Nesta...
WORKING PAPER SERIES

Nesta Working Paper No. 15/13

# Joining up the dots: Using social data to measure the effects of events on innovation

Bruce Cronin
Riccardo De Vita
Guido Conaldi
*Greenwich Business School*

# Joining up the dots: Using social data to measure the effects of events on innovation

Bruce Cronin
Greenwich business school

Riccardo De Vita
Greenwich business school

Guido Conaldi
Greenwich business school

Nesta Working Paper 15/13
August 2015

www.nesta.org.uk/wp15-13

**Abstract**

The paper studies the effects of the LeWeb tech conferences using data collected from the social media platform Twitter and the code sharing website GitHub. The extent to which attendance at the conference and other factors determined the patterns of tweeting among participants are examined. A group of attendants of the London LeWeb conference who did not attend the subsequent Paris event is used to assess the effects of LeWeb Paris. Conference attendees are matched to their corresponding profiles on GitHub to allow the effect on code collaboration to be examined. Permutation regression and Stochastic Actor Orientated Modelling (SAOM) are used to undertake a statistical evaluation of the changes in network.

**JEL Classification**: D85 Network Formation and Analysis: Theory, L14 Transactional Relationship, Contracts and Reputation, Networks, L17 Open Source Products and Markets, C39 Other, C59 Other

**Keywords**: Social networks, data, social media, Twitter, conferences, events, GitHub, evaluation

**Report prepared for Nesta**

**August 2015**

**University of Greenwich, a charity and company limited by guarantee, registered in England (reg. no. 986729).**
**Registered office: Old Royal Naval College, Park Row, Greenwich, London SE10 9LS.**

## Executive Summary

This study extends earlier Nesta research on the relationship between in-person interaction in conferences and events and subsequent collaboration. This research is aligned with a more general research programme pursued by Nesta, exploring the impact of events on innovation and the development of innovative measures to capture such impact, particularly the analysis of social media interactions. The context studied is the European web-tech community engaged in the prominent *LeWeb* conferences, a large and growing phenomenon. To better capture the impact of attendance to *LeWeb* on innovation, the collaboration of a particular subset of participants who are involved in open source software development is analysed.

Achieving these objectives required a specific set of data and the adoption of appropriate methodologies. Social Network Analysis (SNA) has been used to map connectivity and collaboration and longitudinal network models have been employed. The data analysed consists of online connections linking *LeWeb* participants on *Twitter* and *GitHub*. *Twitter* is the most prominent microblogging tool and allows users to follow other people online and to post new information. It therefore offers a simple and inexpensive way to acquire and share information. *GitHub* is a major platform for the sharing of software code, designed to stimulate collaboration. The *GitHub* website claims more than 5 million users. The *GitHub* platform allows the mapping of collaboration networks as developers are linked to all the projects they are working on.

The main findings of the research are as follows:

**Event attendance and interactions on Twitter**

1. Participants in LeWeb'12 London who attended *LeWeb'13* Paris increased their interactions via *Twitter* (i.e. retweets, mentions and replies) with participants in this (and the earlier *LeWeb'12 London)* 13 per cent more than those who attended LeWeb'12 London but did not attend LeWeb'13 Paris.

2. Participants in *LeWeb'13* particularly used *Twitter* to interact with those who attended multiple *LeWeb* events and with those who had interacted via *Twitter* with them (reciprocation).

3. Participants in *LeWeb'13* Paris followed more *Twitter* users and had more Twitter followers globally than a control group, but added fewer followers and followed fewer users on Twitter during the period of the event. There was no significant change in the number of Twitter followers when looking exclusively at following between LeWeb'13 Paris participants, however this analysis is based on a sample of fewer than 20 attendees.[1]

[1] Owing to data limitations it has unfortunately not been possible to fully study the patterns of Twitter following among only those that attended the events, as opposed to the total numbers that those who attended the event followed on Twitter overall.

4. Participants classified as entrepreneurs gained more followers on Twitter and followed more users during the period than any other identified subgroup.

5. *LeWeb* has a large core of participants attending multiple *LeWeb* events, around eight per cent of the total number of people who have attended LeWeb had been to more than two events. The members of the core that are most influential in the network of multiple participation in *LeWeb events* are predominantly associated with media organisations, corporations, and venture capitalists.

***Event attendance and effects on software collaboration***

6. *GitHub* developers who participate in *LeWeb* events have a distinctive pattern of collaboration on *GitHub*, with dense regions of significant interaction on software projects (strong ties) combined with smaller numbers of sparse regions of weak ties. This is characteristic of high performing networks in a wide variety of situations. *LeWeb* participants' projects had four times as many *GitHub* users subscribing to updates and were used to seed new projects by other users four times more often than those of a randomly selected control group of *GitHub* participants.

7. *GitHub* developers who were central to the network of participants in multiple *LeWeb* events consistently attracted much greater code contributions to their projects than *GitHub* users in the control group. In the final year of code contributions examined, 2013, LeWeb participants had 112 per cent more contributions to their *GitHub* projects than the random GitHub users.

8. Thus, activity and performance in *GitHub* is associated with particular patterns of collaboration in *LeWeb*. But *Twitter* is not a significant channel in this relationship; *GitHub* users participating in *LeWeb'13* Paris did not gain more *Twitter* followers or tweet more than other participants.

9. The positive association between *LeWeb* participation and distinctive collaboration patterns, project performance and greater open source contributions on *GitHub* may be a symptom of exposure, in that these developers and their projects are more visible through this prominent activity in the web-tech sector. It may also be that participation in the *LeWeb* events is a source of novel projects that are intrinsically attractive to code contributors. And it may indicate a selection effect; high status developers within the *GitHub* community may be more motivated and financially able to participate more extensively in high profile events such as *LeWeb*.

In summary, we found that participation in the *LeWeb* face-to-face events had a demonstrable impact on subsequent online communications and collaboration. The broad relationship between participation in face-to-face events and subsequent online communication and collaboration identified in this study of *LeWeb* is likely to be found in conference events in general. With increasing use of social media, it is likely that the interactions forged in the face-to-face event will be translated into interactions in social media as well as other forms of collaboration.

# Contents

## Glossary

| Term | Explanation |
|---|---|
| **Alter** | A node directly connected to a focal node (ego). |
| **ARD centrality** | Weights two-step centrality by discounting steps that are further away. |
| **Beta centrality** | A weighted version of degree centrality, weighting by the degree centrality of the actors it is connected to, the weighting determined by a user-set beta. |
| **Betweenness centrality** | The number of times a node is on the shortest path between each pair of nodes in a network, so is an indicator of the node's role in bridging the less connected parts of the network. |
| **Breadth** | The extent to which a network is not compact. |
| **Centrality** | The extent to which a node contributes to the structure of a network by virtue of its position in the network. |
| **Centralization** | The extent to which a network centres on a single node. |
| **Cohesiveness** | The extent to which a network is structurally cohesive, structural cohesion defined as "the minimum of actors whose removal would not allow the group to remain connected or would reduce the group to but a single member." (White and Harary 2001: 334-35) |
| **Collaborator** | A *GitHub* user who has been provided read and write access throughout a repository by the owner. |
| **Commit** | A contribution to a *GitHub* repository. |
| **Compactness** | The mean of the reciprocal of the shortest paths between each pair of nodes. |
| **Connectedness** | The proportion of pairs of nodes that can reach each other. |
| **Contributor** | A *GitHub* user who has provided software code suggestions to a repository that the owner may or may not decide to use. |
| **Closeness centrality** | The sum of shortest paths of connections from one node to each other node in a network, an indicator of the extent to which the node is at the centre of a cohesive group. |
| **Closure** | The extent to which a network is comprised of transitive triple nodes. |
| **Degree centrality** | The number of ties a node in a network has with other nodes. |
| **Density** | The total number of ties divided by the total number of possible ties in a binary network. |

| | |
|---|---|
| **Ego** | A focal node i.e. a node that is being considered in a given network. For example, how many people are connected to focal node X? |
| **Eigenvector centrality** | Weights degree centrality by the number of ties each connected node has, the weighting determined by the cohesiveness of the network |
| **Follower** | A *Twitter* user who has subscribed to view tweets of a particular *Twitter* user. |
| **Following** | Subscribing to view the tweets made by a particular *Twitter* user. |
| **Fork** | Instances of software code on one *GitHub* project repository being used to develop a separate software project on the site. |
| **Fragmentation** | The proportion of pairs of nodes that cannot reach each other. |
| **Hashtag** | A “#” prefix to a word in a tweet, facilitating grouping and searching tweets by group. |
| **Indegree centrality** | The number of ties towards the focal node. |
| **Mention** | Citing a specific *Twitter* user-name in a tweet. |
| **Node** | A social entity in a social network, typically represented visually by a circle or square. |
| **Outdegree centrality** | The number of ties from the focal node. |
| **Outcloseness centrality** | Closeness centrality in terms of ties from a focal node towards other nodes. |
| **Out 2-step centrality** | Two-step centrality in terms of ties from a focal node towards other nodes. |
| **Paired tweet** | A mention, reply or retweet involving a tweeting pair. |
| **Reply** | Responding to a specific *Twitter* user’s tweet. |
| **Repo** | A repository on *GitHub* holding code and commentary relating to a software project |
| **Retweet** | Republish a *Twitter* user’s tweet. |
| **Strong tie** | An inter-personal relationship characterised by high amounts of time, emotional intensity, intimacy and reciprocation. |
| **Tie** | An inter-personal relationship, typically represented visually by a line or an arrow. |
| **Transitivity** | A transitive relationship is where three nodes are related A- B, B-C and C-A. Networks with many transitive relationships, high transitivity, tend to be “clumpy” (Borgatti et al. 2013). |
| **Tweet** | The format for posts on Twitter.com, limited to 140 characters, typically containing links to online documents or extended commentaries |

| | |
|---|---|
| **Tweeting pair** | A pair of tweeters A and B where A or B mentions, replies to or retweets B or A respectively more than once during the Twitter data collection period. |
| **Two-step centrality** | A measure of how many actors are within two steps of the focal actor, as where a focal actor A retweets actor B who has retweeted actor C. |
| **Watcher** | *GitHub* users that have subscribed to be alerted when a project repository is updated. |
| **Weak tie** | An inter-personal relationship characterised by low amounts of time, emotional intensity, intimacy and reciprocation. |
| **Weighted overall clustering coefficient** | The weighted average of the density of ties among immediate connections of each node. Equates to network transitivity. |

# 1. Introduction

In a recent Nesta research project (Bakhshi, Davies & Mateos-Garcia, 2013) participants in the *LeWeb'12* London conference were found to connect to each other via *Twitter* faster than they connected with non-participants, with many international reciprocal connections made. Nesta commissioned the current research to establish a causal effect of event attendance by controlling for common demographic and behavioural characteristics among connectors that might independently explain this greater interaction. The research also aims to assess the effects of event participation beyond information sharing to other forms of collaborative activity, such as joint projects.

