

---

# EVALUATING STRUCTURAL EDGE IMPORTANCE IN TEMPORAL NETWORKS

---

A PREPRINT

Isobel Seabrook<sup>1,2</sup>, Fabio Caccioli<sup>1, 3, 4</sup>, and Paolo Barucca<sup>1</sup>

<sup>1</sup>Department of Computer Science, University College London, London, United Kingdom

<sup>2</sup>Financial Conduct Authority, London, United Kingdom

<sup>3</sup>Systemic Risk Centre, London School of Economics, London, United Kingdom

<sup>4</sup>London Mathematical Laboratory, London, UK

December 10, 2020

## ABSTRACT

To monitor risks in temporal financial networks, we need to understand how individual behaviours affect the global evolution of networks. We define a structural importance metric which we refer to as  $l_e$ , based on perturbing the adjacency matrix and observing the resultant change in eigenvalue. We then propose a model for network evolution where this metric can control the probabilities of subsequent edge changes. Experiments on networks generated according to our model, and real financial and social networks show that i) a perturbation based approximation of  $l_e$  is effective in the case of isolated changes ii) network evolution can be parameterised by the extent to which  $l_e$  influences the propensity and scale of changes, and iii) predictability of change from the value of  $l_e$  can be understood from these parameters. These methods have application in financial regulation whereby it is important to understand how individual changes to systems will impact the global behaviour. It also provides fundamental insights into spectral predictability in networks, and demonstrates how spectral perturbations can be a useful tool in understanding the interplay between micro and macro features of networks.

## 1 Introduction

Understanding how individual edges in a network influence its structure and evolution is important in a range of applications. Considering financial networks, network structure has implications for financial stability [1], market efficiency [2] and consumer safety [3]. Identification of players to monitor more closely is of paramount importance to regulators and policy makers, with many attributing the severity of the 2008 crisis to systemic flaws in the banking ecosystem [4]. Our research focuses on understanding how individual edges affect the structure of networks, and how this relates to network stability and evolution. First, we present a brief review of related literature, first considering individual effects on network structure, then those that link network structure to stability and systemic risk, before considering how network structure relates to temporal evolution. We then define a measure for structural edge importance  $l_e$ , and propose a model for network evolution in which an edge's importance can be indicative of future changes. Our results show that  $l_e$  values are higher for edges which appear to play a more important structural role, and that subsequent changes occurring in the real networks analysed depend to some extent on the value of  $l_e$ .

### 1.1 Individual effects on network structure

The effect an individual node or edge can have on a network's structure depends not only on the scale of its activity, but also on its position within a network, and the activity of neighbouring nodes and edges. Understanding these interrelations remains one of the key challenges in network science.

Recently, structural node importance has gained a large amount of attention due to its relevance in use cases across a wide range of fields [5]. In networks of epidemic processes, it is intuitive that the removal of a node that acts as a bridge between communities can be used to stem the spread of a disease, and significant effort has been taken to understand the influences of community structure on epidemic spreading [5, 6, 7]. Other applications include preventing network based attacks [8, 9] and understanding and actioning on the spread of gossip in society [10]. Wang et. al. [11] make use of the observation that the spectrum of the adjacency matrix gives an indication of community structure. In noting that for a network with  $c$  strong communities, the  $c$  largest eigenvalues of the adjacency matrix are significantly larger, they define node importance as the relative change in the  $c$  largest eigenvalues of the upon the node's removal. By making use of perturbation theory, they present a computationally efficient measure of node importance, which performs well in the case of homogeneous community sizes, however they note that their method fails to identify important nodes in a small community when a much larger community is present. Similar to Wang et. al., Lü et. al. [12] propose a universal structural consistency index for a network based on perturbing the adjacency matrix, and demonstrate that this index is a good index for link predictability. Our work considers the same central concept of applying perturbations to the adjacency matrix, however differs in that we are working at a lower level of granularity by proposing an edge based measure, rather than node or network based.

Similarly, many works in the financial literature focus on node specific influence on stability. For example, Battiston et. al. [13] define a novel node ranking coined DebtRank, which takes recursively into account the impact of distress of an initial node across the whole network. Their measure amounts to the fraction of the total economic value in the network that is potentially affected by the distress or default of a specific node. They applied their method to a network of loans from the Federal Reserve to financial institutions between 2008 and 2010, enriched with equity investment relations, and found a strongly connected core of 22 institutions which all became too systemically important to fail at the 2008 crisis peak. They demonstrated the effectiveness of their node ranking in comparison to other centrality measures, and found that it was the only measure to deliver a clear response well before the crisis peak. However, their method specifically considers the case of distress propagation, and does not explicitly measure how an individual node or edge affects the structure of the network in general. Barucca et. al. [14] investigate whether a change to few selected banks in the network of the e-MID<sup>1</sup> market can affect the large scale structure of a network through node removal or degree mutation, and comparing the network structure that results to the original.

Although the bulk of the attention has focused on importance of actors in networks, Helander et. al. [15] propose a method for characterising the relative importance of an edge, which they refer to as edge gravity. Edge gravity measures how often an edge occurs in any possible network path. They show that important edges are not necessarily adjacent to nodes of importance as identified by standard centrality metrics, and they also observe that high centrality nodes often have their centrality over-represented by being adjacent to ‘edges to nowhere’. Similar path based methods include the *BCC<sub>MOD</sub>* (Betweenness Centrality and Clique Model) proposed by [16], which weights the importance of the two nodes forming the endpoints of the edge with the number of cliques containing the edge. Their method outperforms several well-known methods including Jaccard coefficient and betweenness centrality in identifying critical edges both in network connectivity and spreading dynamic. Our methods contrast these path based approaches by instead considering a network spectrum based approach, however both approaches show strong connections to node centrality measures; as shown in section 2, an approximation to the network eigenvalue derivative is proportional to the product of the constituent nodes’ centralities. In addition, our research focuses on the temporal behaviour of the network in relation to structural importance, for which future work could consider using alternative measures of structural importance to understand the expected temporal behaviour.

## 1.2 Network structure in relation to stability and systemic risk

Increasing complexity and stability are inextricably linked, with works as early as May’s investigations into ecosystems with increasing biodiversity highlighting the relationship [17]. In the context of financial markets, although market integration and diversification are widely believed to play a stabilising role [18, 19], Bardoscia et. al. [20] demonstrated that two factors of increasing complexity, namely increasing the number of institutions (nodes) and contracts (edges) in an interbank network can drive the system to instability. By studying the behaviour of the leading eigenvalue of the leverage matrix, they first of all define pathways to distress in terms of the eigenvalue trajectory, and then demonstrate that these two mechanisms result in the emergence of cycles in the network, which amplify distress if the weights in the cycle sum to greater than 1. Similarly, Markose et. al. [21] present the idea of institutions being ‘too interconnected to fail’ through an exploration of the structure of the US CDS market. They consider an empirical network constructed from market shares, and make use of the May-Wigner condition for stability <sup>2</sup> in comparison to a random network.

---

<sup>1</sup>e-MID is the Italian electronic market for interbank deposits, a platform for trading unsecured money-market deposits.

<sup>2</sup>The May-Wigner condition for stability is a critical threshold below which any random network has a high probability of stability, and is defined as  $D < \frac{1}{Ns^2}$  where  $D$  is the network diameter,  $N$  is the number of nodes and  $s$  is the strength of average interactions between nodes

They show that although the CDS structure shows better outcomes than a random network when subject to shocks, the demise of any one big player will bring down other big players. Caccioli et. al. [22] showed in a theoretical exploration that uncontrolled proliferation of financial instruments can lead to large instability in markets, and suggest potential interventions such as the introduction of a Tobin tax [23], which is shown by Bianconi et. al. to have a stabilising effect [24]. Related to this, Brock et. al. [25] used ‘arrow securities’ as a proxy for more complicated hedging instruments, and found that these incentivise construction of larger positions, resulting in a reinforcement effect due to large gains/losses as a result of being on the ‘right’ or ‘wrong’ side of the market. They showed that this is associated with greater instability, and also that the primary bifurcation parameter, marking the onset of instability, occurs earlier when there are more arrow securities. In contrast to the majority of the data centric financial literature which focuses on interbank trading, Bardoscia et. al. [26] analysed UK Trade Repository data, which includes all transactions occurring through a Central Counterparty clearing house (CCP) in the UK. Looking at a snapshot of the open positions on a single day for interest rate derivatives, FX derivatives and credit default swaps, they compared a ranking derived from the centrality measures to a ranking derived from modelling the network’s response to liquidity contagion, looking at how shocks propagate across the network and translate into payment deficiencies across the different markets. The model considers the stress faced by an institution - the difference between all payments it is required to make and all payment inflows from counterparties, and allows stress to spill over between the layers. They found that centrality measures can be used as a proxy for the vulnerability of financial institutions.

### 1.3 Network structure in relation to temporal evolution

To understand network stability across time, many researchers have focused on studying the mechanisms for network growth, and defining network models to understand the origin of observed properties of real networks [27, 28, 29, 30, 31]. These include the Barabasi-Albert model [32], which demonstrates that scale-free degree distributions observed in real networks can be explained by the presence of growth and preferential attachment in the network evolution. Although the majority of these methods assume time independence in the growth mechanism, Falkenberg et. al. [33] present a simple adaptation to the Barabasi-Albert model , in which new nodes attach to nodes in the existing network in proportion to the number of nodes one or two steps from the target node. This results in an implicit time dependence, which arises from a node’s attractiveness being dependent on its local environment which changes as the network evolves. Central to their model is the idea that network structure and temporal evolution are inherently linked, however their model is limited to the influence of local environment. Others focus on considering temporal networks as multilayer networks, in which one can account for the fact that connectivity patterns in different layers can depend on each other. Bazzi et al. [34] proposed a generative model which explicitly incorporates a user-specified dependency between layers that is flexible enough to incorporate complex interlayer relationships such as dependencies between a layer and all layers that follows, incorporating memory effects into the model. A handful of studies have attempted to link global network structure to temporal evolution, such as Peixoto et. al. [35], who suggest dynamical variation of the degree-corrected stochastic block model that is capable of finding meaningful large-scale temporal structures in real-world systems and predict their temporal evolution. Their method works with both discrete and continuous time representations, making it versatile to a range of applications. Watts et. al. [36] consider semi-random ‘small world’ networks and show that the dynamics are an explicit function of the network structure, and also show find an enhanced propagation speed for small world networks.

A common and general framework for network growth is the fitness model, in which each node has associated with it a time independent ‘fitness’ which represents its propensity to attract links, as proposed by Barabási and Bianconi [37]. They find that different fitnesses results in multiscaling in the dynamic evolution, or in other words that the time dependence of a node’s connectivity depends on the fitness. Attempts have been made build on this model in order to understand the origins of network dynamics, such as a recent study by Kobayashi et. al. [38]. They find that population and activity dynamics are sufficient to explain two types of scaling empirically observed in real networks, however their methods do not explicitly allow for of different roles to be captured within a network, by assuming a uniform distribution of fitness parameters. In our research, we explore instead how an edge level quantity derived from the spectrum of a network can similarly be used to determine which edges change in the network. We present methods for estimation of parameters which control both the overall activity in the network, as well as the bias to change for edges with a larger structural importance, and show how these reproduce behaviours observed empirically.

In the following sections, we look to address two questions: Can we quantify the extent to which an edge affects the overall network structure, and does this provide information on the network’s temporal evolution? We focus on looking at the derivatives of the network’s leading eigenvalue with respect to individual nodes or edges, to assess their ‘importance’. We further present evidence that this measure could be a useful indicator in understanding temporal changes in network structure, and present the results of application to a selection of relevant real network datasets. Our main results demonstrate that the elementwise derivative of the leading eigenvalue ( $\lambda_e$ ) is predictive of subsequent change for five different networks analysed, and that predictability can be related to the specific realisation of two parameters,  $\alpha$

and  $\rho$  network evolution model in which edges change with probability  $\alpha l_e \rho$ . This has potential implications for stability, as a system experiencing more changes to edges of structural dominance could see a reinforcing effect, leading to an unstable system. These methods could be useful in classifying financial asset systems to inform regulation activities and policy making. We further show that the scale of resultant changes can be related to the realisation of two additional parameters  $\beta$  and  $\gamma$ , again with potential stability implications.

## 2 Methodology

### 2.1 Definition of temporal networks

Traditionally, network analytics has focused on static representations of networks, either looking at single snapshots in time, or considering a projection of the time dimension onto a static view by aggregating the links in a time window. In doing so, some, or all, of the temporal information about the network is lost.

However, recently, there have been developments in modelling of systems as temporal networks, for which the system is represented by a contact sequence  $(i, j, t)$  where  $i$  and  $j$  constitute the vertex set  $V$  at time  $t$ . This representation also allows for edges which take time to traverse, or contracts completed after a duration  $\delta t$  by representing the contact sequence as  $(i, j, t, \delta t)$  [39]. Since we are considering transactions as instantaneous, we are not interested in transmission time for edges, and we are considering applications where time is discretised, we can formally define a temporal graph  $G_t^w(t_{min}, t_{max})$  as in [40] as the ordered sequence of graphs

$$G_{t_{min}}, G_{t_{min}+w}, \dots, G_{t_{max}} \quad (1)$$

where  $w$  is the size of the time aggregation (e.g. daily). The adjacency matrix  $A_{ij}^s$  between nodes  $i, j$  at time  $s$  is 1 if and only if there exists a link between  $i$  and  $j$  in  $G_t$ ,  $t \leq s \leq t + w$ .

This differs from a time sequence of static graphs in that the edges in the temporal network need not be transitive, i.e.  $A \rightarrow B, B \rightarrow C \not\Rightarrow A \rightarrow C$ , and also allows for time to be continuous, meaning the full topological structure and correlations are captured. Temporal networks can be extended to include weights of the edges.

### 2.2 Central concept - eigenvalue derivatives as a measure of importance

For a given graph  $G_t(V, E)$  with adjacency matrix  $A_{ij}^t$ , the eigenspectrum of  $A_{ij}^t$  is the set of eigenvalues  $\lambda$  that satisfy the equation

$$\mathbf{Ax} = \lambda \mathbf{x} \quad (2)$$

By observing changes in the eigenspectrum of a graph, we can gain an insight into structural changes. As we are looking at network snapshots across time, we have a ‘time series’ of graphs and we can consider the change in the leading eigenvalue between successive time snapshots,

$$\Delta \lambda = \lambda(A^{(t+1)}) - \lambda(A^{(t)}) \approx \sum_{ij} \frac{\partial \lambda}{\partial A_{ij}} \Delta A_{ij} \quad (3)$$

where we have made a first order approximation, and the derivative is with respect to the  $(i, j)$ th entry of the matrix, as opposed to the entire matrix. Here  $\lambda$  refers to the leading eigenvalue of the adjacency matrix.

The two parts of equation 3 can be seen as a play off between the potential of an edge to influence the structure ( $\frac{\partial \lambda}{\partial A_{ij}}$ ) and the actual change in the network structure ( $\Delta A_{ij}$ ). Our experiments with synthetic networks look to assess the extent to which our derivation below, which makes approximations and assumptions, captures the true behaviour.

The first term here measures the sensitivity of the eigenvalue to changes in an individual edge, which we refer to as the structural importance of an edge, which we refer to as  $l_e$ . We derive approximations for  $l_e$  in equation 4 for the undirected case by taking a perturbation theory approach. Although not explicitly explored in this paper, we also present equation 5 for the directed case:

$$l_e = \frac{\partial \lambda}{\partial A_{ij}} = 2\lambda_{0,i}\lambda_{0,j} \quad (4)$$

$$\frac{\partial s^A}{\partial M_{ij}} = \frac{\lambda_{0,i}^M \lambda_{0,j}^M}{2s^A} \quad (5)$$

where  $\lambda_{0,i}$  refers to the  $i$ th component of the eigenvector corresponding to the leading eigenvalue,  $s^A$  refers to the leading singular value of the adjacency matrix and  $\lambda_{0,i}^M$  refers to the  $i$ th component of the eigenvector corresponding to

the leading eigenvalue of  $\mathbf{M}$ . Full derivations for these can be found in appendix B, and we validate the approximation for the undirected case in results section 4.1. We see here that both equations 4 and 5 are proportional to the product of the eigenvector centralities of the nodes involved in the edge. Our definitions are defined in terms of the eigenvector corresponding to the largest eigenvalue, which usually has non-zero values only for the largest connected component of a network. For this reason, in this paper we restrict ourselves to exploring the giant component of the networks, however generalising these to allow for disconnected components will be considered in future work.

We can capture the relationship between  $l_e$  and subsequent edge changes by observing the distributions of  $P(\Delta A = 0 | \ln(l_e))$  and the joint probability  $P(\Delta A, l_e)$ , which we explore in detail in the results sections 4.2.5 and . Our findings from these are compared to our model for the temporal evolution of networks which we propose in section 3, to assess the extent to which our model captures the true behaviour observed.

The second term considers the changes that subsequently occur in response to the value of  $l_e$ . This is of significance from a stability perspective; edges that are structurally important could cause a system to become unstable by changing frequently or by a large amount. Conversely, they may also act to stabilise a system if it begins to move towards a regime of instability. This can be explored by considering temporal graphs as described by equation 1, and assuming that the evolution is Markovian. We consider this first of all in the proposal of a model for network evolution, parameterised by the extent to which  $l_e$  is indicative of the propensity of an edge to change, and the scale of the resultant changes. We further assess the predictability of changes from the value of  $l_e$  through the use of a logistic regression classifier, and relate the performance of this to the model parameters.

### 3 Model for network evolution

In order to understand the relation between structural importance and stability of a network over time, we need a model that captures two behaviours. The first of these is that the value of  $l_e$  is indicative of the probability for an edge to change, and the second is that the size of a resultant change can be related to  $l_e$ .

