The Measure of the Archive: The Robustness of Network Analysis in Early Modern Correspondence

Yann C. Ryan and Sebastian E. Ahnert

Yann C. Ryan. Department of English, Queen Mary, University of London, Mile End Road, London E1 4NS, United Kingdom. y.c.y.ryan@qmul.ac.uk.

Sebastian E. Ahnert. Department of Chemical Engineering & Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge CB3 0AS, United Kingdom, and The Alan Turing Institute, British Library, 96 Euston Road, London NW1 2DB, United Kingdom.

ABSTRACT

Network analysis of historical correspondence can be a fruitful way to address historical research questions, and has been increasingly used in historical studies over the past decade. As with many areas of quantitative humanities research, the reliability of the results are often called into question, given that such approaches require 'hard data' as input, yet almost inevitably use datasets with partial or missing records. Other disciplines using network analysis have conducted robustness experiments designed to test the impact of data loss or error on their results. In order to test how this missing data might affect our own area of research, we conducted a number of experiments designed to simulate the impact of the kinds of loss often seen in historical correspondence data, including random document loss, missing years, and errors in the disambiguation and de-duplication process. The results show that most network centrality measures maintain robustness until a very large proportion of the data (60% or more) is removed. Some measures showed a linear change in robustness, while others remained high and then fell off sharply. Only one, transitivity (local clustering coefficient) was significantly impacted throughout. We tested a range of data loss scenarios (random single letters, folio books of manuscript letters, catalogues, and entire years) and a range of commonly-used network metrics. In addition, we tested the robustness of more complex network analysis results in the literature that combine several network metrics to highlight individuals in the network, and found that the same types of individuals would have likely been highlighted even with 50% random letter loss. Alongside the article is a web application, built using Shiny, which will calculate robustness measures for a user-uploaded network dataset. We conclude that researchers working with similar historical correspondence datasets might be able to consider network analysis results to be robust in most cases, rather than work on the assumption that missing data would lead to very different findings or results.

The archives that historians work with are filled with traces of damaged, lost, and destroyed letters, which, had they survived, might have profoundly changed established historical narratives. Sometimes these archival gaps are created during the lifetime of the individual in question: fleeing persecution in Catholic-run Bohemia, the theologian, pedagogue and minister Jan Comenius took refuge in

Leszno in Poland, where he administered a school and was made the leader of the Moravian and Bohemian churches. In 1656, during the Swedish invasion of Poland, Comenius had declared his support for the Swedish side. In retaliation, the Polish Catholic partisans burned down Comenius's town, including his school. According to John Pell, both Comenius and the town's inhabitants had resolved not to remove any goods from the town but rather to weather the coming storm. This was a mistake, wrote Pell in a letter to Samuel Hartlib: the army burned Lesna to the ground, along with Comenius' writings and even the town's archives. In the letter he wrote:

> They did not fear that he would abandon them as long as his books and writings were not sent from thence. Thus they have lost both his manuscripts and their own records &c which might have deserved an exception from their general resolution of sending nothing out of Lesna.[1]

This was not the first time Comenius had witnessed the destruction of his records: he had a similar loss in Fulnek, Moravia, in 1623, not to mention the less violent loss of documents associated with a life of exile.[2] This quotation above illustrates the contemporary perception of the value of written records, and the corresponding loss when they were destroyed. What might have been lost in this invaluable cache of letters, books and documents? It would almost certainly cause us to revise our understanding of Comenius's network, and the intellectual exchange of which he was a part.

In other cases, a correspondence archive is affected by missing data long after the death of the individual at its centre. Since the death of the Anglo-Prussian intelligencer Samuel Hartlib in 1660, users and custodians have, according to Leigh Penman, 'reorganised, subtracted from, and added to [Hartlib's] archive, fundamentally altering its physical and textual make-up.'[3] These additions and subtractions lead to the revision or re-writing of histories, in often substantial ways. Sometimes the additions are on a smaller scale, such as Noel Malcolm's

discovery and publication, in 2001, of six new letters from the French intelligencer Marin Mersenne, but even a single added connection might force us to re-visit our understanding of an individual's network.[4] In other cases, the data is partial rather than missing: a piece or group of correspondence data may be missing the name of a sender or recipient, or have some uncertainty because it is unknown whether the Julian or Gregorian calendars were used, for example.

Traditional histories are often rewritten or revised in the light of newly-discovered evidence, but this does not generally hold back existing scholarship: it is rarely argued that one should avoid a topic because there might exist, in unknown quantities, some undiscovered part of their archive—a comment which is sometimes levelled at histories written using quantitative methods. We might think that quantitative measurements, based, as they necessarily are, on seemingly immovable 'hard data', are more at the mercy of missing or uncertain parts of archives. Is this the case in practice? In this paper we explore these effects through one such quantitative method: historical network analysis using correspondence data. What effect might events like those above—leading to the destruction or discovery of letters—have on the quantitative methods and results we use in our writing of histories today? To what extent—if any—should we be cautious of conclusions drawn from quantitative results, given that most archives are partial?

Historical network analysis

Recent years have seen a growth in the use of historical data to construct and analyse complex networks. These networks are constructed from various historical data. The Six Degrees of Francis Bacon project uses the co-occurrence of individuals in entries of the Oxford Dictionary of National Biography (ODNB) to create a probabilistic social network of figures between 1500-1700, and employs degree scores to find influential individuals who do not have their own separate entry in the source material.[5] Ladd (2021) has constructed networks of those associated with the early modern book trade by looking at co-occurence

in book dedications.[6] In many other studies, historical correspondence has been used as the basis for network construction.[7] Some of the consequential analysis of these datasets rests on network visualisation, which typically takes the form of the 'force-directed' network diagram. While such diagrams can provide useful insights for smaller networks they quickly become unhelpful 'hairballs' as the size of the network increases. A more quantitative approach is to use a set of network metrics designed to understand the 'centrality' of a given actor within a network, often with the aim of understanding relationships, roles, or influence.[8]

These metrics operate across a spectrum of complexities. The simplest is the degree of a node, which is its number of connections. This metric can highlight well-connected nodes. There are other more complex metrics which can be used. Ahnert and Ahnert use betweenness centrality and eigenvector centrality and compare them to the degree, in order to reveal hidden influencers, who were not necessarily well-connected nodes but who bridged communities and exerted their influence in other ways.[9] Betweenness centrality considers all shortest paths between two nodes in the network and counts how often a given node or edge lies on these shortest paths. Eigenvector centrality recursively scores a node on how well it is connected to well-connected nodes: individuals which 'have the ear' of a powerful person may also exert influence in a network. In a follow-on project, Ahnert and Ahnert take an additional five key network measurements to create a 'profile' for each individual, which can then be clustered together to find those with similar roles and even predict likely spies or intelligencers.[10]

What these studies have in common is that they all, by the nature of their sources, make inferences based on a partial perspective of the network, which results from partial or fragmentary data. Critique around the importance of understanding partial data in the digital humanities is not new, but there are relatively few studies which actually seek to ascertain its impact.[11] One might think that network analysis would be particularly sensitive to missing data: for example, a measurement of a node's betweenness centrality, which relies on the ability to measure unbroken paths through a network, could potentially produce very

different results if a single node in a crucial structural position were removed. Here we model the removal of archival data in a variety of simulated ways from three large correspondence datasets, and measure the sensitivity of the results to this removal. This follows on from work in other areas using complex networks, namely archaeology and the social sciences. We find that the patterns in our datasets broadly align with other disciplines, and suggest that many of the anxieties surrounding the use of metrics in historical network analysis may be unfounded.

A common misconception is that the incompleteness of network data is particularly pertinent in the context of historical data. Network analysis is applied in a vast range of interdisciplinary settings and data sets, from neuroscience, ecology, and molecular biology to computer science, physics, and engineering. Only in very few circumstances is the analysed network complete and accurate. Most of the time connections are either inferred from noisy data or derived from a partial snapshot of the system. Nevertheless network analysis has provided many useful insights into these systems. This is because some results are not affected by missing data, some are affected but can still be interpreted usefully in the light of missing data, and lastly, some results in fact tell us more about the biases and gaps in the data. Our aim in this paper is to establish the extent to which the results of quantitative network analysis are affected by absent data in the particular context of historical correspondence networks.