This project extends the earlier Nesta project with a more comprehensive longitudinal study of the *LeWeb* conferences, combining data on participation in a series of these events with data on collaborative projects in *GitHub* (a major site for open source software collaborations). As highlighted in the earlier research, the *LeWeb* conferences bring together many major figures in technological innovation in Europe so constitute an exemplar of a potentially influential event. The conferences have been held annually in Paris since 2004, supplemented by London conferences in 2012 and 2013. A major group among attendees, the third largest category by organisation type, are developers involved in open source software projects, potentially interested both in the frontiers of the evolution of the web discussed at these events and in potential partners for their businesses, particularly venture capitalists. The collaborative and brokering emphasis of *LeWeb* underpins important access to data for this research; participant lists are publically available for each event. Similarly, *GitHub's* collaborative emphasis provides public access to a great volume of data on participant interaction.

In this project we give particular attention to the effect of *LeWeb'13* Paris on attendees. This event provided the opportunity to capture *Twitter* data on attendees before and after its occurrence. This study provides a link with the earlier research on the *LeWeb'12* London event . Studying the collaborative behaviour of participants in *LeWeb'13* Paris adds a further event to that analysed in the original research, while the participation of some of the *LeWeb'12* London attendees in the Paris 2013 event provides comparator groups with otherwise similar characteristics. Finally, we extend the earlier research by also considering collaboration among a further subgroup of *LeWeb'13* Paris attendees, software developers engaged in the *GitHub* platform, for which considerable data is also available.

This report is structured as follows. Section 2 provides an overview of literature on the interface between in-person and online collaboration, with a focus on open source software development. Section 3 introduces the research questions and methodology for this project. Sections 4 to 8 present the results of the investigation and section 9 discusses these results.

## 2. Theoretical Framework

As Bakhshi, Davies & Mateos-Garcia (2013) have identified, events such as trade fairs and conferences are potentially critical nexuses in the innovation process. Events provide opportunities for the development of both weak and strong ties among participants. Weak ties comprise transitory interactions where participants are exposed for small periods to ideas, activities, performances and people that they may not have encountered before; these ties may be pursued through subsequent communication. These events may reinforce previous encounters with such ideas, activities, performances and people, strengthening the engagement and creating a stronger tie; this is more likely in repeated participation in a serial event or on the basis of established communication. Ties are not necessarily strengthened in such events, however; participants may not like what they see.

*Twitter*, and microblogging in general, is a rapidly growing communication channel structured to cultivate weak ties but which also has mechanisms to consolidate stronger ties. Tweets, short comments typically containing links to online documents or extended commentaries, are publically browsable and users can subscribe to view the tweets made by a particular individual by following them, that is, adding the tweeter to the user's following list (a list of users the user follows). A tweeter can view those who have subscribed to their tweets on their follower's list. Large followers' lists provide social influence and so the mechanism facilitates the formation of weak ties, where the costs of maintaining the relationship is minimal and linking the focal individual with very diverse groups. But because it is possible to cite a specific user-name in a tweet (a mention), republish a user's tweet for the particular attention of one's followers (retweet) and to respond to a specific user's tweet (reply), *Twitter* also provides mechanisms for the growth of stronger, partially reciprocated ties where individuals follow each other and communicate via *Twitter*.

*Twitter* is the paramount example of a microblogging tool and its use is growing in association with conferences (e.g. Letierce, Passant, Breslin & Decker, 2010b; Sopan, Rey, Butler & Shneiderman, 2012). Both conference organisers and participants can benefit in several ways from the use of *Twitter* (Reinhardt, Ebner, Beham & Costa, 2009). According to research carried out in 2007 (Java, Song, Finin & Tseng, 2007) users mainly use *Twitter* for sharing purposes, reporting news, information, communicating their daily activities or to have conversations. By analysing the connections of users, Java et al (2007) distinguished different and potentially overlapping categories of *Twitter* users: "information seekers" are characterised by limited posting activity, but following several other users; "information sources", on the other hand, are those users with many followers, interested in their post – but with frequency of posting not necessarily a key determinant of following; finally "friends" are those users replicating on *Twitter* their off-line social networks, and thus having, for example, many interactions with personal acquaintances, colleagues and relatives.[2]

---

[2] The term 'friends' has also got a separate technical meaning within the context of Twitter, it is the people that a person is following on Twitter.

When considering scientific communities and conference participants the information sharing motivation becomes predominant: *Twitter* is particularly effective for spreading scientific knowledge to different communities and to share information by disseminating url links to substantive resources (Letierce, Passant, Breslin & Decker 2010a; Reinhardt et al., 2009). An analysis of three distinct conferences (Letierce et al., 2010a) revealed a number of interesting patterns about the use and importance of *Twitter* during conferences. Firstly, prominence in the online network is associated with physical participation to the conference, as well as with having played specific roles in it, such as being an organizer or a keynote speaker. Even more interestingly, the study found a much higher level of retweeting and a higher use of hashtags when comparing *Twitter* activities of conference participants to a random sample. This result shows the much stronger attention to the content and spread of specific information by conference attendees. Qualitative data collected through interviews emphasised how participants can use tweets to expand their networks.

The limited text characters available for a tweet, together with the tendency to prioritise communication to peers in the same professional community and the use of well-known hashtags can however limit the potential of *Twitter* to reach out to new communities or people not physically attending a specific event (Ebner et al., 2010; Letierce et al., 2010a). Moreover, qualitative studies on the use of *Twitter*, even if limited in scope, have found evidence of the importance of factors related to individual characteristics and context (including presence at relevant conferences) in influencing the use of *Twitter* by scientists (Kieslinger, Ebner & Wiesenhofer, 2011).

In other words, *Twitter* is a powerful tool both for conference participants and organizers to manage information networks and access new resources and can be fruitfully used during conferences. At the same time, the use of *Twitter* and its suitability to connect to different audiences is also influenced by contingent factors.

Despite some limitations and room for further improvement, the potential for *Twitter* to impact on collaboration is evident (Honey & Herring, 2009; Zhao & Rosson, 2009). Considering *LeWeb* attendees, it is therefore possible to expect that participants developing different types of relationships through *Twitter* will show differing collaborative behaviours.

The interplay between conference attendance and online interactions can be easily understood considering the extensive research carried out by Caroline Haythornthwaite (e.g. 2005). The author, investigating interaction among a group of researchers and of online distance learners, found evidence of linked online and offline interactions. Firstly, stronger ties are associated with the use of more media, both online and offline. Secondly, the creation of a technological infrastructure such as *Twitter* generates the opportunity for the creation of new links, which, however, do not necessarily emerge.

Such infrastructures make a connection available technically, even if not yet activated socially. These technical connections support latent social network ties, used here to indicate ties that are technically possible but not yet activated socially. They are only activated, i.e. converted from latent to weak, by some sort of social interaction between members, e.g. by telephoning someone, attending a group-wide meeting, reading and contributing to a web-board, emailing others, etc. (Haythornthwaite, 2005: 137).

In the case of a *LeWeb* conference, therefore, it is plausible to assume that participation in the event triggers an interaction facilitated by the existence of online networking and collaboration tools (e.g. *Twitter*, *GitHub*), resulting in distinctive relational patterns. Considering the web-tech focus of *LeWeb* participants, a relevant form of collaboration may include one aimed at the production of new software. So we might expect changes to interactions mediated by *Twitter* and *GitHub* following in-person interaction at *LeWeb*.

The software industry is a paramount example of a knowledge intensive industry (McKelvey, 2001) where problem solving is crucial (Conaldi & Lomi, 2013) and different communication network structures are associated with different types of innovation pursuit (Conaldi & De Vita, 2011). Literature on innovation emphasises the benefits associated with different network structures (Ahuja, 2000), but their effectiveness seems to vary with contingent factors such as the type of knowledge exchanged and the type of innovation pursued. The strength of social interaction in terms of amounts of time, emotional intensity, intimacy and reciprocation provide varying channels for different types of information and knowledge exchange (Granovetter, 1973). Weak ties are particularly useful for exploration while strong ties facilitate transfer of complex information and exploitation of knowledge. While both strong and weak ties are required to sustain the complementary learning process of exploration and exploitation, an optimal balance is difficult to achieve with investment in one at the expense of investment in the other. But combinations of dense regions of strong ties interlinked by numerous redundant weak ties are associated with high performance outcomes in a wide range of social situations from knowledge transfer within business to scientific collaboration and corporate strategic alliances (Cronin, 2007). While it is not possible to expect *a priori* a specific network structure to characterising *GitHub* projects involving *LeWeb* participants, it is expected that in some of these typical high performing morphologies will be found.

# 3. Methodology

## 3.1 Research Objectives and Hypotheses

In this project, our objective was to explore the particular collaborative-oriented environment of the technology sector around the *LeWeb* events and the Twitter and *GitHub* platforms in order to understand the relationship between the social connections established in conferences and subsequent collaborative activity. Specifically, we sought to answer the following questions:

1. What was the additional effect of attending *LeWeb'13* Paris on *Twitter* connections and communications?
2. What was the additional effect of attending past *LeWebs* on *GitHub* connections and engagement?
3. What were the effects of changes in *LeWeb* attendees' interconnectedness on *GitHub* on project success?

The literature review developed in the previous chapter allows the development of more specific research hypotheses, summarised here and which will be discussed in the following paragraphs.

