

Networks of Influence in Scottish Enlightenment Publishing

Yann Ryan^{1*}, Mikko Tolonen²

¹Katholieke Universiteit Leuven,
Leuven, Belgium

*E-mail: y.c.ryan@hum.leidenuniv.nl;
yann.ryan@gmail.com

²University of Helsinki

Abstract

The development of the “public sphere” and the Enlightenment are closely related to the networks of publishers involved in print culture. In Britain, a key question is whether diversity (in terms of nationality) in publishing increased over time. In this paper, we use large-scale library catalog data and a class of models known as Exponential Random Graph Models (ERGMs) to investigate the extent to which an 18th-century network of publishers in London was homophilous with respect to nationality, and specifically investigate whether those with close publishing connections to Scotland formed a separate group or if they were simply integrated into the London book market. As there is little external information on the majority of publishers, we generate node and edge attributes from the catalog data itself. The results suggest that social processes were deeply involved in the decisions behind copublishing and collaborations and that there is indeed a significant positive effect on tie formation if both nodes are “Scottish” publishers, though this lessens over the century. We find that other important factors in tie formation are edgewise shared partners and similarity in patterns of genres published.

Keywords

network analysis, homophily, historical network research, publisher networks, book history

Introduction

A foundational question with respect to the Enlightenment and the birth of modernity has been whether the public sphere was born during the 18th century as argued by Habermas (2003) and Himmelfarb (2004). This question has often been studied in connection to public places such as clubs, associations, and coffee houses (Cowan, 2004; Pincus, 1995). Some scholars have also linked the development of the public sphere to a print culture that undermined the relevance of church and court (see especially, Robertson, 2020; St. Clair, 2007; Worden, 2001). Besides the liberty of the press (Martin, 2001; Ross, 2018), the element of the book industry that drove the development of public discourse was access to book production. In this

paper, we will study the ability of Scottish publishers to access the British book market, at that time controlled by London publishers.

Thus, what we aim to uncover are some of the main processes behind what is called the “Scottish Enlightenment” (Robertson, 2000). The Scottish Enlightenment is the name given to a period, generally considered the second half of the 18th century and the early 19th, where Scottish thinkers and writers, including Adam Smith, David Hume, James Boswell, James Ferguson and many others, flourished and had an outsized impact on the intellectual output of Britain, making valuable contributions to history, economics, science, literature, and political discourse (Broadie, 2006, 2007; Campbell & Skinner, 1982). However, the processes by which this came about are contested and opaque: historians have suggested reasons

ranging from urbanization to Scottish Calvinist doctrine. For some, the Enlightenment was the age of even radical progress (Bonner, 2004; Buchan, 2003), while other scholars find nothing particularly special in the Scottish Enlightenment as such but underline that it should be seen as a continuum of developments that started earlier than in the later part of the 18th century (Allen, 1993; Jackson-Williams, 2020; Mann, 2000).

More recent work has argued that we should consider the role of publishers alongside authors and their printed texts (e.g. Hesse, 2004; Johns, 1998; Munck, 2019). Scottish Enlightenment output, Richard Sher has shown, was supported by a network of publishers, and only by understanding these relationships can we fully grasp the processes that led to it (Sher, 2006). These women and men financed book production, paid popular authors in advance, and even suggested topics to meet the demands of readers, yet their role is still not fully understood in comparison to the attention to which we give Enlightenment thinkers and authors—in part because we know very little about most of them. Across Britain and Ireland, publishers formed complicated and changing partnerships and established complex networks, all in the service of the production and dissemination of the Enlightenment's intellectual efforts. It has been well established in extensive studies of book trade that the relationship between London and Edinburgh booksellers in the 18th century was a tug-of-war that had different legal (MacQueen, 2014; Ross, 1992; Saunders, 1994; Walters, 1974), practical (Raven, 2005, 2014; Sher, 1998), and economical (Woodmansee, 1984) features that were a natural hindrance also for the development of public sphere. As mentioned, while the practical studies of the book trade are many, systematic studies of publishing diversity are nonexistent. Our aim is to start filling this gap to show the intellectual relevance of 18th-century publishers (McKenzie et al., 2002; Whitehouse, 2015).

Many of the most important 18th-century publishers, such as Andrew Millar and William Strahan, were Scots who had settled (or established branches of their Edinburgh businesses) in London following the Act of Union in 1707 and were instrumental in the publication of many of the key works of the Scottish Enlightenment, often copublished with publishers based in Edinburgh (Brown & Kennedy, 2018, Sher, 2007). Publishers often worked in groups both small and large, ranging from individual partnerships to monopolist organizations known as “congers,” and Scottish publishers in London were embedded within this system (Belanger, 1975). Publishers

formed tightly knit groups, and there were strict rules governing who could invest in and therefore become part of the network. They formed dynasties, anointed successors, and took on or recommended apprentices to other friends or family, apprentices who later often became business associates, and through this process, established links and business partnerships.

One bookseller who embodies this transitive nature is William Creech: the Scottish-born Creech was apprenticed to an Edinburgh-based publisher, Alexander Kincaid, before moving to London and befriending the Scottish printer William Strahan, and his partner Thomas Cadell (who himself was English but the successor to the business of Andrew Millar, a very prominent Scottish publisher in London). We know from the historical record that Creech was later able to use these connections to influence his old master, Kincaid, to take him on as a partner (Benedict, 2004; Sher, 2006, pp. 401–410). This, in network language, would be termed *triadic closure*: the tendency, in social networks, for two adjacent nodes to be connected if they have a mutual connection in common. In this case, Kincaid's separate connections to Creech and Cadell/Strahan resulted in Creech's eventual connection to Cadell/Strahan, closing the triad.

This is all to say that the system of publishers in London looked very much like a network, for instance, showing clear evidence of triadic closure, as in the example above. However, some of the precise ways in which these collaborations formed is still unclear. Thanks to historical scholarship, we know a great deal about a very small number of publishers, but very little about the overall process and the thousands of other publishers responsible for book production across the 18th century. Specifically, there is little empirical evidence on the extent to which Scottish publishers in London tended to form partnerships with each other more than would be expected simply by their overall prominence. To what extent did they form separate groups? Were they shut out of London book publishing circles, and was publishing large volumes of Scottish works likely to make one less or more successful in finding and keeping business relationships? Uncovering this would be a valuable piece of the puzzle in understanding the processes that led to the Scottish Enlightenment.

The tendency of individuals to connect to others with the same attributes (for example, social, cultural, or geographic similarities) is known as homophily, and has been demonstrated in many real-world social networks (for an overview, see McPherson et al., 2001). In this paper, we use a class of models known

as Exponential Random Graph Models (ERGMs) to investigate the extent to which the network of London publishers was homophilous with respect to Scottish emigres, and therefore whether they formed a separate group or whether they were simply integrated into the London publisher milieu. To do this at scale, we have insufficient external information about the publishers (such as apprenticeships, social background, or nationality), because it only exists for very small proportions of the data. In many cases, little is known about individual publishers other than their names on the title page. As the ERGM method relies on known attributes of the nodes and edges, we construct them based on the best information we do have for each publisher: statistics relating to the books they publish.