Missing data

Missing network data is a common problem in many other fields. Archaeologists are often confronted with highly fragmented data, yet the use of network analysis in their field is established and growing. Leidwanger *et al.* outlined some of the pitfalls of network analysis in archaeology, warning that it can lead to an illusion of objectivity.[12] Archaeologists often use objects as proxies for relationships between nodes: for example two nodes, island communities, say, might share an edge if the same style of pottery is found in each. This can lead

to networks based on very partial data, but it has been shown that most metrics in these networks are robust to node removal.[13]

Social network data is also often incomplete, because of 'boundary specification' problems (difficulties in deciding the boundary of a network), non-response to surveys, or inaccuracy.[14] In order to evaluate the effect of this incompleteness, Galaskiewicz (1991) compared different types of sampling techniques by evaluating their effect on the in-degree of nodes and on 'popular' versus 'unpopular' actors in the network, and found that some results remained largely unchanged when data was removed, and that the choice of sampling technique did not have much of an effect.[15] More recently Smith and Moody measured six metrics across twelve datasets and found that some measurements were particularly sensitive to missing data, and that large centralised networks were more robust.[16] In a second study, the effect of non-random subsamples on the network metrics is considered, and it is shown that the removal of more central nodes has a larger effect in general, again with some dependence on the particular metric and the network type.[17]

The advantage of sampling in social network analysis is clear - it can reduce the time or cost of a study by reducing the number or length of interviews, for example, and as such measuring its impact on results is seen as crucial. Costenbader and Valente used bootstrap sampling—a method for estimating a statistic by continually resampling and replacing observations in a dataset— of survey responses to measure the stability of 11 measurements of centrality on 59 different networks, and found that although there were variations across networks, in general in-degree (a count of a node's incoming connections) and eigenvector centrality scores were relatively stable even when large portions of the networks were removed.[18]

Compared with the humanities and social sciences the natural sciences can sometimes give an impression that data incompleteness is less of a problem, due to the control that scientists have over the design and execution of particular experiments. This is illusory however, specifically in the context of network

science, which aims to study a wide range of real-world networks, most of which come in the form of highly incomplete and often unreliable datasets. The effect of missing data on network metrics has therefore also received attention in the sciences, and particularly in biology, which often aims to infer networks of physical interaction between proteins as well as networks of regulatory interactions between genes (among other types of networks). This network data is highly incomplete, but in many cases also contains spurious links that do not exist in the living cell. This is because protein-protein networks and gene regulatory networks are often inferred from circumstantial evidence, rather than direct measurements. For example, a DNA 'promoter' sequence close to a particular gene may imply that the 'transcription factor' protein of another gene is able to bind to the DNA at this position, representing a regulatory interaction between these two genes, but this binding event may never actually happen in the cell because the promoter sequence may be physically inaccessible due to the spatial organisation of the DNA. Because of these kinds of uncertainty, the edges of biological interaction networks are typically assigned confidence scores. This poses another problem, namely that such 'weighted' networks are challenging to analyse.

Most network metrics can be generalised to weighted networks, but the weighted counterparts are often difficult to interpret. As a result weighted networks are often subjected to a threshold in order to turn them into unweighted networks that lend themselves more readily to conventional network metrics. The choice of threshold however is somewhat arbitrary, which is why recent work examines the consequences of this choice of threshold and the resulting variation in the amount of missing data in these networks.[19] The study finds that some measures (including degree and PageRank, a form of eigenvector centrality) are robust and vary little for different thresholds, others (betweenness) are moderately robust, and again others (local clustering coefficient) are highly sensitive. The authors conclude that a subset of network metrics yields similar results for a variety of thresholds and recommends the use of these when dealing with this particular type of biological network data.

In order to study the effect of missing data on the quantitative analysis of historical correspondence networks, neither the removal of nodes nor the introduction of thresholds are the most informative ways to consider incomplete networks. Node removal is unlikely to be useful, because large correspondence networks are very likely to contain at least some data for 'popular' nodes - even if their own archives have not been digitised or are not available. An example of this is the archive of John Thurloe, who between 1651 and 1660 was a highly influential figure in Oliver Cromwell's government, and Secretary to the Council of State during the Protectorate. His personal archive has not been added to a centralised repository, yet because letters by him appear in a number of other archives, he nevertheless features in other quantitative correspondence network studies and even ranks highly in some network metrics. Correspondence networks are weighted networks if we regard the number of letters sent from one person to another as a weight of that directed connection. Thresholding may therefore in principle offer some insight into incomplete networks, but it does not represent a particularly realistic representation of the kind of data loss found in historical correspondence archives. In order to model the effect of missing data in historical correspondence networks, then, the sampling methods we employ must emulate the varied reasons why particular letters or letter collections are absent from the historical record.

Networks based on historical correspondence archives have their own patterns of missing data. Often they contain very large numbers of letters between particular pairs of individuals. Missing data will either be the result of the inevitable random loss of individual documents or be of particular events that lead to a more systematic absence: the missing data may consist of entire folios (the large volumes into which correspondence is often bound for preservation) that are missing, or a catalogue of correspondence (here meant as a discrete, self-contained set of correspondences, often containing all the collected letters of a single individual) which has not been digitised yet. If large numbers of letters are destroyed on purpose, this will likely happen in a systematic way. If a person or an institution archives their correspondence in date order in boxes or folios,

entire blocks of years might be lost, rather than just a random set of letters. In other cases, letters or catalogues may be added to the existing archive, which raises the question of how much we should infer from existing archives, given that more data may be added in the future.

This paper takes various likely patterns of 'missingness' into account, and models missing nodes, letters, folios, and catalogues. Unlike social network data, calculating missing historical correspondence data cannot be done by inferring completeness from the 'response rate' of a survey, but rather must work on the assumption that the correspondence we have is incomplete, and a subset from a much larger body. In some cases, historical network data has been inferred using statistical text mining of biographical source material, and is therefore inherently probabilistic. The project mentioned above, Six Degrees of Francis Bacon, uses the co-occurence of individuals in ODNB entries to infer relationships between historical actors, and as such is likely to be missing many ties as well as inferring some relationships that did not in fact exist.

As discussed above, historical network data is at the mercy of collection, archival and digitisation practices, and as such its 'missingness' has a particular structure. Historical correspondence networks may be incomplete because individual letters have been lost, because entire folios or even catalogues have not been digitised, or because whole years of material are absent, having been destroyed or never been archived as a result of of revolutions, wars, or sieges. Sometimes political disruption is the catalyst for a transfer rather than an absence in archives. Correspondence held by the Bodleian between 1644—1649, and forming the Bodleian card catalogue data used here, contains substantial correspondence for the Civil War period which otherwise would have likely ended up in the State Papers. This paper attempts to reflect the effect of these historical contingencies by simulating the removal of data along these categories through sampling methods that mirror the different types of absences found in our data. We find that in large historical correspondence datasets the robustness of many widely used network measures is high even when large random samples are removed, and that these results are largely independent of the specific

type of sampling.

Method

Datasets

We study three large early modern historical correspondence networks, which rely on three collections that were amassed in very different ways, and for very different reasons. The first of these, Early Modern Letters Online (EMLO), is a collection of individual catalogues of correspondence—currently 135 in total—contributed to EMLO by hundreds of academic scholars over the past twelve years.[20] Most of these are the correspondence of a single individual, with one major exception, the Bodleian card catalogue (BCC), which we separate out, as a second dataset. This card catalogue was the product of work by two twentieth-century employees and one volunteer in the Bodleian and ultimately based on individual and idiosyncratic acquisitions by the Library over time, resulting in an 'ad hoc' and 'iterative' set of metadata.[21] EMLO and the BCC span almost four centuries of European history between 1500 and 1900, but most of the data in EMLO covers the seventeenth and eighteenth centuries, and British and Dutch diplomatic and intellectual history in particular. Our third dataset is the State Papers of the Stuart period of British history (1603-1714), derived from the nineteenth century catalogues (called 'calendars') of the collection, which were digitised by the company Gale. The State Paper Office was established in 1610 with the aim of collecting the English State's private and working manuscripts in a single place, and was to become the principal archive and working library for the parliamentary executive, while essentially being the private papers of the monarch.[22] We might therefore expect this 'official' record of the English State to present a more unified or coherent worldview than EMLO. In fact, the State Papers are also full of partial or shifting perspectives: individual secretaries often viewed their official documents as 'private' and kept them as their possessions on leaving office.

Together, these represent about 320,000 individual pieces of correspondence, and 60,000 individuals, spread over a timespan of four hundred years (table 1). Both EMLO and SPO have had extensive reconciliation and disambiguation carried out: The Networking Archives project team uses a custom-designed 'disambiguation engine' for the State Papers, built as an aid to manual disambiguation. With the help of this tool and two years of work by team members, more than 54,000 name strings have been reduced to 30,000. All letters added to EMLO are processed through EMLO Recon - bespoke software which suggests matches to existing people records in the dataset.

