movements for which the final destination was not a farm (i.e., cattle moved to slaughter or exported) were excluded as these movements constituted endpoints or sinks within the network and thus not likely to play a role in between-farm transmission chains. Auctions in Uruguay increasingly occur online, and animals sold in such transactions move directly between the buyer and seller. Data was not available for traditional auctions that still occur, but these tend to be small scale and local. Changes of ownership involving markets, auctions, or dealers were represented in these data as the seller and buyer with no intermediate step due to lack of information on movements through markets/auctions. While non-reported movements of animals may occur with very low frequency, we do not expect such movements to have an impact on this description of cattle movement networks, given the large sample sizes described in this study (>60,000 farms over four years) and that all cattle must be registered in the first six months of life.

Population description

Over the course of the study period, 62,767 unique livestock premises were recorded, though not all were in existence for all four years. For consistency with official government definitions (Molina, 2012), premises were classified into six production types using a decision tree, which was developed in consultation with the Uruguayan Ministry of Livestock, Agriculture and Fisheries. First, farms that appeared on the milk production

registry were defined as dairy farms. For non-dairy premises, those with five or fewer animals in the demographic attribute database were designated as small farms. For the remaining premises, the ratio of steers to breeding cows was used to classify farms as breeding (ratio < 0.3) and fattening (ratio > 1.0) operations. Farms with intermediate ratios (0.3 – 1.0) were considered to be engaged in both breeding and fattening and termed complete cycle operations (Molina, 2012). Farms with steers, but no breeding cows, were also considered fattening farms. Farms with neither steers nor breeding cows were primarily composed of heifers and calves. These were designated growing farms. The number of farms and frequency of each production type are listed in Table 1. The distribution of herd sizes by production type is also listed in Table 1. Some small farms reported a herd size of zero because fluctuations in herd size across time may mean that they may have no cattle at the time of reporting. The largest observed herd size was between 13,000 and 17,000 head for all production types except growing (maximum 3,831 head) and small farms (maximum 5 head).

Table 1. Proportion of production types and total number of cattle farms in years with 12 full months of data (2009-2012), as well as median herd sizes with inter-quartile ranges (IQR).

	2009	2010	2011	2012	Median herd size (IQR range)	
Dairy	0.10	0.10	0.10	0.09	106	(53 – 227)
Small farm	0.08	0.08	0.09	0.09	2	(0 – 4)
Breeding	0.42	0.42	0.43	0.45	69	(25 – 205)
Complete cycle	0.16	0.16	0.15	0.14	127	(34 – 464)
Fattening	0.22	0.22	0.21	0.20	131	(37 – 392)
Growing	0.02	0.02	0.02	0.03	28	(12 – 73)
Total farms	45854	46218	45643	45605		

158

159 ***Network analysis***

160 Cattle movements were represented as a network comprised of nodes (farms)
161 interlinked by edges representing discrete shipments of animals. A path in the network can
162 be defined as a sequence of edges and nodes by which two farms are indirectly connected,
163 and along which a pathogen can be transmitted. The shortest path between two nodes is
164 the minimum number of nodes that must be passed through to reach one node from
165 another (Wasserman and Faust, 1994). Paths can either be undirected or directed, where
166 directed paths account for the direction in which animals flow between farms. Uruguay
167 cattle networks were directed in that the source and destination of each movement of
168 animals was known, and the weight of the edge was defined as the number of cattle moved.
169 Two nodes are considered reachable if a directed path exists connecting them. If no such
170 path exists, these nodes are considered unreachable. Given the directed nature of
171 movements, farm A may be reachable by farm B, but no path exists from in the reverse
172 direction.

173 Network- and farm-level summary statistics were based on movements aggregated
174 over a seasonal or annual timeframe. Annual networks were constructed for each of the
175 four years for which complete data were available (2009 – 2012). As many farms moved
176 cattle less than once per year, an annual timeframe allowed us to capture a sufficient
177 number of movements to characterize network structure while still allowing for temporal
178 variation in movement patterns. In addition, within each of these years, four seasonal
179 networks were constructed. Seasons were defined as summer (January – March), autumn
180 (April – June), winter (July – September), and spring (October – December). In addition, a

full network was constructed using all available movement records (July 2008 – May 2013). Networks aggregated over longer time scales may be relevant for chronic diseases, such as bovine tuberculosis, while shorter timeframes capture short-term variation that may be critical for understanding the spread of acute, rapidly spreading diseases such as FMD (Cross et al., 2005; Craft, 2015). Higher levels of both global and farm-level connectivity are expected to be important indicators of the vulnerability of trade networks to infectious diseases (Keeling and Eames, 2005; Martínez-López et al., 2009c).

Measures of individual farm connectivity were quantified using in-degree, out-degree, the total number of cattle sent or received by the farm, in- and out- neighbourhood size, and betweenness centrality (defined in Table 2 and Supplementary Material). All farm-level metrics were calculated on annual networks. Connectivity measures were summarized for each production type and we report the mean, median, 95th quantile, and maximum. As node-level metrics are often correlated, we report the Spearman's correlation coefficient (ρ) among node-level metrics.

We focus on these specific farm-level metrics because their association with the spread of infectious disease has been demonstrated in previous analyses (Kiss et al., 2006; Ortiz-Pelaez et al., 2006; Natale et al., 2009; Volkova et al., 2010d). Degree, number of cattle sent/received, and neighbourhood size are measures that directly quantify the extent to which individual farms and their immediate neighbours engage in movements.