Data

The ESTC

To investigate this network of Scottish publishers, we use the data from the English Short Title Catalogue (ESTC), which is a record of every known publication in the English language or published in Anglophone countries, from 1473 to 1800: a total of ~466,000 records (Tolonen et al., 2015, 2021). More specifically, we use the records from the years 1700 to 1800, which amount to 346,568 records. The majority of these (173,005) are pamphlets, 73,524 are book length, and another 86,248 are in between (Figure 1). Over 200,000 of these records are listed with London as a place of publication, but there are also

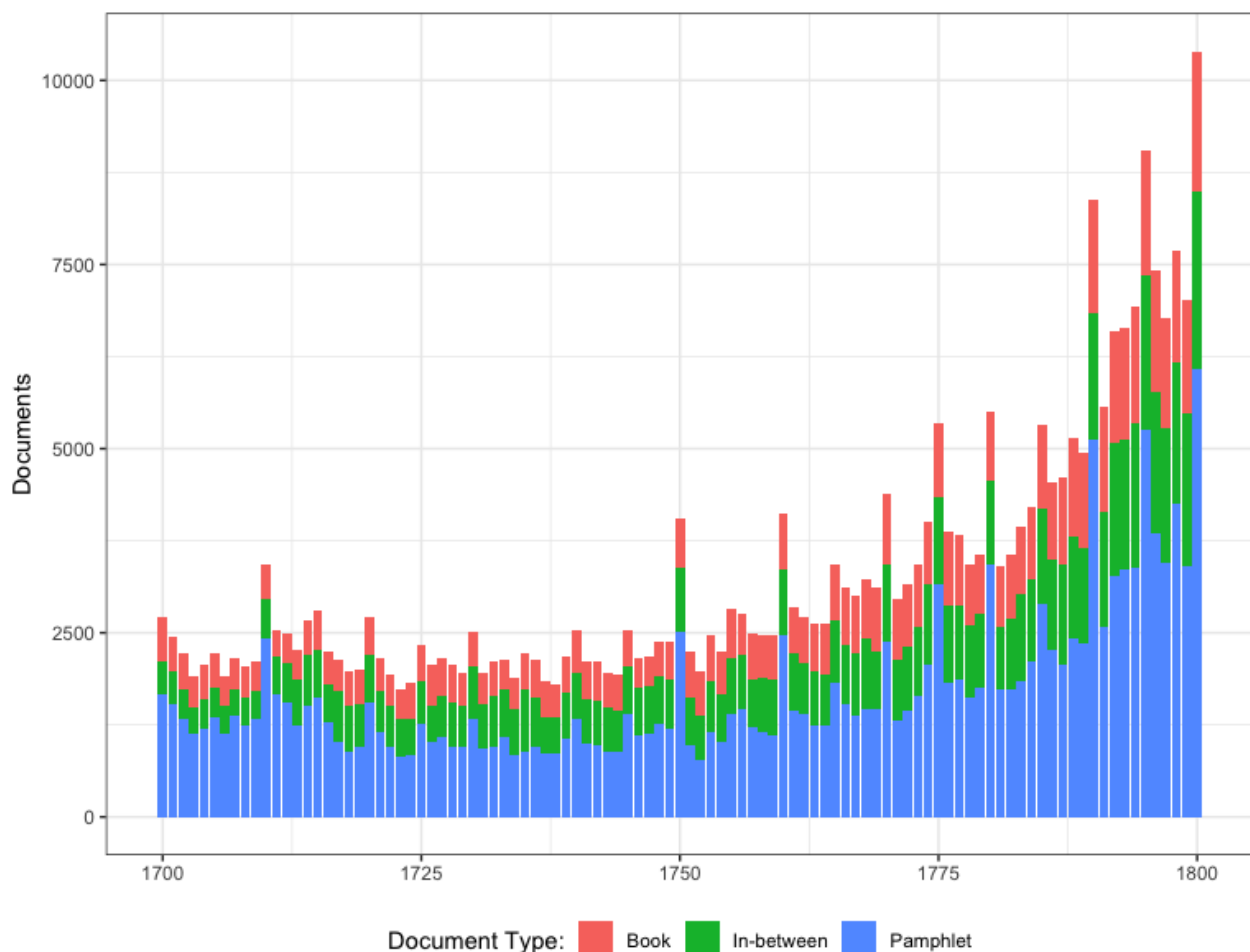


Figure 1: Graph of the number of records per year for each document type in the ESTC, from 1700–1800. The definitions of these document types have fuzzy boundaries. For our purposes, we count any document of between 1 and 32 pages as a pamphlet, between 33 and 128 as an “in-between,” and over 128 as a book. ESTC, English Short Title Catalogue.

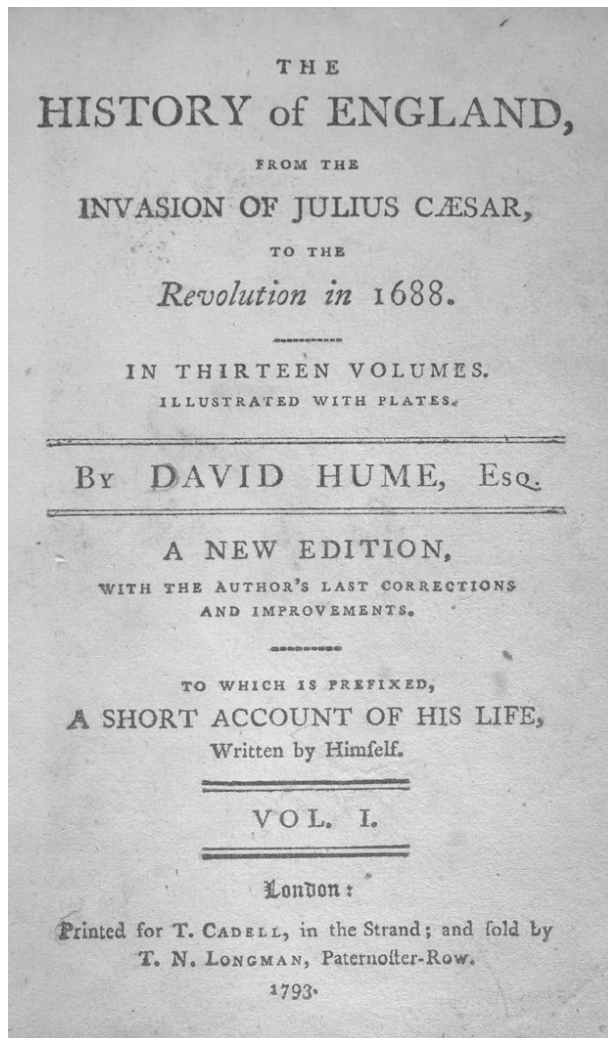


Figure 2: An example of a typical 18th century book title page. Most included information about the author, date, place of publication, as well as its publishers, booksellers, and printers, which has been recorded as structured data in the ESTC. Image from <https://digital.library.mcgill.ca/hume/search/large.php?id=8832>. ESTC, English Short Title Catalogue.

substantial numbers of works from elsewhere in England, Scotland, Ireland, and North America.