*Hyp#1 – Participants in LeWeb'13 Paris gained more Twitter followers and increased the number of users followed more than non-participants with otherwise similar characteristics.*

*Hyp#2 – Participation in LeWeb'13 Paris increased tweeting among participants.*

*Hyp#3 – The extra connections gained at LeWeb'13 Paris varied among distinct subgroups by organisation type, nationality and participation in multiple events.*

*Hyp#4 – Different patterns of collaboration in GitHub projects were associated with different patterns of participation in LeWeb events.*

*Hyp#5 – Participants in LeWeb'13 Paris* who were on *GitHub gained more Twitter friends and followers than those that did not.*

*Hyp#6 – These particular GitHub collaborative patterns that participants in LeWeb'12 involved in were associated with higher project performance.*

*Hyp#7 – These GitHub collaborative patterns involved network structures typical of high-performance collaboration in other contexts.*

## 3.2 Data Collection

We compiled a dataset of participants in each *LeWeb* event 2009-2013 (including the 2012 and 2013 London events).[3] We derived preliminary demographic information from an analysis of names, location and organisational affiliation. Participant gender was derived by comparing first names to name-gender dictionaries. Organisation types were coded on the basis of website domains, descriptors within the organisation name (bank, university, agency etc.) and a manual examination of websites listed as part of the participant information.

We used three sources to analyse social interaction among *LeWeb* participants. First, the participant lists themselves provided data for an analysis of serial participation in the events 2009-13. Second, we used the *Twitter* Application Programming Interface (API)[4] to download data on friends and followers from *Twitter* of all participants in the London 2012 and Paris 2013 *LeWeb* events who had listed their *Twitter* IDs; typically, this excluded the headline speakers, whose Twitter IDs were not always published.[5] The download was undertaken at two time points: immediately before the start of Paris 2013 and six weeks after its end, providing data to assess the effect of the Paris 2013 event on the number of friends and followers. Then, we downloaded all tweets by participants on a daily basis during the two time points, providing data to analyse the pattern of interaction in terms of mentions, replies and retweets.

We then identified a subset of those *LeWeb* participants who were involved in *GitHub* projects during the same period, using queries in the *GitHub* API.[6] While data is publically available for *GitHub* projects before 2009, *LeWeb* participation data is not, so the same starting point for both datasets was used for comparative purposes. The search strategy was based on the rationale that an active user of several social media platforms is likely to use the same public ID for *Twitter* and *GitHub* as this serves as an identifier in the community in which they communicate. However, this will not always be the case and so a full name search will isolate some of these cases. The way that names are entered and recorded on different platforms varies, however, and a last name search can provide data for further investigation of variations in first names. We accepted a match where:

i. the *Twitter* user name listed on *LeWeb* and the *GitHub* user name were the same ***and*** one of the three following conditions also applied:
ii. the last name was the same on the two lists;[7] or
iii. the affiliated organisation name was the same on the two lists; or
iv. the affiliated website was the same on the two lists.

---

[3] Participant information for earlier events, which commenced in 2004 were not publicly available.
[4] Api.twitter.com/console
[5] Due to download limitations set by the Twitter api, it proved difficult to download data listing specific user-follower and user-following relationships. However, aggregate totals of followers and friends (users followed) were systematically collected.
[6] Api.github.com
[7] The likelihood of the userids on both lists and the last names being the same on both lists is very low even where the last name is a common name.

We created a further dataset of randomly selected *GitHub* participants of equal size to the *GitHub* subset, to act as a control group.

Demographic information was derived from a name, location and affiliation analysis. For each of the two *GitHub* groups we identified project contributors and characteristics such as length of collaboration and group role, and performance indicators of each project, principally numbers of watchers and forks. Watchers are users that have subscribed to be alerted when a project site is updated. Forks are instances of the software code on one project repository being used to develop a separate software project on the site.

## 3.3 Data Analysis

We examined three major relationships over time:

i. The pattern of strong and weak ties among *LeWeb* participants compared to a control group, indicated by:
   a. repeated event participation (Section 4);
   b. *Twitter* interaction in terms of following (Sections 5 and Appendix B); and
   c. communication on *Twitter* (Section 6);
ii. The pattern of strong and weak ties among *LeWeb* participants' *GitHub* contributors and collaborators indicated by repeated project participation, compared to results for a counterfactual control group (Section 7); and
iii. The relationship between the two networks and collaborative project outcomes (Section 7).

We undertook descriptive analyses by comparing characteristics of the participants of each *LeWeb* to determine trends over time (Section 4). We used Social Network Analysis to map and visualise the pattern of relationships, categorise the structural properties of the network of relationships and identify distinctive characteristics of the most central actors.

Since only 59.6 per cent of the *LeWeb'13* Paris participants and 63.7 per cent of the *LeWeb'12* London participants listed Twitter IDs on registration, there is a question as to the extent to which the network analysis represents actual online interaction among participants. But because these events are major collaborative opportunities for participants in the web-tech industry and because Twitter IDs were explicitly asked for on registration and were publicised on the conference websites it is unlikely that many participants would have omitted to provide these details. So we assume that the networks analysed are effectively complete. In any case, the principal network metrics employed are quite robust in the face of small amounts of missing data (Costenbader & Valente, 2003).

To provide a rigorous analysis of changes to *Twitter*-based interaction associated with participation in the latest LeWeb (*LeWeb'13* Paris), we divided the *LeWeb'12* London participants into two groups. The first group (the participant group) comprised those *LeWeb'12* London attendees who went on to take part in *LeWeb'13* Paris. The second group (the control group) comprise those *LeWeb'12* London attendees who did not participate in *LeWeb'13* Paris. This provided two groups for comparison with otherwise very similar characteristics. It also provides a link with the dataset analysed in the earlier Nesta research.

A simple investigation of changes to Twitter-based interaction among *LeWeb'13* Paris participants as compared to changes to Twitter-based interaction among *LeWeb'12* London participants over the same period would not have been as rigorous as the method above. The events while thematically similar have different characteristics, the London event being less established than the Paris events and would not control for the strong French locational character of the latter (see section 4). With separate populations, it would not be possible to use participation in *LeWeb'13* Paris as a dummy variable in regression analysis; analysis would be restricted to analysis of variance, where methods for analysing network analysis are not well developed.

Specialised methods of regression analysis were employed to analyse the network data because of the very high levels of interdependency in the data. For each relationship, we employed randomised permutation-based statistical models to control for this interdependency, which arises because the subject of analysis is connections among participants, with the dependencies reinforced over time; a connection is likely to be both maintained and reinforced through network effects such as reciprocation and popularity. Conventional statistical analysis is not useable in this context as it assumes independence among observations and known distributions of connections to compare these to. By contrast, very little is known about the distributions of network data, other than that small differences in configurations have large effects on observed patterns. Application of conventional statistical techniques to network data generates numerous spurious relationships between variables (Borgatti, Everett & Johnson, 2013; Snijders, van de Bunt & Steglich, 2010).

Permutation-based regression overcomes the limitations of conventional statistical analysis of network data as it makes no assumptions about the independence or distribution of observations. Rather, observations are compared to a large number of random permutations (typically 10000) within the vector of values for each variable being compared. The statistical significance of each coefficient among the observed data is determined by the extent to which it persistently differs from the randomised data. The coefficients themselves are derived from an ordinary least squares approach (Borgatti, Everett & Freeman, 2002).

In the case of following and tweeting, because these involve individual decisions, actor-based dynamics can be explored. Changes to these networks were analysed with a class of Stochastic Actor-Oriented models (SAOM) for social networks (Snijders, van de Bunt & Steglich, 2010).

The regression modelling allowed us to assess the relative effects of the network structures compared to particular participant and project characteristics. The longitudinal design and SAOM modelling allowed us to determine some likely endogenous determinants of change in the network.

We start by assuming that the network structure observed at any one time develops as a result of interdependent individual decisions. Actors are only allowed to change the ties under their direct control and no single actor has control over the entire network structure. Statistically, this assumption leads to a representation of the network structure that is observed at any moment as a realization of a continuous-time stochastic process $Y(t)$ – where observed realizations are $y(t_\tau)$ (with $\tau = 1,2,\ldots,T$) (Snijders, 2001). At any point in time the process produces the observed network $Y(t) = y$. In the specific case we discuss $y$ is a unipartite network of size *NxN* – that is, a network composed of one set of *N* actors and their interactions – with tie variables $y_{ij}$= 1 if participant i communicates via *Twitter* with participant j. Formally, the model is a continuous time Markov process, whose state space is defined in terms of all the possible combinations of network ties (Snijders, 2001).

Linking SAOM to data requires specification of two main components. The first is a *rate function* $\lambda_i(\alpha, y)$ which controls how quickly opportunities for changing network ties arise. In our case, the relevant change is in the communication exchanges via *Twitter*. Participants get opportunities to change their targeted recipients at the rate:

$$\lambda(Y(t)) = \exp\left(\alpha_0 + \sum_k \alpha_k x_{ik}\right), \qquad (1)$$

The rate may be constant between observations (when $\alpha_k$ = 0, for all k), or it may change depending on actor-specific covariates ( $x_{ik}$ ). In our model specification, the rate is constant between observation moments (with $\alpha_k$ = 0, for all k); $\alpha_k$ being the rate at which actors are given the chance to consider whether to tweet or not. The rate function is used to parameterise the frequency with which actors are assumed to consider potential changes to their outgoing network ties. In its most basic specification with only $\alpha_0$ being estimated, the rate parameter can be used to gauge the amount of change decisions that actors need to consider for the network as a whole to evolve between to observation points. Since we do not have information regarding the underlying propensity of actors to consider the possibility of a tweet, we will adopt this basic specification in the analysis.

The second component of SAOM is the individual decision of participants to change their connections. This decision is controlled by an *evaluation function ($f_i$)* representing the relative attractiveness for participant *i* of moving from network state *y* to *y'*, where y and y' are successive network configurations differing in terms of only one tie – or none in case participants opt not to change any tie. In this case y = y'. Among the possible *m* changes that a participant can make at any one time, he/she is assumed to choose so that $f_i(\beta, y(i \to j), X) + U_i(t, y, j)$ is maximized. In this formulation *f* is a deterministic evaluation function, X is a set of covariates, and $U_i$ is a random variable that captures the component of an actor's preference that is not systematic. In this respect $U_i$ can be conceived as disturbance in the utility maximisation process. $U_i$ are assumed to be independent and identically distributed for all $t, y, j$ . The deterministic part of the evaluation function assumes the typical linear form (Snijders, van de Bunt & Steglich, 2010):

$$f_i(\beta, y, x) = \sum_k \beta_k \, s_{ik}(y, x) \qquad (2)$$

where $f_i(\beta, y)$ is the value of the evaluation function for participant *i* depending on the state of the network (y), and the term $s_{ik}(y)$ encompass three categories of "effects" which may be associated with (i) network *motifs*, or recurrent patterns of interconnections, *(where k is the sum over the different effects entering the function)*; (ii) actor-specific covariates representing "exogenous" characteristics of the participants, and (iii) interactions between network *motifs* and exogenous covariates; these three categories encompass all the types of variables that can be specified in an evaluation function. Finally, $\beta_k$ are parameters that may be estimated from the data and represent the weight, or strength, of these effects.