We thus propose a model in which we can control the extent to which  $l_e$  influences a subsequent edge change, both in probability of occurrence and resultant scale. Specifically, we propose a model in which the network evolution exhibits the Markovian property as in as in [28, 35]:

$$A_{ij}^t = V_{ij}^t A_{ij}^{t-1} U_{ij}^t + (1 - V_{ij}) A_{ij}^{t-1} \quad (6)$$

where  $V_{ij} \sim \mathbb{B}(\alpha(l_{eij})^\rho)$  and  $U_{ij}^t$  is the distribution of edge changes. Here we introduce two parameters which control the probability of an edge to change -  $\rho$  which controls the level to which the value of  $l_e$  influences the probability for an edge to change, and  $\alpha$  scales  $V_{ij}$  to ensure that it is a valid probability. A positive value for  $\rho$  indicates that more important edges are more likely to change, and a negative  $\rho$  would indicate the opposite.

The simplicity of this model means that we are unable to account for edges appearing and disappearing in the network. We will look to incorporate this in future research.

#### 3.1 Parameter estimation in real networks

Assuming that our data evolves according to the model in equation 6, we can use observations from real networks to estimate the most likely values of  $\alpha$  and  $\rho$  from the data. Following a maximum likelihood approach, we can derive estimations for these parameters, by maximimising the following log-likelihood as proposed in appendix E:

$$\ln(L(\mathbf{x}|\theta)) = \sum_e k_e \ln(\theta_e) + (1 - k_e) \ln(1 - \theta_e) \quad (7)$$

We note here that since  $\alpha$  and  $\rho$  are constrained to result in a valid probability calculated from  $\alpha l_e^\rho$ , the minimisation is subject to constraints and must satisfy the Karush-Kuhn-Tucker conditions[41]. In practice, numerical optimisation of the log-likelihood in equation 7 was used to estimate  $\alpha$  and  $\rho$ .

#### 3.2 Structural influence and network predictability

Depending on the values of the parameters for a given dataset, we might expect the observed values of  $l_e$  to be predictive of subsequent change. Specifically, since  $\rho$  controls the relationship between  $l_e$  and the propensity for an edge to change, a high value of  $\rho$  would suggest that  $l_e$  would be more predictive of future change. Similarly for  $\alpha$ , within the constraints for  $\alpha l_e^\rho$  to give the probability of an edge to change, a larger  $\alpha$  factor will increase the distance between

change probabilities for edges with different  $l_e$ , thus also strengthening the relationship between the value of  $l_e$  and the propensity for an edge to change. In order to evaluate these effects, we make use of logistic regression for classification of edges into changing vs. unchanging from the values of  $l_e$ , and compare the results to a null model consisting of the average over multiple trials in which edges randomly change with probability equal to the fraction of observed changes. The data is split into training and test sets in a stratified manner, with 20% used to test the model on unseen data. The predictions are compared according to balanced accuracy, defined as the average of recall obtained on each class, and Area Under Curve scores for both Receiver Operating Characteristic curves and Precision Recall curves.

## 4 Results

### 4.1 Validation of $l_e$ using toy networks

Here we assess the extent to which the approximations made in calculating  $l_e$  hold. We do this by approximating the change in eigenvalues as the coefficient weighted sum of the edge weight changes,  $\Delta\lambda = \sum_e l_e \Delta A_e$ , and comparing the gradient of this to the value of  $l_e$ . Our derivation of  $l_e$  makes the simplification in assuming that edge changes occur independently of each other. Our first test thus considers the case of an individual edge changing at each timestep, and we consider perturbations applied to a barbell graph, to observe the effects of network structure, a ring graph, to observe the effects of weight with structural equivalence, and a Erdős–Rényi (ER) graph as a baseline. The results in figures 1, 2 3 show the line of constant  $l_e$ , overlaid with the observed  $\Delta A_e$  and corresponding  $\Delta\lambda$  values.

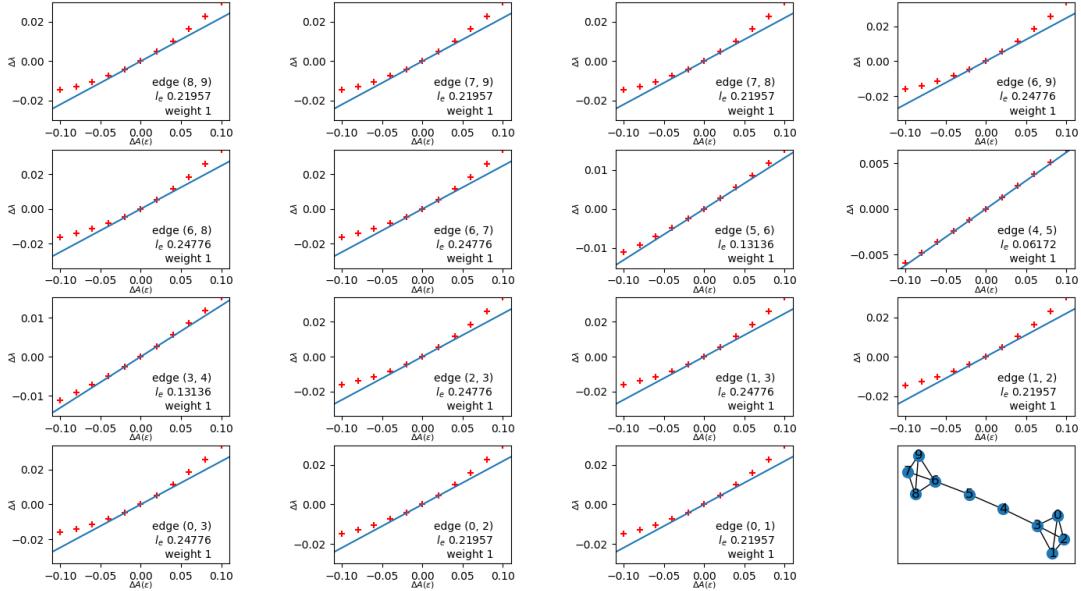


Figure 1: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta\lambda$ , compared to line of constant  $l_e$ . Barbell graph, with equal initial weights

We see here that our linear approximation generally holds for relative edge changes less than  $\Delta A = 0.05$ . We also see for the barbell graph that  $l_e$  captures the structural role of the edges, with edges in the cliques having higher values of  $l_e$  than those in the bridge. For the ring graph, we observe a poorer fit for edges with low values of  $l_e$ , and the larger  $l_e$  edges tend to be adjacent to edges with similar  $l_e$  values. Although the edge with the largest weight also has the largest value of  $l_e$ , in general there does not appear to be a simple relationship between edge weight, or weight of neighbouring edges, and the value of  $l_e$ . For the weighted random network we see similar observations are made for the weighted ER graph, with the lowest  $l_e$  values observed for more peripheral edges, and the two edges with the largest weights also having the highest  $l_e$  values. Further results for the case of a weighted barbell, and unweighted ring and random networks are shown in the appendix A.

Results for the case of two edges changing are shown in appendix A. In these we observe for the barbell graph better fit is observed for higher values of  $l_e$ . For the ring networks and random networks, we see that our model performs well if the observed edge has a larger value of  $l_e$  than the other changing edge, but performs poorly when the value of  $l_e$  is

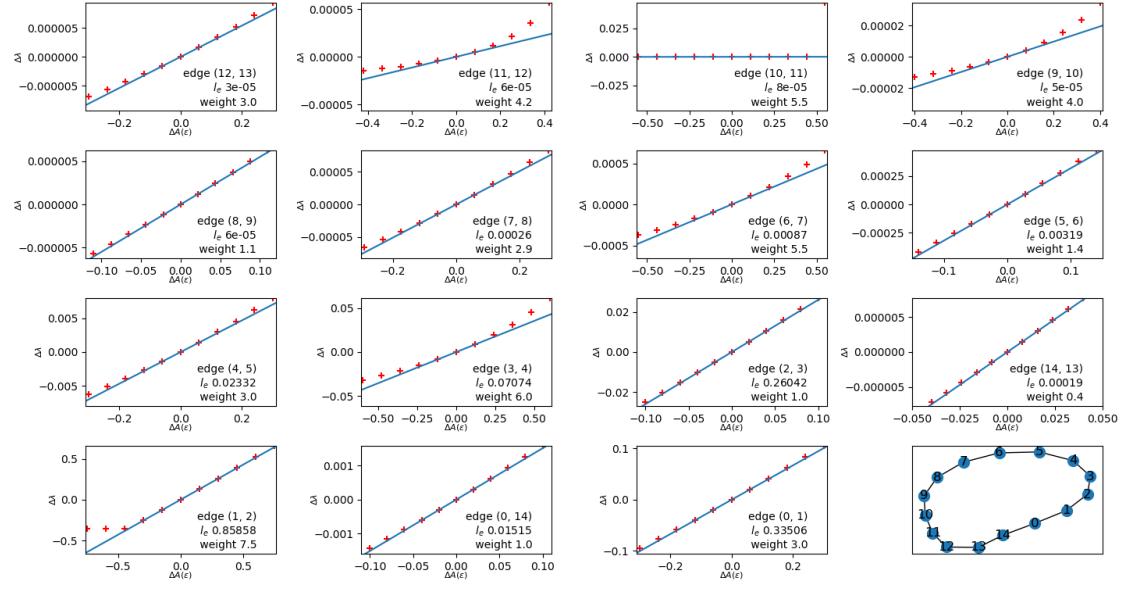


Figure 2: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta\lambda$ , compared to line of constant  $l_e$ . Ring graph with each edge independently assigned a random integer between 1 and 10.

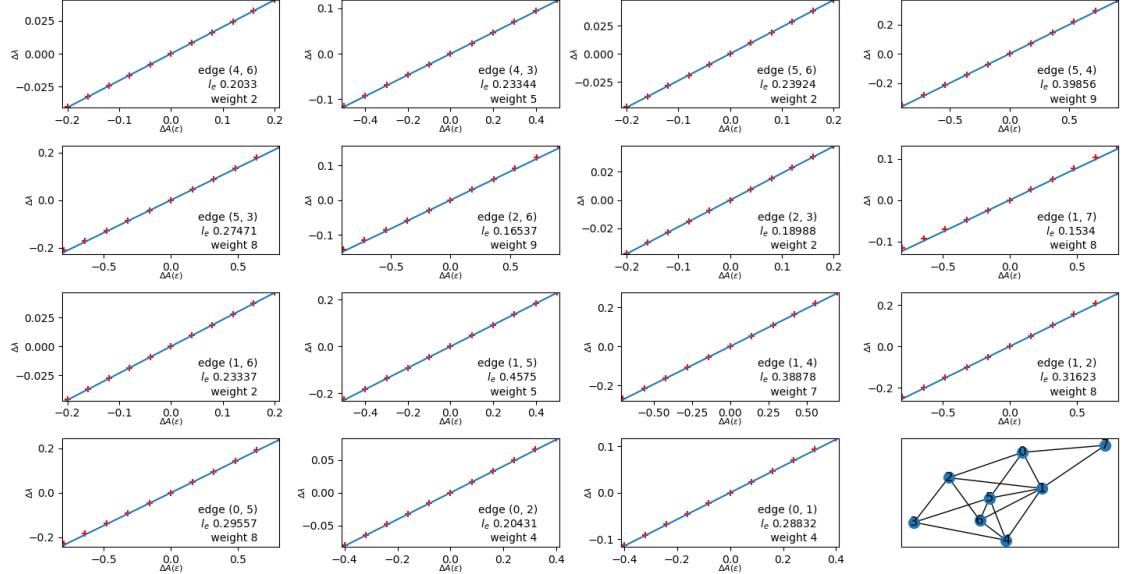


Figure 3: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta\lambda$ , compared to line of constant  $l_e$ . Erdős–Rényi graph with each edge independently assigned a random integer between 1 and 10.

smaller. The case of complete structural equivalence and equal weights in the unweighted ring network shows good performance for all edges.

The breakdown of the method when there are multiple changes occurring between snapshots suggests that our approximation for  $l_e$  may be better suited to a continuous or pseudo-continuous representation of a temporal network, which can be seen as the limit of a discrete temporal network in which each snapshot captures an individual edge change occurring at an infinitesimally different time to the neighbouring snapshot changes.

## 4.2 Relationship between $l_e$ and the presence of edge changes

We can understand the role of the parameters  $\alpha, \rho$  by observing the effect of varying the parameters on the distributions of the values of  $l_e$  for changing vs. non-changing edges,  $P(\Delta A = 0 | \ln(l_e))$ . We first consider this for data generated according to our model in equation 6, first keeping  $\rho$  fixed and varying  $\alpha$ , then fixing  $\alpha$  and varying  $\rho$ .

### 4.2.1 Model with varying $\alpha$

Figures 4 and 5 show the resulting distributions for varying values of  $\alpha$ . We see that an increase in  $\alpha$  results in a decrease in the probability of an edge to change for all values of  $l_e$ , and for larger values of  $\alpha$ , the rate of increase of change probability with  $l_e$  is slightly larger.

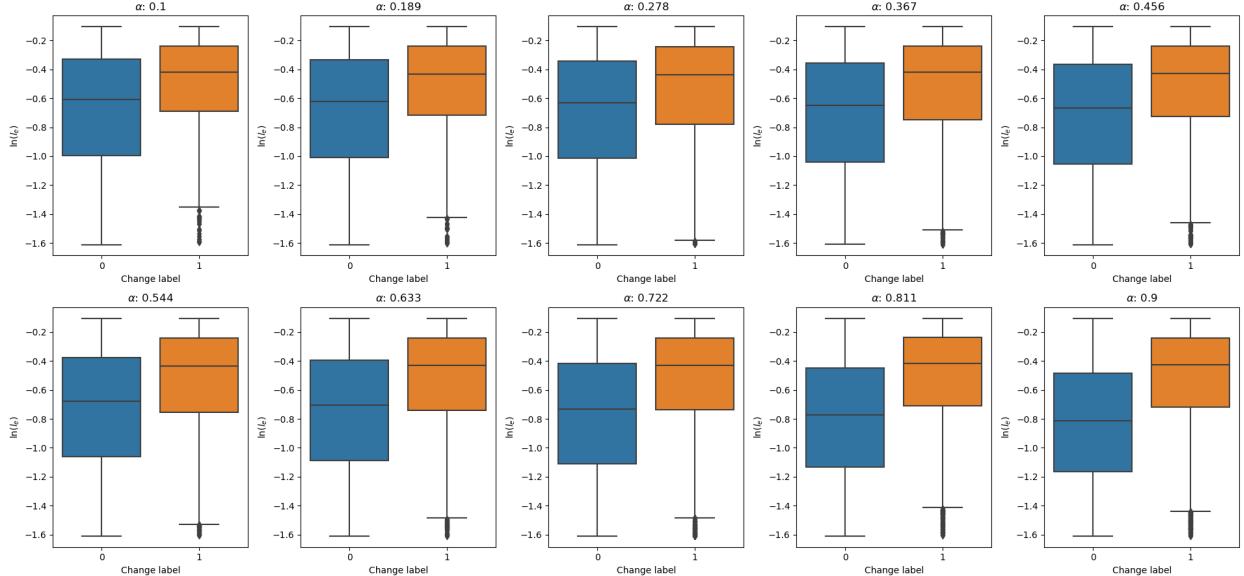
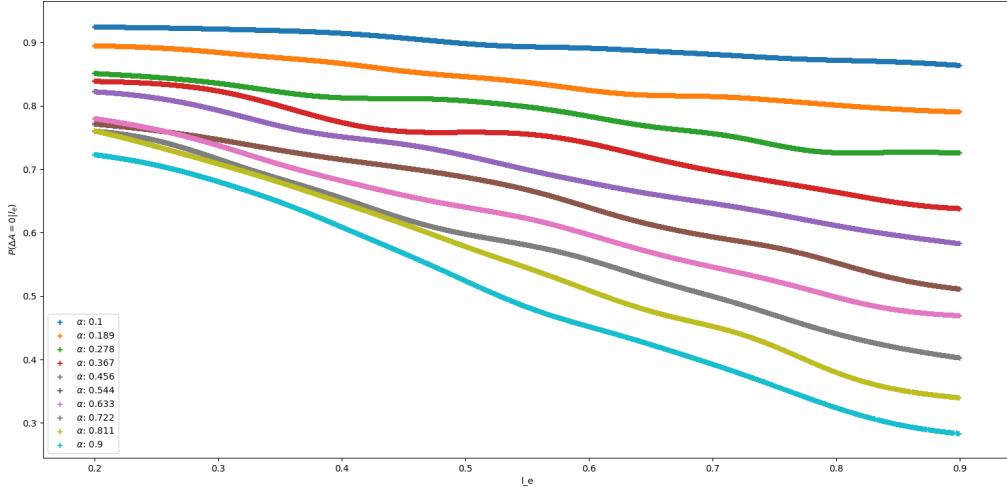
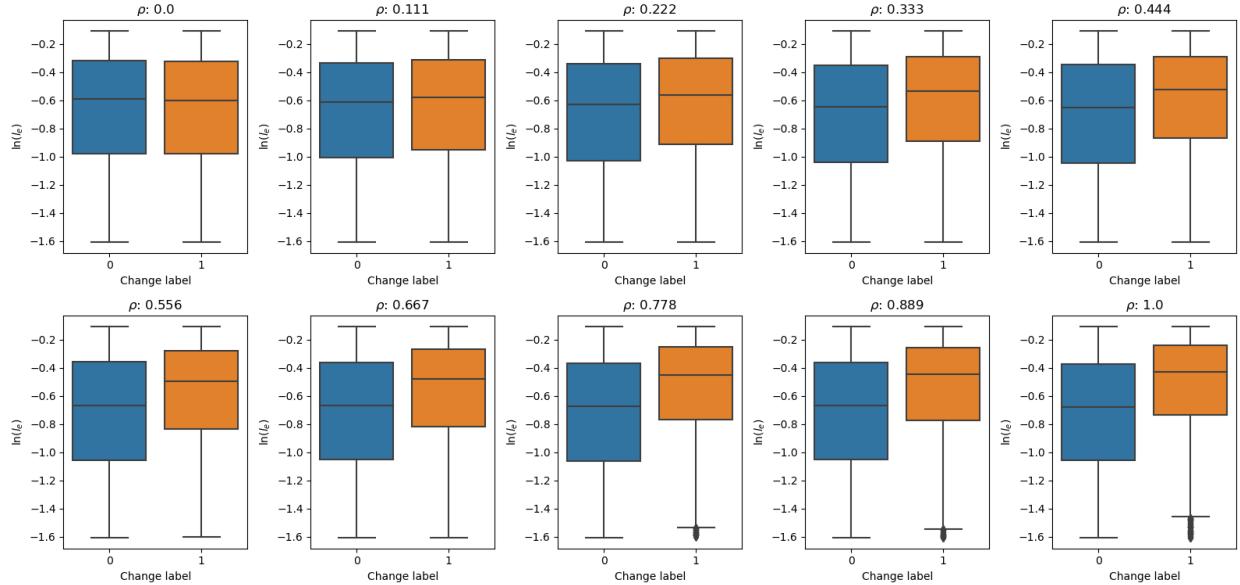


Figure 4: Distributions of  $l_e$  for edge changes vs. no changes, when varying  $\alpha$ .