The information on these records is drawn mainly from the title pages. Early modern book title pages had a standardized form, which included the *imprint*: essentially a record of who printed, published, and sold each book (Figure 2). This information, alongside information on the author, price, format, and place of publication for each book, has been standardized and

turned into structured data (Hill et al., 2019). There are a few caveats: not all books have surviving copies, and the ESTC is not entirely complete (Raven, 2007). While the records have been harmonized as far as possible, it is not always possible to differentiate between two publishers of the same name. Furthermore, not all of these agents are always listed on the imprint, and in addition, the role of the printer and bookseller changed substantially over the 18th century (see Ozment, 2022; Treadwell, 1982). For this analysis, we do not attempt to mitigate these facts.

Methods

Network from Imprints

Recent years have seen the increasing application of network and social network analysis methods in historical fields. In our own field—book history—the most fruitful use of network analysis has been based on co-occurrence or copublishing (Gavin, 2016; Ladd, 2021). To construct a network of publishers, we used the information on the imprints to create a two-mode network graph, where the first node type is publishers and the second is books. As this study was specifically interested in networks of Scottish publishers in London, we filtered the data to only include records with London listed as the place of publication. This network was then projected to one-mode, so that publishers were directly connected to each other based on shared co-occurrences on imprints. This method has been used previously by several studies, including in the historical domain, and is similar in practice to a co-occurrence or co-citation network (for examples, see Qiu et al., 2014; Tang et al., 2021). The network was then divided into five time slices of 20 years each. This time slicing was done for three reasons. First, the temporal distribution of the data is very uneven (see Figure 1) and individual node statistics would be highly skewed toward the end of the century; second, it allows for the study of change over time; and third, if the entire publisher data was modeled together, the network would be artificially complete, joining paths across multiple generations of publishers who could not have been alive at the same time. At the same time, it should be noted that the decision to use time slices is to a certain extent problematic. The choice of fixed periods means that, in some cases, genuine network paths will have been artificially severed, if the links occur on the end of one and beginning of another slice. Other methods, such as a sliding “time window” or periods selected by pre-established periods rather than a fixed number of years may give better results. However, it was

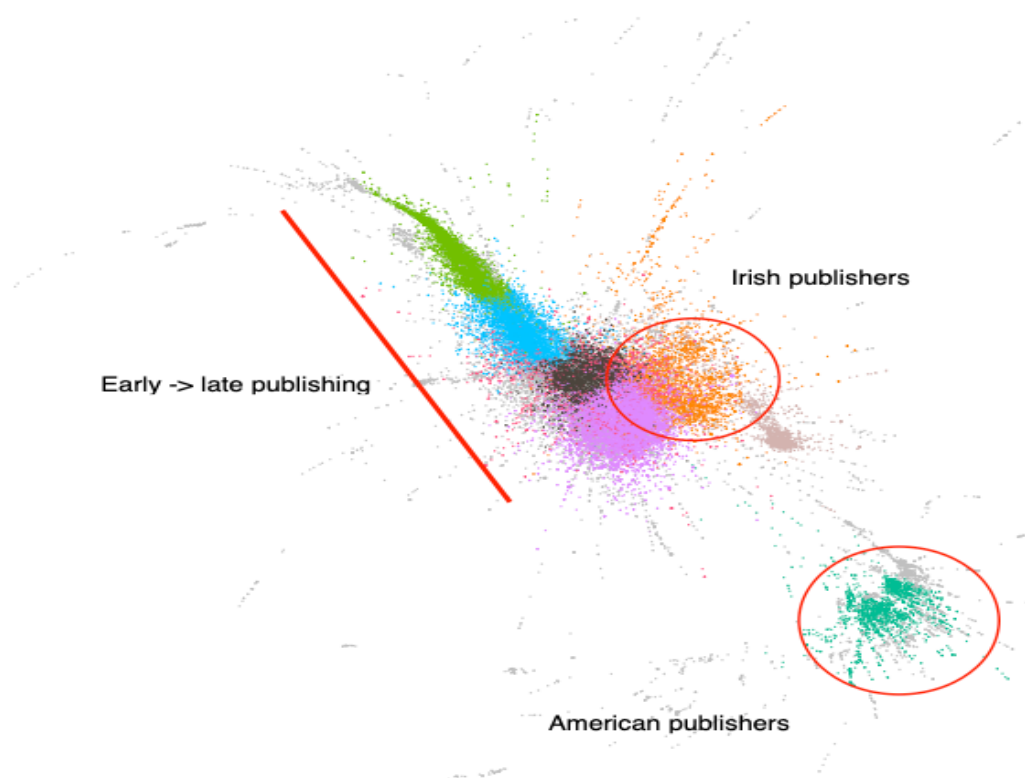


Figure 3: Force-directed network of the entire ESTC dataset (1473–1800) with Louvain community detection labels. Edges have been removed for readability. As the data is longitudinal and passes through multiple generations of individuals, its connections are mostly temporal, resulting in the elongated shape of the main component, consisting of a series of subgroups that are connected in series rather than all to each other. Scottish publishers are mostly distributed along this part of the network: the green, blue, brown, and pink clusters. Also visible is an Irish cluster (in orange), and a more distant American cluster (in green). ESTC, English Short Title Catalogue.

thought that the benefits of the fixed, lengthy period, namely simplicity and reducing the number of total calculations, outweighed the drawbacks.

As many pairs of publishers co-occurred on multiple books, projecting the network resulted in weights associated with the new edges. We filtered the network to retain edges with a weight ≥ 2 , to consider only more “intense” or perhaps genuine collaborations, and otherwise disregarded the weight information. This filtering step also had a second benefit in that it reduced the network’s complexity and had a positive effect on the convergence of the ERGM models.

Network Description

The global attributes of the five time-sliced networks are given in the table below. Our observations show

that they increase in size over time, which matches the distribution of data in the ESTC (there are many more records toward the end of the century, because more books were published and because more records survived). The networks have only moderately high clustering and modularity, and the density *decreases* over time, mirroring the inverse of the increase in overall size. They are mostly connected, consisting of a main “giant” component, but there are a number of smaller components in each.