Individual decisions to change network ties are based on a comparison of the values of the evaluation function computed across the permissible choice options. More specifically, if *y'= y(i→j)* denotes the network that would be observed if participant *i* changed his connections to participant *j* (creation of a new tie or termination of an existing tie, dependent on the existence of a tie), and if *y'= y(i→i)* denotes the network that would be observed if no change was made (and hence y = y'), then the probability to observe one change would be:

$$\Pr(y(i \to j, x)) = \frac{\exp\left(\sum_k \beta_k s_{ik}(y(i \to j), x)\right)}{\sum_h \exp\left(\sum_k \beta_k s_{ik}\left(y(i \to h), x\right)\right)}. \quad (3)$$

Where the sum over h is a sum over all the states of the network caused by a single change in person i's connections.

This characterisation of $\Pr(y(i \to j) | x)$ is possible as the non-systematic component $U_i(t, y, j)$ in the evaluation function is assumed to follow an extreme value distribution with mean = 0 and scale parameter = 1. Under this assumption the model just described is consistent with interpretations of revealed preferences (see, for example, McFadden 1974). According to this interpretation when participant *i* changes his or her connections (his/her row in the network) producing a change from configuration *y* to *y',* then he/she is acting as if he/she prefers *y' to y.*

## 4. Participation in *LeWeb* events

In a high-profile series of events such as *LeWeb*, there are likely to be subgroups formed by interaction through repeated co-participation in events. To examine this, we downloaded the publically available participant lists from each *LeWeb* event from 2009 to 2013 (these being the *LeWebs* with public data on attendance), including the separate London and Paris events that occurred in both 2012 and 2013. We considered the distribution of participation in the events. We mapped participation as relationships between an individual and one or more events and isolated the characteristics of the most central participants.

As presented in Table 4.1, the number of participants in *LeWeb* increased from 2106 in Paris in 2009 to 3151 in the Paris 2013 event. In total there were 11536 unique participants, most participating in a single event, 2462 in more than one event and 973 (8% of the unique participants) had been to more than two events (Table 4.2).

**Table 4.1. Participants in *LeWeb* events 2009-2013**

| Event | 2009 | 2010 | 2011 | 2012 London | 2012 Paris | 2013 London | 2013 Paris | Total | Total Unique |
|---|---|---|---|---|---|---|---|---|---|
| **Participants** | 2106 | 2098 | 3019 | 1254 | 2827 | 1194 | 3151 | 15649 | 11536 |

Note: prior to 2012 LeWeb was only held in Paris

**Table 4.2. Distribution of participants by number of events**

| Number of Events attended | Participants N | Participants % | Cumulative Participants |
|---|---|---|---|
| 1 | 9074 | 78.7% | 11536 |
| 2 | 1489 | 12.9% | 2462 |
| 3 | 556 | 4.8% | 973 |
| 4 | 238 | 2.1% | 417 |
| 5 | 118 | 1.0% | 179 |
| 6 | 40 | 0.3% | 61 |
| 7 | 21 | 0.2% | 21 |

We mapped the network of interaction arising from co-participation in the same events, with a line between two nodes representing participation in an event. Those individuals repeatedly present in the same events as other individuals have a greater likelihood of interacting and so are more central in the network. As visualised in Figure 4.1, the core of this network, defined as participation in three or more events,[8] is centred on *LeWeb*'11, *LeWeb'12* Paris and *LeWeb'13* Paris, a symptom of both the number of participants and the number of repeating participants. The earlier and London events are less central.

[8] 3-core (3 or more links); 973 participants in seven events.

**Figure 4.1. Network of core *LeWeb* participants 2009-2013 by gender**[9]

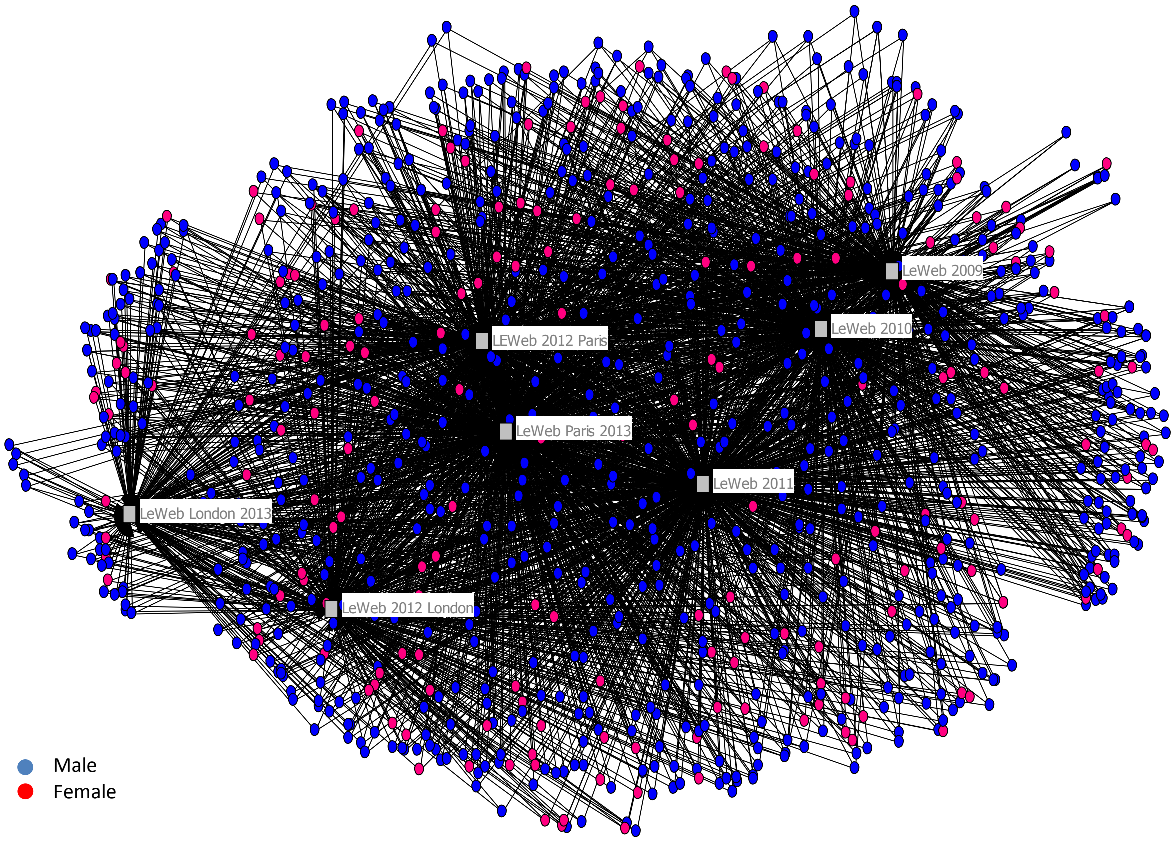



Figure 4.1 also includes the gender composition of the core, derived from a name analysis. Women are generally at the periphery of the core with only four women within the central circle of *LeWeb'10, LeWeb'11, LeWeb'12* Paris and *LeWeb'13* Paris.

The national location of participants, as given in their *Twitter* account settings, is presented in Figure 4.2; the great majority of core participants are located in France. Figure 4.3 discriminates participants by organisation type, derived from an examination of company names and websites listed on the participant lists, as described in the section 3. Core participants are predominantly affiliated with corporations (generally software companies or high value consumer goods producers), news media and other publishers, and venture capitalist firms as well as individual developers. The events mix observers of developments in web technologies with the brokering of new businesses.

---

[9] This and subsequent visualisations, unless otherwise stated, were created in Netdraw 2.131 using a spring-embedded algorithm starting from a Gower Scaling.

**Figure 4.2. Core *LeWeb* participants (those that had participated in at least three events) by location**

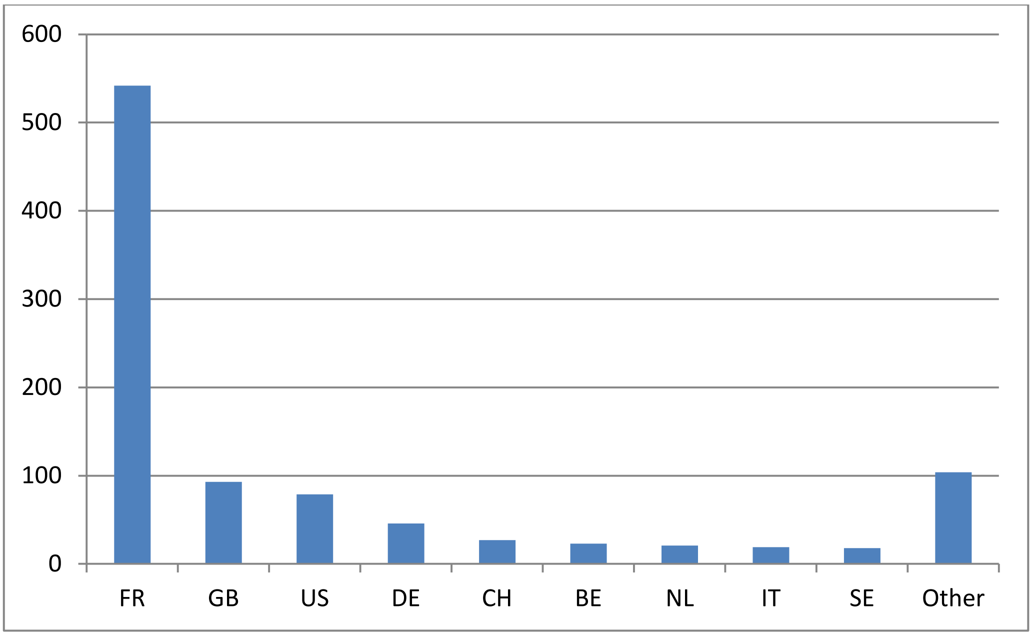
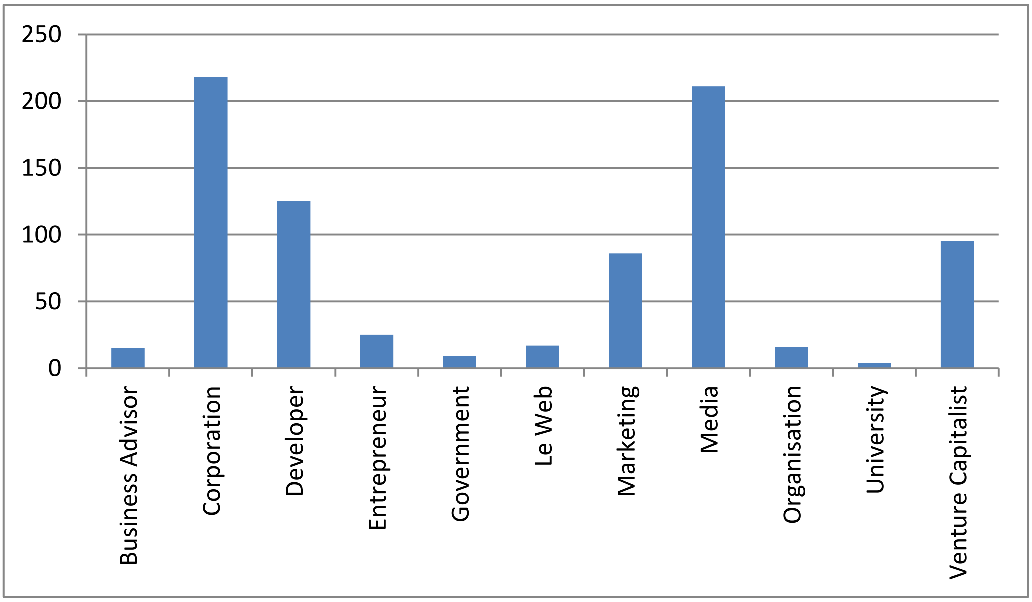