Figure 5:  $P(\Delta A = 0 | \ln(l_e))$  as a function of  $\ln(l_e)$  for  $0.1 < \alpha < 1$ 

#### 4.2.2 Model with varying $\rho$

Figure 7 shows the distributions of  $P(\Delta A = 0 | \ln(l_e))$ . We see here that for increasing  $\rho$ , the probability of observing no change increases, and also for increasing  $l_e$ , the probability decreases for a given  $\rho$ , at a rate that shows a significant dependence on  $\rho$ .

Figure 6: Distributions of  $l_e$  value for the case of edge changes vs. no changes.

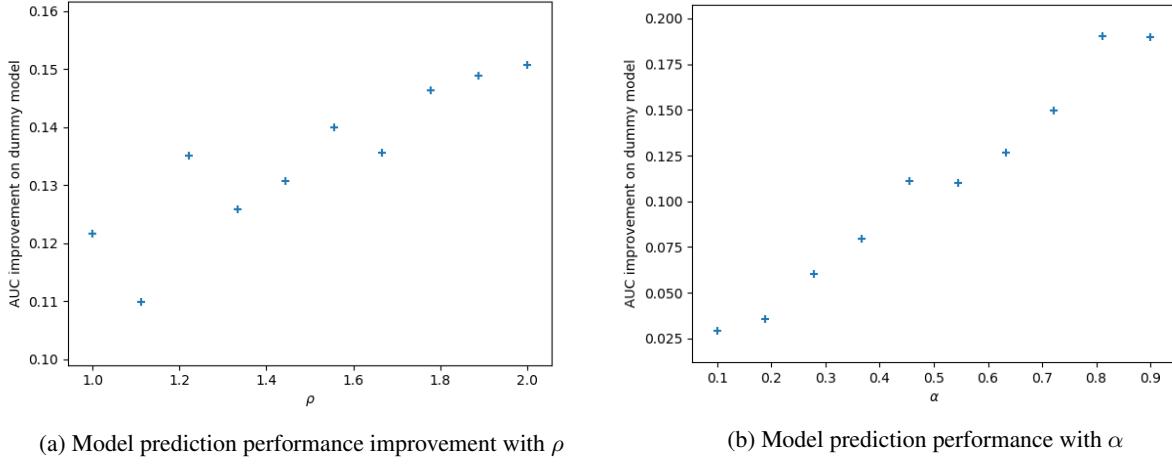
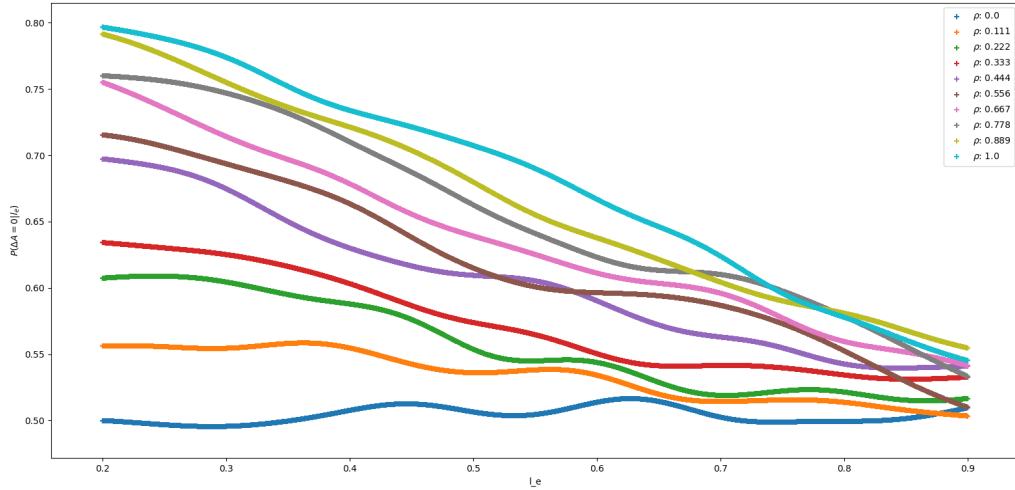


Figure 8: Improvement from dummy model of Precision-Recall AUC scores

Figure 7:  $P(\Delta A = 0 | \ln(l_e))$  as a function of  $\ln(l_e)$  for  $0 < \rho < 2.5$ .

#### 4.2.3 Predictability improvement with $\alpha$ and $\rho$

As detailed in section 3.2, here we apply a logistic regression classifier with single feature  $l_e$ , to datasets with varying  $\alpha$  and  $\rho$ . Figures 8a and 8b show the improvement in the test set Precision-Recall Area Under Curve scores for increasing values of each parameter. We see from these that increasing both parameters improves the predictability of changes given the value of  $l_e$ , consistent with our observations of the rate of increase of change probability being positively correlated with both  $\alpha$  and  $\rho$ .

#### 4.2.4 Static observations in real data

We have seen in the above application to synthetic networks that our model behaves as expected, with networks with a large  $\rho$  (and  $\alpha$ ) being more predictable. Now we explore the performance of our structural influence metric and model through the application to five real datasets. Firstly, given that our research has been motivated by a need to monitor risks in a financial setting, we considered a network of country level bilateral trade [42] and three different capital markets transaction datasets reported under MIFID II regulations were considered. However, our methods can

be applied more generally to any temporal networks, and due to the availability and high volume of research conducted into social networks (see [10]), we also considered a network of messages sent between College students [43]. A full description of these can be found in appendix D.

In order to understand the usefulness of  $l_e$  as a metric for structural importance, we first examine the edges that rank the highest according to their values of  $l_e$  for the bilateral trade dataset, the historical context of international trade can give us an idea of which edges we might expect to be ‘important’. For the bilateral trade dataset, we see the largest values of  $l_e$  for the edge between Portugal and Spain in 1872, and considering the sum across all time, for Greece and Turkey. These are examples of edges with both nodes having large eigenvector centrality; edges involving only one central node are seen to have lower values of  $l_e$ . This means that inter-European edges almost exclusively make up the top 100 ranked edges, whereas the lowest ranked  $l_e$  edges occur when one, or both, of the nodes have very low centrality scores. Similarly, for the other datasets, the highest values of  $l_e$  were also observed for edges involving nodes with high eigenvector centrality. In general, we see that the rankings of  $l_e$  are uncorrelated with the rankings of edges according to their betweenness centrality, or their mean value of  $\Delta A$ , however do for some cases correlate with the product of the participating node’s degrees and strengths.

Dataset	$\text{Corr}(l_e, \Delta A)$	$\text{Corr}(l_e, \text{EBC})$	$\text{Corr}(l_e, \deg_{n1} \times \deg_{n2})$	$\text{Corr}(l_e, S_{n1} \times S_{n2})$
Bilateral Trade	-0.061	-0.397	0.352	0.786
College Messaging	-0.169	0.035	0.581	0.434
Equity-1	-0.104	0.135	0.717	0.320
Equity-2	-0.047	0.166	0.580	0.265
Equity-3	-0.010	0.041	0.923	0.763

Table 1: Spearman’s rank correlations for  $l_e$  with the rank by edge weight, edge betweenness centrality and product of nodes’ degrees.

As these datasets contain large numbers of edges (the smallest contained 2785 edges), we cannot fully explore all of the individual observed values of  $l_e$  as for the toy networks. Instead, we consider the probabilities of observing values of  $l_e$  by making use of Kernel Density Estimation to estimate the probability density functions from the data.

Figure 9 shows the estimated Probability Density Functions of the logarithm of the value of  $l_e$ . We see from these that for all networks, the values observed for  $l_e$  tend to be very small. Omitting the tails of the distributions for diminishingly small values of  $l_e$ , we see a similarity in the values of  $l_e$  observed across 3 similar equity datasets, and although across all 5 datasets analysed, the distribution is found to be approximately lognormal, the social network shows a much broader distribution of  $l_e$ . The peak of the distribution for the college messaging dataset is also much lower, observed at approximately  $\ln(l_e) = -8.8$ , whereas the bilateral trade dataset shows a peak at -3.3 , and the equity datasets at -3, -2.5 and -4.2.

#### 4.2.5 Dynamic observations in real networks

We now address the central concept of the relationship of  $l_e$  observed for our real networks and the probability of an edge to change. Figure 10 shows the distributions of the  $\ln(l_e)$  values observed for non-changing edges in comparison to changing edges. We see that in all cases, there is a shift in the mean value of  $\ln(l_e)$  towards higher values for edges which do change, which would be suggestive of a positive  $\rho$  parameter, and an ability to predict the presence of changes given the value of  $l_e$ . The smallest shifts are observed for the Bilateral Trade dataset and Equity-3, which show negligible differences in the mean and quartiles of the values of  $l_e$  for changes and no changes. In all cases, the differences in the mean values of  $l_e$  for change vs. no change is significant, with a two-sided t-test showing  $p < 0.05$  for all datasets.

To further understand how the value of  $l_e$  relates to the probability for edges to change, we look at the distributions of  $P(\Delta A = 0 | l_e)$  as shown in 11. Here we see a decreasing probability of  $\Delta A = 0$  for the bulk of the distribution for increasing  $l_e$  for the bilateral trade and Equity-3 datasets, however the rarely observed edges with  $l_e > 0.3$  for these datasets show larger probabilities to remain unchanged. We again see a slight initial decrease for Equity-1 and 2 datasets however the relationship is clearly non-linear for large  $l_e$ . The college messaging dataset shows a much larger probability in general for edges to remain unchanged, and shows a very slight decrease in probability to remain unchanged for very small  $l_e$  values, however is dominated by noise for  $l_e > 0.05$ .

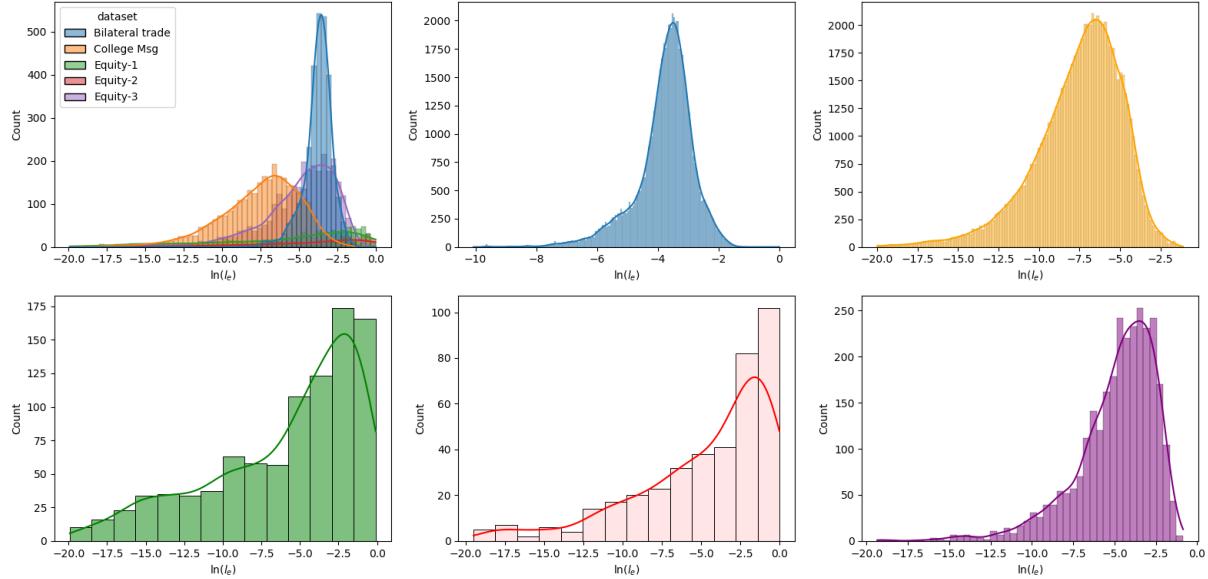


Figure 9: Probability distribution of the values of  $\log(l_e)$  for different networks.

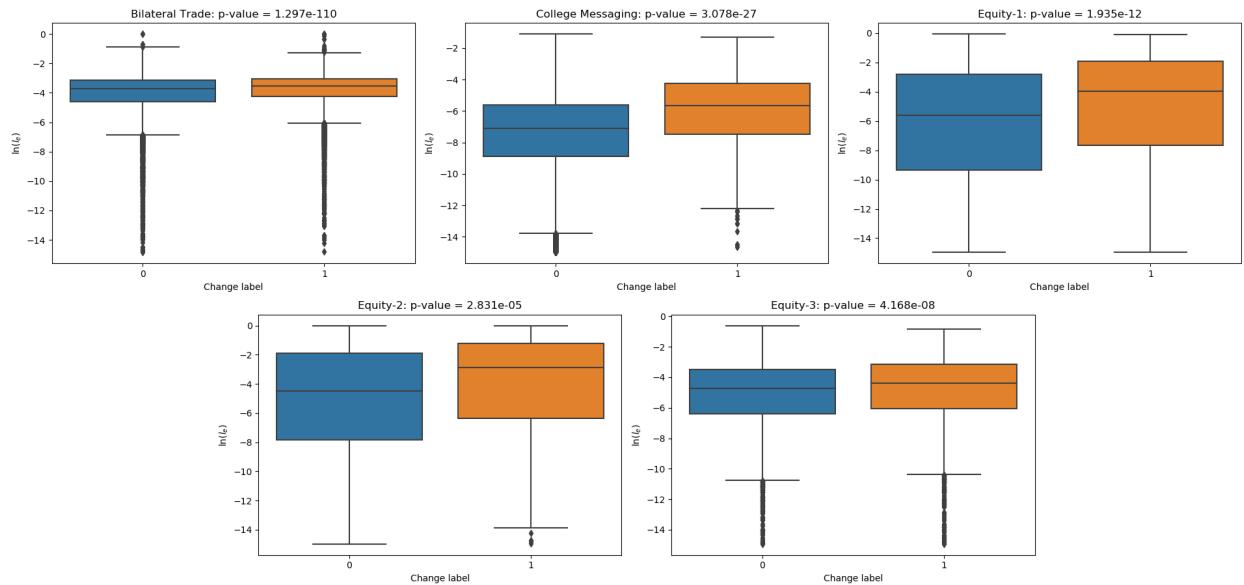


Figure 10: Boxplots showing the distribution of  $l_e$  values observed according to the presence or absence of an edge subsequently changing.

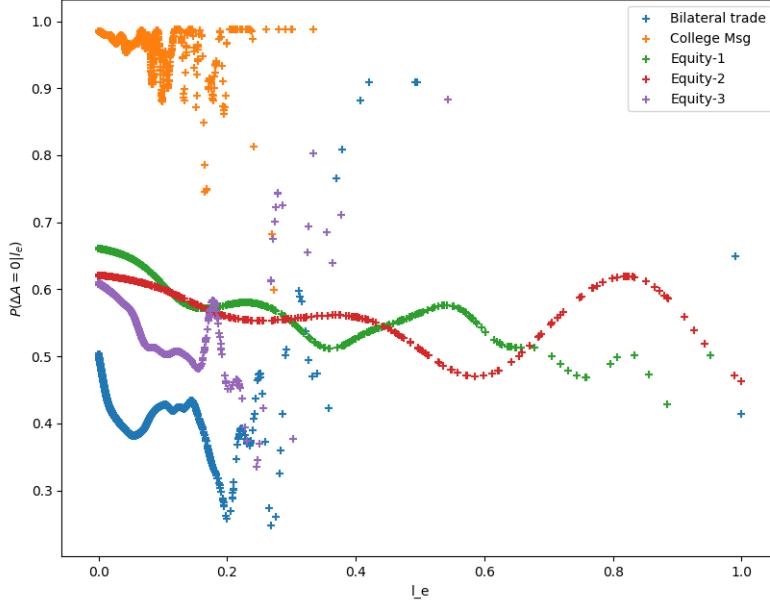


Figure 11:  $P(\Delta A = 0 | \ln(l_e))$  as a function of  $\ln(l_e)$  for the 5 real datasets.

Comparing to section 4.2, we considered the ideal cases of linear positive, neutral and negative relationships between  $l_e$  and the probability of edge changes. In reality, as shown in figure 11, we see things are more complex, with different relationships apparent for different  $l_e$  ranges. In particular, for edges with lower values of  $l_e$ , the negative relationship between the value of  $l_e$  and the probability of an edge to remain unchanged suggests that a parameterisation of our model with positive value of  $\rho$  would be effective in capturing the behaviour of the bulk of the network. However changes to the small handful of edges with the largest values of  $l_e$  are less likely. These observations could suggest that there are a few structurally important edges which act to stabilise a system which would otherwise move towards a regime of instability.

#### 4.2.6 Estimation of $\alpha$ and $\rho$ from data

In table 2, we present the values of  $\alpha$  and  $\rho$  estimated for our 5 different datasets. The errors on these estimations are given by the inverse hessian of the Log-Likelihood, which is found by numerical approximation. In comparison with figures 5 and 11, we see the ordering of the estimated value of  $\alpha$  appears to agree with the positions of the college messaging dataset and the equity datasets. The parameter  $\rho$  appears to correspond with the overall gradients observed in figure 11 for the bulk of the distributions observed for low values of  $l_e$ . These observations suggest that our model is mostly capturing the imbalance of observed changes in the parameter  $\rho$ , and the overall average change probability for each dataset in the parameter  $\alpha$ .