Despite this work, and the size of the datasets, both EMLO and SPO should be considered partial. State Papers Online, for example, contains very little from the State Papers: Foreign series past 1660, and the number of documents for the years of the English Civil Wars (1642-1651) are minimal, as Charles II left London and moved the centre of his administration to Oxford and York. In other cases data is 'missing' even though it has survived, because it has not yet been digitised. The catalogue of Johan de Witt, though one of the largest in EMLO, is still embryonic and is currently only one-fifth of his total surviving correspondence. [23] In other cases, such as that of Jan Comenius, mentioned in the introduction, the majority of an archive has been destroyed and will never be recovered. The correspondence data may be imbalanced, too: archives, historically, are usually a collection of an individual or family's incoming letters. Where large numbers of outgoing letters are found in a catalogue, it is usually the result of a modern effort to 'reassemble' a single individual's correspondence. A large connected dataset such as EMLO mitigates this imbalance to a certain extent, as letters sent by one individual may be found in letters received of another, but it is another source of data 'loss' to be noted.

For the purposes of the experiments carried out in this article, as well as on the *Networking Archives* project more generally, all three of these datasets were converted into directed, unweighted networks, with each author/recipient pair represented as an edge. Letters with multiple authors and recipients were separated out into multiple edges. Though the archives often contain many mul-

Table 1: Network Summaries

	EMLO	BCC	SPO
Letter Records	100,277	54,620	174,278
Nodes	13,333	9,747	32,269
Unique Edges	21,678	12,705	56,850
Diameter[*]	15	18	13
Avg path length	4.38	5.55	4.02
Clustering Coefficient	0.00553	0.00519	0.00575
Earliest date	1502	1508	1485
Latest date	1853	1900	1793

[*] Global metrics were calculated on an undirected version of each network

tiples of letters between pairs of authors and recipients, the use of unweighted networks was a deliberate choice: based on the assumption that rankings which depend on the existence of an edge rather than its strength are less susceptible to error in the case of missing letters. Though a study of weighted network robustness would likely be informative, it is outside of the scope of this paper.

Despite their different origins, the datasets exhibit some surprising similarities. The distribution of the number of connections, or 'degree distribution', follows a similar pattern in all three, which can be described by a power law, indicated by a straight line when plotted on a double logarithmic scale (figure 1). The networks are said, then, to be 'scale free': meaning that a similar pattern, of a small number of highly-connected nodes and a large number of nodes with very few connections, can be observed in any given sub-section of the network. The implications of scale-free networks have been discussed at length elsewhere, but this structure is also important for understanding the ways in which different categories of measures are affected by missing data.[24]
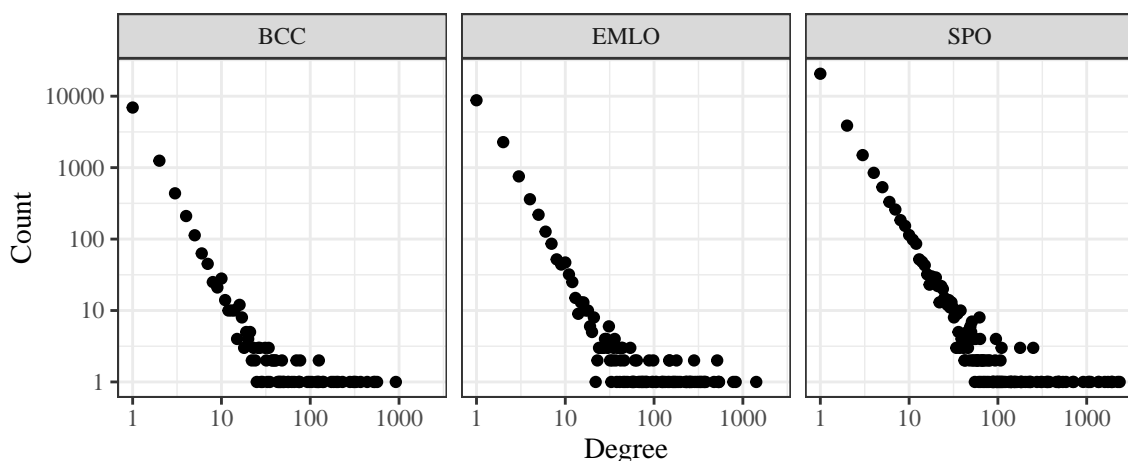
Figure 1: Degree distribution plot for each of the networks. The downward-sloping diagonal line on a log-log axis plot is an indication that the networks are scale-free.

Sampling techniques

For each of the three networks, we created subsamples of the network by removing a) letters, b) nodes, c) years, d) folios, and e) entire correspondence catalogues, reflecting common forms of absences in the historical record, as discussed above. In addition to this we performed a final experiment, which simulated another likely source of data error, namely erroneous de-duplication or disambiguation. In the first five cases we removed a particular portion of the entity in question, e.g. 10% of letters, or 30% of catalogues, rounding up where necessary. The samples were produced in 1% steps, from 99% to 1% of data remaining. In the final experiment the disambiguation and de-duplication were undone for subsections of the data, in steps of 10% (due to the considerable computational cost). As the volume of correspondence for individual years and in individual folios and catalogues varies considerably we expected larger fluctuations between samples in these cases than when removing a randomly selected set of individual letters.

Next, we compared the values of a set of standard metrics (total degree, betweenness centrality, eigenvector centrality, closeness centrality, and transitivity) in the full network to the equivalent values in the sampled networks for each node, disregarding any nodes that did not appear in the sample, similar to previ-

ous approaches in the literature.[25] To compare the values we used Spearman's rank correlation (referred to ρ hereafter), because rankings are a more useful way for interpreting network metrics in many contexts, as absolute values may a) fluctuate due to larger historical developments and changes in archival practice, and b) are difficult to interpret in the case of betweenness centrality, closeness centrality, and eigenvector centrality, because the absolute values for the highest-ranking nodes can be orders of magnitude larger than the lowest-ranking for these measurements. In all cases, the relationship between the original and sample values is monotonic and therefore appropriate for a Spearman's rank correlation. We collected 100 independent samples at each 1% interval, from 99% to 1% of the full network, for each category of removed entity (letter, folio, year, catalogue, and node). This process was repeated 40 times to get a realistic average value and to measure variability. In other words, we simulated progressively increasing amounts of missing data, and measured how this affected quantitative results.

Results

We find that the measures are remarkably robust to many types of data removal, across all of the networks. The figures 3 to 7 below display some of the key findings, and full results are in summarised form in table 2. In each case we display the changes in correlation as larger parts of the network are removed as well as the variability of that change. The former is visualised in the charts below as a single blue line, representing the mean correlation for each sample (from 99% to 1%) and the latter is measured as the standard deviation of the forty observations for each sample, visualised as a gray shaded area. In general, most Spearman correlations remained high (many with a ρ above 0.7) until 50% of the network was removed. Some metrics and sampling methods showed more variability than others, which is shown here through the standard deviation: for example, sampling catalogues resulted in high variance (a standard deviation of 0.328, where the possible values ranged from -1 to +1), and to a lesser extent, sampling years and nodes did too. Most network measures were robust to

letter removal, and showed very little variability. This may be because of the particular structure of historical correspondence networks. In general, many connections between pairs of individuals will be marked by many letters, written over a number of years. This means that the most prominent connections in the network are also the most robust ones with regard to letter removal. Node removal produced similar results, though it resulted in more variation in measures that are calculated by using the entire network (closeness and eigenvector centralities). Surprisingly, even removing entire catalogues—essentially removing at random some of the top-scoring degree nodes—had little effect on the degree of the remaining nodes though it did have a more substantial impact on closeness and eigenvector centralities. We find that one metric, local clustering, or transitivity, is consistently sensitive to the removal of any type of data. This highlights that some research questions will be affected by incompleteness much more than others. It also underlines how remarkably robust most metrics actually are: the sensitivity of local clustering is the sensitivity that skeptics of quantitative analysis might expect to see across the board. We find that it is the exception that proves the rule, as local clustering demonstrates that some network metrics *may* be very sensitive to data removal, but most are *not*.

Letter Removal

Letter removal had surprisingly little effect on any of our three networks - the correlations stayed high (ρ above 0.6 in all three datasets) even when just 10% of the total letters remained, and the variation between random samples was low (figure 2). This is probably because of the aforementioned nature of correspondence data, in which many network edges are marked repeatedly, by many letters. Removing random letters when there are large numbers of them between two people makes it very unlikely the corresponding network edge disappears completely. There were some subtle differences between the metrics, in two broad patterns: some measurements, chiefly degree and betweenness centrality, became less correlated in an almost linear fashion as progressively larger parts of the network were removed, whereas eigenvector centrality showed very lit-

Table 2: Summary statistics for the removal correlations at fifty and ninety percent removed, showing the mean Spearman's rho (ρ) and standard deviation across forty iterations.