**Figure 4.3. Core *LeWeb* participants by organisation type**




We compared a range of the characteristics of individual participants with the centrality of their positions within the network, using a multivariate permutation based regression analysis. We excluded organisation types with low frequencies in Figure 4.3. But we added a second category of developer, those participants who we matched through a search of registered users of *GitHub*.

Four common measures of network centrality were considered. Degree centrality, in this case, is the number of events a participant took part in. Eigenvector centrality weights degree centrality by the number of participants attending each event. Closeness centrality examines the path of connections that would be needed to reach all other participants in all events, so is an indicator of reach through the network as a whole. Betweenness centrality measures the number of times a participant is on the shortest path between each pair of nodes, so is an indicator of bridging the less connected parts of the network. In these last two cases, reach can be thought of as a flow of interaction from the first event to the last; in retrospect an individual with high levels of betweeness and closensess centrality can be seen to have been central to the totality of the flow of interactions over time.

The results of the regression are presented in Table 4.3. The dependent variable in each model is one of the four measures of centrality. The first three independent variables are dummy variables for gender and location and whether a participant was matched on *GitHub* (a subset of developers). The remaining independent variables comprise our classification of affiliated organisation by type, as described in section 3.2. The values reported are the unstandardized coefficients of the association between the dependent and independent variables and can be interpreted in the same way as a standard regression.

The models account for only a very small proportion of variation in the data (R-square), indicating that we understand very little of the variation in why people attended multiple *LeWeb* events. But the four organisational types are persistently significant across three different measures of centrality so highlight the important features of those network influences that are present.

**Table 4.3. Multivariate regression of core network centrality and participant characteristics[10]**

| | Model 1 | | Model2 | | Model 3 | | Model 4 | |
|---|---|---|---|---|---|---|---|---|
| **Dependent Variable** | Degree Centrality | | Eigenvector Centrality | | Betweenness Centrality | | Closeness Centrality | |
| **Male** | 0.039210 | | -0.000170 | | 0.506910 | | 13.14085 | ** |
| **France location** | 0.061201 | | 0.002139 | *** | 0.538064 | | -24.55841 | |
| ***GitHub* match** | 0.081250 | | 0.001510 | | 1.602715 | | -0.131536 | |
| **Developer** | 0.053647 | | 0.000142 | | 0.603706 | | 9.850981 | |
| **Venture Capital** | 0.511226 | *** | 0.003559 | *** | 7.467112 | *** | -15.79849 | |
| ***LeWeb* staff** | 1.054596 | *** | 0.004196 | ** | 13.58761 | ** | 4.403875 | |
| **Marketer** | -0.045520 | | 0.000216 | | -0.556880 | | 3.399135 | |
| **Media** | 0.238792 | ** | 0.002335 | ** | 3.354986 | ** | -13.52859 | |
| **Corporation** | 0.152603 | * | 0.001295 | ** | 2.269397 | * | -7.019051 | |
| **R-square** | 0.041 | | 0.042 | | 0.035 | | 0.036 | |
| **Adj. R-square** | 0.031 | | 0.032 | | 0.025 | | 0.026 | |
| ***p*** | 0.000 | *** | 0.000 | *** | 0.002 | ** | 0.002 | ** |
| **Observations** | 973 | | 973 | | 973 | | 973 | |
| **Permutations** | 10000 | | 10000 | | 10000 | | 10000 | |

*p < 0.1; ** *p* < 0.05; *** *p* < 0.000

[10] Centrality metrics and node-level randomized permutation regressions were calculated with UCINET 6.474. The coefficients reported are unstandardized, that is, the relationships are presented in terms of the variables' original, raw units. So a change in the male variable from 0 to 1 (male as opposed to not male) is associated with an increase in Degree Centrality of 0.039, although this is not a statistically significant relationship. Regression statistics are reported in more detail in the Appendix, Tables A1 to A4.

In Model 1, the dependent variable is degree centrality; participation in more events is associated unsurprisingly with members of the *LeWeb* organising team, but also venture capitalists, media organisations and corporations.  Model 2 takes into account the popularity of each event, finding similar characteristics of participants but also disproportionate involvement of French nationals. Model 3 finds the same basic pattern associated with betweenness. Model 4 finds no characteristic associated with reach across the network other than maleness; this is a global measure of centrality likely to differ from the other indicators of centrality more local to individual participants.

Putting the event organisers and other variables aside, then, on a number of indicators the most central participants at the series of *LeWeb* events were venture capitalists, media organisations and corporations, indicating a more transient participation by smaller organisations such as developers and marketing agencies.  French location and being male were also central on particular dimensions.

So in summary, *LeWeb* is a large and growing phenomenon, with a core of repeating participants accounting for around eight per cent of the total. The central participants are predominantly France-located males from media organisations and corporations, most likely present to digest trends in web-tech, and venture capitalists seeking new opportunities.

## 5. *Twitter* Followers and Following

A face-to-face event such as *LeWeb* provides the opportunity to encounter new contacts, new ideas and sources of inspiration. Given that *LeWeb* promotes participants' *Twitter* user IDs as a contact mechanism, participants in a *LeWeb* event are likely to add *Twitter* users attending the event to their following lists and are likely to gain additional *Twitter* users following them among those attending the event.

To measure the impact of such an event on the participants following and followers, we queried the *Twitter* Application Programming Interface (API) to download the following and followers lists of those participants in *LeWeb'12* London and *LeWeb'13* Paris who provided their *Twitter* User IDs when they registered. We divided the *LeWeb'12* London attendees into two groups: those that also participated in *LeWeb'13* Paris and those that did not. We examined changes to the number of users they followed and the number of followers they had during the period before and after *LeWeb'13* Paris.

As presented in Table 5.1, the participant group (*Le Web'12* London attendees who participated in *LeWeb'13* Paris) followed more *Twitter* users and had more followers than the control group (*Le Web'12* London attendees who did not participate in *LeWeb'13* Paris) but was of similar order; low-1000s in the case of following and around 10,000 in the case of followers.

**Table 5.1. Change in *Twitter* following and followers by *LeWeb'12* London attendees after *LeWeb'13* Paris**

| | *LeWeb'12* London Participants | | *LeWeb'13* Paris Participants |
|---|---|---|---|
| | Participants in *LeWeb'13* Paris | Non-Participants in *LeWeb'13* Paris | All Participants |
| **Sample size (N)** | 370 | 884 | 3151 |
| **N with Twitter User IDs** | 238 | 561 | 1878 |
| **Mean Users Followed before Paris 2013** | 1268 | 1035 | 1415 |
| **Mean Users Followed after Paris 2013** | 1341 | 1101 | 1265 |
| **Change in Mean Users Followed** | 5.8% | 6.4% | -10.6% |
| **Median Users Followed before Paris 2013** | 529 | 450 | 489 |
| **Median Users Followed after Paris 2013** | 589 | 512 | 585 |
| **Change in Median Users Followed** | 11.3% | 13.8% | 19.6% |
| **Mean Followers before Paris 2013** | 10711 | 9094 | 9672 |
| **Mean Followers after Paris 2013** | 11617 | 10423 | 9346 |
| **Change in Mean Followers** | 8.0% | 14.6% | -3.4% |
| **Median Followers before Paris 2013** | 1511 | 665 | 802 |
| **Median Followers after Paris 2013** | 1778 | 922 | 667 |
| **Change in Median Followers** | 17.7% | 38.6% | -16.9% |
| **Followers/Followed Ratio before Paris 2013** | 8.4 | 8.8 | 6.8 |
| **Followers/Followed Ratio after Paris 2013** | 8.7 | 9.5 | 7.4 |

The ratio of followers to followed was also similar for both groups, the magnitude being a ratio typical of popular thought-leaders.[11] The change to the mean number of followers and followed was less for the participant group than for the control group, a symptom of skewing within the participant group by some users with very large numbers of followers and followed. Changes to the median were similar in all cases except followers among the control group. In both cases there was an increase in the followers to followed ratio, particularly among the control group. There was no clear association between participation in *LeWeb'13* Paris and changes to the numbers followed or following.

To examine whether the comparison of these averages hide underlying dynamics, we undertook a randomised permutation test regression analysis to identify statistically significant effects at the individual user level. Both the change to the number of *LeWeb'12* London followers and users followed during the period immediately preceding *LeWeb'13* Paris and six weeks after the event were examined. The independent variables tested are listed in Table 5.2. These include whether a *LeWeb'12* London attendee participated in *LeWeb'13* Paris or not, to determine if participation in the event was associated with a change in online activity.[12] Other independent variables include whether they were identified as a *GitHub* user and our classification of their affiliated organisation. The remaining six independent variables considered whether the attendee was a member of the *LeWeb* core discussed in Section 4 and various measures of their centrality within that core.