Betweenness centrality, in contrast, quantifies the position of a farm within the entire network, with high values indicating that a farm is important for creating linkages and maintaining connectivity between other farms. Neighbourhood size is the number of farms that can be reached from a focal farm through a direct movement or indirectly via a single

intermediate farm (path length <2 steps)(Ortiz-Pelaez et al., 2006; Dubé et al., 2008; Nöremark et al., 2011), as neighbourhoods based on path lengths of two were demonstrated to be associated with infection risk in the early stages of the UK 2001 FMD epidemic (Ortiz-Pelaez et al., 2006). This metric takes into account the temporal sequence of edges, since it is not only important that a path exists, but that the edges occur in such an order that would allow flow of pathogens from one farm to another. The in-neighbourhood provides a metric of the number of farms that could potentially transmit a pathogen to the focal farm, whereas the out-neighbourhood represents the number of farms that could potentially become infected by the focal farm (Dubé et al., 2008; Nöremark et al., 2011).

Network-level metrics (defined in Table 2 and Supplementary material) were calculated on the full, annual, and seasonal networks to assess temporal trends in connectivity. Metrics included network density, percent isolates, size of the giant weakly and strongly connected components (GWCC and GSCC, respectively), in- and out-degree centralization, betweenness centralization, clustering coefficient, reciprocity, diameter, and mean path length (Table 2). A component is a section of the network where all nodes are connected to each other either through direct linkages or via intermediate farms. The giant component is the largest component in which all nodes are connected via directed paths (strong component, GSCC) or paths with the direction ignored (weak component, GWCC)(Wasserman and Faust, 1994). The GSCC and GWCC are often considered an indicator of the minimum and maximum possible size of an epidemic, respectively, if no interventions are implemented (Kao et al., 2006). Centralization of any farm-level metric essentially quantifies the extent to which the network is structured around a minority of nodes (Wasserman and Faust, 1994). High levels of centralization are indicative of the

presence of super-spreaders in the network. However, pathogen spread can be deterred if infections are prevented in the central core of highly-connected farms (Valente, 2010). We also tested the network for scale-free properties by fitting a linear regression to the log-log cumulative degree distribution. A linear relationship indicates that a small number of farms are responsible for maintaining a large proportion of contacts (Barabasi and Albert, 1999).

All analyses were conducted in *R* version 3.1.1 using the *igraph* and *EpiContactTrace* packages (Csardi and Nepusz, 2006; Dubé et al., 2008; Nöremark et al., 2011).

Table 2. Definitions of farm- and network-level metrics.

	Definition
<u><i>Farm-level metrics</i></u>	
• In / Out-degree	Total number of farms from which the focal node received/sent cattle annually ^{1,2}
• Number of animals sent/received	Total number of cattle sent or received by a farm annually
• In / Out-neighbourhood size	Total number of farms that can be reached in two steps along paths that are directed inwards/outwards toward the focal farm, taking into account the temporal sequence of edges ³
• Betweenness	Number of directed paths that pass through the focal node when the shortest path is traced between all other pairs of nodes. The temporal sequence of edges is ignored ^{1,2}
<u><i>Network-level metrics:</i></u>	
• Network density	Number of edges in the observed network relative to total number possible in a completely connected network ¹
• Percent isolates	Percent farms that engage in no movements during a given period
• Giant weakly connected component (GWCC)	Proportion of nodes that are connected in the largest component when directionality of movement is ignored ^{1,4}
• Giant strongly connected component (GSCC)	Proportion of nodes that are connected in the largest component when directionality of movement is considered ^{1,4}
• In / Out-degree centralization	Quantifies the extent to which a minority of farms are responsible for a majority of incoming/outgoing movements ¹

• Betweenness centralization	Quantifies the extent to which paths flow through only a small subset of influential nodes ¹
• Clustering coefficient	For each node, the clustering coefficient is the probability that two nodes connected to the focal node are also connected to each other. Isolates are ignored, and the average clustering coefficient across all nodes is reported ¹
• Reciprocity	Proportion of edges where an edge in the opposite direction also exists ¹
• Diameter	Longest geodesic path in the network ¹
• Mean path length	Mean number of edges that must be traversed to connect two nodes. The temporal sequence of edges is ignored ¹

¹(Wasserman and Faust, 1994); ²(Martínez-López et al., 2009c); ³ (Dubé et al., 2008; Nöremark et al., 2011); ⁴ (Kao et al., 2006).

Statistical analysis

To determine if the observed structure in the cattle movement networks could have emerged by chance, we compared structural features of the cattle movement networks to 1000 random Bernoulli networks of the same size and density. Network-level metrics were considered unlikely to emerge by chance if less than 5% of the randomly-generated networks had metric values were more extreme than the observed values.

Possible associations between network metrics and production type, herd size, and geography were also investigated. Significance of differences in centrality metrics among production types was tested using Kruskal-Wallis tests with Bonferroni corrections. All correlations among network metrics and between network metrics and herd size were reported with a Spearman's correlation coefficient (ρ).

We calculated assortativity by production type and size to determine whether farms preferentially moved animals with farms of the same type or size. Essentially, this metric quantifies the extent to which movements occur among premises with similar versus dissimilar characteristics. The expected number of movements between each pair of production types was calculated based on the total observed movements by each

production type. Deviations from expectations are reported as observed-to-expected ratios. This reflects the extent to which movements occurred more or less frequently between two production types than predicted if movements occurred randomly.

Temporal and spatial patterns

We examined correlations across time (Pearson's correlation coefficient) in node-level metrics and the number of times edges recurred across years.

The Euclidean distance between each pair of farms engaging in movement was calculated using the geolocation database. To assess whether distance in space predicted path length within the network, we first randomly selected 5,000 farms and, for each pair, calculated the Euclidean distance and the length of the shortest path through the full network. Correlations between distance and path length were summarized using Pearson's correlation coefficient.