		Catalogues		Letters		Nodes		Years	
	Portion removed	Mean ρ	SD[*]	Mean ρ	SD	Mean ρ	SD	Mean ρ	SD
Betweenness Centrality[†]	50	0.772	0.035	0.794	0.009	0.865	0.019	0.792	0.031
Betweenness Centrality	90	0.508	0.067	0.556	0.027	0.568	0.021	0.551	0.042
Closeness Centrality	50	0.922	0.030	0.943	0.010	0.947	0.022	0.934	0.023
Closeness Centrality	90	0.669	0.175	0.789	0.038	0.717	0.091	0.716	0.095
In-Degree	50	0.917	0.017	0.909	0.009	0.965	0.006	0.913	0.016
In-Degree	90	0.751	0.059	0.724	0.012	0.845	0.034	0.744	0.032
Out-Degree	50	0.878	0.028	0.881	0.016	0.933	0.011	0.883	0.023
Out-Degree	90	0.630	0.087	0.637	0.046	0.684	0.038	0.638	0.059
Total Degree	50	0.843	0.030	0.832	0.010	0.903	0.011	0.839	0.023
Total Degree	90	0.636	0.064	0.650	0.022	0.663	0.033	0.646	0.030
Eigenvector Centrality	50	0.913	0.071	0.948	0.013	0.934	0.045	0.933	0.033
Eigenvector Centrality	90	0.639	0.224	0.810	0.054	0.643	0.183	0.706	0.147
Local Clustering Coefficient	50	0.739	0.044	0.734	0.014	0.775	0.044	0.736	0.044
Local Clustering Coefficient	90	0.397	0.126	0.337	0.029	0.298	0.059	0.380	0.068

[*] Standard deviation

[†] Explanations of this and other standard network terms can be found in a glossary at the end of this article.

tle difference until most of the network was removed, at which point it dropped substantially. In all cases, the standard deviation was small: the values generated by each iteration stayed very close to the mean.

Node Removal

We assume that this method of data removal for historical networks is the least reflective of real-world missing historical data—most of the 'important' actors within the networks are found across a number of archives, and it is unlikely (though not impossible, in some historical contexts) that an individual would be systematically erased across all archives, as both a sender and recipient of letters. There is existing literature which looks at the global tolerance of scale-free networks to both random error and coordinated attack, concluding that networks are robust to the former but vulnerable to the latter.[26] This may partially explain why when random nodes are removed the individual network metrics are also resilient. Degree and betweenness sensitivity to node removal are very similar
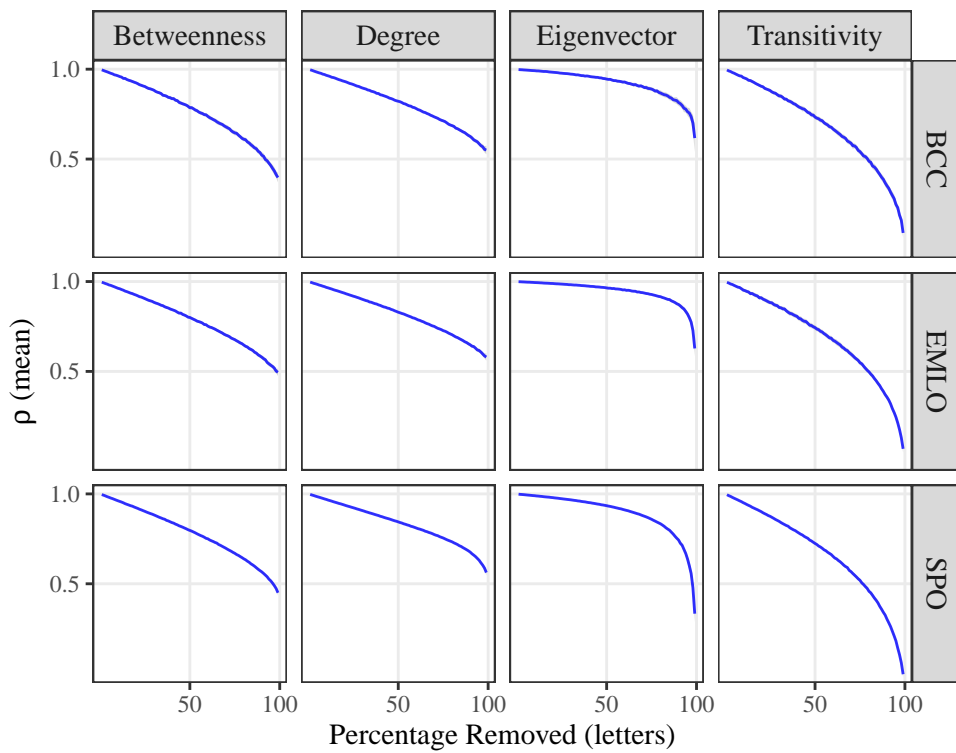
Figure 2: Robustness of Key Measures to Letter Removal. The blue line indicates the mean Spearman correlation (ρ) mean across all forty iterations. The standard deviation at each sample is represented as a shaded gray area, however it is usually not visible in this mode of removal because the variation between iterations was so small.
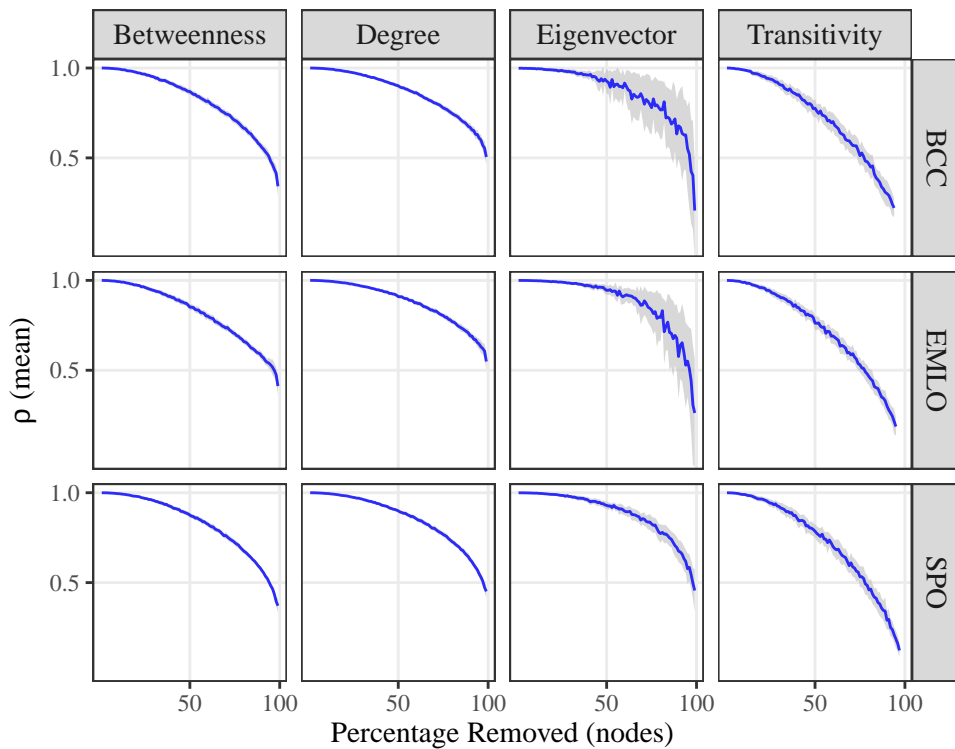
Figure 3: Robustness of Measures to Node Removal. The blue line indicates the mean Spearman correlation (ρ) across forty iterations. The standard deviation at each sample is represented as a shaded gray area. Generally, correlations remained high until 70% of the network was removed and then declined sharply. Degree and betweeness sensitivity were similar to letter removal but closeness (not shown) and eigenvector centralities displayed more variation.

to the case of letter removal, with subtle differences. Node removal correlations are very high (ρ above 0.8) until 70% of the network remains, and then decline more sharply than with letter removal (figure 3). For closeness and eigenvector centrality, there is much more variation in the sensitivity across each iteration, as increasing portions of the network are removed. This may be because eigenvector and closeness centrality are more dependent on high-degree nodes, which are removed in larger numbers in these simulations than when letters are removed. The variation in eigenvector centrality scores is particularly striking for the BCC and EMLO networks, but less so for SPO.

Catalogue removal

One of the three archives studied here, EMLO, is divided into individual cata-logues.[27] Each of these catalogues is generally the correspondence of a single individual, often collected by an individual scholar. Because EMLO is organ-ised, and new data is added to it, at a catalogue level, removing varying numbers of these catalogues may provide a realistic simulation of the impact of non-random missing (or added) data on historical scholarship. A common concern in historical network research is that a) any historical correspondence record is inevitably highly incomplete, and b) that many of the records that do survive have not yet been digitised, or even archived. In addition, the ongoing digi-tisation efforts mean that any analysed dataset will change over time as more correspondence is added to the digital archives. The question whether quantita-tive results obtained with current data will still hold after these future additions is therefore a further concern. The removal of catalogues can be used to exam-ine the validity of the above concerns. While the results reveal larger variation between the independent simulations (due to the broad distribution of catalogue sizes), all network metrics are remarkably robust up to 50% of removed cata-logues (ρ above 0.8 in all cases except transivity and betweenness centrality) and on average not much less robust than for letter or node removal even for 90% of removed catalogues (though in some simulations the correlation plum-mets for this percentage).