The regression results are presented in Table 5.2. The first two models examine participant characteristics associated with changes to followers after *LeWeb'13* Paris, the last two models consider changes to following over the same period. Models one and three consider participant occupations with participation in *LeWeb'13* Paris. Models two and four introduce whether participants hold core positions within the network. No model explains follower or following activity in total but they do highlight some variables associated with this activity, though not necessarily causing it.

*LeWeb'12* London attendees classified as entrepreneurs were more likely to gain more *Twitter* followers during the period (40.8) than any other identified subgroup. And L*eWeb'12* London attendees classified as developers increased the number of people they followed on *Twitter* during the period (19.7) than any other identified subgroup. This did not extend to users who were also identified as *GitHub* users, however. Neither changes to followers or changes to following were statistically associated with participation in *Le Web'13* Paris. Core positions within *LeWeb* as a whole did not influence the results.

---

[11] See tffratio.com

[12] This approach was chosen over a comparison of changes to following/follower activity by *LeWeb'12* London-only participants versus *LeWeb'13* Paris-only participants as it controlled for the national differences between participants in the two events and was likely to capture more central participants in *LeWeb.*

**Table 5.2. Changes in *LeWeb'12* London followers and following after *LeWeb'13* Paris by subgroup**[13]

| | Model 1 | | Model 2 | | Model 3 | Model 4 | |
|---|---|---|---|---|---|---|---|
| **Dependent variable:** | ΔFollowers | | ΔFollowers | | ΔFollowing | ΔFollowing | |
| **LeWeb'13 Paris** | -4.38 | | 11.63 | | -4.99 | 1.38 | |
| **GitHub match** | 2.64 | | -9.50 | | -7.99 | -20.09 | |
| **Developer** | 8.20 | | 9.89 | | 17.93 | 19.73 | ** |
| **Government** | 12.58 | | 11.37 | | -2.05 | -1.33 | |
| **Media** | -5.22 | | -4.47 | | 0.016 | 0.24 | |
| **Venture Capitalist** | -6.96 | | -3.90 | | -1.00 | 0.33 | |
| **Marketing** | -4.66 | | -4.02 | | 0.47 | 0.41 | |
| **Entrepreneur** | 40.82 | ** | 40.86 | ** | 0.94 | 0.82 | |
| **University** | -5.89 | | -8.37 | | -0.94 | -1.68 | |
| **Participant Core** | | | 1729838 | | | -361425 | |
| **Participant Core Degree** | | | 998.04 | | | -199.01 | |
| **Participant Core Betweenness** | | | 0.02 | | | 0.16 | |
| **Participant Core Closeness** | | | -294.91 | | | 61.62 | |
| **Participant Core Eigenvector** | | | -2480.22 | | | -355.41 | |
| **Participant Core 2-Local Eigenvector** | | | 0.03 | | | 0.00 | |
| **R-square** | 0.019 | | 0.023 | | 0.006 | 0.007 | |
| **Adj R-square** | 0.007 | | 0.001 | | -0.006 | -0.016 | |
| **F** | 1.691 | | 1.105 | | 0.560 | 0.338 | |
| ***P*** | 0.141 | | 0.542 | | 0.608 | 0.819 | |
| **Observations** | 799 | | 799 | | 799 | 799 | |
| **Permutations** | 10000 | | 10000 | | 10000 | 10000 | |

** $p < 0.05$

An interpretation for these sectoral differences in following and follower activity might be that entrepreneurs use *Twitter* as one of many channels for publicising their activities in a general search for weak ties (followers) that may be beneficial in some combination in the future. By contrast, developers, as distinct from those involved in large corporations, may be start-ups, who are engaged in more specific projects and therefore are more purposeful in their search, seeking stronger ties and adding users to their following lists that could be more immediately beneficial. Participation in *LeWeb'13* Paris did not have any detectable independent impact on this behaviour, however; the control group did not have a statistically different growth of contacts than the participant group. This suggests the accumulation of contacts on follower and

[13] The regressions (of node-level randomized permutation type) were calculated with UCINET 6.474. Unstandardised coefficients; in Model 4 participation in LeWeb'13 Paris (a change in the variable from 0 to 1) is associated with an increase in users followed by 1.38, though this is not statistically significant relationship.

following lists is a widespread activity and, with a median of 1400-2400 contacts each, new contacts added at an individual event do not make a significant difference.

In summary, then, the participant group of *LeWeb'12* London attendees who participated in *LeWeb'13* Paris on average followed more *Twitter* users and had more followers than the control group. But the control group added proportionately more followers and users they followed after *LeWeb'13* Paris. There was no statistically significant effect from *LeWeb'12* attendees London participating in *LeWeb'13* Paris on the number of followers or users followed, however. Thus our first hypothesis, that participants in *LeWeb'13* Paris gained more Twitter followers and increased the number of users followed more than non-participants with otherwise similar characteristics, was not supported in this analysis.

A more rigorous test of the first hypothesis would have been to examine changes to followers and users followed solely among participants in the two events. However, difficulties in collecting this data limited this. An examination of changes to followers and users followed among a small group of LeWeb participants is presented in Appendix B. No measurable effect of participation in *LeWeb'13* Paris was found on *Twitter* followers among the sample examined, however the sample size is small (fewer than 20 participants).

In terms of all *Twitter* followers and users followed, *LeWeb'12* London attendees classified as entrepreneurs gained more *Twitter* followers during the period than any other identified subgroup. And *LeWeb'12* London attendees classified as developers increased the number of *Twitter* users during the period they followed more than any other identified subgroup. This provides some support for hypothesis 3, that the extra connections gained at *LeWeb'13* Paris varied among distinct subgroups by organisation type, nationality and participation in multiple events. However, hypothesis 4, that participants in *LeWeb'13* Paris involved in *GitHub* projects gained more Twitter friends and followers than those who were not, was not supported.

## 6. Tweeting among *LeWeb* Participants

To better understand the relationship between face-to-face events and online interaction we collated tweets from participants in two *LeWeb*s (London 2012 and Paris 2013) for a period before and after *LeWeb'13* Paris. These were participants in *LeWeb'13* Paris itself and participants in *LeWeb'12* London, some of whom took part in both events. The London event thus provided a participant group and control group of individuals with similar characteristics but with different exposure to the face-to-face event.

After collating the tweets, we used social network analysis to examine the patterns of interaction of the two groups before and after the *LeWeb'13* Paris event. We compared the network structures of interaction and examined the relative positions of all participants within these networks. We undertook statistical analysis of the relationship between individuals' participation in *LeWeb'13* Paris and their relative position in the subsequent network of *Twitter* interactions. And we isolated the principal drivers of the change in interactions between them.

Between early December 2013 and the end of January 2014, we mined *Twitter* on a daily basis for tweets issued by London 2012 and Paris 2013 participants. We collated 430,564 unique tweets, a small proportion dating as early as 2009 as *Twitter* makes available through its API the last 800 tweets of each user. There is no method of determining definitively what proportion this represents of the total tweets undertaken by a user. But It seems unlikely that many users would issue more than 800 tweets in a day, so this is likely to be close to a comprehensive account of the period surveyed.

We selected two periods for comparison, on the basis of the distribution of tweets (See Appendix Figure A1). The first (Period 1) was the five weeks from 1/11/13 to 7/12/13, the day before the start of *LeWeb'13* Paris. The second (Period 2) was from 2/1/14 to 13/1/14, immediately following the holiday period ending on New Year's Day, where there was a considerable volume of holiday-related tweeting. After New Year's Day, we reasoned tweeting would be more representative of "normal business" and so more comparable with the earlier period. At the same time, while shorter, the second period encompassed four times as many tweets (97,362) than the first (24,600). But selecting an even shorter second period would make the latter more vulnerable to particular news or other unusual events. So this seemed a reasonable compromise. As there is no other viable second window in these terms, there was little scope for sensitivity testing.

As most tweets are isolated broadcasts with no record collected of who reads these, the analysis was restricted to relational tweets, that is where a tweet mentions, replies to or retweets a tweet posted by another user. We only examined relationships among *LeWeb'12* London or among *LeWeb'13* Paris attendees. Table 6.1 summarises the number of tweets from each sample.

**Table 6.1. Mentions, replies to or retweets by sample**

| | London 2012 Period 1 | London 2012 Period 2 | Paris 2013 Period 1 | Paris 2013 Period 2 |
|---|---|---|---|---|
| **Mentions, replies to or retweets referring to other attendees** | 329 | 754 | 517 | 1261 |
| **All tweets by attendees** | 11527 | 53060 | 12553 | 44302 |

For each event and time period, we identified the network of relationships among *LeWeb* participants formed by mentioning, replying and retweeting. We measured standard structural characteristics of each network and compared the networks on these dimensions (See Appendix Table A1). Each group and period exhibited very similar structural characteristics and the change from the first period to the second was very similar in each case. The cohesiveness of the network, that is, how closely or "knittedly" participants interact with each other,[14] increased on all indicators other than transitivity,[15] which increased among all *LeWeb'12* London participants but decreased among *LeWeb'13* Paris participants. Similarly, the distribution of reciprocated and unreciprocated communication remained remarkably constant across all groups and periods and these were not statistically distinct (See Appendix Table A2). So, participation in *LeWeb'13* Paris had no measurable impact on the structure of mentioning, replying and retweeting in general.

In order to undertake a more detailed examination of the changes in tweeting interaction, we focused on a more cohesive, repeated set of interactions. This was to allow some modelling of the dynamics of network change, which cannot be done with very sparse networks. We extracted a subset of the most intensive tweeters, those who mentioned, replied to or retweeted the same person more than once during the whole period that the tweeting data was collected for ("paired tweets"; see Appendix Figure A2 for the distribution of this repeat activity). We identified 74 participants in *LeWeb'12* London who were involved in 59 distinct tweeting pairs (one or more tweets) in the month before *LeWeb'13* Paris and the same pairs of individuals were involved in the same tweeting pairs in the month after *LeWeb'13* Paris. The 74 participants were involved in 893 paired tweets (be they replies, retweets, or mentions among the paired participants) in total, encompassing 149 participants in total.[16] 90 of these took part in *LeWeb'13* Paris. Among the *LeWeb'13* Paris participants, a total of 560 participants were involved in 1778 paired tweets. While this is a small proportion of all tweeting activity in the event, it comprises a persistent social structure (as opposed to broadcasting or one-off weak ties), and thus is amendable to stochastic actor oriented modelling.