Missing data

Over the entire study period, a total of 14,780 farms were missing demographic data for at least one year (overall 18,248 farm-years, 6.0% of the total). These farm-years were involved in 14.2% of movements. As these farms still may contribute to overall connectivity and the centrality metrics of other farms, they were retained in the network, but not included in the analysis of farm-level metrics when stratified by production type and herd size. 8,539 farm-years (2.5%), involving 3498 unique farms, were missing geolocation data and were involved in 7.3% of movement records. These farms were retained in the network, but excluded from all spatial analyses. 9,460 farm-years (3.1%), involving 8,703

unique farms and accounting for 3.2% of all movements, were missing both geolocation and demographic data. These farms were excluded from all spatial analyses and comparisons by production type and herd size.

RESULTS

Annual networks of cattle movement within Uruguay were characterized by extremely low density, with <0.01% of the all possible edges actually occurring. However, 68 – 73% of farms were connected to one another if directionality of movements was ignored. Each year, 20-25% of farms were isolates and engaged in no movements whatsoever. Observed networks were less connected and more structured than randomly generated networks of the same size and same density (Table 3, where $P < 0.001$ for all measures). Although random networks exhibited significantly higher levels of connectivity, as demonstrated by fewer isolates and a greater number of nodes involved in the GWCC and GSCC, connections were relatively uniformly distributed across nodes. In contrast, movements in observed networks were centralized through or around a more limited subset of nodes (reflected in higher betweenness and degree centralization). Random networks never achieved the same level of reciprocity or clustering of movements.

Table 3. Network-level measures of connectivity of cattle movement networks summarized by year and for the entire study period.

2009	2010	2011	2012	All years

Density	5.3E-05	5.3E-05	5.5E-05	5.6E-05	7.2E-05
Percent isolates	0.25	0.22	0.21	0.22	0.01
Proportion in Giant Weakly Connected Component (GWCC)	0.68	0.72	0.73	0.73	0.97
Proportion in Giant Strongly Connected Component (GSCC)	0.15	0.17	0.18	0.18	0.51
In-degree centralization	0.04	0.02	0.03	0.02	0.04
Out-degree centralization	4e-3	3e-3	3e-3	3e-3	0.13
Betweenness centralization	0.11	0.09	0.10	0.05	0.10
Clustering coefficient	0.09	0.10	0.10	0.10	0.13
Reciprocity	0.11	0.09	0.10	0.10	0.13
Diameter	33	33	31	26	20
Mean path length	8.97	9.11	8.78	9.01	6.1
Total number of edges	63,083	69,198	72,072	71,505	278,563
Power-law exponent (In-degree)	1.73	1.68	1.68	1.79	1.72
Power-law exponent (Out-degree)	2.48	2.69	2.56	2.45	2.45

Less than 50% of farms engaged in outgoing cattle movements annually, though our data did not include movements to slaughter. Approximately 55% of farms received cattle annually. The mean, median, 95th percentile, and maximum and minimum values of farm-level measures are reported in the Supplementary Information (Tables S1 – S3). In general, the distributions of farm-level metrics of network connectivity were extremely right-skewed, with the majority of farms having only a few connections and a small minority of farms having a large number of connections. Each year, 72-74% of all cattle moved were moved from the farms that scored in the top 10% for total animals sent. 82-83% of all cattle moved were moved into farms scoring in the top 10% for total animals received. Indeed, when plotted on a log-log scale, the distributions of in-degree and out-degree approximated a power-law distribution with exponents ranging from -1.68 to -1.79, and -2.44 to -2.69, respectively (Figure S1).

The likelihood of engaging in a movement was dependent on production type (Figures 1 - 2). Fattening farms had the highest number of in-movements, with over 60% engaging in inward movements annually. On average, fattening farms were connected to

three other farms via inward movements (median = 1 farm), though some fattening farms received cattle from up to 1,240 other premises. Fattening farms received an average of 75 cattle per year, which was more than twice the mean number of cattle received by any other farm type. Dairies, in contrast, were the production type engaging in the most outward movements, with nearly 75% of dairies connected to at least one other farm via outward movements. Dairies also exhibited the largest out-neighbourhoods, with a median of seven farms reached within two steps. Dairy farms moved an average of 46 cattle annually.

Fattening and dairy farms had significantly higher betweenness than other production types (Kruskal-Wallis test, $p < 0.01$, Figure 2). Indeed, fattening and dairy farms were more than twice as numerous than expected in the 1,000 farms with the highest betweenness centrality. Although there was a larger proportion of growing farms than expected among the 1,000 farms with the highest observed betweenness, the overall betweenness of growing farms was not significantly different from complete cycle or breeding farms.

In general, all farm-level network measures were correlated with each other (Supplementary material, Table S4). Notably, correlations existed between farm in- and out-degree ($\rho = 0.31$, $p < 0.01$), the total number of animals sent and received ($\rho = 0.33$, $p < 0.01$), and in- and out-neighbourhood size ($\rho = 0.47$, $p < 0.01$). Betweenness was correlated with both in-degree ($\rho = 0.73$, $p < 0.01$) and out-degree ($\rho = 0.63$, $p < 0.01$).

Farms with larger herd sizes were more likely to send and receive cattle; the mean number of cattle sent and received annually for farms in the lowest quartile for herd size (<20 cattle) was 6.8 and 7.5 animals, respectively, whereas farms in the fourth quartile for

herd size (>240 cattle) sent and received a median 86 and 35 animals. A farm's in- and out-degree and betweenness were all correlated with herd size ($\rho = 0.40 - 0.49$, $p < 0.01$ in all cases), though the size of in- and out-neighbourhoods were somewhat less correlated with herd size ($\rho = 0.33-0.36$, $p < 0.01$ in all cases). For example, only 25% of small farms (<5 cattle) engaged in any movements at all, and those that did had much smaller neighbourhood sizes and lower betweenness centrality than other production types (Figure 2). Farms were also somewhat more likely to trade animals with farms of similar size ($\rho = 0.36$, $p < 0.01$).