A force-directed diagram of the entire network shows that the clustering is primarily along temporal lines: if we include places other than London in the data, it would be geographically clustered too (Figure 3). Figure 3 suggests that there are separate subgraphs of Irish and North American publishers, but that Scottish publishers are difficult to distinguish from their London counterparts. This simple analysis

suggests that there was significant overlap between publishers based in Scotland and their London counterparts.

Node and Edge Attributes

Measurements of homophily rely on shared node or edge covariates. Because of the historical time gap, the large number of publishers involved, and their obscurity, for the vast majority, we have little or no exogenous information such as gender, age, social class, and so forth. A portion (about 12%) have basic records on reliable historical sources such as the British Book Trade Index, but EGRM models rely on data that is relatively complete. Our solution was to construct exogenous attributes from the book records themselves. To begin with, we constructed two base datasets: (1) a dataset of Scottish authors and (2) genre information for each title. To create the former, we extrapolated a set of Scottish authors from the dataset, defined as all authors where at least 50% of their works were first published in Edinburgh, Glasgow, or Aberdeen. To this, we added a dataset of 1,400 Scottish authors, which we manually put together from a number of sources and research. For the latter genre, information was generated using a neural network classifier, which we describe in detail in a separate paper (Zhang et al., 2022).

From these datasets, we defined a set of node and edge attributes, given in Table 3.

ERGMs

To measure the degree of homophily within these publisher groups, we used ERGMs. ERGMs are a class of models that aim to identify the process behind link formation in a network (van Der Pol, 2017). ERGMs allow a researcher to construct a model consisting of a set of node and edge attributes, and receive as an output a probability, in the form of a conditional log odds, of the importance of each to the probability of a tie within the network. ERGMs have become increasingly popular in social network and political science research (see Cranmer & Desmarais, 2011; Maejima, 2020; Robins et al., 2007) although their use with historical data remains limited (examples include Breure & Heiberger, 2019; Brughmans et al., 2014). Using ERGMs helps us to move beyond the exploratory analysis of networks and use statistical methods to make empirical claims about the processes behind their formation. Our aim in this paper was to investigate which node attributes were the most important factors in predicting the existence of a tie between two publisher nodes—if an

attribute such as Scottish nationality was important, this would indicate that Scottish publishers tended to form homophilic networks—essentially that “birds of a feather flock together.”

The key reason to use ERGMs here is because much of the probability of the existence of a tie in a network can be explained through fundamental structural properties. Most importantly, it has been repeatedly shown that much of a network’s formation can be explained through *triadic closure*, meaning that the probability of a tie between two nodes is highly dependent on whether they share a neighbor (Snijders et al., 2006). In the publisher network, this translates to the fact that two publishers are much more likely than chance to be listed on an imprint if they also share an imprint with a mutual associate. This dependence, or confounding, means that ordinary statistical methods (for example, multivariate regression, where the outcome is either 1 for a tie or 0 for no tie) cannot be used. ERGMs allow us to control for this fact by generating a large number of networks, allowing us to model structural network attributes related to triadic closure and reciprocity, from which probability statistics can then be drawn.

There are a number of implementations of the ERGM method, including one within the R package *statnet*, which was used here (see Hunter et al., 2008; Krivitsky et al., 2022 for updates). We constructed an ERGM model for each of the five time-sliced networks from Table 1 and a group of node and edge covariates based on the attributes described in the previous section. These are three “nodematch” covariates: (1) whether or not the node is a “Scottish publisher,” calculated separately for Scottish and non-Scottish nodes, (2) whether both nodes are +25% Edinburgh publications, and finally (3) gender. Then two “absdiff” attributes are: (1) the absolute difference between the first publication date of each pair of nodes and (2) the difference between the percentages of works by Scottish authors. We also used an edge covariate: the cosine similarity—ranging from 0 to 1—between the vectors of genre counts for each publisher. Finally, we used Geometrically Weighted Edge Shared Partners (GWESPs), which is a measurement similar to counting the number of triangles, which essentially controls for the fact that shared neighbors is a strong predictor of a tie in a network (Hunter, 2007). GWESP is used instead of triangles because the latter has been shown to lead to model degeneracy (Snijders et al., 2006).

We made one ERGM model for each time slice, each with the same basic parameters. The MCMC interval and the sample size were set to 1,000, and we ran a maximum of 400 iterations. We tested a

Table 1. Network statistics for each time slice.

	1700	1720	1740	1760	1780
Total nodes	744	810	808	1,207	1,850
Number of components (size of largest)	16 (708)	12 (785)	12 (784)	26 (1,103)	25 (1,782)
Number of edges	4,906	7,058	6,668	12,053	15,519
Average degree	13.19	17.43	16.5	19.97	16.78
Global clustering coefficient	0.31	0.34	0.36	0.38	0.33
Average local clustering	0.56	0.68	0.69	0.74	0.71
Density	0.0177	0.0215	0.0205	0.0166	0.0091
Modularity (with Louvain)	0.34	0.27	0.31	0.27	0.4
Average distance	3.09	2.88	2.95	2.9	3.15
Book ^a	15,810	19,519	17,644	26,250	30,263
In-between	10,730	12,916	12,882	16,787	23,927
Pamphlet	15,822	8,544	10,203	10,611	17,513
NA	1,378	2,323	2,549	4,370	5,345

^aMeaning the number of books, pamphlets, and in-betweens in total in that slice, not how many used in the eventual network.

number of combinations of covariates and selected a model that had the highest goodness of fit and also a lower Akaike information criterion (AIC). AIC quantifies how well a model explains the data while penalizing it for having too many parameters. Lower AIC values indicate better models, making it a valuable tool for choosing the most appropriate model among a set of candidates.

Results

The ERGM output mirrors the typical output of a logistic regression—the log-odds of a tie for a change statistic of one. Therefore, the probability (for a change statistic of one) can be calculated with the formula $u/(1 + u)$, where u is the exponential of the log-odds. To report the results, in most cases, we give the change in probability of a tie, given a particular attribute. This is done by adding the relevant probabilities to the base probability for an edge formation. In some cases, this is multiplied by a change statistic. For example, the results for the first year of publication give the log-odds for every 1-year difference between the first year of publication for a pair of nodes. For covariates relating to a match between two nodes, the probability if such a match is present is calculated by multiplying by a change statistic of 2. The coefficient results for each time slice model are given below.

To make the results easier to interpret, consider the base likelihood of an edge formation, all other parameters in our model being equal—a value which is essentially another way of measuring a network's density. As is typical in a sparse network of this kind, the “base” probability is low: 0.00132 in 1700–20, 0.00101 in 1720–40, 0.00070 in 1740–60, 0.00062 in 1760–80, and 0.00021 in 1780–1800 ($p < 0.001$ in all cases). It also reduces over the century, because, as seen from Table 1, the network increases in size and becomes less dense over time—primarily because there are many more book records toward the end of the century. From this “base,” we can calculate the percentage change in probability of a tie, when we add in a range of node and edge attributes.