---

[14] White and Harary (2001: 334-35) define structural cohesion of a network as "the minimum of actors whose removal would not allow the group to remain connected or would reduce the group to but a single member." There are a variety of network metrics that indicate structure relative to this state (See the discussion in Borgatti, Everett and Johnson 2013).

[15] A transitive relationship is where three nodes are related A- B, B-C and C-A. Networks with many transitive relationships, high transitivity, tend to be 'clumpy' (Borgatti et al. 2013).

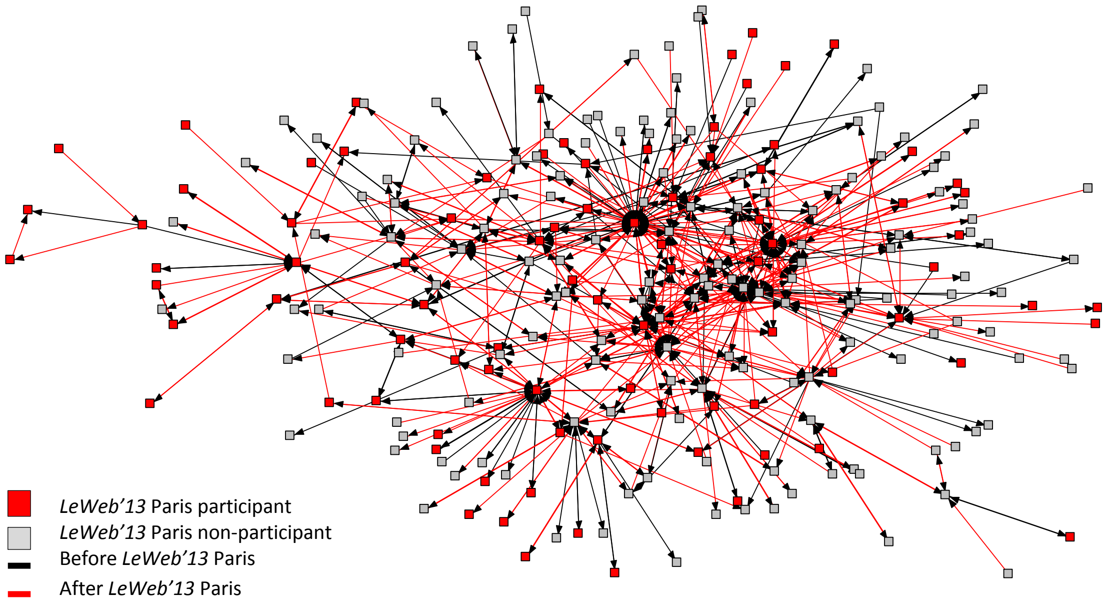
[16] See glossary for precise definitions of tweeting pairs and paired tweets.

Figures 6.1-6.3 represents tweets among *LeWeb'12* London participants involving the subgroup of 74 participants who on at least one occasion mentioned, replied to or retweeted the same participant before and after Paris 2013. We examine the totality of the tweeting activity among this group both before and after *LeWeb'13* Paris. Figures 6.4 and 6.5 repeat this for participants in LeWeb'13 Paris.

In each figure, the black lines represent tweets before *LeWeb'13* Paris and the red lines tweets after *LeWeb'13* Paris. Pairs where tweets occurred both before and after LeWeb Paris are included in the "after" category and so are red. Grey squares represent actors who attended *LeWeb'12* London and red squares represent actors who attended *LeWeb'13* Paris only. The arrows point to the participant who was mentioned, replied to or retweeted.

In Figure 6.1 the most frequent participants cited on *Twitter* are distributed in a central ring within the two periods, though the most central participants are those who did not attend *LeWeb'13* Paris. Repeated tweeting after *LeWeb'13* Paris was concentrated among the most central participants in the network, whether they took part in the event or not.

**Figure 6.1. *LeWeb'12* London attendees' mentions, replies or retweets of other *LeWeb'12* London attendees pre- and post- *LeWeb'13* Paris** [17]




Note: This visualisation (and those which follow) show only participants in paired tweets, that is those who mentioned, replied or retweeted another delegate or who were mentioned, replied to or retweeted more than once.

[17] Main component. This and the next four visualisations were created in Netdraw 2.131 using a spring-embedded algorithm starting from a Gower Scaling.

**Figure 6.2. *LeWeb'12* London attendees with nodes scaled by who was most frequently mentioned, replied to or retweeted by before *LeWeb'13* Paris**

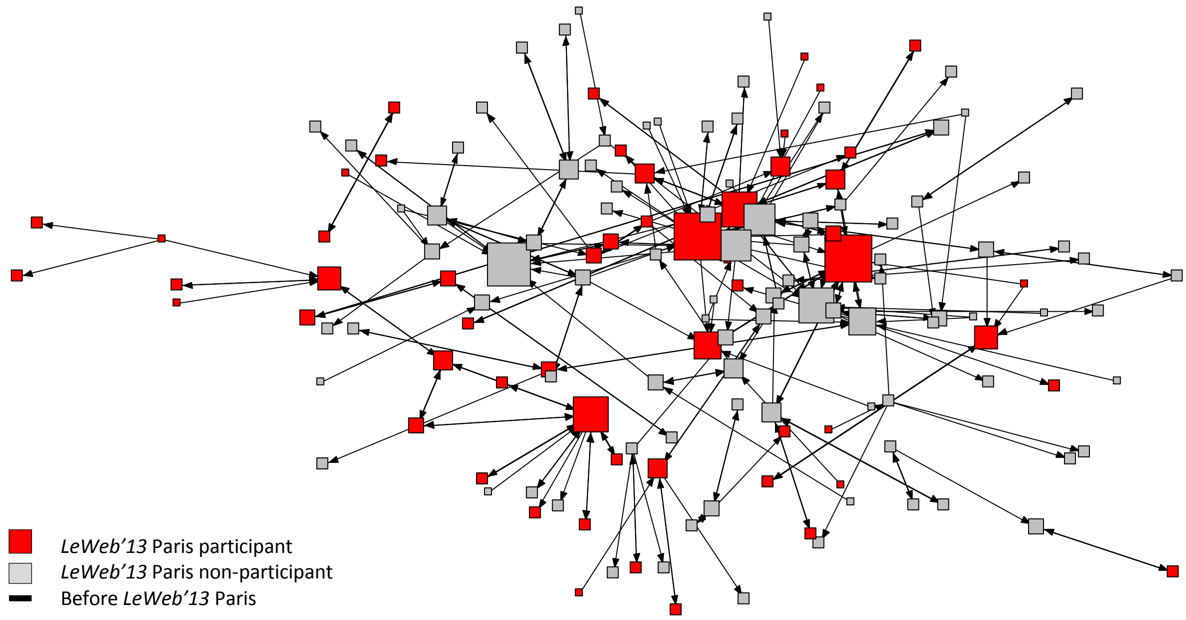



**Figure 6.3. *LeWeb'12* London attendees with nodes scaled by who was most frequently mentioned, replied to or retweeted by after *LeWeb'13* Paris**

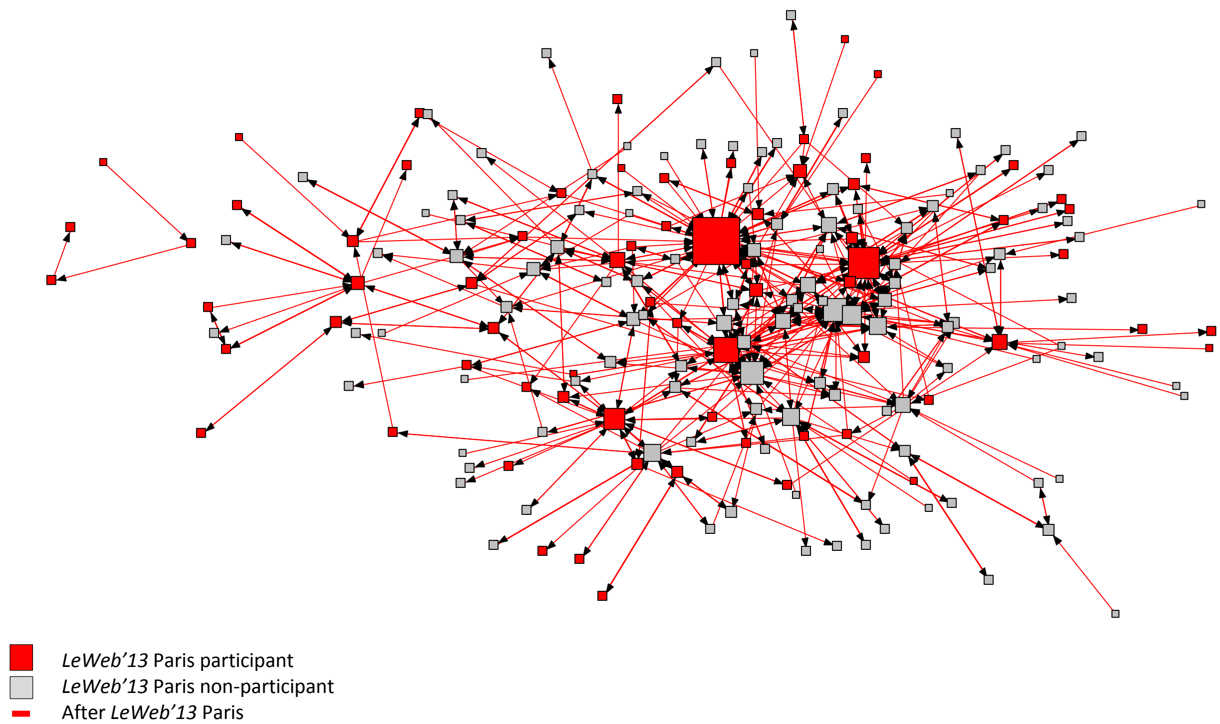


**Figure 6.4. *LeWeb'13* Paris attendees with nodes scaled by who was most frequently mentioned, replied to or retweeted by before *LeWeb'13* Paris**