Table 4 shows the assortativity of movements by production type. Values greater than or less than one indicate production types between which movements were more or less common, respectively, than predicted if movements occurred randomly. Unsurprisingly, fattening farms sent animals to breeding and dairy farms less frequently than expected due to chance, though they traded animals among themselves somewhat more often than expected. Small farms were more likely to trade animals with other small farms. Dairy farms were almost four times more likely to move animals among themselves than expected, and were also more likely to send and receive animals from growing farms. Dairies received cattle from breeding and complete cycle farms less frequently than expected, and also did not send cattle to breeding, complete cycle, and fattening farms as often as expected. This suggests that the dairy industry is somewhat compartmentalized from the rest of the cattle industry, though examining the neighbourhoods of dairies reveals that there is substantial movement of animals between dairy and non-dairy premises (Figure 3).

Table 4. Assortativity of movements by production type. Values >1 indicate movements that are more frequent than expected **by chance**, whereas values <1 indicate movements that are less frequent than expected **by chance alone**.

		Destination					
		Breeding	Complete cycle	Dairy	Fattening	Growing	Small farm
Origin	Breeding	1.14	1.04	0.57	0.96	0.87	1
	Complete cycle	0.98	1.15	0.54	1.06	0.81	0.94
	Dairy	0.72	0.71	3.74	0.79	1.98	1.07
	Fattening	0.83	0.97	0.8	1.19	0.88	0.87
	Growing	0.91	0.75	1.93	0.96	1.35	1.04
	Small farm	1.08	1.09	0.94	0.87	0.87	2.13

Annual networks of cattle movements showed substantial year-to-year consistency. At the network-level, the frequency of each production type, number of movements per season, and other measures of connectivity varied little across years. At the individual farm level, in- and out-degree, the total numbers of animals sent or received, and betweenness were highly correlated across years (Pearson's $\rho = 0.63 - 0.75$, $p < 0.01$ in all cases). In addition, farm size was almost perfectly correlated across years (Pearson's $\rho = 0.94$, $p < 0.01$).

Although farm-level connectivity values were conserved across years, suggesting a consistent tendency to engage in movements over time, the consistency of movements between two specific farms across years was very low. 70% of edges occurred only once during the study period, and 86% of edges occurred only in one year (though they may have occurred multiple times within that year). Only 8.3% of edges recurred in two different years, 2.2% in three years, and 1.0% in four years. The maximum number of times any edge recurred was 249 times, though the mean was only 1.7 ± 2.6 SD. Total rates

of reciprocity were relatively low, though significantly higher than random expectations. About 10% of movements were reciprocal between the source and destination.

Cattle movements within Uruguay followed a distinct seasonal pattern. Total number of movements peaked in June (autumn), a pattern which was largely driven by movements of non-dairy calves born in spring. Due to increased frequency of movements and a larger number of farms engaging in movement, the size of the GWCC within the 3-month seasonal networks peaked in the autumn. Over 30% of farms were connected via indirect paths during this time period (Figure 4).

Movements of animals were mostly local, and the median distance of a movement was ~40 km. However, long distance movements of up to 550 km did occur with low frequency, which is approximately the width of the country. Longer distance movements were more likely among farms with larger herd sizes. The median distance for movements in which both the sending and receiving farm fell in the fourth quartile of herd size (>240 cattle) was 66 km, while the median distance when both farms fell in the first quartile of herd size (<20 cattle) was only 11 km.

In Uruguay, approximately 40% of dyads (pairs of farms) were unreachable either because they occupied disconnected components of the network or because the direction of movements did not allow for a connecting path. For the 60% of dyads where a connecting path existed, the average path length connecting two farms was approximately eight to nine steps. Thus, a farm could reach most other farms via eight to nine intermediate farms. We found that, overall, the path length between two nodes was not correlated with their geographic distance from each other (Pearson's $\rho = 0.16$). When only considering dyads with direct connections (path length of 1) or connected through one

intermediate farm (path length of 2), movements were geographically localized, with a median distance of 40, 77, and 132 km for path lengths of 1, 2, and 3, respectively. However, the distribution of geographic distances between dyads connected by paths of more than three steps was similar to the distribution of distances between farms where no path in the network existed (Figure 5).

DISCUSSION

Cattle movement networks in Uruguay were highly fragmented, with the majority of farms engaging in little to no movement annually. Like other studies of livestock movement, farm-level connectivity was extremely right-skewed (Christley et al., 2005; Bigras-Poulin et al., 2007; Natale et al., 2009; Volkova et al., 2010d, a), as can be seen through a comparison of median, mean, and maximum values of degree, number of animals sent/received, and neighbourhood size (Figure 1, Tables S1-S3). In many cases, mean connectivity scores were greater than the interquartile range of the data, highlighting the presence of outliers. In terms of total cattle moved, the top 20% of farms accounted for 88% of all animals moved outwards, and 94% of all animals moved onto farms. This is even more extreme when we examine the 10% of farms that moved the most animals. These farms accounted for more than 70% of all animals moved off of farms and more than 80% of all animals moved onto farms. Thus, the cattle industry within Uruguay represents a more extreme case of the so-called 20-80 rule in that 20% of farms account for >80% of transmission potential (Woolhouse et al., 1997). Furthermore, the network exhibits significant scale-free properties (Table 3, Figure S1), with the power-law exponents

characterizing the degree distribution in the same range as reported in other livestock networks (1.5– 2.6)(Christley et al., 2005; Natale et al., 2009; Rautureau et al., 2011; Mweu et al., 2013; Grear et al., 2014; Lal Dutta et al., 2014). These patterns mirror what has been observed in other livestock movement networks (Woolhouse et al., 2005; Volkova et al., 2010d; Lal Dutta et al., 2014).