Dataset	Estimated $\alpha$	Estimated $\rho$
Bilateral trade	$0.783 \pm 1.08 \times 10^{-4}$	$0.072 \pm 1.21 \times 10^{-5}$
College messaging	$0.033 \pm 3.03 \times 10^{-6}$	$0.270 \pm 2.76 \times 10^{-6}$
Equity-1	$0.392 \pm 2.74 \times 10^{-4}$	$0.030 \pm 2.82 \times 10^{-5}$
Equity-2	$0.401 \pm 3.13 \times 10^{-5}$	$0.016 \pm 4.93 \times 10^{-5}$
Equity-3	$0.465 \pm 2.57 \times 10^{-4}$	$0.036 \pm 4.21 \times 10^{-5}$

Table 2: Estimated  $\alpha$  and  $\rho$  for the 5 real different datasets

Figure 31 in the appendix shows the result of generating distributions of  $P(\Delta A = 0 | l_e)$  for the estimated parameters, in comparison to the real datasets, restricted to  $l_e < 0.3$  in order to observe the bulk of the distributions. We see

that the dataset generated according to the parameters estimated appear to show a reasonable agreement to the actual distribution, and differences here can be attributed to the differences in the initial network conditions.

#### 4.2.7 Edge change predictability

Given the non-zero estimated values of the parameters  $\alpha$  and  $\rho$ , it is natural to assess the performance of using the value of  $l_e$  to predict a subsequent change. Figures 12 and 13 show the Receiver Operating Characteristic and Precision-Recall Curves for the 5 different datasets.

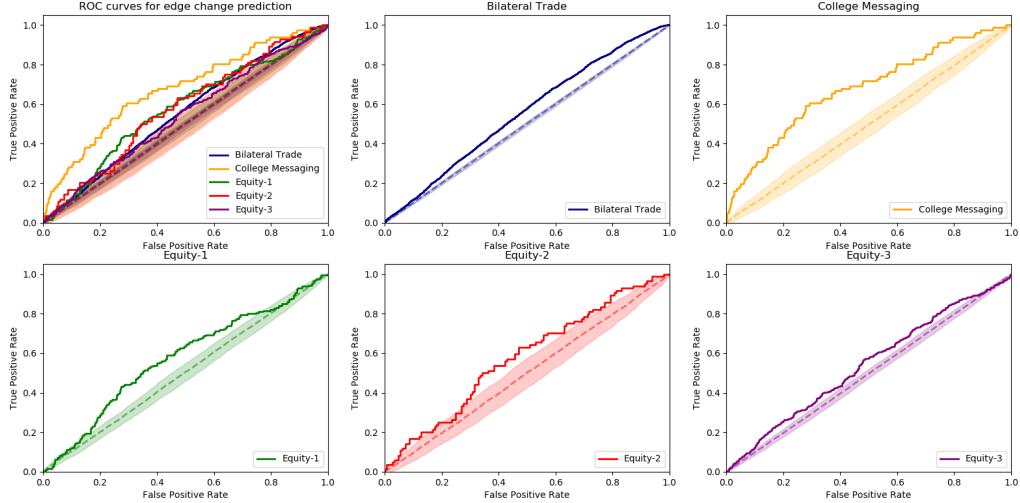


Figure 12: ROC curves for a logistic regression classifier making use of  $\ln(l_e)$  to predict  $\Delta A = 1$ . The dashed lines and shaded areas represent the mean 95% confidence intervals for the dummy model.

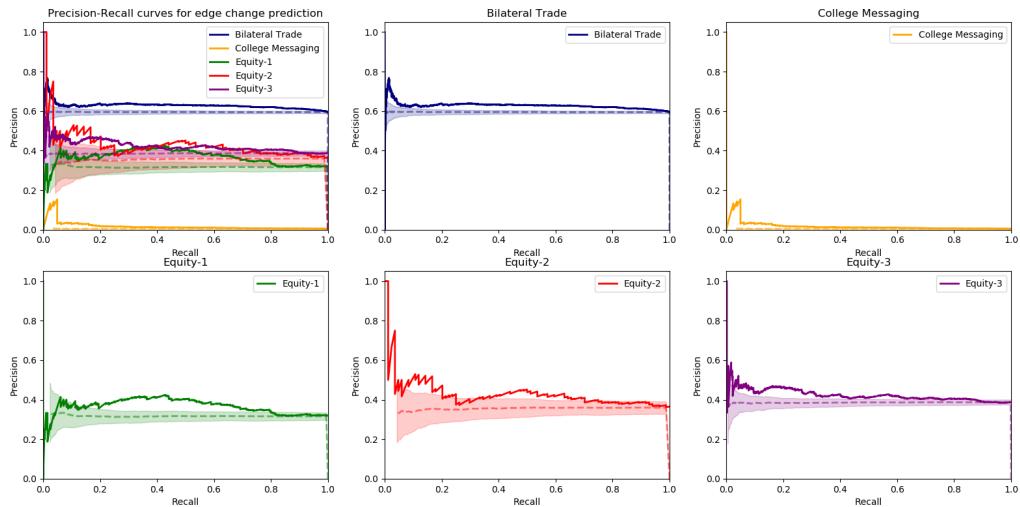


Figure 13: ROC curves for a logistic regression classifier making use of  $\ln(l_e)$  to predict  $\Delta A = 1$ . The dashed lines represent the results for a stratified random allocation of labels.

All datasets are seen to perform slightly better than the dummy model, with better performance seen for the College Messaging dataset and Equity-1 and 2, which also show larger differences in the distribution of  $l_e$  across change vs. no change in figure 6. Poorer performance is seen for the bilateral trade and Equity-3 datasets, which show similar shaped distributions in figure 11 with an initial steep decrease in probability to remain unchanged for increasing  $l_e$ , however this trend appears to reverse for  $l_e > 0.3$ . These datasets also show little difference in the distribution of values observed in 6 and are found to have low values of  $\rho$ . Although the college messaging dataset shows the best performance, particularly in the left hand side of the ROC curve, this is driven by the significant class imbalance with only 5% of the observations showing a non-zero  $\Delta A$ , as opposed to the bilateral trade dataset which shows a 20% proportion of non-zero changes.

Dataset	Balanced accuracy	ROC AUC	Precision-Recall AUC
Bilateral trade	0.542 (0.5)	0.554 (0.5)	0.628 (0.595)
College messaging	0.623 (0.5)	0.678 (0.5)	0.017 (0.005)
Equity-1	0.568 (0.5)	0.576 (0.5)	0.365 (0.313)
Equity-2	0.566 (0.5)	0.579 (0.5)	0.430 (0.351)
Equity-3	0.527 (0.5)	0.542 (0.5)	0.424 (0.381)

Table 3: Values of Area Under Curve scores for ROC and Precision-Recall curves. Numbers in brackets represent the score achieved by a model which randomly predicts 1 or 0 in proportion to the dataset prior.

### 4.3 Relationship of between $l_e$ and size of weight changes

We now consider if the value of  $l_e$  is observed to have an affect on the scale of subsequent edge changes. As in section 4.2, we again consider data generated according to the model in 6, and we choose to take  $U_{ij}^t = \mathcal{N}(\mu = 0, \sigma = \beta l_e^\gamma)$ . This introduces two new parameters,  $\beta$  which controls the width of the distribution of edge changes, and  $\gamma$  which controls the level to which  $l_e$  influences the variance of the edge change distribution.

#### 4.3.1 Variation of $\gamma$

Figure 14 shows the distributions of  $P(\ln(1 + \Delta A), l_e)$  for a range of values of  $\gamma$ . We see here that for positive  $\gamma$ , the width of the distribution widens for larger  $l_e$ . For negative  $\gamma$ , we see the opposite, that the width of the distribution becomes narrower for larger  $l_e$ .

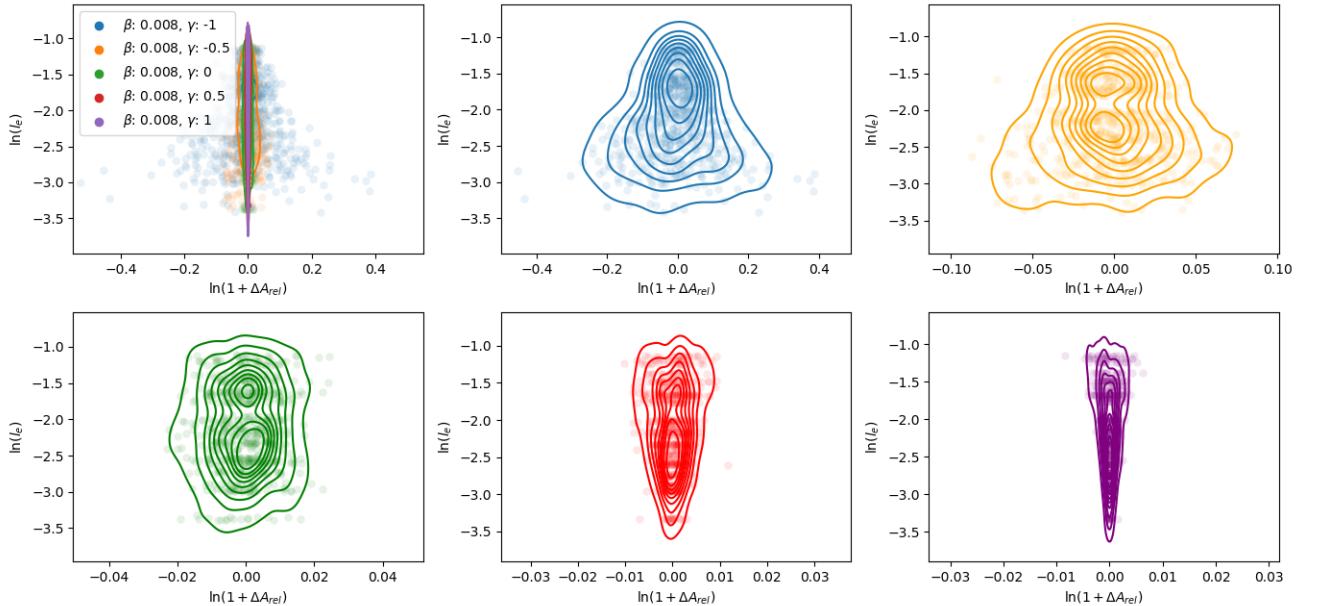


Figure 14: Distributions of  $P(\ln(1 + \Delta A), \ln(l_e))$  for fixed  $\beta = 0.008, -1 < \gamma < 1$

### 4.3.2 Variation of $\beta$

Figure 15 shows the distributions of  $P(\ln(1 + \Delta A), l_e)$  for a range of values of  $\beta$ . We see here that as  $\beta$  increases, the width of the distributions increase.

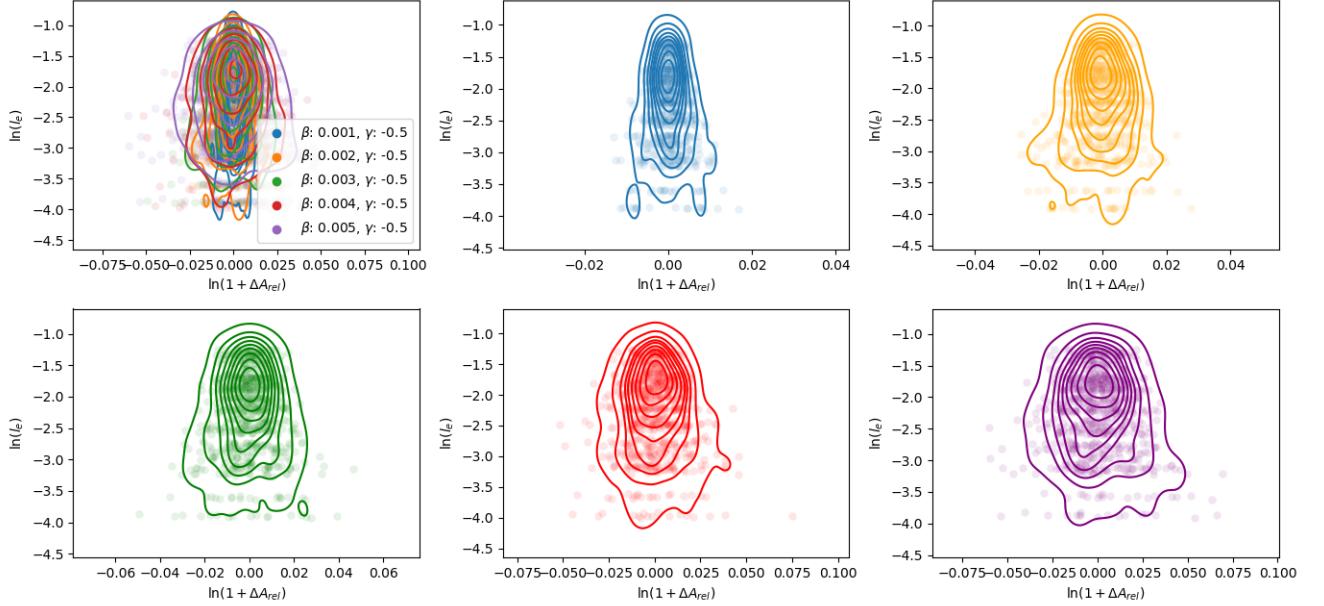


Figure 15: Distributions of  $P(\ln(1 + \Delta A), \ln(l_e))$  for fixed  $\gamma = -0.5, 0.001 < \beta < 0.005$ .

### 4.3.3 Weight distributions for real networks

We now consider the same 5 real datasets considered in section 4.2.5. Figure 16 shows the distributions of  $P(\ln(1 + \Delta A), \ln(l_e))$  for the case of edges that do change, i.e.  $\Delta A \neq 0$  for the five real networks. Here  $\Delta A$  refers to the relative change in the value of the edge weight from  $t_0$  to  $t_1$ , which takes values in the interval  $[-1, \infty]$ , and  $l_e$  is measured at time  $t_0$ . Infinite values for  $\Delta A$ , corresponding to the case of a new edge appearing, were observed but are not captured in the plots. The prominence of these across the different datasets are 4.7% of the bilateral trade dataset, 0.086% of the college messaging dataset, 0.012%, 0% and 0.0028% of the equity datasets<sup>3</sup>. We see a slight widening of the distributions for larger values of  $l_e$  for Equity-1 and 2 datasets, and to a larger extent for the third equity dataset. The bilateral trade dataset shows initial widening as  $l_e$  increases, however narrows again for the largest  $l_e$  edges. The college messaging dataset shows two distinct peaks, corresponding to changes in edge weight of  $\pm 1$ , which are over-represented in this dataset as it is unweighted, and the edge weight solely represents the count of interactions in the time window of consideration. The slight widening for larger  $l_e$  for all datasets is suggestive of a positive relationship between the value of  $l_e$  and the variance of the distribution of subsequent edge changes.

### 4.3.4 Parameter estimation for $\beta$ and $\gamma$

Dataset	Estimated $\beta$	Estimated $\gamma$
Bilateral trade	$1.11 \times 10^{-4} \pm 8.93 \times 10^{-18}$	$1.93 \pm 1.10 \times 10^{-27}$
College messaging	$7.77 \times 10^{-5} \pm 1.10 \times 10^{-16}$	$1.15 \pm 3.07 \times 10^{-27}$
Equity-1	$1.51 \times 10^{-5} \pm 4.55 \times 10^{-24}$	$1.35 \pm 2.55 \times 10^{-36}$
Equity-2	$1.42 \times 10^{-4} \pm 1.36 \times 10^{-26}$	$1.32 \pm 5.81 \times 10^{-37}$
Equity-3	$1.63 \times 10^{-4} \pm 2.20 \times 10^{-26}$	$1.42 \pm 1.43 \times 10^{-36}$

Table 4: Estimated  $\beta$  and  $\gamma$  for the 5 real different datasets

<sup>3</sup>The prominence of new edges has been significantly reduced by focusing on the giant component

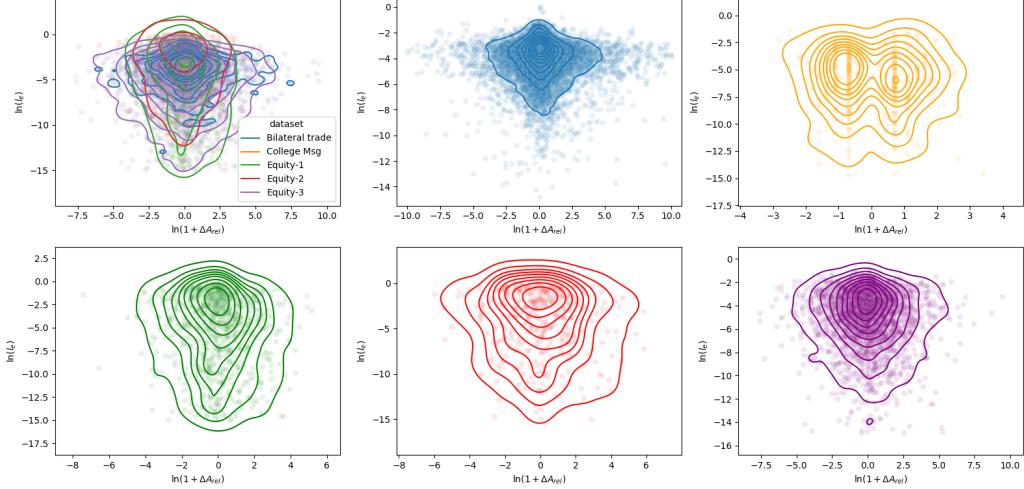


Figure 16: Contours showing the distributions of  $P(\ln(1 + \Delta A), \ln(l_e))$  for the 5 real datasets. Underlying observations of  $\ln(l_e)$  and  $\ln(1 + \Delta A_{rel})$  represented by the dots underlying these.