In all time slices, the same set of attributes make up the most significant predictors of a tie. Key is GWESP: the measure of shared edges or transitivity. Across all five time slices, the impact of this is important: it has a significant and strong positive effect. In the 1700–1720 time slice, for example, having a single shared edgewise partner increases the probability of the existence of a tie almost 1,000%: from 0.00132 to 0.0114, with similar numbers across all periods. The significance of this is that it points to the likelihood that this network, of publishers connected by imprints, in some way formed fundamentally by the same principles as real-world social networks, namely, the idea that “the friend of my friend is also

Table 2. Log-odds for all covariates used by the model.

Time slice	1700	1720	1740	1760	1780
Edges	6.63*** 0.08	6.92*** 0.08	7.29*** 0.08	7.48*** 0.11	8.41*** 0.07
nodematch.is_ed.no	0.55*** 0.03	0.81*** 0.04	0.77*** 0.04	0.80*** 0.07	0.57*** 0.06
nodematch.is_ed.yes	1.14*** 0.06	1.52*** 0.05	1.40*** 0.04	1.67*** 0.06	1.20*** 0.11
nodecov.p_scot	0 0.01	0.02** 0.01	0.01** 0	0.02*** 0	0.01 0.01
gwesp.fixed.0.01	2.17*** 0.06	2.24*** 0.07	2.38*** 0.07	2.61*** 0.09	3.11*** 0.08
nodematch.actor_gender.female	0.25 0.42	0.31 0.33	1.02* 0.41	Inf	Inf
nodematch.actor_gender.male	0.23*** 0.02	0.10*** 0.03	0.18*** 0.03	0.30*** 0.05	0.19*** 0.04
absdiff.p_scot	0.04*** 0.01	0.02** 0.01	0.04*** 0.01	0.03*** 0.01	0.02* 0.01
absdiff.publication_year	0.05*** 0	0.09*** 0	0.08*** 0	0.08*** 0.01	0.07*** 0
edgecov.dist	1.72*** 0.06	2.70*** 0.06	2.63*** 0.06	2.44*** 0.07	2.49*** 0.08

Standard error is underneath.

*** $p < 0.0001$; ** $p < 0.01$; * $p < 0.1$.

my friend,” rather than some other primary principle such as random chance, economics, social class, or nationality. In this case, we might consider it as statistical confirmation that social processes were deeply involved in the decisions behind copublishing and collaborations.

The second consistently significant covariate is the measure of genre similarity. This was intended to capture the extent to which publishers formed networks around similar publishing patterns, by comparing the cosine similarity between the frequency of publications in each genre, for a given pair of edges. In this model, the effect on probability for a cosine similarity of 1 ranges from an increase of 448% to 1315%, depending on the time slice. This suggests that publishers published within certain genres, and that this had an effect on their appearances on imprints.

The third significant covariate is that most relevant to the subject of this paper. In this case, we used a “nodematch” attribute, which measures the impact on a tie if both edges share a categorical attribute. The attribute of interest is whether or not the node is a “Scottish Publisher,” as defined in Table 2 above. For this, we can calculate the different results separately, meaning it is possible to get separate results for both nodes matching on Scottish, and both matching on “not Scottish.” Our results show that if both nodes are labelled as Scottish, the probability of a tie rises from between 946% and 2,870% above the base level, depending on the time slice, meaning Scottish publishers in London display very substantial tendencies toward homophily across all time slices. The significance of this will be discussed below. Conversely, if both nodes are labelled as *not* Scottish, there is a moderate *decrease* in the likelihood of

Table 3. Overview of node and edge attributes derived from ESTC data.

Attribute	Type	Description
Percentage of Scottish work	Node	Proportion of work published by Scottish authors
“Scottish publisher”	Node	Binary attribute, defined as any publisher whose publication list was at least 25% works by the above Scottish authors. The 25% threshold was chosen after some experimentation—because most publishers published mostly English works, a higher threshold meant not including many important Scottish publishers.
Edinburgh publisher	Node	At least 50% Edinburgh publications, regardless of the likely author nationality.
Scottish similarity	Edge	The difference in percentage points of the volume of Scottish works published by each node in a pair.
Genre similarity	Edge	Genre information was turned into an edge attribute: the number of publications for each publisher in each genre was expressed as a vector, and a similarity matrix was made using the cosine distance between each pair.
Publisher gender	Node	Estimated using a rule-based approach on first names
Earliest year published	Node	Year of first record in the ESTC.
Total works	Node	Total count of records in the ESTC

ESTC, English Short Title Catalogue.

tie formation (ranging from -66% to -79%, which suggests that “local” publishers in London did *not* tend to form cliques in the same way (Figure 4).

Other attributes were also included in the model. First, we extracted the first recorded appearance of each publisher in the ESTC data and used the absolute difference between this value as a covariate between each pair of nodes. The results show that for each year of difference, the likelihood of a tie *decreased* by between 5% and 8% across all time slices, meaning that there was a moderately strong clustering effect between publishers of similar “ages.” We also looked at the effect of printing Scottish material on the *overall* connectivity of a node in the network (rather than its impact on the likelihood of a particular tie). This is done by adding the percentage of Scottish work as a node covariate to the model. This shows that for each additional percentage of Scottish publications, a node’s overall connectivity was *reduced* by between 0.64% and 2.4%, though this was only statistically significant for three out of five time slices. A measurement of the difference between two nodes in the percentage of Scottish works showed a similar pattern, with each additional 1% difference reducing the probability of a tie between 2% and 3.4%. Overall, then, we might conclude that being a prominent publisher of

Scottish work in London, both reduced one’s overall connectivity in the network, and at the same time, increased one’s likelihood to connect with others of a similar background, more often than random chance.

Finally, we added gender as an attribute in an attempt to measure any effects, although it should be noted that this data is very imbalanced, as women made up only a very small proportion of all publishers. The effect of both nodes being female does not have a significant effect on the likelihood of a tie; both nodes being male results in a small but significant effect.