In addition, heterogeneity in measures of incoming movement were more extreme than that of outgoing movement, a pattern that has also been observed in other livestock trade networks, including France, Argentina, and the UK (Volkova et al., 2010d, a; Aznar et al., 2011; Lal Dutta et al., 2014). This likely emerges as a result of farms that finish cattle having a high turnover of animals, allowing them to accept multiple batches per year. Outward movements, which here did not include movements to slaughter, may be less variable than inward movements because they are driven, at least in part, by the production of calves through breeding.

The heterogeneity and right-skew observed in farm-level connectivity means that a small number of highly connected farms are candidate super-spreader farms, and may be disproportionately important in spreading infectious diseases (Lloyd-Smith et al., 2005; Natale et al., 2009). Frequent inward movements create opportunities to acquire infected animals, which may make these highly connected farms important sentinels for surveillance. High levels of outward movement means that, if infected, these farms have the potential to spread a disease widely (Natale et al., 2009). For example, epidemiological models show that the number of secondary infected farms produced by an infected farm is correlated with betweenness and is almost linearly related to out-degree (Natale et al., 2009).

Inward and outward movements in Uruguay were slightly correlated ($\rho = 0.24 - 0.34$, Pearson's $\rho = 0.30 - 0.36$), which is higher than observed in some other livestock networks. Volkova et al. (2010) reported a coefficient of $0.12 - 0.33$ for the Pearson's correlation between in-and out-degree, whereas Woolhouse et al. (2005) reported no correlation and Lal Dutta et al. (2014) reported a negative correlation in farm in- and out-movements. Such high correlations in Uruguay create a situation where highly connected farms are both more likely to become infected and to subsequently transmit the pathogen to other farms. Variation in contact rates and covariance in in- and out-movements has been shown to greatly increase R_0 (number of farms infected by a single infectious farm in a naïve population) compared to estimates calculated using only mean contact rates (Volkova et al., 2010a, d).

Farms that have both high in- and out-movements also tend to have high betweenness, and are likely important conduits of pathogen flow in a network. Fattening and dairy farms, for example, had significantly higher betweenness than other production types. This suggests that fattening and, to a lesser extent, dairy farms occupy crucial roles in the network by connecting other farms that would otherwise be more poorly connected. Intervention measures targeted at farms with high betweenness may potentially fragment transmission chains and reduce the size and extent of infectious disease outbreaks. In addition, we argue that targeting sampling in highly connected farms (high degree and betweenness) would increase the sensitivity of active surveillance systems in Uruguay, such as those used for foot-and-mouth disease, as opposed to random sampling of farms.

It is important to analyze the repeatability of network measures at the node-level if we are to use estimates of past movements as an indicator of future risk. Farm-level

connectivity was conserved across years, with highly connected farms remaining so for at least several years in a row. Therefore, control strategies targeted at key farms can be based on the identification of potential super-spreaders from past years' movement records, which, again, may have important implications in the implementation of active surveillance systems. However, caution should be exercised in applying past movement data to future risk, especially if metrics are based on movement data aggregated at shorter than annual timeframes (Bajardi et al., 2012; Mweu et al., 2013).

Despite no overall temporal trend in network connectivity across years, seasonal networks exhibited a distinct peak annually during the autumn months, especially June. Not only were there more movements during this period, but the size of the GWCC was also larger. Due to increased movements during this time period, any pathogen introduced to Uruguay during or immediately preceding this period (March – June) would have the potential to spread much more rapidly than a pathogen introduced at a different time of year. Anecdotally, two outbreaks of FMD occurred in 2000 and 2001. The first outbreak occurred in October, 2000 and only impacted one cattle farm and one swine farm (OIE, 2001). The second outbreak occurred from April to August, 2001, a season that our data indicates as a period of high movements. This outbreak affected over 2000 farms (OIE, 2001). While this observation only provides anecdotal evidence, the larger size of the 2001 epidemic is consistent with the predictions that increased levels of movement in the preceding months would generate more extensive between-farm transmission. Seasonal trends and temporal variation in animal movements are not uncommon in livestock networks, and have been demonstrated in a number of other countries (Martínez-López et al., 2009a; Natale et al., 2009; Aznar et al., 2011; Nöremark et al., 2011; Mweu et al., 2013).

The dairy industry was somewhat insular in that dairy farms were nearly four times as likely to send and receive cattle from other dairies than expected by chance. Dairies are often considered high risk for pathogen transmission largely because of differences in management, such as congregating animals twice daily for milking, that may be conducive to rapid, within-herd transmission (Goodchild and Clifton-Hadley, 2001; Madex and Seegers, 2010). While it may at first appear that the dairy industry may be potentially compartmentalized from non-dairy farms, there was still a significant flow of animals between dairies and non-dairies (Figure 3). In addition, dairy farms had the largest out-neighbourhoods, which means that an infected dairy has a large number of potential contacts to which it could transmit infection (Figure 2). Therefore, dairies have the potential to be key players in disseminating pathogens. (Green et al., 2006; Ortiz-Pelaez et al., 2006)

Comparisons of network metrics across countries can be difficult due to differences in population size, data recording, production types, inclusion of movements to slaughter or through markets, and the time frame of analysis. Uruguay farms appear to move cattle less frequently than cattle industries in other countries, though our dataset did not include movements to slaughter. (Ortiz-Pelaez et al., 2006; Mweu et al., 2013; Lal Dutta et al., 2014) The reported mean in- and out-degree annually in Scottish beef herds was ~16 and 7.5, respectively (Woolhouse et al., 2005), which is far higher than the degree of 1-2 typical of Uruguay. Uruguayan dairy farms appear to move animals at similar rates to Canadian dairy farms, although the Canadian network did not include movements of animals to non-dairy premises, so movements were likely underreported (Dubé et al., 2008). The degree

reported in cattle movement networks in Sweden was 1-2 contacts per year, which is similar to what is reported here.