Folio Removal

When we think of missing archival data, perhaps the first image that comes to mind is missing folios from shelves - these folios, as material objects, may go missing, be borrowed and not returned, be unavailable for digitisation because of conservation concerns, or destroyed, by accident or deliberately. Two of the archives here are organised by folio. The Bodleian card catalogue data comes from just over 500 individual manuscript folios - each containing anywhere from one to five hundred pieces of correspondence and the State Papers are
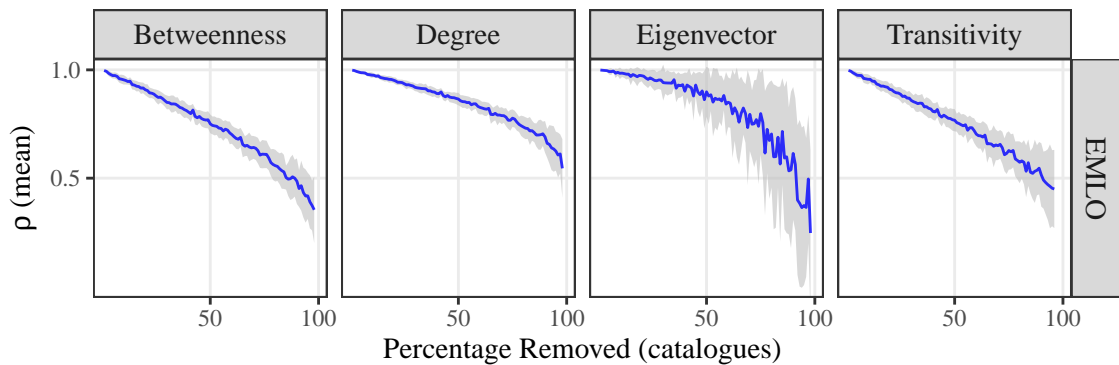
Figure 4: Robustness of Measures to Catalogue Removal in EMLO dataset. Only EMLO has the catalogue as an organisational unit and so is displayed separately here. Removing full catalogues still did not have a large effect on correlations, though more variation between iterations was displayed - most likely because catalogue vary by size. The blue line indicates the mean Spearman correlation (ρ) across all forty iterations. The standard deviation at each sample is represented as a shaded gray area.

organised similarly. Our algorithm simulates, essentially, walking the shelves of the archives at random and removing individual manuscript volumes, and calculates the effect on the resulting network measures.

When removing random samples of folios we find similar results to removing letters or nodes, with some results that highlights the organisation of the data. The eigenvector centrality measure of the BCC network shows high variability for random folio removal, which mirrors the pattern for node removal and may suggest that the way in which folios are organised affects the sensitivity of network measurements to random removal. For example, Bodleian card catalogue folios are more likely to contain the correspondence of a single individual, whereas in the State Papers, correspondence for important individuals is spread across a number of volumes. Historical network analysis, then, may be particularly sensitive to folio removal when an archive has been arranged by individual rather than, say, topic or date, though further investigation of this is needed.

Year removal

The three sample datasets, as seen in table 1, span large ranges of dates, sometimes three or four hundred years (though it should be noted that in all cases, the majority of records are found in a more limited timespan). Intuitively, it
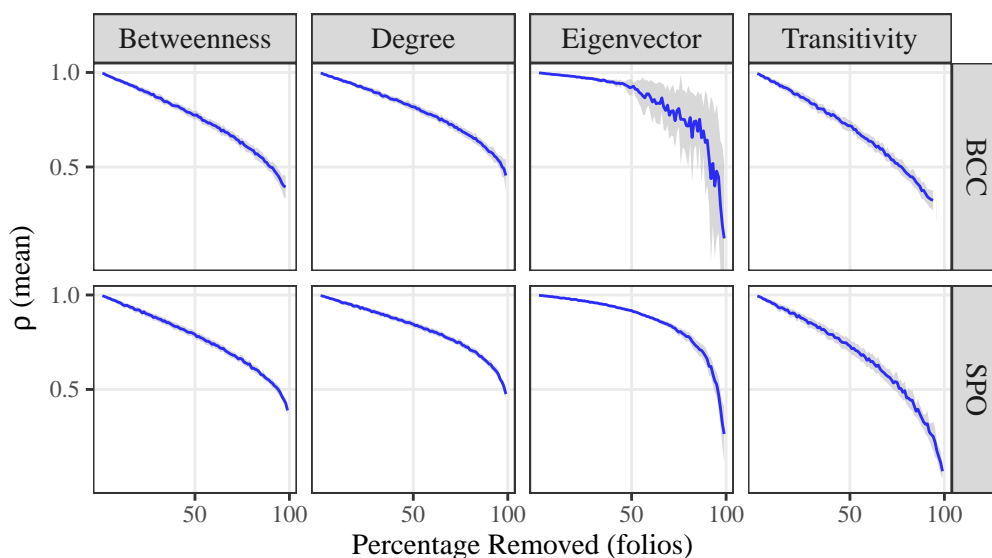
76

Figure 5: Folio Removal results for BCC and SPO datasets. Blue line indicates the mean correlation score across all forty iterations. The standard deviation at each sample is represented as a shaded gray area. SPO and BCC include information on the source folio, and we used this as a unit to remove in sampling. Results differed as a result of the type of material collected in these folios: BCC folios are often person-based, and therefore show more variability, whereas SPO folios are often archived by subject or chronologically, and seem to display less variability.

is entirely plausible that longitudinal historical datasets may be missing single or multiple years of data. EMLO, being a series of curated catalogues, is very much concentrated in a subset of the years it spans, but even a dataset like the State Papers has peaks and troughs across time. Diplomats often travelled for extensive periods of time, resulting in the dispersal of their correspondence. Joseph Williamson, the Under-secretary of State for the Southern Department of England between 1660 and 1674, travelled to Cologne to represent the State at a diplomatic conference for much of 1673 and 1674, and responsibility for incoming government correspondence passed to a clerk in his office.[28] Because of this, there is little correspondence involving Williamson for these years, and his dominance of the archive, due to his unprecedented personal effort in maintaining and archiving his working papers when in Whitehall, is such that this means there is a substantial dip in the volume of data during these years. In this case, the 'missing' data is in fact dispersed elsewhere and therefore not a part of the State's archive, but one can imagine a similar scenario where the material could be simply lost. Data for particular years can be missing or relatively

sparse for other more drastic reasons. During the English Civil Wars from 1642 to 1651, the complications arising from the split in the State's administration into parliamentary and royalist factions, compounded by the natural chaos of war meant that part of it moved its working papers elsewhere, and for these years again there is a substantial gap in the State Papers correspondence. A similar profile of variability in the volume of letters per year is also found in the EMLO dataset. Despite containing records spanning from approximately 1500 to 1800, the database coverage is uneven: two ten-year periods (1634-1644 and 1664-1673) contain 22.5% of the total volume of letters. Because historical datasets are often longitudinal, ranging over a span of decades or centuries, substantial gaps in the temporal coverage become more likely and might be thought to have a great deal of effect on the resulting network. Results for this type of removed data largely mirror those for letter removal, with more variability for some metrics (as some years contain much more correspondence than others).

Reconciliation and Disambiguation Errors

Another common source of data error or corruption found in correspondence archives is that which can be attributed to errors in names. Doing network analysis from correspondence relies on individuals with the same name being split where appropriate, and ambiguous or spelling variants being merged in other cases. One of the datasets used here has been cleaned as part of the project from which this work arises. In order to test the impact of network analysis on uncleaned or partially-cleaned data, we used the records of this process to produce versions of the network where increasing random portions of the data had been returned to its uncleaned state, and again ran a series of correlations between the original and sample data network metric ranks. Figure 7 (a) shows that the differences between the original and sample are broadly comparable to those found when removing letters: there is little variation between samples, and in most cases the trajectory is linear.