All 5 datasets show positive values of  $\gamma$ , suggestive of a relationship between the width of the distribution of edge changes and the value of  $l_e$ . The dataset with the highest values for  $\gamma$ , the bilateral trade dataset dataset, also shows the largest level of bias towards larger change distribution width for higher  $l_e$  in figure 16. Correspondingly, the lowest  $\gamma$  value is seen for the college messaging dataset, which shows the least bias towards larger changes occurring for larger values of  $l_e$ . The values for  $\beta$  are similar across the 5 datasets, and all relatively low. It is difficult to draw conclusions from these, as the behaviours controlled by the two parameters cannot be separated and observed alone in the distributions in figure 16.

## 5 Discussion & Conclusion

The ability to understand how microscopic changes in networks affect the macroscopic evolution across time is one of the key challenges in dynamic network analysis. In this study we have begun to explore the use of derivatives of network spectra to capture this. We derive a measure of edge based structural influence,  $l_e$ , and explore the extent to which the value is indicative of future changes. We first of all demonstrated that for small and isolated perturbations applied to the network, the eigenvalue derivative is approximated well by equation 3. However, we observe the approximation breaks down for multiple changes happening during the same time snapshot, suggesting that the measure may be more suited to a continuous or pseudo-continuous representation of network evolution, in which each time snapshot contains a single edge change.

Considering the 5 real datasets, we observe lognormal distributions of the values of  $l_e$ , indicating structural influence dominated by a small handful of edges. We propose a model in which the probability for an edge to change is given by  $\alpha l_e^\rho$ . This model allows us to control the extent to which  $l_e$  dictates the propensity for an edge to change, and also controls the scale of a subsequent change. Focusing on the former, we observe similarities in the shapes of the distributions of  $P(\Delta A = 0, \ln(l_e))$  when generating synthetic networks according to this model and those observed in the data, and the values observed for  $\alpha$  and  $\rho$  are suggestive of a relationship between the value of  $l_e$  and the subsequent presence of change. In using  $l_e$  in a logistic regression classifier to predict change, we see that  $l_e$  is slightly predictive of change in all cases, but only marginally so for the case of the bilateral trade and Equity-3 datasets. This corresponds with our observations of small values of  $\rho$  for these datasets, along with similar, non-linear distributions shapes for the probability of no change for increasing  $l_e$ . These observations indicate that the static structural importance is indicative of the presence of a subsequent change, however more work is needed to understand the shape of the distribution and the identification of different  $l_e$  regimes. We will also consider taking a similar approach with other measures of edge importance, for example edge gravity [15]. More work is also needed to understand the subsequent impact on the global network structure of an edge changing. It may be that a change influential edge could act to destabilise a system;

conversely, the change could move the system towards a state of stability. We will look to investigate this in future analyses.

We note here that  $\alpha$  and  $\rho$  themselves are useful parameters that could be used to classify networks according to their growth stability. A large value of  $\alpha$  would be an indicator for larger levels of overall network activity. A network with very large  $\rho$  would be characterised by changes occurring to the nodes with the largest  $l_e$ , conversely, a network with very small  $\rho$  would see changes distributed across all edges, regardless of the value of  $l_e$ . In the context of financial markets, these contrasting situations would require different approaches, and  $\rho$  could be used by policy makers to inform which asset classes should be monitored as a whole (for the case of small  $\rho$ ) or following an approach targeting those nodes with the highest  $l_e$ .

Our model doesn't account for edges appearing and disappearing in the network, and assumes that edge changes are independent of each other. For the first limitation, we note that edge appearance and disappearance would be unlikely to heavily influence the behaviour of the Equity networks, as we observed very low percentages (0.012%, 0% and 0.0028%) of new edges appearing<sup>4</sup>, but for the other two networks this behaviour is much more prominent at 4.7% for the bilateral trade network and 0.086% for the college messaging network. The measure  $l_e$  itself is able to assess the importance of an edge that subsequently disappears, and also those that appear between two existing nodes, so understanding how these appearances and disappearances can be captured in a model for network growth would be highly beneficial for future work. On the second point, we noted in our exploration of toy networks that the ability of our approximation of the eigenvalue derivative breaks down for multiple edge changes present. Conversely, many works such as Bandi et. al. have noted that predictability is aggregation scale specific. In future work we will thus investigate the trade-off between improved approximation of  $l_e$  for the quasi-continuous limit in which each time snapshot contains a single edge change, and improved predictability for larger aggregation scales. In addition to this, further analysis is needed to assess the effectiveness of  $l_e$  as an indicator for risk, as so far we understand that the value of  $l_e$  bears some relationship to how the network subsequently changes, but we have not yet considered the resultant changes of edges with high values of  $l_e$ , and how these have an effect on the rest of the network in terms of risk and stability. This is another area we will pursue in future work. We will also consider extending our methods to consider structural node importance, which is of use to policy makers who may wish to monitor which players could have an adverse impact on markets. It is also worth noting that although using raw transaction data gives us the lowest granularity view of the data, our work has so far not considered the higher order effects of trading behaviour on price. Such an effect results in the influence of edges reaching disconnected components, which cannot be captured by our methods, so we will consider generalising our methods to allow for networks with disconnected components. Finally, we will consider using our methods for classification of a large number of networks, and also extend our methods to understand the parameters which control the resultant weight changes.

## 6 Data availability

The datasets referred to as Equity-1, Equity-2 and Equity-3 in this paper, were extracted from a dataset of transaction reports collected by the FCA under MIFID II regulations. The datasets were used under agreement from the data owners at the Financial Conduct Authority for the current study, and are not publicly available. A data note describing the data is shown in appendix D, and comparable analysis is conducted for all investigations for two open source datasets (see below).

The dataset referred to as Bilateral Trade are hosted by Katherine Barbieri, University of South Carolina, and Omar Keshk, Ohio State University, available at <http://correlatesofwar.org>. The dataset referred to as College Messaging are available in the Stanford Large Network Dataset Collection repository, <https://snap.stanford.edu/data/CollegeMsg.html>

An implementation of the methods referenced in this paper can be found at [44].

## References

- [1] E. Nier, J. Yang, T. Yorulmazer, and A. Alentorn. Network models and financial stability. *Bank of England Working Paper no. 346*, 2008.
- [2] T.C. Silva, S. M. Guerra, B. M. Tabak, and R.C. de Castro Miranda. Financial networks, bank efficiency and risk-taking. *Banco Central do Brasil Working Paper n. 428.*, 2016.
- [3] H. J. Allen. Financial stability regulation as indirect investor/consumer protection regulation: Implications for regulatory mandates and structure. *Legal Studies Research Paper Series Research Paper 17-5*, 2017.

---

<sup>4</sup>This is largely due to the preprocessing steps applied to the network, since we are only considering the giant component.

- [4] A. Haldane and R. May. Systemic risk in banking ecosystems. *Nature* 469, 351–355 <https://doi.org/10.1038/nature09659>, 2011.
- [5] X. Ren and L. Lu. Review of ranking nodes in complex networks. *Chinese Science Bulletin* 59, 1175. 10.1360/972013-1280., 2014.
- [6] C. Stegehuis, R. van der Hofstad, and van Leeuwaarden. Epidemic spreading on complex networks with community structures. *Sci Rep* 6, 29748 <https://doi.org/10.1038/srep29748>, 2016.
- [7] Guangming Ren and Xingyuan Wang. Epidemic spreading in time-varying community networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 24(2):023116, 2014.
- [8] Lai Y.C., Motter A.E., and Nishikawa T. Attacks and cascades in complex networks. In: Ben-Naim E., Frauenfelder H., Toroczkai Z. (eds) *Complex Networks. Lecture Notes in Physics*, vol 650. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-540-44485-5\\_14](https://doi.org/10.1007/978-3-540-44485-5_14), 2004.
- [9] H. Xu, J. Zhang, J Yang, and L. Lun. Identifying important nodes in complex networks based on multiattribute evaluation. *Mathematical Problems in Engineering*, vol. Article ID 8268436, <https://doi.org/10.1155/2018/8268436>, 2018.
- [10] Y. Moreno, M. Nekovee, and A. F. Pacheco. Dynamics of rumor spreading in complex networks. *Phys. Rev. E*, 69:066130, Jun 2004.
- [11] Y. Wang, Z. Di, and Y. Fan. Identifying and characterizing nodes important to community structure using the spectrum of the graph.
- [12] E. Stanley, L. Lü, L. Pan, T. Zhou, and Y. Zhang. Toward link predictability of complex networks. 112 (8) 2325–2330; DOI: 10.1073/pnas.1424644112, 2015.
- [13] S. Battiston, M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli. Debtrank: Too central to fail?financialnetworks, the fed and systemic risk. *Sci. Rep.* 2,541; DOI:10.1038/srep00541, 2012.
- [14] P. Barucca and F. Lillo. The organization of the interbank network and how ecb unconventional measures affected the e-mid overnight market. *Comput Manag Sci* (2017), arxiv: 1511.08068, 2017.
- [15] M.E. Helander and S. McAllister. The gravity of an edge.
- [16] E. Yu, D. Chen, and J. Zhao. Identifying critical edges in complex networks. *Sci Rep* 8, 14469 <https://doi.org/10.1038/s41598-018-32631-8>, 2018.
- [17] R. M. May. Will a large complex system be stable? *NATURE VOL. 238 AUGUST 18 1972*.
- [18] F. Restoy. Market integration: the role of regulation. *Speech by Chairman of the Financial Stability Institute, Bank for International Settlements, at the IIF Market fragmentation roundtable, Washington DC, United States, 10 April 2019.*, 2019.
- [19] P. A. Samuelson. General proof that diversification pays. *The Journal of Financial and Quantitative Analysis*, 2(1):pp. 1–13, 1967.
- [20] M. Bardoscia, S. Battiston, F. Caccioli, and G. Caldarelli. Pathways towards instability in financial networks. *Nature Communications* 8:14416 | DOI: 10.1038/ncomms14416, 2016.
- [21] S. Markose, S. Giansante, GatkowskiM., and Shaghaghi A. R. Too interconnected to fail: Financial contagion and systemic risk in network model of cds and other credit enhancement obligations of us banks. *COMISEF WORKING PAPERS SERIESWPS-033*, 2010.
- [22] F. Caccioli, M. Marsili, and P. Vivo. Eroding market stability by proliferation of financialinstruments. *arXiv:0910.0064 [q-fin.TR]*, 2009.
- [23] J. Tobin. The new economics one decade older. *The Eliot Janeway Lectures on Historical Economics in Honour of Joseph Schumpeter 1972 (Princeton University Press, Princeton, US);Eastern Economic Journal IV153-159*, 1978.
- [24] G. Bianconi, T. Galla, M. Marsili, and P. Pin. Effects of tobin taxes in minority game markets. *J. Econ. Behavior and Organization, Volume 70, Issues 1–2*, 2009.
- [25] W.A. Brock, C.H. Hommes, and Wagener F.O.O. More hedging instruments may destabilize markets. *Journal of Economic Dynamics Control* 33 1912–1928, 2009.
- [26] M. Bardoscia, G. Bianconi, and G. Ferrara. Multiplex network analysis of the ul otc derivatives market. *Bank of England Working Papers*, 2018.
- [27] T. Aste, W. Shaw, and T. Di Matteo. Correlation structure and dynamics in volatile markets. *New Journal of Physics*, 2010.

- [28] P. Mazzarisi, P. Barucca, F. Lillo, and D. Tantari. A dynamic network model with persistent links and node-specific latent variables, with an application to the interbank market. *European Journal of Operational Research Volume 281, Issue 1, 16, Pages 50-65*, 2020.
- [29] P. Barucca and F. Lillo. Disentangling bipartite and core-periphery structure in financial markets. *Chaos, solitons and fractals, arxiv: 1551.08830*, 2015.
- [30] B. Barucca, M. Bardoscia, F. Caccioli, M. D'Errico, G. Visentin, S. Battiston, and G. Caldarelli. Network valuation in financial systems. *Mathematical Finance.30:1181–1204.*, 2020.
- [31] K. Soramäki, M. Bech, J. Arnold, R. Glass, and W. Beyeler. The topology of interbank payment flows. *Physica A 379 (2007) 317–333*, 2007.
- [32] A. Barabasi and R. Albert. Emergence of scaling in random networks. *arXiv:cond-mat/9910332 [cond-mat.dis-nn]*, 1999.
- [33] M. Falkenberg, J. Lee, S. Amano, K. Ogawa, K. Yano, Y. Miyake, T. Evans, and K. Christensen. Identifying time dependence in network growth. *PHYSICAL REVIEW RESEARCH 2, 023352*, 2020.
- [34] M. Bazzi, L. Jeub, A. Arenas, S. Howison, and M. Porter. A framework for the construction of generative models for mesoscale structure in multilayer networks. *arXiv:1608.06196v5 [cs.SI]*, 2019.
- [35] T. P. Peixoto and M. Rosvall. Modelling sequences and temporal networks with dynamic community structures. *Nat Commun 8, 582, https://doi.org/10.1038/s41467-017-00148-9*, 2017.
- [36] D. Watts and S. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature 393, 440–442 https://doi.org/10.1038/30918*, 1998.
- [37] Barabási A.L. Bianconi G. Competition and multiscaling in evolving networks. *Europhysics Letters. 54 (4): 436–442. arXiv:cond-mat/0011029*, 2001.
- [38] T. Kobayashi and M. Génois. Two types of densification scaling in the evolution of temporal networks. *Physical Review E 102, 052302* 2020, 2020.
- [39] P. Holme and J. Saramaki. Temporal networks. *Phys. Rep. 519, 97-125 (2012). arXiv:1108.1780 [nlin.AO]*, 2010.
- [40] J. Tang, M. Musolesi, C. Mascolo, V. Latora, and V. Nicosia. Analysing information flows and key mediators through temporal centrality metrics. *SNS '10: Proceedings of the 3rd Workshop on Social Network Systems Article No.: 3 Pages 1–6 https://doi.org/10.1145/1852658.1852661*, 2010.
- [41] H. W. Kuhn and A. W. Tucker. Nonlinear programming. *Proceedings of the Seconds Berkeley Symposium on Mathematical Statistics and Probability, pp. 481-492*, 1950.
- [42] K. Barbieri and Omar K. Correlates of war project trade data set codebook, version 4.0, 2016.
- [43] P. Panzarasa, T. Opsahl, and K. M. Carley. Patterns and dynamics of users’ behavior and interaction: Network analysis of an online community., 2009.
- [44] I. Seabrook. Structural edge importance - python implementation. *Github repository*, 2020.

## A Results of perturbing two nodes in toy networks

The plots shown in figures 17 - 19 show additional results for the case of perturbing single edges for weighted barbell networks, and unweighted ring and Erdős–Rényi networks.

In figures 20 - 24 we present the results of changing two individual edges, and observing the resulting change in  $\lambda$  for the range of perturbations applied. We overlay this with a line of constant  $l_e$ , to assess the performance of our approximation.

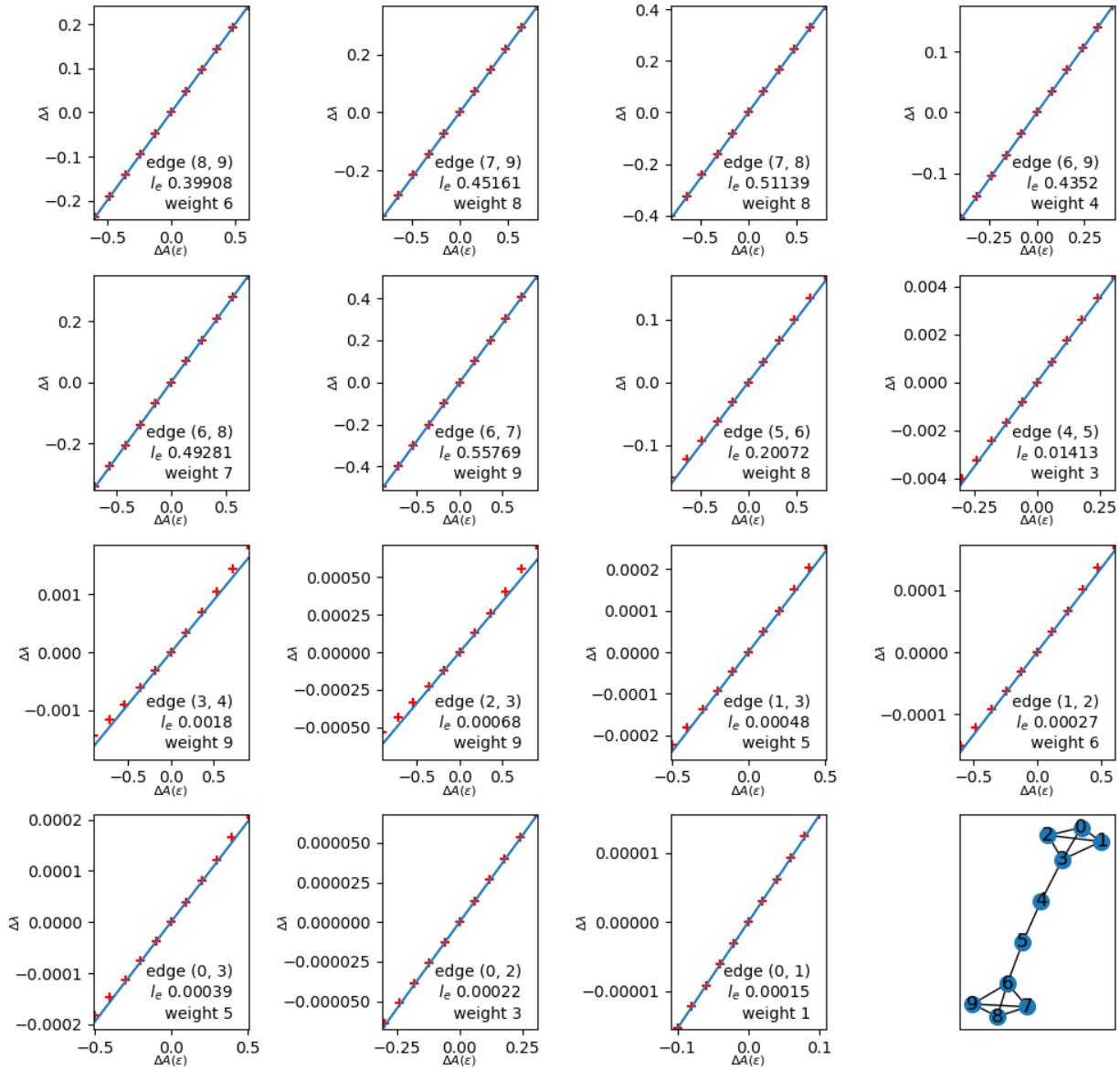


Figure 17: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta \lambda$ , compared to line of constant  $l_e$ . Weighted barbell network