Movements of cattle were predominantly local, though larger farms were more likely to engage in long distance movements (Figure 5). Thus, movements made by large farms may be important conduits by which pathogens jump to new geographic regions. During an outbreak, targeted surveillance of larger farms in unaffected regions may be more likely to detect the introduction of the pathogen. Similarly, movement restrictions targeted at larger farms may be more likely to prevent wider spatial spread of a pathogen. In addition, connectivity via movements between two farms was related to geographic distance only if farms were connected directly or via movements through only one to two intermediate farms (Figure 5). Longer paths in the network could span large geographic distances, which indicates that infectious diseases transmitted via animal movements may disseminate broad geographic distances with just a few between-farm transmission events. This suggests that control strategies based on regionalization may not be adequate to capture all at-risk farms.

Our analysis has three potential shortcomings. First, we used a static network (seasonal or annual) to represent a dynamic system in that all movements within a specified time period are binned and the temporal order is lost. However, static networks are regularly used in epidemiological models of between-farm transmission, and they provide a good metric of epidemic size (Vernon and Keeling, 2009), though they may over-represent the network's connectivity. Second, our data did not include movements through auctions and markets, though these increasingly occur online in Uruguay. Inclusion of data on movements through traditional (not online) markets would likely enhance the

connectivity of the Uruguayan network, given that they occupy highly connected positions in networks in other countries (Ortiz-Pelaez et al., 2006; Mweu et al., 2013; Lal Dutta et al., 2014). Markets have been implicated as playing key roles in the dissemination of pathogens in livestock movement networks (Green et al., 2006; Kao et al., 2006; Ortiz-Pelaez et al., 2006), thus analysis of movement data for even small-scale, local markets should be prioritized to improve predictions of disease transmission dynamics in Uruguay. Third, our analysis focused on movement of cattle among farms. However, there are additional pathways that may create opportunities for between-farm pathogen transmission, including movement of other livestock, personnel, and equipment (Green et al., 2006; Ortiz-Pelaez et al., 2006). While other potential routes of transmission create additional complexity to predicting disease dynamics, an understanding of the underlying patterns of cattle movement elucidate one dimension of the complex web of interactions among farms.

At a population level, the structure and connectivity of annual movement networks differed significantly from random networks of the same size and density. Despite having an equal number of edges, observed networks had lower levels of connectivity than random networks. This pattern was primarily driven by the fact that, in random networks, edges were distributed relatively uniformly across nodes. Consequently, the size of the GWCC and GSCC were significantly larger in random networks. GSCC and GWCC have been suggested to be correlated with the minimum and maximum size of an epidemic, assuming no interventions, and serve as a representation of how widely an infectious disease could spread (Kao et al., 2006). Thus, the larger sizes of the GSCC and GWCC in random networks means that any predictions made from random networks may overestimate the size of an

epidemic. Observed networks also exhibited higher levels of centralization both in betweenness and degree, indicating large amounts of heterogeneity in farm-level connectivity. This suggests that a small subset of farms were disproportionately important in maintaining connectivity, while most farms occupied more peripheral positions in the network. High centralization may either facilitate or deter pathogen spread, depending on which farms become infected early in an outbreak. Infectious disease outbreaks may be limited if introduced into a periphery farm given that a disease may have few opportunities to spread. However, transmission may be explosive if the disease reaches a highly connected farm, making centralized networks vulnerable to large and rapid outbreaks (Barrat et al., 2004; Keeling and Eames, 2005; Lloyd-Smith et al., 2005).

Epidemiological models of disease spread have historically assumed that contact among nodes in a population is homogenous, and that the probability of contact is equal for every pair of nodes in the population. This pattern of contact can be approximated using random networks, in which contact occurs with a constant probability across farms. However, network structure often differs significantly from random mixing, and complex structure within networks has implications for epidemiological predictions (Keeling and Eames, 2005; Turner et al., 2008; Ames et al., 2011). Without knowledge of the structure of Uruguay's cattle movement network, network-level connectivity would likely be overestimated, whereas centralization would be underestimated, both of which would impact predictions and control strategies for disease. This highlights the importance of collecting livestock movement data in order to accurately assess system-level vulnerability to infectious disease threats. Analysis of such data improves risk assessments both at the

farm and industry level, and the identification of high-risk farms provides guidance for the development of targeted surveillance and control strategies.

Acknowledgements

Data were provided by the Directory of Agricultural Statistics, Ministry of Livestock, Agriculture, and Fisheries, Montevideo, Uruguay. This research was supported by USDA-NIFA AFRI Foundational Program grant #2013-01130. M.C. was funded by National Science Foundation (DEB-1413925) and the University of Minnesota's College of Veterinary Medicine (Animal Health Formula Funds). J.A. was partially supported by Global Food Venture MnDrive initiatives. This material is based upon work supported by the Cooperative State Research Service, U.S. Department of Agriculture, under Project Number MINV 62-044. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author and do not necessarily reflect the view of the U.S. Department of Agriculture.

REFERENCES

- Ames, G.M., George, D.B., Hampson, C.P., Kanerek, A.R., McBee, C.D., Lockwood, D.R., Achter, J.D., Webb, C.T., 2011. Using network properties to predict disease dynamics on human contact networks. *Proc. R. Soc. Lond., Ser. B: Biol. Sci.* 278, 3544-2550.
- Aznar, M.N., Stevenson, M.A., Zarich, L., Leon, E.A., 2011. Analysis of cattle movements in Argentina, 2005. *Prev Vet Med* 98, 119-127.

610 Bajardi, P., Barrat, A., Savini, L., Colizza, V., 2012. Optimizing surveillance for livestock
611 disease spreading through animal movements. *J R Soc Interface* 9, 2814-2825.