There is however an interesting difference between in/out degree and total degree scores - the robustness of the latter is slightly but noticeably lower (Figure
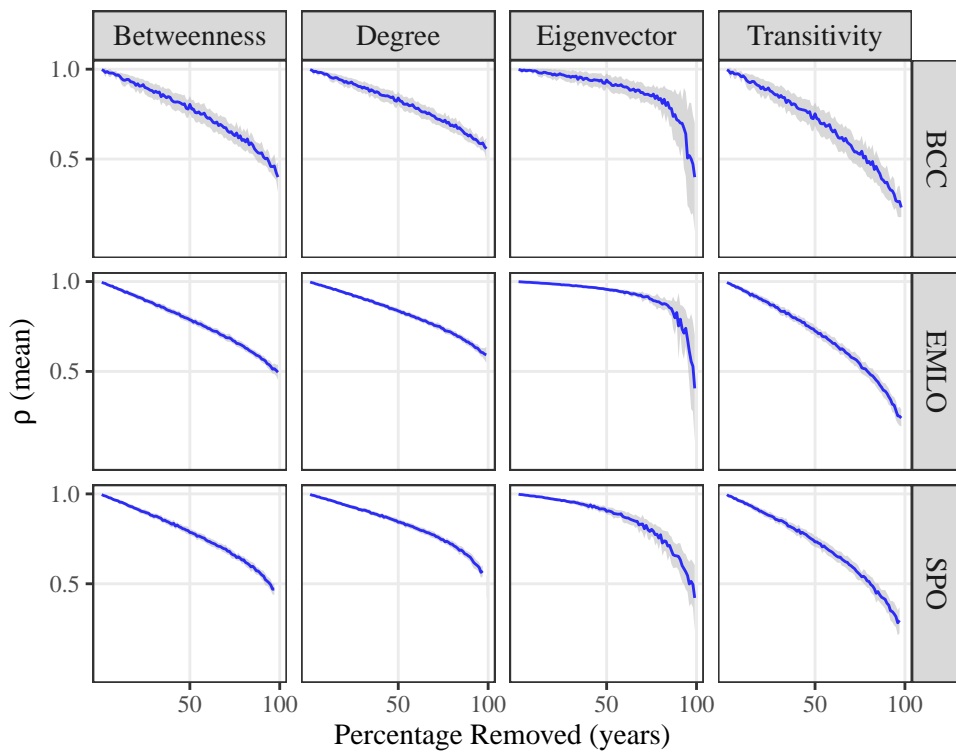
Figure 6: Robustness of Measures to Year Removal. The blue line indicates the mean Spearman correlation score across all forty iterations. The standard deviation at each sample is represented as a shaded gray area. The results mirrored those of letter removal, with some more variability due to variations in the amount of material per year.)
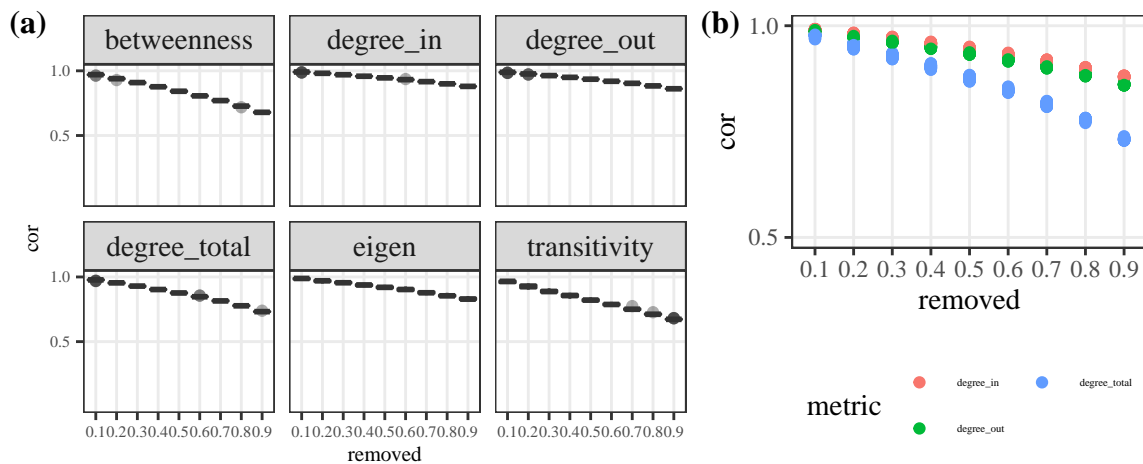
Figure 7: Robustness results from disambiguation error experiment. To calculate the impact of no or unfinished data cleaning, we reverted progressively larger samples of the data to its uncleaned state, using the log files generated by the data cleaning process. Results were similar to other methods of sampling. In (b), a noticeable difference between in/out degree and total degree robustness can be seen. This is because most of the data cleaning is merging, and often the secondary name variants are mostly either in or outgoing letters rather than a balance between the two.

7 (b)). This results from the nature of the uncleaned data and the shape of the State Papers archive and calendars. Much more of the data cleaning occurred in the form of merging variant names rather than splitting common names into multiple. When letters catalogued under a different name are found and merged a master record, they are often either incoming *or* outgoing—for example the master record for William Cecil, 2nd Earl of Salisbury is a mixture of incoming and outgoing letters, but those catalogued under his earlier title, Viscount Cranborne, are mostly incoming. When letters are merged, in/out degree scores seem to be affected less than total degree.

Effect on real-world results

How might the results described above impact a real-world analysis of an historical network? To understand the effect these numbers might have on actual findings, we used the same sampling strategy but looked at its effect on individual nodes. Two examples are presented here. First, a historical argument that the project team of Networking Archives are making with the help of this data rests in part on the specific rankings of three key Secretaries of State: Joseph
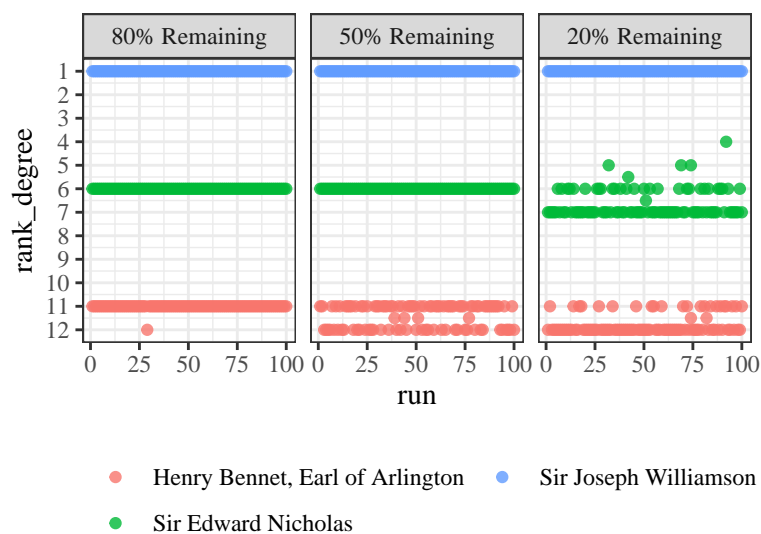
Figure 8: To demonstrate the effect of missing data on individual network results, we calculated the degree ranks of three Secretaries of State at 20, 50 and 80 per cent of the network remaining, run with 100 random samples. The rankings of all three stayed remarkably similar, only changing significantly when 80% of the network was removed.

Williamson, Lord Arlington, and Edward Nicholas. Using the same sampling methods as above, we calculated the ranks for these nodes with 80% and 50% of the network remaining, 100 separate times. Through this we see that the effect on basic network measures on some of the highest-ranking nodes was remarkably minimal: with 80% of the network removed the rank changes for all three were at most plus or minus 1 (figure 8).

In order to fully assess robustness we ran an experiment specifically tailored towards findings using lower-ranking individuals. Ahnert and Ahnert (2019) compared betweenness centrality and degree to highlight a group of often-overlooked bridging nodes: those whose importance lay not in their total connections, but in their capacity to bridge separate parts of the network together.[29] In that paper individuals were highlighted by plotting degree and betweenness centrality ranks and looking for those significantly below the trend line.

Doing this for SPO reveals James Butler, Duke of Ormond as a classic example of one of these bridges. From an 'Old English' family born in London, Butler bridged several networks both politically and temporally: not only was he a
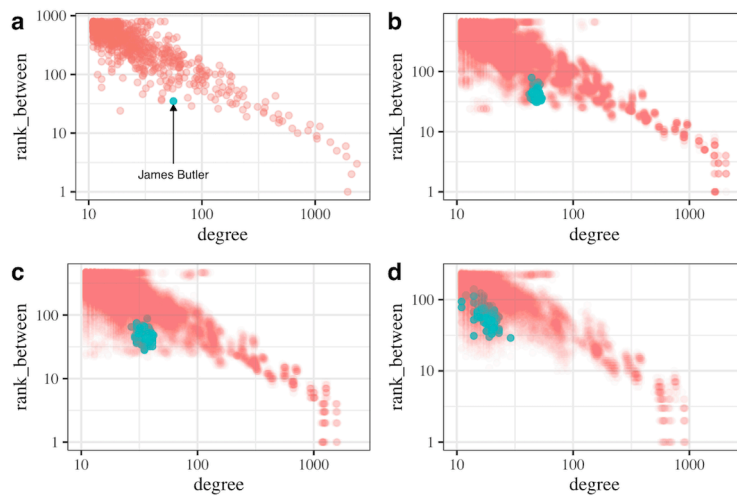
Figure 9: Assessing the effect of data loss on a downstream task. James Butler, Earl of Ormond can be described as a 'bridge' node in State Papers Online: neither his degree or betweenness centrality scores are within the highest-ranking, but his betweenness centrality is proportionately high in comparison to his degree. This can be seen from his position well into the lower half in a scatterplot of the rankings of the two measurements in (a), which plot the rank of his betweenness centrality against his degree, on a log-log plot. In (b), (c) and (d) we calculated his position with 20%, 50% and 80% of the network removed, respectively. This was done 100 times and the results plotted one over another, with Ormond highlighted. As can be seen here, he remains roughly in the same area of the scatterplot, except in some instances when 80% of the network is removed.

connecting link between the English and Irish nobility, he was one of the few high-ranking politicians with a significant career both before the Civil War *and* following the Restoration of Charles II in 1660.[30] Thus, despite his relatively low degree (ranked 226th), his betweenness centrality is proportionately high (ranked 38th), indicating his value as a 'bridge' in the network. This is clearly seen in the scatterplot below (figure 9).