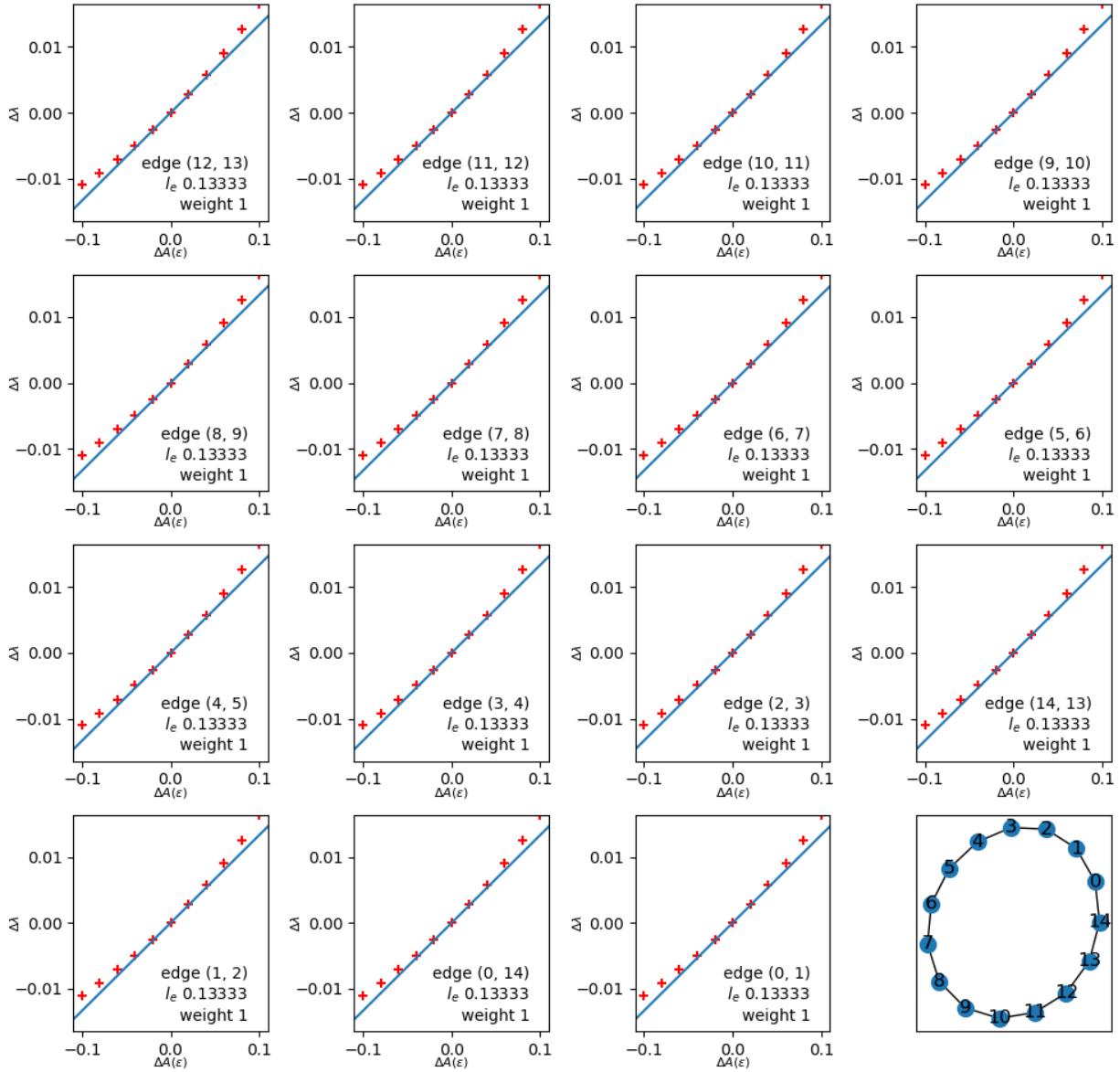


Figure 18: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta \lambda$ , compared to line of constant  $l_e$ . Unweighted ring network.

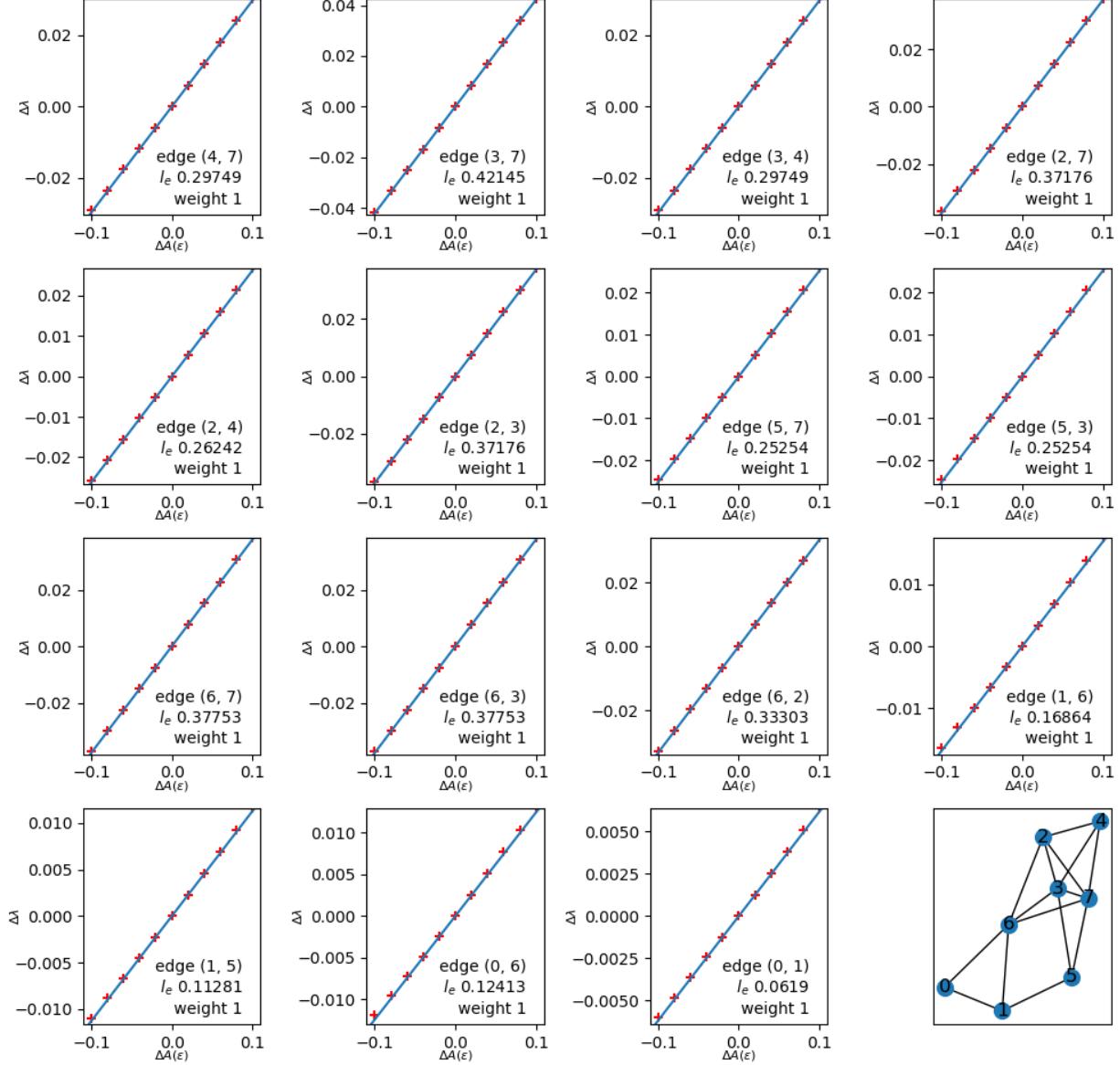


Figure 19: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta \lambda$ , compared to line of constant  $l_e$ . Unweighted Erdős-Rényi network

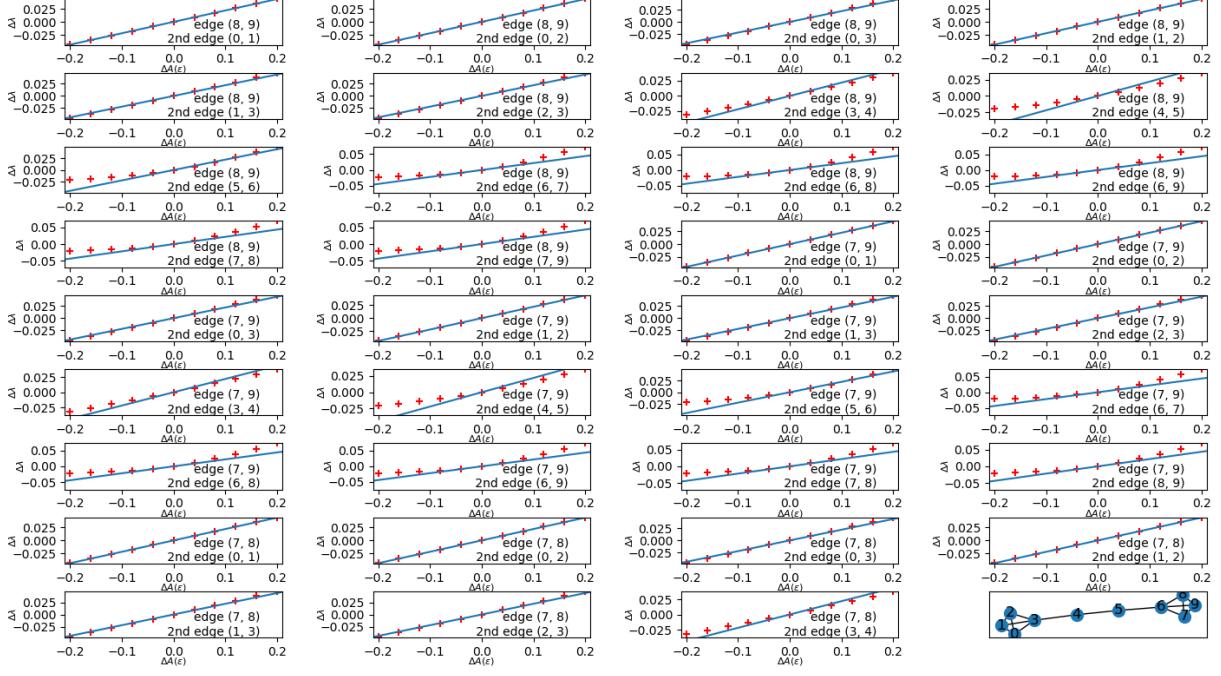


Figure 20: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta \lambda$ , compared to line of constant  $l_e$ , for the case of two edges changing. The plots consider one of the two perturbed edges. Barbell network, with equal initial weights

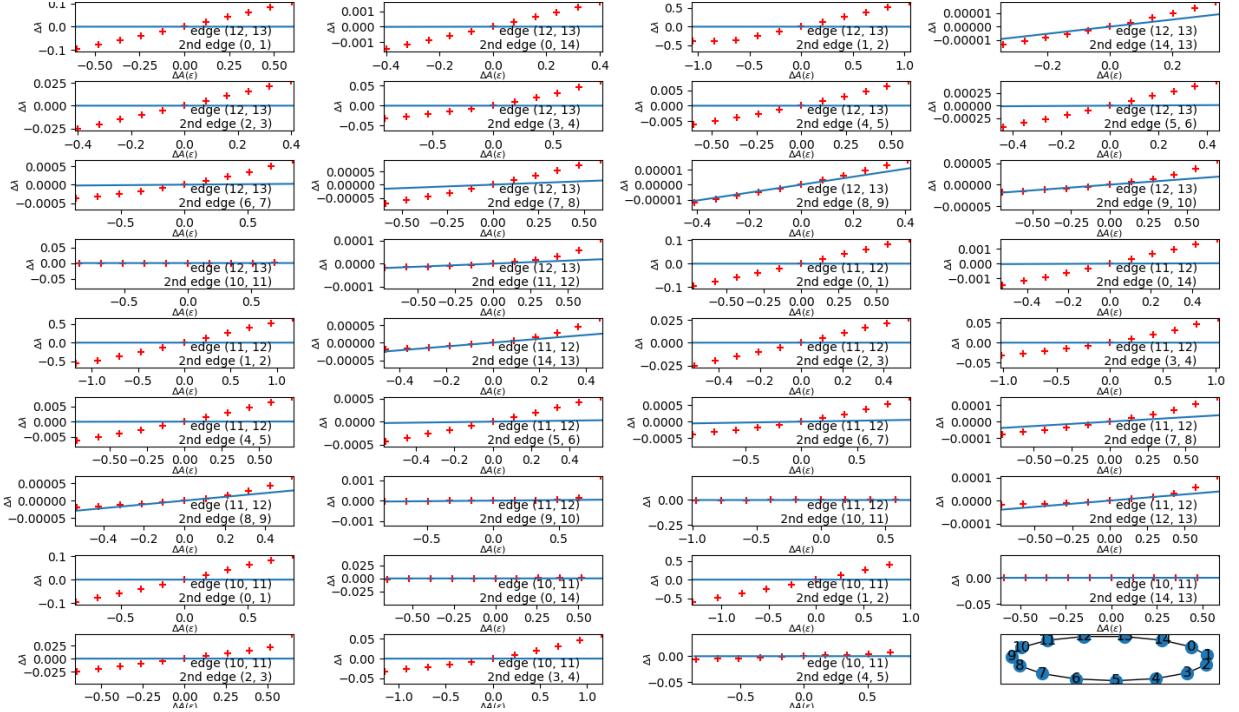


Figure 21: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta \lambda$ , compared to line of constant  $l_e$ , for the case of two edges changing. The plots consider one of the two perturbed edges. Ring network with randomly assigned weights.

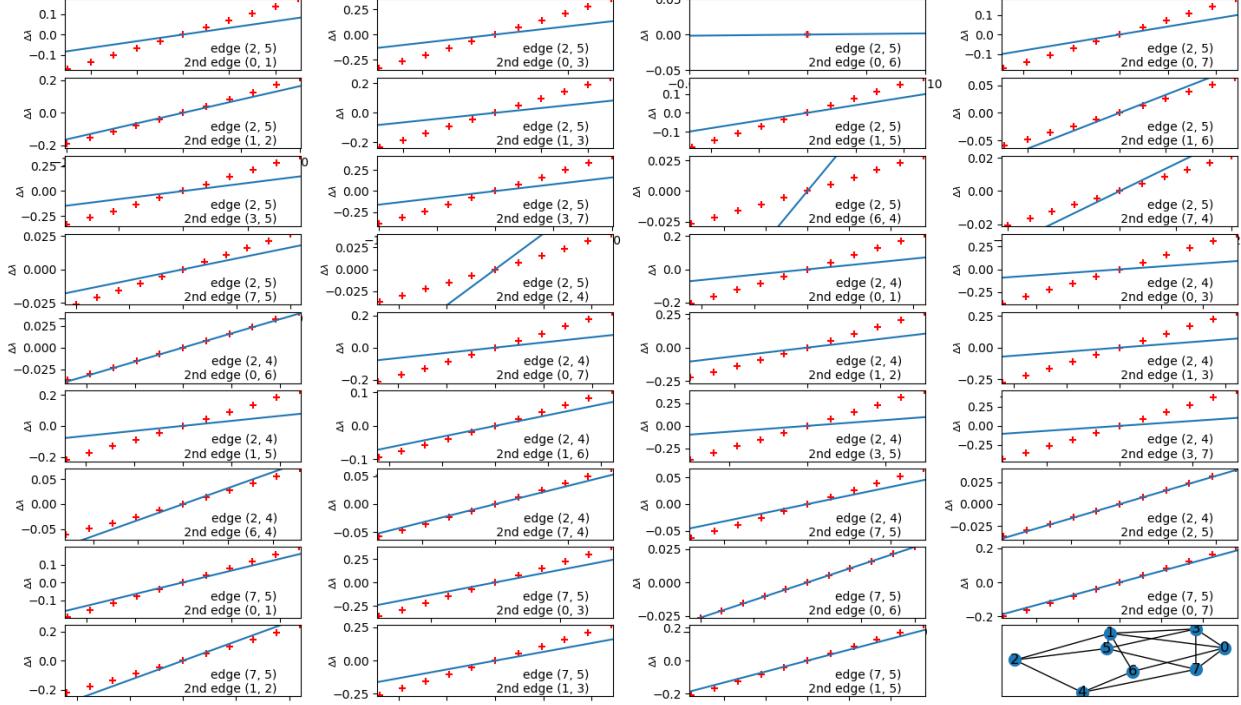


Figure 22: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta \lambda$ , compared to line of constant  $l_e$ , for the case of two edges changing. The plots consider one of the two perturbed edges. Erdős–Rényi network with randomly assigned weights

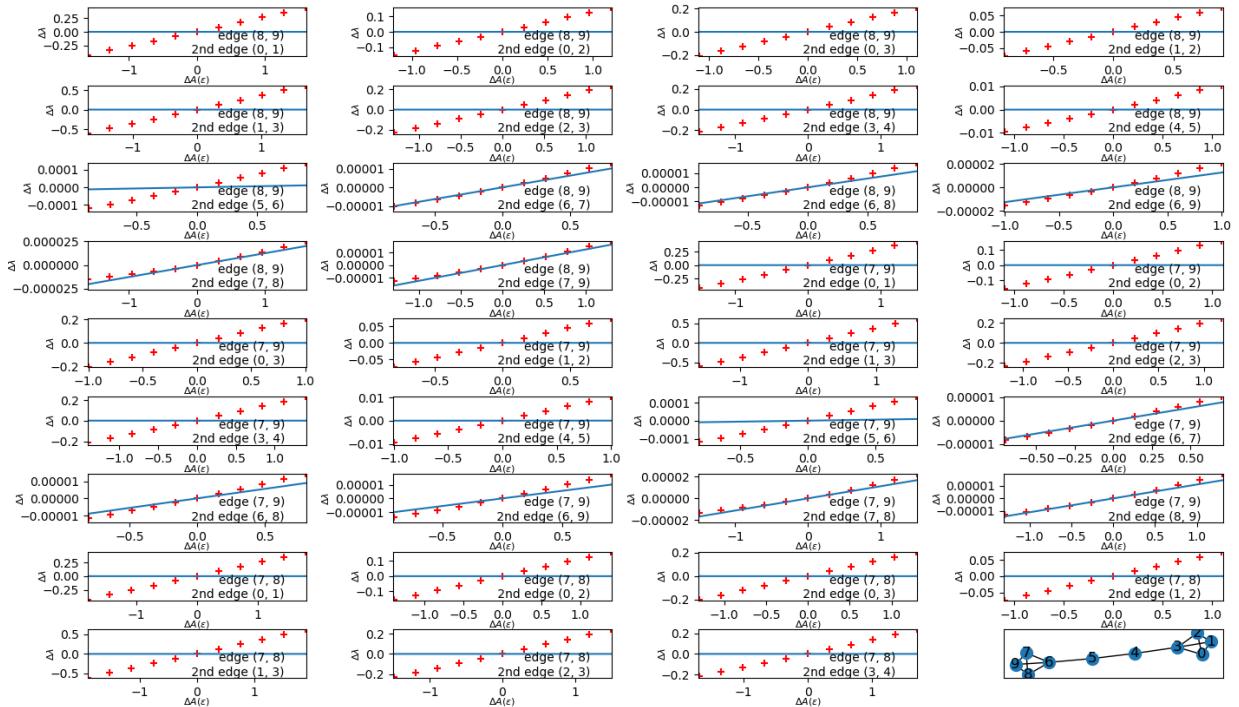


Figure 23: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta \lambda$ , compared to line of constant  $l_e$ , for the case of two edges changing. The plots consider one of the two perturbed edges. Barbell network, with randomly assigned weights.