Change Over Time

As well as looking at the time slices as discrete networks, we can use them to understand changes in the formation of the network over time. Here, we have plotted the probabilities for a tie for a range of attributes: when both nodes are Scottish, both not Scottish, when both have a single shared edge partner, and for a genre cosine similarity of 1. In general, probabilities fall over time—this is because the network becomes much less dense (see Table 4) and therefore the overall probability of a tie is lower. The level of homophily (measured as the probability of a tie if both share the “Scottish” attribute) is always

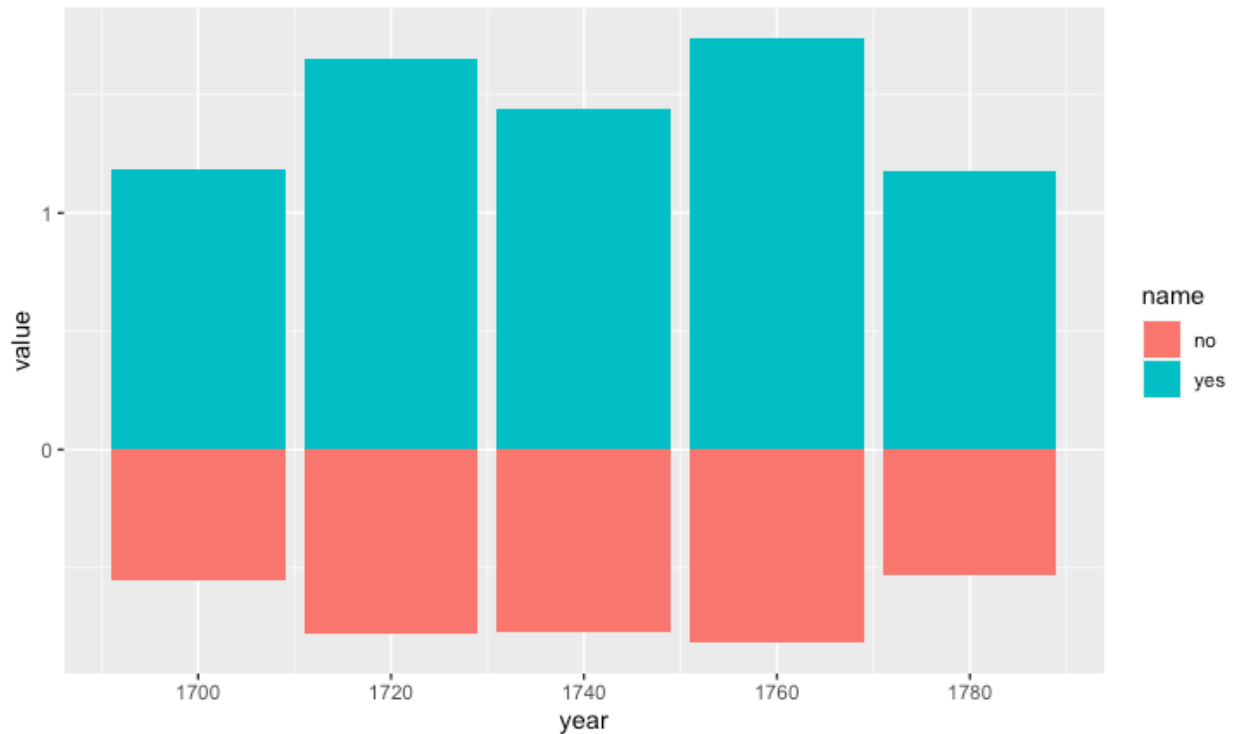


Figure 4: Coefficients (not probabilities) for Scottish and non-Scottish publishers, over time. Non-Scottish is consistently below zero, meaning that if both nodes were not Scottish, the probability of a tie moderately decreased.

substantial, but what is interesting is its changing impact, and the rise of the importance (proportionally) of the structural GWESP property. The impact of Scottishness on the network lessens and is replaced by the straightforward social network principle of transitivity, which by the final time slice is more than double as important a factor in edge formation than Scottishness. The impact of genre similarity also loses ground to GWESP over time.

What implications does this have for our understanding of the processes guiding the formation of this network? The evidence suggests that over the century, the London market became more homogenous and access became easier, with publishers willing to collaborate across a range of topics and with those with different backgrounds to their own. We know that as Scottish publishers became more established in London, they likely made more business contacts, and subsequently worked with a wider range of other publishers than just their compatriots. When the time came for this first generation to pass on businesses, as happened toward the end of the century, we might speculate that their successors became even more fully

integrated within the wider London market. This is reflected in the statistical analysis of the network. This is important also in terms of showing that Scotland was not a mere satellite for England and the entry of the Scottish publishers to the British market was important also in the economic setting (Devine, 2004).

Discussion

The aim of this article was to use statistical network methods, namely ERGMs, to examine the social practices and mechanisms that explain the formation of 18th century publisher networks. The first finding is confirmation of the extent to which the network can be explained by typical social network dynamics such as triadic closure. If nothing else, this points to the validity in treating the data on book imprints as a formal network and the potential value of results gained from doing so. More interestingly, it has shown predominantly Scottish publishers in London had a much higher likelihood than chance to be present together on a book imprint, but having this attribute also led to a small overall decrease in connectivity. At the same time, both being “local” (i.e., not

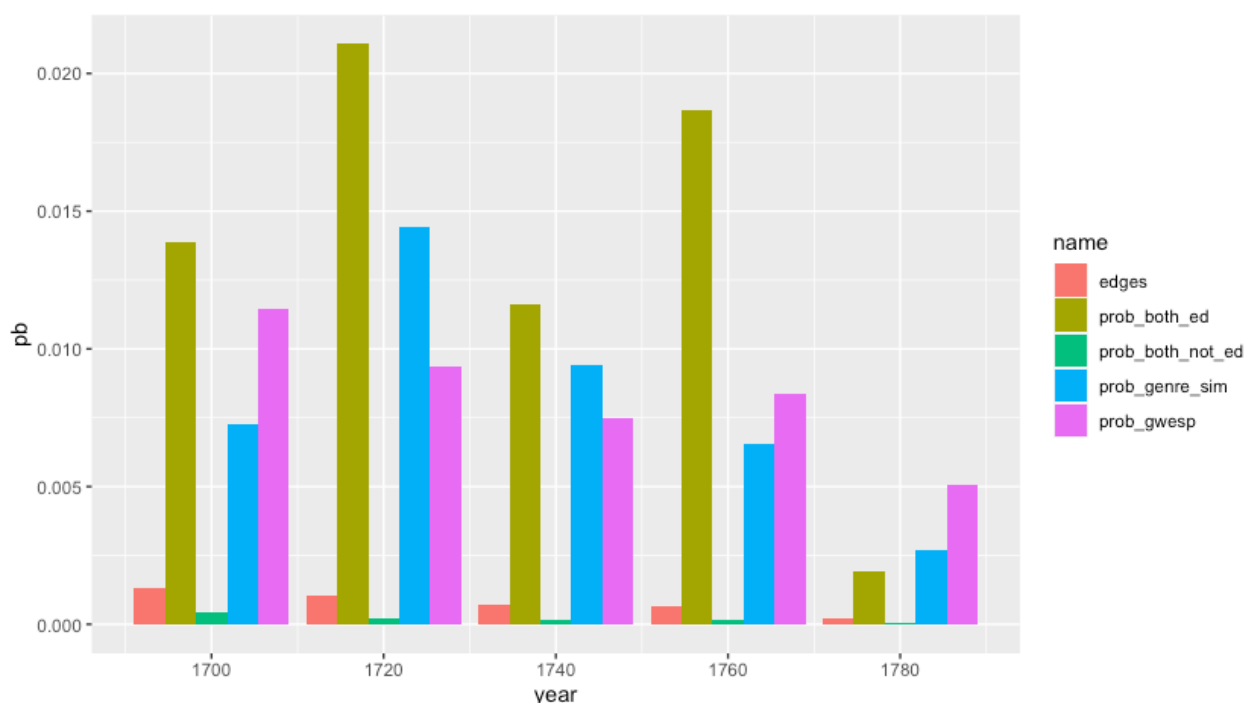


Figure 5: Probabilities per decade for a number of key coefficients. These suggest that in general the impact of two nodes being Scottish publishers on the probability of a tie lessened over the century, in favor of genre similarity and network effects (edgewise shared partners).