To estimate the effect of missing data on Ormond's status as a 'person of interest' with this visual method, we again simulated removing 20%, 50% and 80% of the data, 100 times, and reproduced the scatterplot above for each run. The result shows that Ormond's position remains in the same average 'area' of the scatter plot each time, particularly when 20% or 50% of the network is removed. With 80% of the network removed, Butler's position as an outlier is consistent across most of the 100 runs.
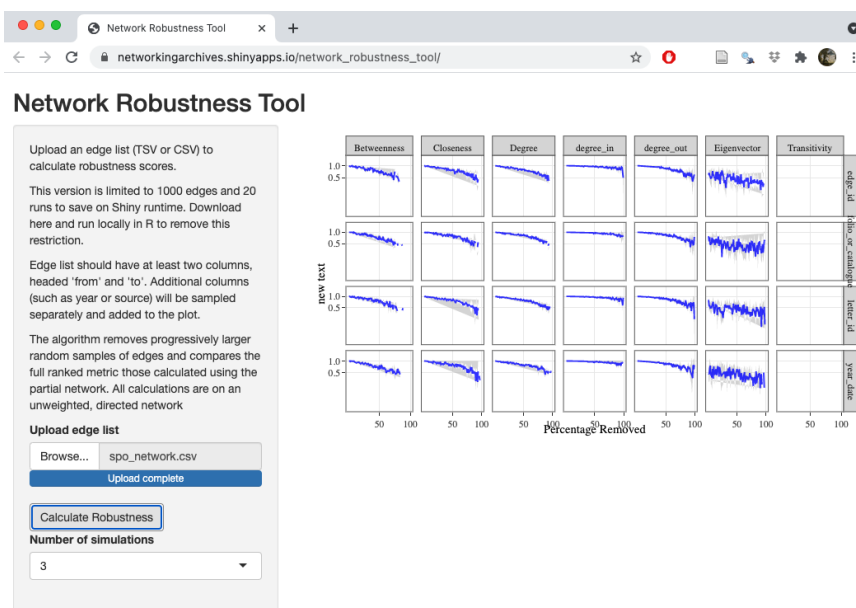
Figure 10: Screenshot of Network Robustness Application, made with Shiny

Shiny Application

To help other researchers working with historical correspondence archives to assess the impact of data loss, we have made available a user-friendly implementation of the code used in this article, which can be used to assess and compare the robustness of a set of network measures of any network.[31] This application, developed using R and Shiny, allows a user to upload a simple network structure—an edge list—and run the same analyses as in this paper, specifying the number of iterations to run (figure 10). If the edges have further attributes (for example folio names or other source information) the impact of removing samples of these attributes on robustness results will also be calculated by the application.

Conclusions

It is worth noting that the use of historical network analysis is multi-faceted and it should be stressed that the results that apply to the networks studied

here—based on correspondence data and with a 'long tail' distribution of letters weighted heavily towards a small number of nodes—may not apply to other sorts of network data such as co-occurence networks. Furthermore, many historical network studies have used global measurements such as assortivity, global clustering coefficients and so forth.[32] The robustness of these metrics and network types would require further research. Historical correspondence network data (as indeed most network data - see Introduction) is always incomplete: this may be because of gaps in archives, the lack of digitisation, or simply because most communication is face-to-face and therefore leaves little record. There is a subset of more mundane 'knowable' missing archival data, related in the first instance to its materiality as well as its editorial and digital afterlives: the missing data within it is from lost letters, burned books, uncalendared State papers, undigitised editions, uncrackable ciphers, and so forth which we can model and understand its impact on findings. Modelling the effect of missing letters is an enterprise which, as we have shown, can help to deal with the problem in the absence of reconstructed archives.

The results here show that even with very large numbers of letters missing, there is suprisingly little effect on the overall network structure—or rankings—of many key network metrics. The experiments above show that different modes of removal have different effects, and that missing correspondence data which affects edges (i.e. missing letters) is more robust than those affecting nodes (such as missing individuals, folios or catalogues). This should also be taken into account when considering the impact of their partial data.

In terms of specific metrics, we conclude that researchers using similarly-shaped datasets for historical network analysis might use caution in the interpretation of eigenvector centrality or transitivity measures if it is thought that there is significant relevant data which are yet to be found or digitised, and use another, more robust measurement instead—particularly if the 'missingness' is node-based. The experiments carried out here show that for historical network analysis on correspondence networks, the 'shape' of the loss has also been shown to have surprisingly little effect on common real-world downstream tasks. Many of

the key findings we as authors have relied on using network analysis would be relatively unchanged even with significant data loss. The rankings of three key nodes hardly change position in some key network centrality rankings even with 80% of the network removed. Furthermore, this robustness is not just applicable to those at the very top of the rankings, either: in a second test, an outlying though not particularly high-ranking node, James Butler, Duke of Ormond maintained his outlier status despite significant data loss as an individual whose betweenness centrality was proportionately high when plotted against his degree in the SPO network.

The 'archival turn' suggests that we should consider archives as multi-dimensional textual objects which need to be interpreted rather than neutral silos of documents to be mined for useful information[33] If this is the case, we should bring our quantitative toolsets and 'read' them at scale much like the more traditional texts which are often the subject of digital humanities. Digital Humanities as a field has been much concerned with representativeness: Andrew Piper, arguing that because culture is 'never finished', there should be a move away from thinking about 'samples' and 'bias' and towards what he terms representativeness: a mode which says that every part is a representation, and in which we focus on the curation of data rather than its quantity (or completeness). We suggest that an important facet of this data curation is to understand its missingness, and, moreover, where possible, the effect that this might have on resulting quantitative results, whether they be, as in this case, network metrics, but also more generally: the same technique might be applied to measures derived from work in Computational Literary Studies or Spatial Humanities. The robustness of the arguments that a scholar builds upon such results depends much more on the historical scholarship employed in their interpretation than on the incompleteness of the underlying data. It follows that historical argument must take archival absences into account, regardless of whether its foundations are quantitative or not.

Notes

[1] John Pell to Samuel Hartlib. Zürich, 27 July 1656. BL Add MS 4280, 37r-v. Digital version of printed transcription consulted, in Jana Amosa Komenského korrespondence. Ed. Adolf Patera. Prague: 1892. No. 159. Found in Early Modern Letters Online [http://tinyurl.com/yaumwclk]

[2] Vladimír Urbánek, "Comenius, the Unity of Brethren, and Correspondence Networks," *Journal of Moravian History* 14, no. 1 (2014): p. 31.

[3] Leigh T. I. Penman, "Omnium Exposita Rapinæ: The Afterlives of the Papers of Samuel Hartlib" [in en], *Book History* 19, no. 1 (2016): p. 4. https://doi.org/10.1353/bh.2016.0000.

[4] Noel Malcolm, "Six Unknown Letters from Mersenne to Vegelin," *The Seventeenth Century* 16, no. 1 (2001): p. 95. https://doi.org/10.1080/0268117X.2001.10555485.

[5] Christopher N. Warren et al., "Six Degrees of Francis Bacon: A Statistical Method for Reconstructing Large Historical Social Networks," *Digital Humanities Quarterly* 010, no. 3 (2016).

[6] John R. Ladd, "Imaginative Networks: Tracing Connections Among Early Modern Book Dedications," *Journal of Cultural Analytics*, March 30, 2021, https://doi.org/10.22148/001c.21993.

[7] For example see Ruth Ahnert and Sebastian E. Ahnert, "Protestant Letter Networks in the Reign of Mary I: A Quantitative Approach" [in en], *ELH* 82, no. 1 (2015): 1–34, https://doi.org/10.1353/elh.2015.0000; Ruth Ahnert and Sebastian E Ahnert, "Metadata, Surveillance and the Tudor State" [in en], *History Workshop Journal* 87 (April 2019): 27–51, https://doi.org/10.1093/hwj/dby033; Evan Bourke, "Female Involvement, Membership, and Centrality: A Social Network Analysis of the Hartlib Circle" [in en], *Literature Compass* 14, no. 4 (2017), https://doi.org/10.1111/lic3.12388.