612 Barabasi, A., Albert, R., 1999. Emergence of scaling in random networks. *Science* 286, 509-
613 512.

614 Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A., 2004. The architecture of
615 complex weighted networks. *Proc. Natl. Acad. Sci. USA* 101, 3747-3752.

616 Bigras-Poulin, M., Barford, K., Mortensen, S., Greiner, M., 2007. Relationship of trade
617 patterns of the Danish swine industry animal movements network to potential
618 disease spread. *Prev Vet Med* 80, 143-165.

619 Brooks-Pollock, E., Roberts, G.O., Keeling, M.J., 2014. A dynamic model of bovine
620 tuberculosis spread and control in Great Britain. *Nature*.

621 Caporale, V., Giovannini, A., Di Francesco, C., Calistri, P., 2001. Importance of the traceability
622 of animals and animal products in epidemiology. *OIE Scientifique et Technical*
623 *Review* 20, 372-377.

624 Christley, R.M., Pinchbeck, G.L., Bowers, R.G., Clancy, D., French, N.P., Bennett, R., Turner, J.,
625 2005. Infection in social networks: using network analysis to identify high-risk
626 individuals. *American Journal of Epidemiology* 162, 1024-1031.

627 Craft, M.E., 2015. Infectious disease transmission and contact networks in wildlife and
628 livestock. *Philos Trans R Soc Lond B Biol Sci* 370, 20140107.

629 Cross, P.C., Lloyd-Smith, J.O., Johnson, P.L.F., Getz, W.M., 2005. Duelling timescales of host
630 movement and disease recovery determine invasion of disease in structured
631 populations. *Ecol. Lett.* 8, 587-595.

632 Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research.
633 InterJournal, Complex Systems 1695.

634 Dubé, C., Ribble, C., Kelton, D., McNab, B., 2008. Comparing network analysis measures to
635 determine potential epidemic size of highly contagious exotic diseases in
636 fragmented monthly networks of dairy cattle movements in Ontario Canada.
637 Transboundary and Emerging Diseases 55, 382-392.

638 Gilbert, M., Mitchell, A., Bourn, D., Mawdsley, J., Clifton-Hadley, R., Wint, W., 2005. Cattle
639 movements and bovine tuberculosis in Great Britain. Nature 435, 491-496.

640 Goodchild, A.V., Clifton-Hadley, R.S., 2001. Cattle-to-cattle transmission of Mycobacterium
641 bovis. Tuberculosis 81, 23-41.

642 Grear, D.A., Kaneene, J.B., Averill, J.J., Webb, C.T., 2014. Local cattle movements in response
643 to ongoing bovine tuberculosis zonation and regulations in Michigan, USA. Prev Vet
644 Med 114, 201-212.

645 Green, D.M., Kiss, I.Z., Kao, R.R., 2006. Modelling the initial spread of foot-and-mouth
646 disease through animal movements. Proceedings of the Royal Society of London B
647 273, 2729-2735.

648 Humblet, M.-F., Boschioli, M.L., Saegerman, C., 2009. Classification of worldwide bovine
649 tuberculosis risk factors in cattle: a stratified approach. Vet. Res. 40, 50.

650 Joseph, K., 2013. Uruguay Livestock and Products Annual GAIN Report. United States
651 Department of Agriculture Global Agricultural Information Network, 1-7.

652 Kao, R.R., 2002. The role of mathematical modelling in the control of the 2001 FMD
653 epidemic in the UK. Trends Microbiol. 10, 279-286.

654 Kao, R.R., Danon, L., Green, D.M., Kiss, I.Z., 2006. Demographic structure and pathogen
655 dynamics on the network of livestock movements in Great Britain. *Proceedings of*
656 *the Royal Society of London B* 273, 1999-2007.

657 Kao, R.R., Green, D.M., Johnson, J., Kiss, I.Z., 2007. Disease dynamics over very different
658 time-scales: foot-and-mouth disease and scrapie on the network of livestock
659 movements in the UK. *J R Soc Interface* 4, 907-916.

660 Keeling, M.J., Eames, K.T.D., 2005. Networks and epidemic models. *J R Soc Interface* 2, 295-
661 307.

662 Kiss, I.Z., Green, D.M., Kao, R.R., 2006. The network of sheep movements within Great
663 Britain : network properties and their implications for infectious disease spread. *J R*
664 *Soc Interface* 3, 669-677.

665 Lal Dutta, B., Ezanno, P., Vergu, E., 2014. Characteristics of the spatio-temporal network of
666 cattle movements in France over a 5-year period. *Prev Vet Med* 117, 79-94.

667 Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., Getz, W.M., 2005. Superspreading and the effect
668 of individual variation on disease emergence. *Nature* 438, 355-359.

669 Madex, F., Seegers, H., 2010. Epidemiology of infectious disease in livestock raised in
670 intensive farming systems. In: Lefèvre, P.-C., Blancou, J., Chermette, R., Uilenberg, G.
671 (Eds.), *Infectious and Parasitic Diseases of Livestock*. Lavoisier, Paris, 34-52.

672 Martínez-López, B., Perez, A.M., Sánchez-Vizcaíno, J.M., 2009a. Combined application of
673 social network and cluster detection analyses for temporal-spatial characterization
674 of animal movements in Salamanca, Spain. *Prev Vet Med* 91, 29-38.

675 Martínez-López, B., Perez, A.M., Sánchez-Vizcaíno, J.M., 2009c. Social network analysis.
 676 Review of general concepts and use in preventive veterinary medicine.
 677 Transboundary and Emerging Diseases 56, 109-120.

678 Molina, C., 2012. La gestion predial como herramienta de apoyo para la toma de decisions.
 679 Plan Agropecuario.

680 Mweu, M.M., Fournié, G., Halasa, T., Toft, N., Nielsen, S.S., 2013. Temporal characterisation
 681 of the network of Danish cattle movements and its implication for disease congrol:
 682 2000-2009. Prev Vet Med 11, 379-387.