predominantly publishing Scottish work) meant a small net *decrease* in the probability of a tie. This suggests that perhaps the network looks as it does partially because Scottish publishers formed cliques in London, rather than because of their deliberate exclusion. The results also suggest that as the century went on, there was more mixing between publishers of different backgrounds, as Scottish publishers became more integrated within the London publisher community and eventually passed businesses on to heirs and successors, a second generation, in many cases born in London, who may have been more integrated still. The increasing importance of the structural GWESP statistic suggests that as Scots became more integrated in the London publishing world, they became less reliant on a community of compatriots for their business success, and increasingly their new contacts came from that most simple of network-formation principles: by having the right friends.

Conclusion

ERGMs have proven a valuable resource in understanding these historical networks. Using network modeling means we can provide empirical evidence to understand network formation and how

the overall macro structure of the network is influenced by its micro-structure, that is, individual connections. Highlighting the significance of methods like ERGMs for the future of traditional humanities research, we uncover an intriguing narrative within the realm of the Scottish Enlightenment. Initially, occupying a prominent position in Scottish publishing in 18th-century London meant reduced connectivity within the broader publishing network while simultaneously increasing collaboration with fellow Scots beyond what random chance would predict. Over time, this pattern evolved as Scottish publishers diversified their partnerships, engaging with a more varied group of colleagues. This transformation challenges the perception of Scottish publishers' peripheral role and underscores their substantial contribution to British intellectual and publishing history, thus emphasizing the importance of innovative methodologies like ERGMs in reshaping our understanding of the past.

At the same time, it should be noted that this approach necessarily works on probability rather than describing historical reality, and as such suggests reasons for network formation rather than definitively explains them. Additionally, this process of deriving pseudo-exogenous properties from the data means there may be other confounding factors

at play, which are invisible to us. Despite this, the methods may be useful in understanding precisely why particular publishers or groups of publishers appeared on imprints together, which in many cases is still unknown. There are a number of future areas to expand this work. The ERGM method could be applied to other types of attributes of the publisher nodes, utilizing the full-text and digital images that are available for a large portion of the data, for example, looking at linguistic similarity or even similar visual properties. We could also exploit the link-prediction functions of ERGMs to estimate or suggest unknown links in the network, allowing us to guess at collaborations not explicitly found on title pages.

This statistical-network approach to historical networks has the potential to make a substantial contribution to our understanding of the makeup and dynamics of historical print networks. By suggesting the key attributes that led to publishers working with one another (not just nationality but other traits, where we have the data), we can come to a better understanding of the networks responsible for the production of cultural objects. Furthermore, by pushing the discipline of historical network analysis past descriptive and toward explanatory statistics, the value of network analysis to the study of the past can be significantly improved.

Acknowledgments

We are grateful to the Academy of Finland for funding this work (Grants 1333716 and 1347706). We would like to also thank members of COMHIS and the RiCEP project for supporting this work.

References

- Allan, D. (1993). *Virtue, learning and Scottish enlightenment*. Edinburgh University Press.
- Belanger, T. (1975). Booksellers' trade sales, 1718–1768. *The Library*, s5-XXX(4), 281–302. <https://doi.org/10.1093/library/s5-XXX.4.281>
- Benedict, B. M. (2004). Creech, William [pseud. Theophrastus] (1745–1815), bookseller and magistrate. In H. C. G. Matthew & B. Harrison (Eds.), *The Oxford dictionary of national biography* (pp. oadn/6662). Oxford University Press. <https://doi.org/10.1093/ref:odnb/6662>
- Bonner, S. E. (2004). *Reclaiming the enlightenment: Toward a politics of radical engagement*. Columbia University Press.
- Breure, A. S. H., & Heiberger, R. H. (2019). Reconstructing science networks from the past. *Journal of Historical Network Research*, 3(1), 92–117.
- Broadie, A. (2006). *The Cambridge companion to the Scottish enlightenment*. Cambridge University Press.
- Broadie, A. (2007). *The Scottish enlightenment: The historical age of the historical nation*. Birlinn Ltd.
- Brown, K. M., & Kennedy, A. (2018). Land of opportunity? The assimilation of Scottish migrants in England, 1603–ca. 1762. *Journal of British Studies*, 57(4), 709–735. <https://doi.org/10.1017/jbr.2018.113>
- Brughmans, T., Keay, S., & Earl, G. (2014). Introducing exponential random graph models for visibility networks. *Journal of Archaeological Science*, 49, 442–454. <https://doi.org/10.1016/j.jas.2014.05.027>
- Buchan, J. (2003). *Capital of the mind: How Edinburgh changed the world*. Berlin.
- Campbell, R. H., & Skinner, A. S. (Eds.). (1982). *The origins and nature of Scottish enlightenment: Essays*. J. Donald.
- Cowan, B. (2004). Mr. Spectator and the coffeehouse public sphere. *Eighteenth-Century Studies*, 37(3), 345–366. <https://doi.org/10.1353/ecs.2004.0021>
- Cranmer, S. J., & Desmarais, B. A. (2011). Inferential network analysis with exponential random graph models. *Political Analysis*, 19(1), 66–86. <https://doi.org/10.1093/pan/mpq037>
- Devine, T. M. (2004). Scotland. In *Cambridge economic history of modern Britain* (pp. 388–416). Cambridge University Press.
- Gavin, M. (2016). Historical text networks: The sociology of early English criticism. *Eighteenth-Century Studies*, 50(1), 53–80. <https://doi.org/10.1353/ecs.2016.0041>
- Habermas, J. (2003 [1962]). The structural transformation of the public sphere: An inquiry into a category of Bourgeois Society. Polity.
- Hesse, C. (2004). Print culture in the enlightenment. In M. Fitzpatrick (Eds.), *Enlightenment world* (pp. 366–380). Routledge.
- Hill, M. J., Vaara, V., Säily, T., Lahti, L., & Tolonen, M. (2019). *Reconstructing intellectual networks: From the ESTC's bibliographic metadata to historical material*. <https://helda.helsinki.fi/handle/10138/302951>
- Himmelfarb, G. (2004). *The roads to modernity: The British, French, and American enlightenments*. Vintage. <https://doi.org/10.25517/JHNR.V3I1.52>
- Hunter, D. R. (2007). Curved exponential family models for social networks. *Social Networks*, 29(2), 216–230. <https://doi.org/10.1016/j.socnet.2006.08.005>
- Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, 24(3), 1–29. <https://doi.org/10.18637/jss.v024.i03>
- Jackson-Williams, K. (2020). *The first Scottish enlightenment. Rebels, priests and history*. Oxford University Press.