[8] See Linton C Freeman, Douglas Roeder, and Robert R Mulholland, "Centrality in social networks: ii. experimental results" [in en], *Social Networks* 2, no. 2 (1979): 119–41, https://doi.org/10.1016/0378-8733(79)90002-9; Phillip Bonacich, "Factoring and weighting approaches to status scores and clique identification" [in en], *The Journal of Mathematical Sociology* 2, no. 1 (1972): 113–20, https://doi.org/10.1080/0022250X.1972.9989806; Phillip Bonacich, "Power and Centrality: A Family of Measures," *American Journal of Sociology* 92, no. 5 (1987): 1170–82, https://doi.org/10.1086/228631; John Scott, "Social Network Analysis" [in en], *Sociology* 22, no. 1 (1988): 109–27, https://doi.org/10.1177/0038038588022001007; Stanley Wasserman and Katherine Faust, *Social Network Analysis: Methods and Applications*, 1st ed. (Cambridge University Press, November 1994), https://doi.org/10.1017/CBO9780511815478.

[9] Ahnert and Ahnert, "Protestant Letter Networks in the Reign of Mary I."

[10] Ahnert and Ahnert, "Metadata, Surveillance and the Tudor State," p. 42.

[11] Katherine Bode, "The Equivalence of "Close" and "Distant" Reading; or, Toward a New Object for Data-Rich Literary History" [in en], *Modern Language Quarterly* 78, no. 1 (2017): for caution, https://doi.org/10.1215/00267929-3699787; exceptions are Matthew Lincoln and Sandra Van Ginhoven, "Modeling a Fragmented Archive:

A Missing Data Case Study from Provenance Research," 2020, which both use imputation to model the effect of missing data, https://doi.org/10.1184/R1/12363059.V1; Koenraad Brosens et al., "Slow Digital Art History in Action: Project Cornelia's Computational Approach to Seventeenth-century Flemish Creative Communities," *Visual Resources* 35, nos. 1-2 (2019): 105–24, https://doi.org/10.1080/01973762.2019.1553444.

[12]Justin Leidwanger et al., "A manifesto for the study of ancient Mediterranean maritime networks" [in en], *Antiquity* Project Gallery 88, no. 342 (2014).

[13]Matthew Peeples, "Network science and statistical techniques for dealing with uncertainties in archaeological datasets," accessed September 16, 2020, http://www.mattpeeples.net/netstats.html.

[14]Gueorgi Kossinets, "Effects of missing data in social networks" [in en], *Social Networks* 28, no. 3 (2006): 247–68, https://doi.org/10.1016/j.socnet.2005.07.002; Edward Laumann, Peter Marsden, and David Prensky, "The Boundary Specification Problem in Network Analysis," *Applied Network Analysis: A Methodological Introduction* 61 (January 1983).

[15]Joseph Galaskiewicz, "Estimating point centrality using different network sampling techniques" [in en], *Social Networks* 13, no. 4 (1991): 347–86, https://doi.org/10.1016/0378-8733(91)90002-B.

[16]Jeffrey A. Smith and James Moody, "Structural effects of network sampling coverage I: Nodes missing at random" [in en], *Social Networks* 35, no. 4 (2013): 652–68, https://doi.org/10.1016/j.socnet.2013.09.003.

[17]Jeffrey A. Smith, James Moody, and Jonathan H. Morgan, "Network sampling coverage II: The effect of non-random missing data on network measurement" [in en], *Social Networks* 48 (January 2017): 78–99, https://doi.org/10.1016/j.socnet.2016.04.005.

[18]Elizabeth Costenbader and Thomas W Valente, "The stability of centrality measures when networks are sampled" [in en], *Social Networks* 25, no. 4 (2003): p. 306, https://doi.org/10.1016/S0378-8733(03)00012-1.

[19]Lyuba V. Bozhilova et al., "Measuring rank robustness in scored protein interaction networks" [in en], *BMC Bioinformatics* 20, no. 1 (2019): 446, https://doi.org/10.1186/s12859-019-3036-6.

[20]See http://emlo-portal.bodleian.ox.ac.uk/collections/?page_id=2259 for a full list of contributors and http://emlo-portal.bodleian.ox.ac.uk/collections/?page_id=1966 for the current list of EMLO catalogues.

[21]Miranda Lewis, "Ghosts in the Machine: (Re)Constructing the Bodleian's Index of Literary Correspondence, 1927-1963.," 2013, http://www.culturesofknowledge.org/?p=295.

[22]For accounts of the organisation of the State Paper Office see Arnold Hunt, "The Early Modern Secretary and the Early Modern Archive" (British Academy, April 26, 2018), https://doi.org/10.5871/bacad/9780197266250.003.0005.

[23]see http://emlo-portal.bodleian.ox.ac.uk/collections/?catalogue=johan-de-witt

[24]See Albert-László Barabási and Réka Albert, "Emergence of Scaling in Random Networks," *Science* 286, no. 5439 (1999): for foundational discussion of scale-free networks. https://doi.org/10.1126/science.286.5439.509; Duncan J. Watts and Steven H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature* 393, no. 6684

(1998): 440–42, https://doi.org/10.1038/30918.

[25]Galaskiewicz, "Estimating point centrality using different network sampling techniques"; Costenbader and Valente, "The stability of centrality measures when networks are sampled"; Stephen P. Borgatti, Kathleen M. Carley, and David Krackhardt, "On the robustness of centrality measures under conditions of imperfect data" [in en], *Social Networks* 28, no. 2 (2006): 124–36, https://doi.org/10.1016/j.socnet.2005.05.001; Smith, Moody, and Morgan, "Network sampling coverage II"; Peeples, "Network science and statistical techniques for dealing with uncertainties in archaeological datasets."

[26]See Réka Albert, Hawoong Jeong, and Albert-László Barabási, "Error and attack tolerance of complex networks" [in en], *Nature* 406, no. 6794 (2000): 378–82, https://doi.org/10.1038/35019019; Petter Holme et al., "Attack vulnerability of complex networks," *Phys. Rev. E* 65 (5 2002): 056109, https://doi.org/10.1103/PhysRevE.65.056109; H. Jeong et al., "Lethality and centrality in protein networks" [in en], *Nature* 411, no. 6833 (2001): 41–42, https://doi.org/10.1038/35075138.

[27]The EMLO catalogues are modern subsets, based on reassembled letters of a single individual (often ultimately from on a printed edition but not in all cases). On the other hand, the mass of SPO data has been reorganised over time, now mostly into 'series' and 'classes', based on either geography or periods of rule rather than individuals. There are a small number of additional individual-based collections, such as the Cecil Papers, which would be broadly comparable to an EMLO catalogue, but most of the records are not organised in this way.

[28]See Alan Marshall, "Sir Joseph Williamson and the Conduct of Administration in Restoration England," *Historical Research* 69, no. 168 (1996): p. 34, https://doi.org/10.1111/j.1468-2281.1996.tb01840.x; These letters have been published in W. D. Christie and Camden Society (Great Britain), *Letters Addressed from London to Sir Joseph Williamson While Plenipotentiary at the Congress of Cologne in the Years 1673 and 1674*, v. 1 (Camden Society, 1874).

[29]Ahnert and Ahnert, "Metadata, Surveillance and the Tudor State," pp. 35-38.

[30]Toby Barnard, "Butler, James, First Duke of Ormond (1610–1688), Lord Lieutenant of Ireland," *Oxford Dictionary of National Biography*, January 2008, https://doi.org/10.1093/ref:odnb/4191.

[31]A version of the application with some usage limits is available at https://networkingarchives.shinyapps.io/network_robustness_tool/. The source code is included in the code repository alongside this article.

[32]For example Matthew Lincoln, "Continuity and Disruption in European Networks of Print Production, 1550-1750," *Artl@s Bulletin* 6, no. 3 (2017); Martin Grandjean, "Les réseaux de la coopération intellectuelle. La Société des Nations comme actrice des échanges scientifiques et culturels dans l'entre-deux-guerres" [in fr] (PhD diss., Université de Lausanne, June 2018).

[33]See Alexandra Walsham, Kate Peters, and Liesbeth Corens, "Introduction: Archives and Information in the Early Modern World" (British Academy, April 26, 2018), https://doi.org/10.5871/bacad/9780197266250.003.0001; Eric Ketelaar, "Tacit narratives: The meanings of archives," *Archival Science* 1, no. 2 (2001): 131–41, https://doi.org/10.1007/bf02435644.