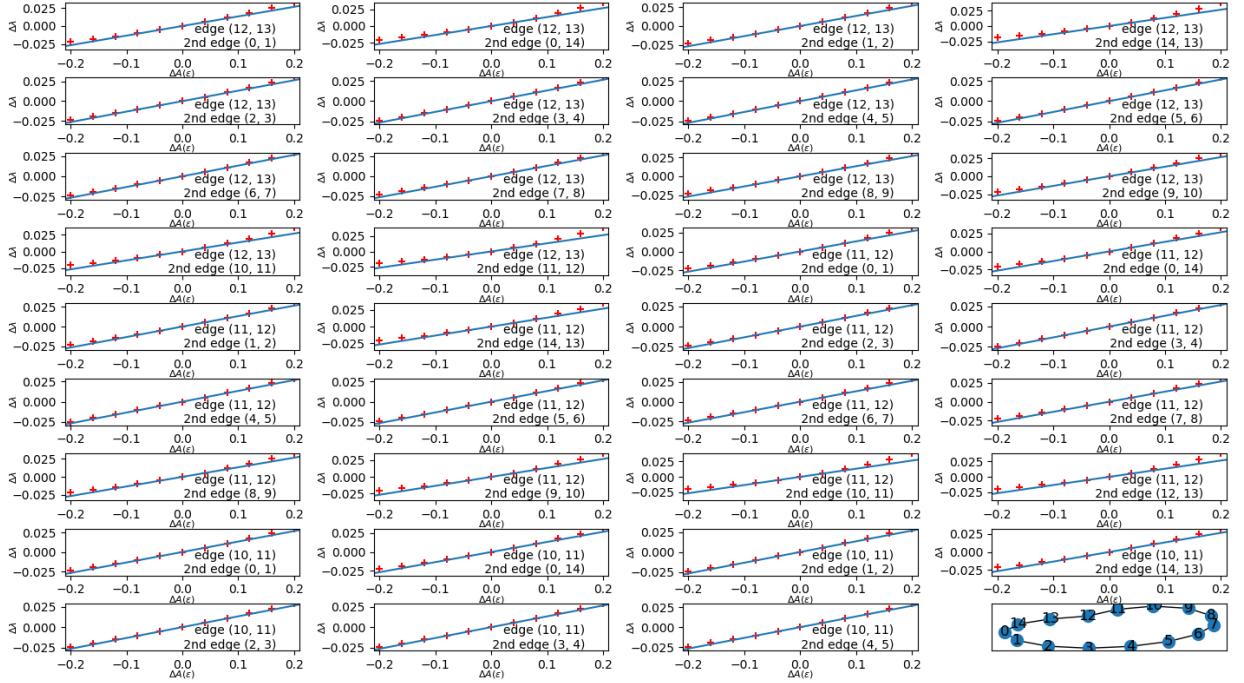


Figure 24: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta \lambda$ , compared to line of constant  $l_e$ , for the case of two edges changing. The plots consider one of the two perturbed edges. Ring network with randomly assigned weights.

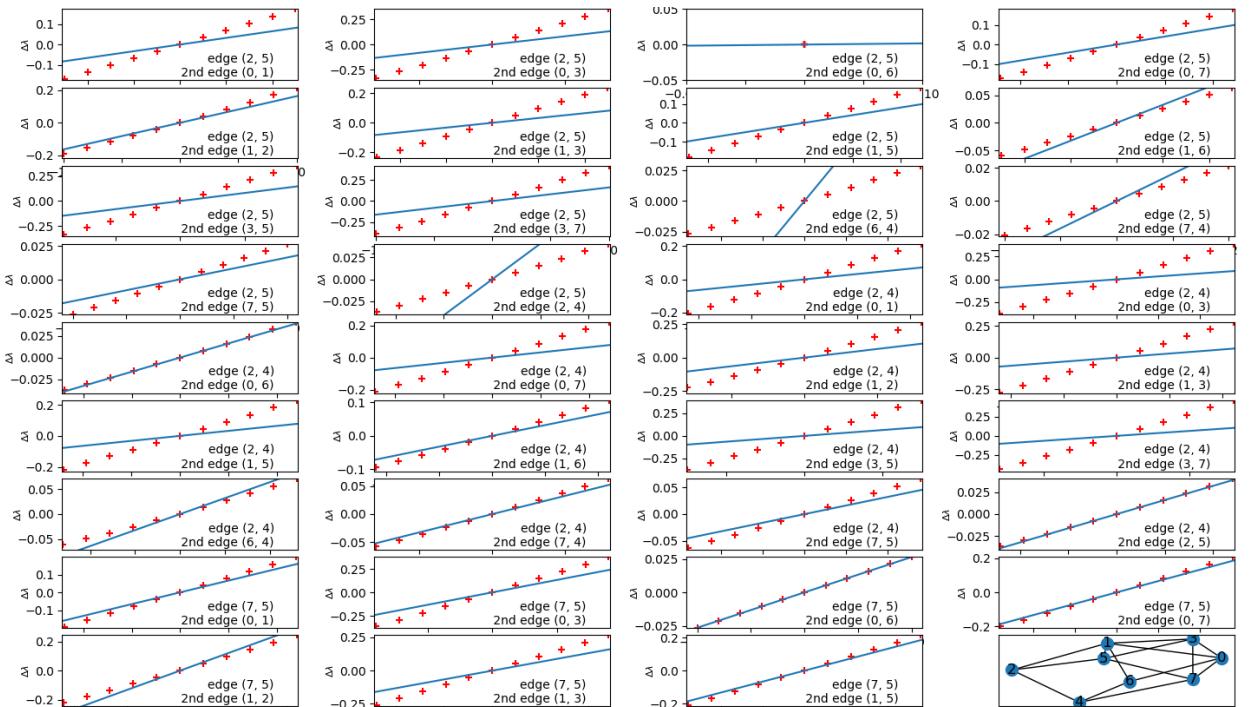


Figure 25: Scatter plot of perturbations  $\Delta A$  and the resulting  $\Delta \lambda$ , compared to line of constant  $l_e$ , for the case of two edges changing. The plots consider one of the two perturbed edges. Erdős–Rényi network with randomly assigned weights

## B Perturbation theory approach to deriving network eigenvalue derivatives

### B.1 Undirected case

Consider a perturbation to the adjacency matrix  $\mathbf{A}$ :

$$\mathbf{A} \rightarrow \mathbf{A} + \epsilon \mathbf{V} \quad (8)$$

and the resulting first order changes to the leading eigenvalue  $\lambda$  and the associated eigenvector  $|\lambda\rangle$ :

$$\lambda = \lambda_0 + \epsilon \lambda \quad (9)$$

$$|\lambda\rangle = |\lambda\rangle_0 + \epsilon |\lambda\rangle_1 \quad (10)$$

Substituting these into our eigenvalue equation

$$(\mathbf{A} + \epsilon \mathbf{V})(|\lambda\rangle_0 + \epsilon |\lambda\rangle_1) = (\lambda_0 + \epsilon \lambda_1 + \dots)(|\lambda\rangle_0 + \epsilon |\lambda\rangle_1 + \dots) \quad (11)$$

and considering terms up to 1st order in  $\epsilon$

$$\mathbf{A} |\lambda\rangle_0 + \epsilon \mathbf{V} |\lambda\rangle_0 + \epsilon \mathbf{A} |\lambda\rangle_1 = \lambda_0 |\lambda\rangle_0 + \epsilon \lambda_1 |\lambda\rangle_0 + \epsilon \lambda_0 |\lambda\rangle_1 \quad (12)$$

Then we can consider each of the terms in  $\epsilon^n$  separately,

$$\epsilon_0 : \mathbf{A} |\lambda\rangle_0 = \lambda_0 |\lambda\rangle_0 \quad (13)$$

$$\epsilon_1 : \mathbf{V} |\lambda\rangle_0 + \mathbf{A} |\lambda\rangle_1 = \lambda_1 |\lambda\rangle_0 + \lambda_0 |\lambda\rangle_1 \quad (14)$$

By multiplying the equation for  $\epsilon^1$  by the left eigenvector  ${}_0\langle \lambda |$ , and making use of the hermitian properties of  $\mathbf{A}$  such that  ${}_0\langle \lambda | \mathbf{A} = \lambda_0 {}_0\langle \lambda |$ , we find

$${}_0\langle \lambda | \mathbf{V} |\lambda\rangle_0 = \lambda_1 {}_0\langle \lambda | \lambda\rangle_0 \quad (15)$$

The perturbation we are considering is changing one row, and one column, i.e. where  $V_{ij} = A_{ij}$  if  $i$  or  $j$  are the row/column we are changing, zero otherwise:

$$V_{ij} = \begin{cases} A_{ij} & \text{if } i = k \text{ or } j = k \\ 0 & \text{otherwise} \end{cases}$$

Then, expanding the indices,

$$\sum_{ij} \lambda_{0,i} V_{ij} \lambda_{0,j} = \sum_{ij} \lambda_{0,i} A_{ij} \lambda_{0,j} \delta_{ik} + \sum_{ij} \lambda_{0,i} A_{ij} \lambda_{0,j} \delta_{jk} = 2 \sum_j \lambda_{0,k} A_{kj} \lambda_{0,j} \quad (16)$$

Where we have re-labelled the indices for the second term and have evaluated the  $\delta$ 's. Leading us to the result:

$$l_e = \frac{\partial \lambda}{\partial A_{ij}} = 2 \lambda_{0,i} \lambda_{0,j} \quad (17)$$

where  $\lambda_{0,i}$  is the  $i$ th component of the eigenvector corresponding to the leading eigenvalue of  $\mathbf{A}$

### B.2 Directed case

For the directed case, we still have that

$$\epsilon_1 : \mathbf{V} |\lambda\rangle_1 + \mathbf{A} |\lambda\rangle_0 = \lambda_1 |\lambda\rangle_0 + \lambda_0 |\lambda\rangle_1 \quad (18)$$

but we can't use the hermitian properties of the matrix  $\mathbf{A}$  as for directed networks  $\mathbf{A}$  is generally not symmetric. We can however consider the matrix  $\mathbf{M} = \mathbf{A} \mathbf{A}^T$  and perturbation  $\mathbf{M} \rightarrow \mathbf{M} + \epsilon \mathbf{W}$ , and use the symmetric result from this. This is useful since the singular values of matrix  $\mathbf{A}$  are defined as the square root of the eigenvalues of  $\mathbf{A} \mathbf{A}^T$ , such that:

$$\partial \lambda^M = 2 s^A \partial s^A \quad (19)$$

where  $\lambda^M$  is the leading eigenvalue of  $\mathbf{M}$  and  $s^A$  is the leading singular value of  $\mathbf{A}$ . We can then make use of our result above for the symmetric matrix,

$${}_0\langle \lambda^M | \mathbf{W} | \lambda^M \rangle_0 = \lambda_1^M \quad (20)$$

Where  ${}_0\langle \lambda^M |$  and  $|\lambda^M\rangle_0$  are the left and right eigenvectors of  $\mathbf{M}$ . For the directed case, our perturbation is changing just a row (or column) independently, i.e.

$$W_{ij} = \begin{cases} M_{ij} & \text{if } i = k \\ 0 & \text{otherwise} \end{cases} \quad (21)$$

Then, expanding the indices,

$$\sum_{ij} \lambda_{0,i}^M W_{ij} \lambda_{0,j}^M = \sum_{ij} \lambda_{0,i}^M M_{ij} \lambda_{0,j}^M \delta_{ik} = \sum_j \lambda_{0,k}^M M_{kj} \lambda_{0,j}^M \quad (22)$$

Leading us to the result

$$\frac{\partial s^A}{\partial M_{ij}} = \frac{\lambda_{0,i}^M \lambda_{0,j}^M}{2s^A} \quad (23)$$

where  $\lambda_{0,i}^M$  is the  $i$ th component of the eigenvector corresponding to the leading eigenvalue of  $\mathbf{M}$ . In both the directed and undirected case above, it is worth noting that the derivations can be generalised to allow new links to be added/removed, however new nodes cannot be added or removed.

## C Kernel Density Estimation for conditional probability estimation

We make use of multivariate conditional Kernel Density Estimation (KDE) to find the probability distributions for the values of  $l_e$ . The functional form of the KDE is

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{(x - x_i)}{h}\right) \quad (24)$$

where

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right) \quad (25)$$

is the kernel function, a non-negative function. The parameter  $h$  is the bandwith, a smoothing parameter. We have used a Gaussian kernel:

$$K = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \quad (26)$$

## D Dataset descriptions

The first dataset considered tracks bilateral trade flows between states from 1870-2014, describing import and export data in current U.S. dollars for pairs of sovereign states [42]. This dataset is interesting not just due to its relevance to our focus on financial markets, but also due to an observed growth across time, apart from two time periods corresponding to the First and Second World Wars.

The second dataset considered was a dataset of private messages sent on an online social network at the University of California. An edge  $(u, v, t)$  means that user  $u$  sent a private message to user  $v$  at time  $t$ . As this network is unweighted, the weights of all of the edges have been set to 1. The network was aggregated to daily snapshots, in which the edge weight is the number of times that edge is active during that day.

Finally, in order to observe the effects of different trading structures on the output of our methods, we applied our techniques to transaction reports relating to three different equity stocks traded on the UK capital markets. The data was aggregated daily, and covers a 2 year period from January 2018<sup>5</sup>. The Equity-3 dataset was analysed for the shorter time range of 03/06/2019 to 05/11/2019. We chose to study networks of transactions for stocks on energy companies due to the high level of trading activity in this sector. The results displayed in this paper consider the giant component networks of 3 different stocks. The first two instruments were traded without the presence of CCPs, one focusing on oil and gas exploration and production and the second focusing on renewable and alternative energy. The third instrument, another oil and gas production stock, shows a network dominated by the presence of a CCP. Due to the sensitivity of the data, these have been referred to as Equity networks 1, 2 and 3 throughout this paper.

---

<sup>5</sup>the Equity-3 dataset was analysed for a 5 month period ending in November 2019

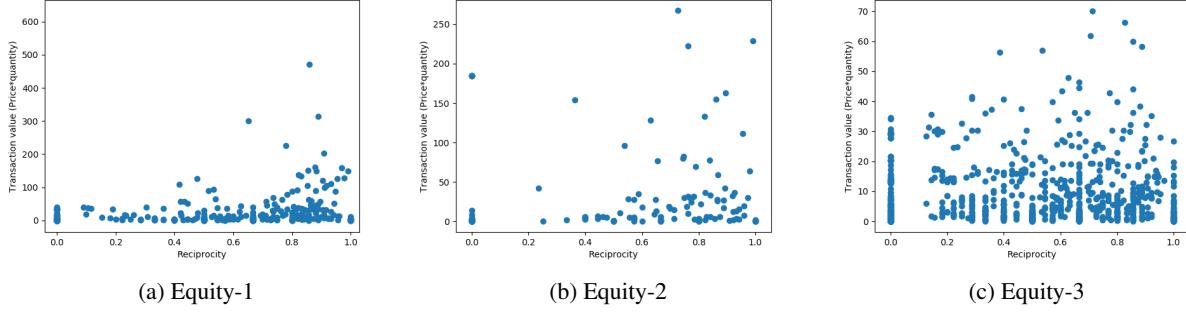


Figure 26: Reciprocity vs. price for the Equity networks

### D.1 Summary statistics for the Equity datasets

The Equity data was made available by the FCA for use in this study, and is not publically available. To provide the reader with additional context, here we include some high level network statistics for these networks. All statistics are based on the networks following the removal of nodes which appear on less than 5 days in the sample, which we classed as ‘inactive’.