- Johns, A. (1998). *The nature of the book: Print and knowledge in the making*. Chicago University Press.
- Krivitsky, P. N., Hunter, D. R., Morris, M., & Klumb, C. (2022). ergm 4: New features. *arXiv*. <https://doi.org/10.48550/arXiv.2106.04997>
- Ladd, J. R. (2021). Imaginative networks: Tracing connections among early modern book dedications. *Journal of Cultural Analytics*, 6(1). <https://doi.org/10.22148/001c.21993>
- MacQueen, H. (2014). The war of the booksellers: Natural law, equity, and literary property in eighteenth-century Scotland. *The Journal of Legal History*, 35(3), 231–257.
- Maejima, N. (2020). Comparing gender homophily among the multilayer media social networks of face-to-face, instant messenger and social networking services: A case study of a high school classroom. *Connections*, 40(1), 77–97. <https://doi.org/10.21307/connections-2019.014>
- Mann, A. (2000). *The Scottish Book Trade, 1500–1720: Print Commerce and Print Control in Early Modern Scotland*. Tuckwell Press.
- Martin, R. W. T. (2001). The free and open press: The founding of American democratic press. *Liberty*. <https://doi.org/10.18574/nyu/9780814759769.001.0001>
- McKenzie, D. F., McDonald, P. D., & Suarez, M. F., (Eds.). (2002). *Making meaning: Printers of the mind and other essays*. University of Massachusetts Press.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415–444. <https://doi.org/10.1146/annurev.soc.27.1.415>
- Munck, T. (2019) *Conflict and enlightenment: Print and political culture in Europe 1635–1795*. Cambridge University Press.
- Ozment, K. (2022). The Women's Print History Project. What Does It Mean to Publish? A Messy Accounting of Anne Dodd. <https://womensprinthistoryproject.com/blog/post/111>
- Pincus, S. (1995). Coffee politicians does create: Coffeehouses and restoration political culture. *The Journal of Modern History*, 67(4), 807–834. <https://doi.org/10.1086/245229>
- Qiu, J.-P., Dong, K., & Yu, H.-Q. (2014). Comparative study on structure and correlation among author co-occurrence networks in bibliometrics. *Scientometrics*, 101(2), 1345–1360. <https://doi.org/10.1007/s11192-014-1315-6>
- Raven, J. (2005). Publishing and Bookselling, 1660–1780. In J. Richetti (Ed.), *The Cambridge history of English literature, 1660–1780* (pp. 13–36). Cambridge University Press.
- Raven, J. (2007). *The business of books: Booksellers and the English book trade, 1450–1850*. Yale University Press.
- Raven, J. (2014). *Publishing business in eighteenth-century England*. Boydell.
- Robertson, J. (2000). The Scottish contribution to the enlightenment. In P. Wood (Ed.), *The Scottish enlightenment: Essays in reinterpretation* (pp. 37–62). University of Rochester Press.
- Robertson, R. (2020). Habermas and the English public sphere reconsidered. *Journal of Interdisciplinary History of Ideas*, 9(17). URL: <http://journals.openedition.org/jihi/991>
- Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (P*) models for social networks. *Social Networks*, 29(2), 173–191. <https://doi.org/10.1016/j.socnet.2006.08.002>
- Ross, T. (1992). Copyright and the invention of tradition. *Eighteenth-Century Studies*, 26, 1–27.
- Ross, T. (2018). *Writing in public: Literature and the liberty of the press in eighteenth-century Britain*. The Johns Hopkins U.P.
- Saunders, D. (1994). *Authorship and copyright*. Routledge.
- Sher, R. (1998). Corporatism and consensus in the late eighteenth-century book trade: The Edinburgh booksellers society in comparative perspective. *Book History*, 1(1998), 32–93.
- Sher, R. (2006). *The enlightenment & the book: Scottish authors & their publishers in eighteenth-century Britain, Ireland, & America*. University of Chicago Press.
- Sher, R. (2007). Scottish publishers in London. In B. Bell, D. Finkelstein, & A. McCleery (Eds.), *The Edinburgh history of the book in Scotland* (pp. 392–408). Edinburgh University Press.
- Snijders, T. A. B., Pattison, P. E., Robins, G. L., & Handcock, M. S. (2006). New specifications for exponential random graph models. *Sociological Methodology*, 36(1), 99–153. <https://doi.org/10.1111/j.1467-9531.2006.00176.x>
- St. Clair, W. (2007). *The reading nation in the romantic period*. Cambridge University Press.
- Tang, K.-Y., Chang, C.-Y., & Hwang, G.-J. (2021). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, 31(4), 2134–2152. <https://doi.org/10.1080/10494820.2021.1875001>
- Tolonen, M., Hill, M. J., Ijaz, A. Z., Vaara, V., & Lahti, L. (2021). Examining the early modern canon: The English short title catalogue and large-scale patterns of cultural production. In I. Baird (Ed.), *Data visualization in enlightenment literature and culture* (pp. 63–119). Springer International Publishing. https://doi.org/10.1007/978-3-030-54913-8_3
- Tolonen, M., Lahti, L., & Ilomäki, N. (2015). A quantitative study of history in the English short-title catalogue (ESTC), 1470–1800. <https://helda.helsinki.fi/handle/10138/228166>
- Treadwell, M. (1982). London trade publishers 1675–1750. *The Library*, s6–IV(2), 99–134. <https://doi.org/10.1093/library/s6-IV.2.99>

van Der Pol, J. (2017). Introduction to network modeling using exponential random graph models (ERGM). <https://hal.archives-ouvertes.fr/hal-01284994>

Walters, G. (1974). The booksellers in 1759 and 1774: The battle for literary property. *Library*, 29, 287–311.

Whitehouse, T. (2015). The textual culture of English protestant dissent, 1720–1800. Oxford University Press.

Woodmansee, M. (1984). The genius and the copyright: Economic and legal conditions of the emergence of the 'author'. *Eighteenth-Century Studies*, 17, 425–448.

Worden, B. (2001). The question of secularization. In A. Houston & S. Pincus (Eds.), *A nation transformed* (pp. 20–40). Cambridge U.P.

Zhang, J., Ryan, Y. C., Rastas, I., Ginter, P., Tolonen, M., & Babbar, R. (2022). *Detecting sequential genre change in eighteenth-century texts*. Computational Humanities Research Conference 2022.