683 Natale, F., Giovannini, A., Savini, L., Palma, D., Possenti, L., Fiore, G., Calistri, P., 2009.
 684 Network analysis of Italian cattle trade patterns and evaluation of risks for potential
 685 disease spread. Prev Vet Med 92, 341-350.

686 Nöremark, M., Hakansson, N., Sternberg Lewerin, S., Lindberg, A., Jonsson, A., 2011.
 687 Network analysis of cattle and pig movements in Sweden: measures relevant for
 688 disease control and risk based surveillance. Prev Vet Med 99, 78-90.

689 OIE, 2001. HANDISTATUS II. FMD report 2000 and 2001. . World Organization for Animal
 690 Health, http://web.oie.int/hs2/zi_pays.asp?c_pays=196&annee=2000

691 OIE, 2014. Terrestrial Animal Health Code. OIE (World Organization for Animal Health),
 692 Paris.

693 Ortiz-Pelaez, A., Pfeffer, D.U., Soares-Magalhães, R.J., Guitian, F.J., 2006. Use of social
 694 network anlysis to characterize the pattern of animal movements in the initial
 695 phases of the 2001 foot and mouth disesae (FMD) epidemic in the UK. Prev Vet Med
 696 76, 40-55.

697 Rautureau, S., Dufour, B., Durand, B., 2011. Vulnerability of animal trade networks to the
698 spread of infectious diseases: A methodological approach applied to evaluation and
699 emergency control strategies in cattle, France, 2005. *Transboundary and Emerging*
700 *Diseases* 58, 110-120.

701 Robinson, S.E., Everett, M.G., Christley, R.M., 2007. Recent network evolution increases the
702 potential for large epidemics in the British cattle population. *J R Soc Interface* 4,
703 669-674.

704 Salathé, M., Jones, J.H., 2010. Dynamics and control of diseases in networks with community
705 structure. *PLoS Computational Biology* 65, e1000736.

706 Scudamore, J.M., Harris, D.M., 2002. Control of foot and mouth disease: lessons from the
707 experience of the outbreak in Great Britain in 2001. *OIE Scientifique et Technical*
708 *Review* 21, 699-710.

709 Turner, J., Bowers, J.A., Clancy, D., Behnke, M.C., Christley, R.M., 2008. A network model of *E*
710 *coli* O157 transmission within a typical UK dairy herd: the effect of heterogeneity
711 and clustering on the prevalence of infection. *Journal of Theoretical Biology* 254, 45-
712 54.

713 Valente, T.W., 2010. *Social Networks and Health: Models, Methods, and Applications*.
714 Oxford University Press Oxford.

715 Vernon, M.C., Keeling, M.J., 2009. Representing the UK's cattle herd as static and dynamic
716 networks. *Proceedings of the Royal Society Series B* 276, 469-476.

717 Volkova, V.V., Howey, R., Savill, N.J., Woolhouse, M.E.J., 2010a. Potential for transmission of
718 infections in networks of cattle farms. *Epidemics* 2, 116-122.

719 Volkova, V.V., Howey, R., Savill, N.J., Woolhouse, M.E.J., 2010d. Sheep movement networks
 720 and the transmission of infectious diseases. PLoS One 5, e11185.
 721 Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications.
 722 Cambridge University Press Cambridge.
 723 Webb, C.R., 2006. Investigating the potential spread of infectious diseases of sheep via
 724 agricultural shows in Great Britain. Epidemiology and Infection 134, 31-40.
 725 Woolhouse, M.E.J., Dye, C., Etards, J.-F., Smith, T., Charlwood, J.D., Garnett, G.P., Hagan, P., Hii,
 726 J.L.K., Ndhlovu, P.D., Quinell, R.J., Watts, C.H., Chandiwana, S.K., Anderson, R.M.,
 727 1997. Heterogeneities in the transmission of infectious agents: implications for the
 728 design of control programs. Proc. Natl. Acad. Sci. USA 94, 338-342.
 729 Woolhouse, M.E.J., Shaw, D.J., Matthews, L., Liu, W.C., Mellor, D.J., Thomas, M.R., 2005.
 730 Epidemiological implications of the contact network structure for cattle farms and
 731 the 20-80 rule. Biol. Lett. 1, 350-352.
 732
 733

FIGURE CAPTIONS

Figure 1. Boxplots of (a) in-degree, (b) out-degree, (c) number of animals received, and (d) number of animals sent by production type. Middle bars in the boxes indicate median values, and the top of the boxes indicate the 75th quantiles. Whiskers denote the 95th quantiles and triangles represent the mean values. All production types were significantly different for in-degree, out-degree, and total animals received (Kruskal-Wallis test, $p < 0.01$). Production types with the same number denoted above the bars in (d) indicate types that did not significantly differ in total animals sent.

Figure 2. (a) Betweenness and neighbourhood sizes for (b) inward movements and (c) outward movements by production type. Middle bars in the boxes indicate median values, and the top of the boxes indicate the 75th quantiles. Whiskers denote the 95th quantiles and triangles represent the mean values. Production types with the same number in 2a indicate types which do not significantly differ in betweenness. All production types significantly differ in in- and out-neighbourhood size (Kruskal-Wallis test, $p < 0.01$).

Figure 3. Examples of the neighbourhoods of four focal nodes. The focal nodes (squares) were drawn randomly from the four most common production types: a) dairy, b) breeding, c) complete cycle, and d) fattening. Example networks include all farms reachable within two steps, taking into account both the direction and temporal sequence of movements.

Figure 4. Seasonality in (a) the total number of animals moved per month and (b) the size of the Giant weakly connected component (GWCC) in seasonal networks. The legend in (a) applies to both figures.

Figure 5. Distribution of geographic distances between farms. Each line represents pairs of farms linked in the network with different path lengths.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5