We see that all three networks have similar connectivities, but Equity-3 is significantly denser than the other two and shows a higher level of reciprocity. All datasets show similar values for the correlation coefficient of the adjacency matrix.

dataset	# nodes	# edges	connectivity	density	reciprocity	correlation coefficient
Equity - 1	232	6, 961	30%	12.99%	67.72	62.9%
Equity - 2	94	3,684	39.2%	42.14%	75.03%	56.84%
Equity - 3	263	9.094	34.58%	13.2%	57.66%	51.23%

Table 5: Network statistics for the three Equity datasets

We can see from figures 26a and 26b that large transaction values are more likely to have a high reciprocity for the first and second dataset. However the same cannot be said for the third Equity dataset, as shown in figure 26c.

The evolution of high level network statistics are shown in figures 27, 28, and 29. Here we see that the networks fluctuate around a relatively stable mean, with no obvious level of growth or decay across the time period.

It is further interesting to note that the third network considered shows the presence of a hierarchy in the network, due to it being an instrument that is traded mainly through the use of Central Clearing Parties (CCP), producing a tiered structure as shown in figure 30. Such a structure can be identified following the identification of a dominant node in the network, i.e. a node with significantly higher degree, and examining its ego network.

## E Parameter estimations

### E.1 Estimation of $\rho$ and $\alpha$

We assume in this section that our networks can be described by a model in which the probability of an edge changing is given by

$$P(\Delta) = \theta_e = \begin{cases} 0 & \alpha l_e^\rho \leq 0 \\ \alpha l_e^\rho & 0 < \alpha l_e^\rho < 1 \\ 1 & \alpha l_e^\rho \geq 1 \end{cases}$$

The maximum likelihood estimate of  $\theta$  then follows the same procedure as in the case of a (potentially biased) coin toss - given a sample of changes  $x_i$ , the likelihood of observing these changes given  $\theta$  is

$$L(x_1, x_2, \dots, x_n | \theta) = \prod_e f(x_e | \theta_e) \quad (27)$$

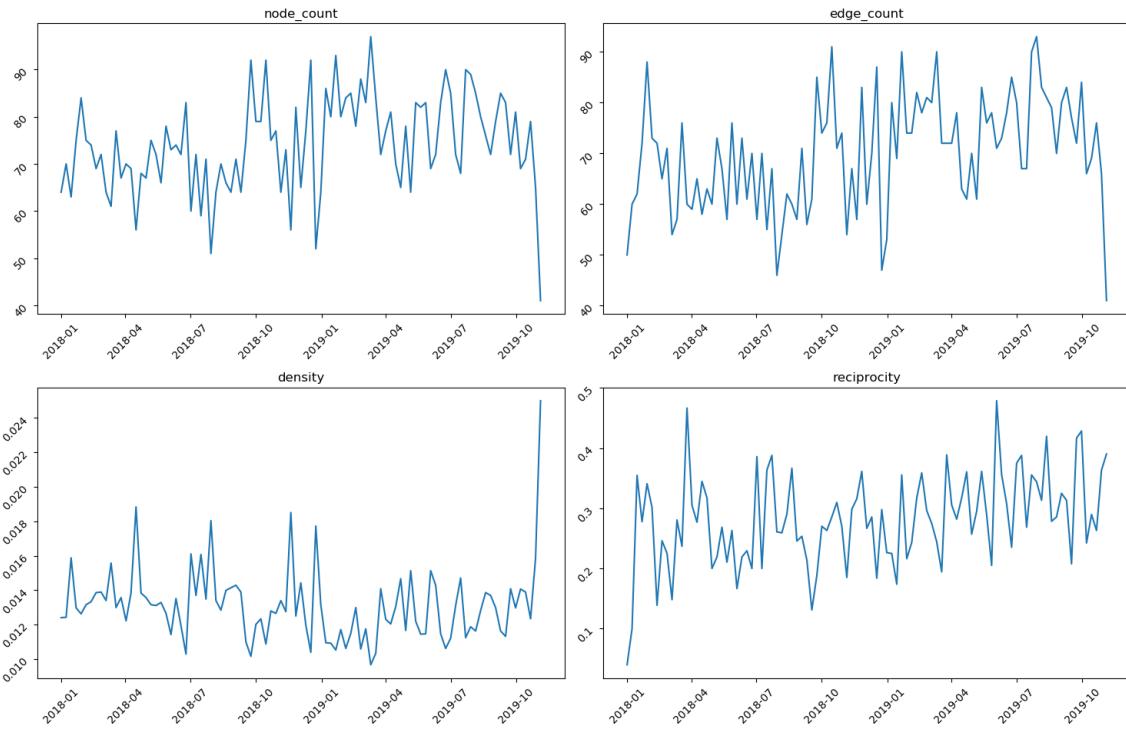


Figure 27: Daily counts of nodes and edges, density and reciprocity across the entire investigation period for Equity-3

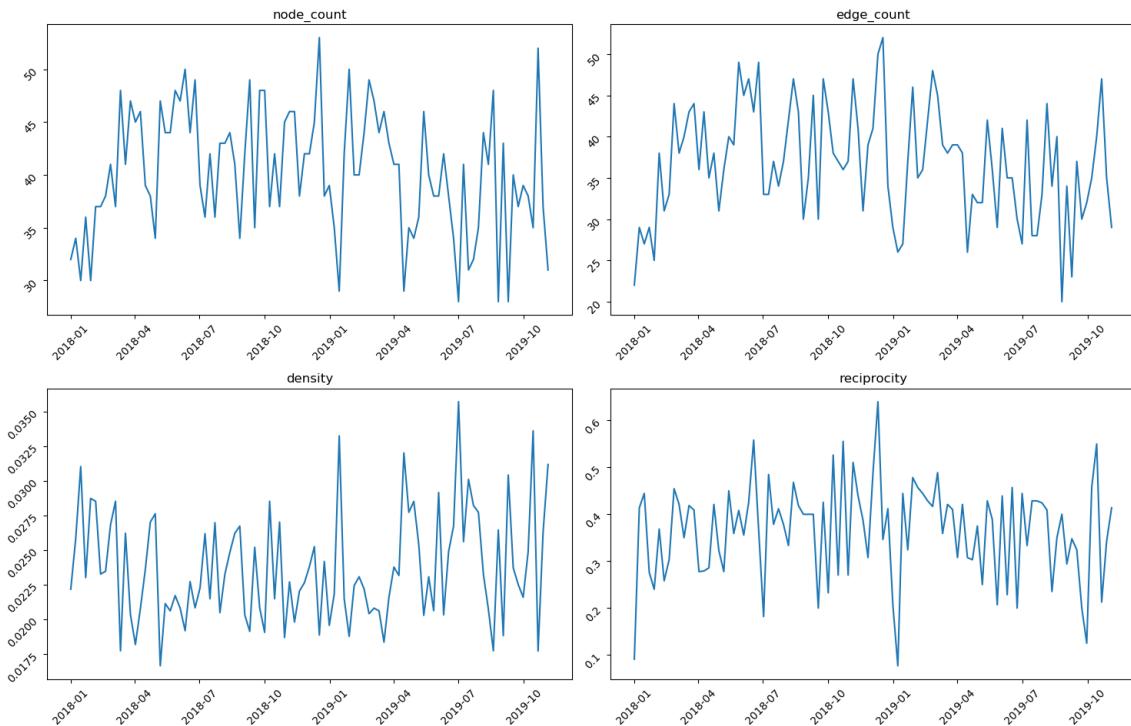


Figure 28: Daily counts of nodes and edges, density and reciprocity across the entire investigation period for Equity-2

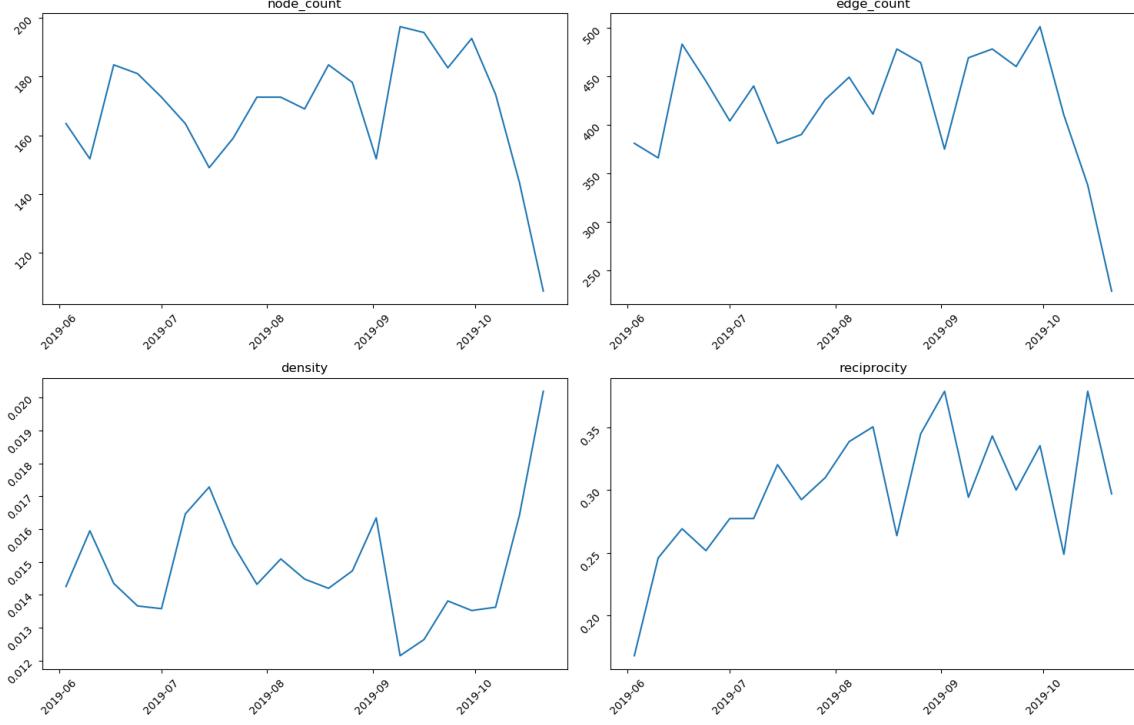


Figure 29: Daily counts of nodes and edges, density and reciprocity across the entire investigation period for Equity-3

where  $f(x_e|\theta_e)$  follows the Bernoulli distribution  $\theta_e^{k_e}(1-\theta_e)^{1-k_e}$  where  $k_e$  is the observed outcome of edge  $e$ . Taking the logarithm of this, our log-likelihood is given by

$$\ln(L(\mathbf{x}|\theta)) = \sum_e^N k_e \ln(\theta_e) + (1 - k_e) \ln(1 - \theta_e) \quad (28)$$

Since  $\alpha l_e^\rho$  is constrained to be a probability, to estimate the parameters which result in the maximum likelihood, we need to minimise the negative log-likelihood with respect to multiple inequality constraints:

$$0 \leq \alpha l_e^\rho \leq 1 \quad (29)$$

Where we have one inequality constraint for each  $l_e$ . To do this, we make use of the Karush-Kuhn-Tucker conditions [41] and numerical optimisation, to find the optimal saddle point which maximises  $L$  with whilst satisfying these constraints.

## E.2 Estimation of $\beta$ and $\gamma$

For the case of the distribution of edge changes drawn from a Gaussian distribution with  $\mu=0$  and  $\sigma = \beta l_e^\gamma$ , the log-likelihood is given by

$$\ln(L) = \sum_e^N \ln\left(\frac{1}{\sqrt{2\pi}\beta l_e^\gamma}\right) \exp\left(\frac{-x_e^2}{2\beta^2 l_e^{2\gamma}}\right) \quad (30)$$

where  $x_e$  refers to the observed relative change of edge  $e$ . Differentiating with respect to  $\beta$ ,

$$\beta = \sqrt{\frac{1}{N} \sum_e^N \frac{x_e^2}{l_e^{2\gamma}}} \quad (31)$$

From which we recover the expected standard deviation for a Gaussian in the case of  $\gamma=0$ .

Differentiating with respect to  $\gamma$ ,

$$\frac{\partial \ln(L)}{\partial \gamma} = \sum_e^N -\ln(l_e) + \frac{\partial}{\partial \gamma} \frac{x_e^2}{2\beta^2} \exp(-2\gamma \ln l_e) \quad (32)$$

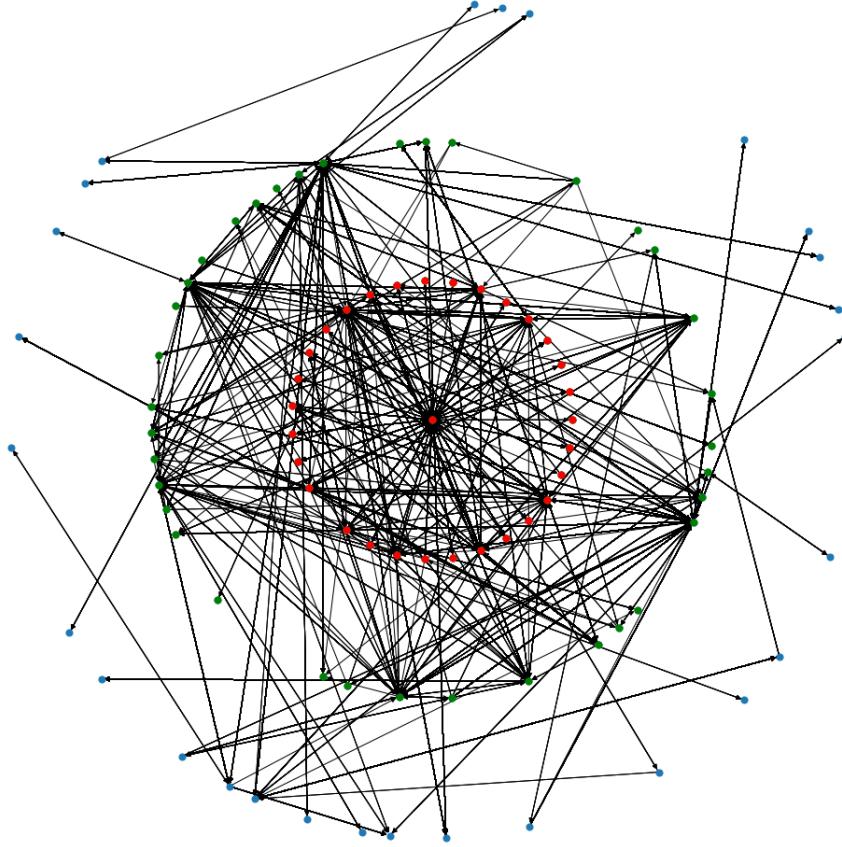


Figure 30: Example of a tiered structure in the GWCC of the trading network for an instrument frequently traded via an individual CCP.

which when set to 0,

$$\sum_e^N \ln(l_e) \left( 1 + \frac{x_e^2 \ln(l_e)}{\beta^2 l_e^{2\gamma}} \right) \quad (33)$$

Substituting 31 for  $\beta$ , and solving numerically allows us to produce an estimate for  $\gamma$ , which can be solved numerically for  $x_e = \Delta A_{rel}$ .

## F Comparison of data distributions to model

Figure 31 shows the bulk of the distributions for  $P(\Delta A = 0 | \ln(l_e))$  for our 5 datasets, in comparison to the equivalent generated from our model for network evolution given by equation 6.

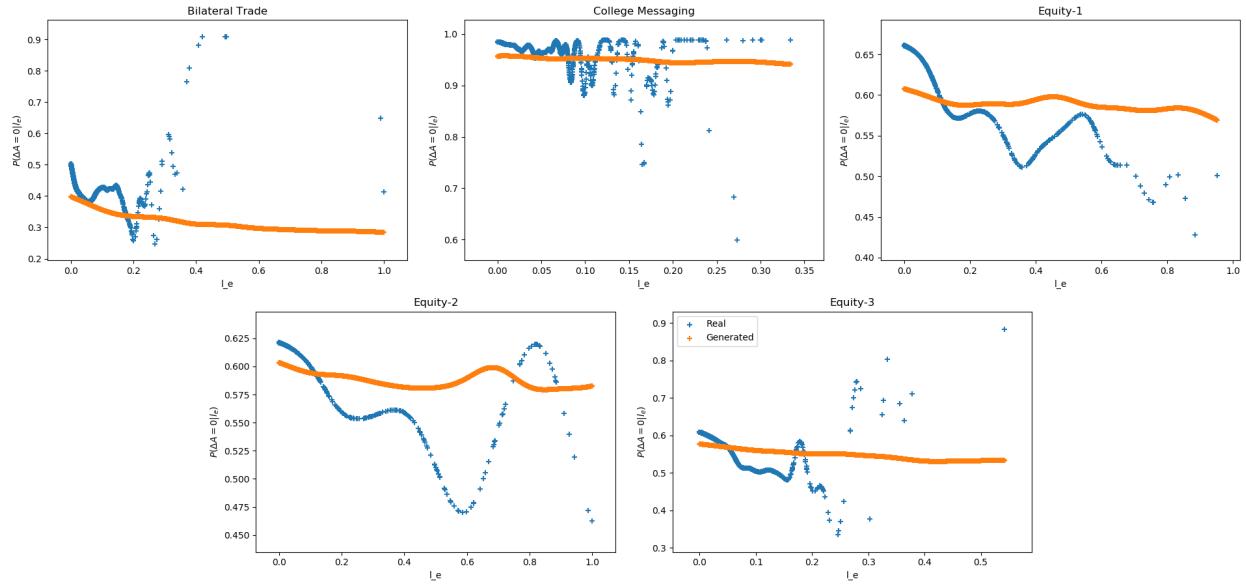


Figure 31:  $P(\Delta A = 0 | \ln(l_e))$  as a function of  $\ln(l_e)$  for the 5 real datasets, overlaid with the distributions for data generated according to the model in 6