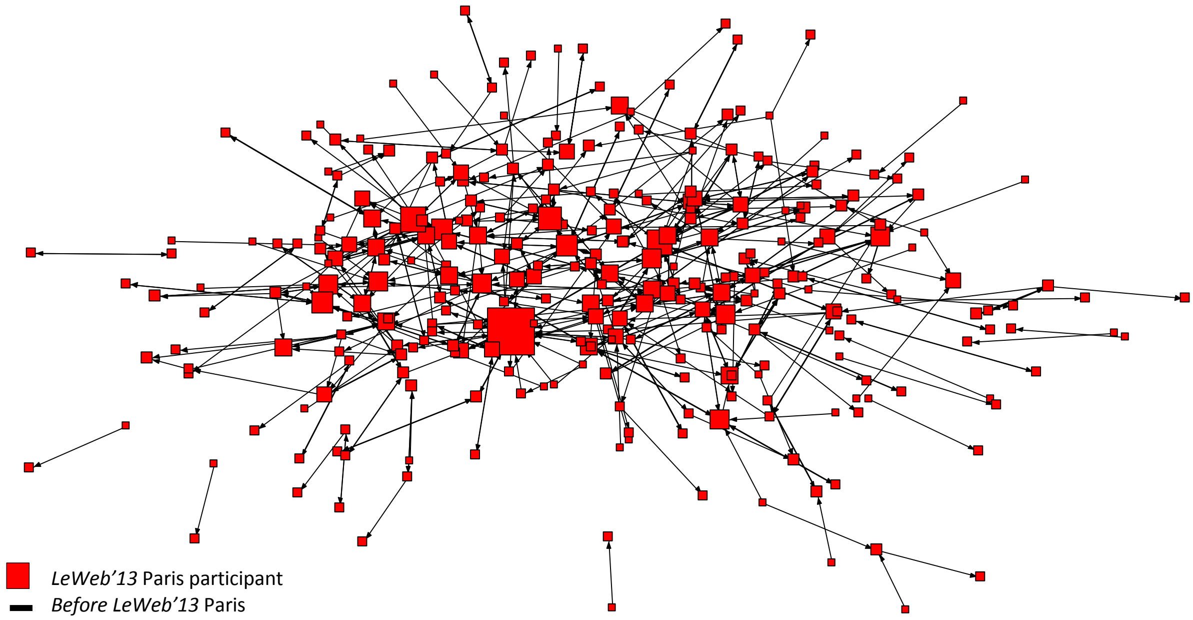
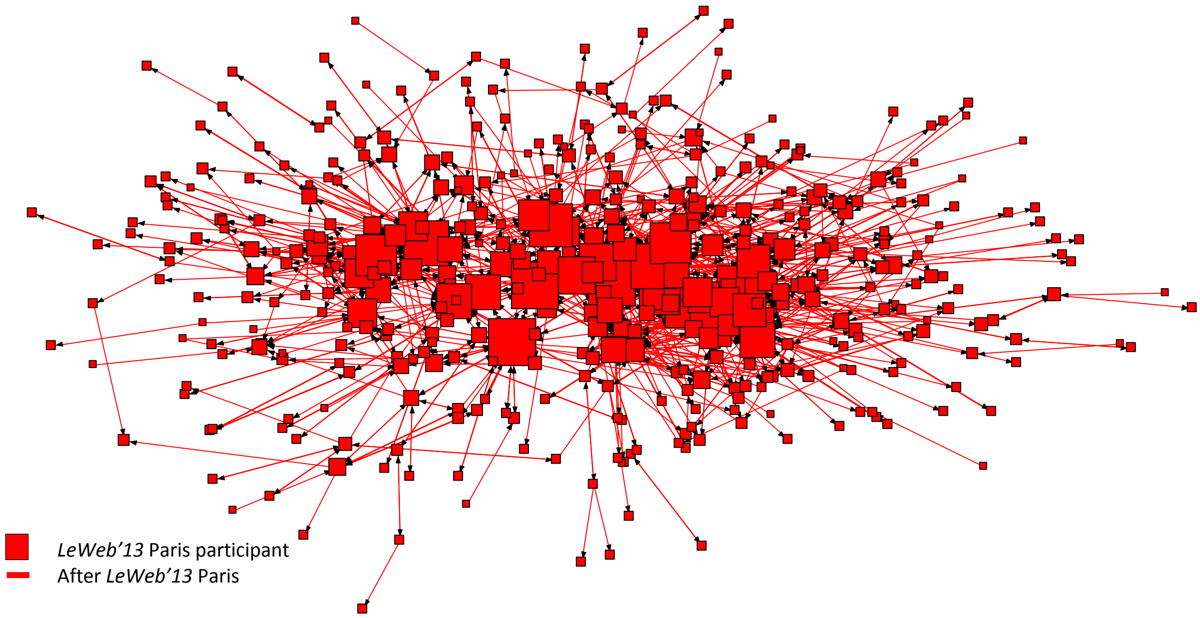



**Figure 6.5. *LeWeb'13* Paris attendees with nodes scaled by who was most frequently mentioned, replied to or retweeted by after *LeWeb'13* Paris**



In Figures 6.2 and 6.3, the larger nodes represent the participants who were most frequently mentioned, replied to or retweeted before and after *LeWeb'13* Paris respectively. While both attendees and non-attendees were central before the event (Figure 6.2), the most central *LeWeb'13* Paris participants were most prominent after the event (Figure 6.3).

The interactions among *LeWeb'12* London repeat tweeters increased in quantity and cohesiveness after *LeWeb'13* Paris, for both the control group and people who went to *LeWeb'13* Paris. Appendix Table A3 reports a range of standard measures of network cohesiveness, all but the last (transitivity) indicating an increase in cohesiveness. This indicates a shift towards global cohesiveness over local cohesiveness as might occur as repeat tweeters became more familiar with the themes being communicated through the network.

Figures 6.4 and 6.5 represent tweets among *LeWeb'13* Paris participants involving participants who on at least one occasion tweeted the same participants before and after the event. The larger nodes represent the participants who were most frequently mentioned, replied to or retweeted before and after the event respectively.

Comparing Figures 6.4 and 6.5, the number of participants frequently mentioned, replied to or retweeted increased broadly after *LeWeb'13* Paris, indicating much more widespread interaction among participants following the event. In part this is affected by the longer timespan of the period after the event (twice the length of the period before) but the increase in this type of tweeting is greater than this.

As presented in Appendix Table A4, like *LeWeb'12* London attendees during the same period, the interactions among *LeWeb'13* Paris repeat tweeters indicate some increase in quantity and cohesiveness after the event, with the exception of closure and the weighted overall clustering coefficient (transitivity). The t-test, however, indicates that the structural characteristics of *LeWeb'12* London and *LeWeb'13* Paris repeat tweet networks were not significantly different.

The final column of Appendix Table A4 indicates that the increased cohesiveness of the repeated tweeting network after *LeWeb'13* Paris among *LeWeb'13* Paris participants was generally greater than that experienced by the *LeWeb'12* London repeat tweeters. However, *LeWeb'12* London repeat tweeters experienced greater connectedness, fragmentation, breadth and betweenness centralization indicators of weaker ties. Thus there is some indication that participation in *LeWeb'13* Paris increased the cohesion among repeat tweeters.

In order to test this impression, we undertook a statistical analysis of the extent to which participation in *LeWeb'13* Paris itself was associated with the positions of all attendees of *LeWeb'12* London within the aggregate community of *LeWeb'12* London and *LeWeb'13* Paris tweeters after *LeWeb'13 Paris*. Table 6.2 presents the results of multivariate randomized permutation test regressions on participation in *LeWeb'13 Paris.* Model 1 gives the relationship between participation in *LeWeb'13* Paris and a set of standard measures of network centrality. Model 2 extends the analysis by including sectoral classification of affiliated organisations.[18]

[18] It is not possible to extend the analysis further by including actor's centrality in multiple *LeWeb* events, as discussed in section 4, because the definition of core membership was participation in three or more

**Table 6.2. Regression of *LeWeb'12* London participation in *LeWeb'13* Paris, on centrality in subsequent mentions, replies to and retweets 2/1/14-16/1/14**[19]

| Dependent variable: *LeWeb'12* London Participant in *LeWeb'13* Paris | Model 1 | | Model 2 | |
|---|---|---|---|---|
| OutDeg | 0.328351 | ** | 0.342061 | ** |
| Indeg | 0.014653 | | 0.015133 | |
| OutBonPwr | 0.000901 | | 0.000889 | |
| InBonPwr | -0.000496 | | -0.000505 | |
| Out2Step | 0.139390 | ** | 0.145355 | ** |
| In2Step | -0.066372 | | -0.066525 | |
| OutARD | -0.557654 | | -0.576018 | |
| InARD | 0.091220 | | 0.088161 | |
| OutEigen | 0.514896 | | 0.513706 | |
| InEigen | -0.497389 | | -0.450610 | |
| Betweenness | -0.001635 | | -0.001451 | |
| InCloseness | -2.358115 | | 2.284094 | |
| OutCloseness | 467.3214 | ** | 483.8352 | ** |
| GitHub match | | | -0.059464 | |
| Developer | | | -0.036862 | |
| Government | | | -0.045065 | |
| Media | | | 0.007556 | |
| Venture Capitalist | | | -0.085786 | |
| Marketing | | | -0.040370 | |
| Entrepreneur | | | -0.038539 | |
| University | | | -0.462004 | |
| R-square | 0.044 | | 0.061 | |
| Adj. R-square | 0.025 | | 0.026 | |
| F | 2.507 | | 1.825 | |
| *P* | 0.216 | | 0.263 | |
| Observations | 607 | | 607 | |
| Permutations | 10000 | | 10000 | |

* *p* < 0.1; ** *p* < 0.05; *** *p* < 0.000

For all but the last independent variable in Model 1, there is an in- and out- version of the measure. This indicates the direction of the relationship. Outdegree is the number of times an actor tweets mentioning or replying or retweets to another within the aggregate community of *LeWeb'12* London and *LeWeb'13* Paris tweeters. Indegree is the number of times an actor is mentioned, replies to or takes a tweet from another. The most central actors are those with the greatest number on each of these indicators.  Other measures locate this activity not simply on

events; there have been only two London events so at least one event for each core member will be Paris based. As *LeWeb'13* Paris has been the largest event, core members are very likely to have participated in this. The correlation between core membership and *LeWeb'13* Paris participation is 43.4%. This definitional overlap would subsume statistically significant relationships.

[19] Unstandardised coefficients.  In Model 1, an increase in Outdegree of 1 is associated with an increase in the probability of being a participant in LeWeb'13 Paris of 0.328.

an individual basis but in the context of the network as a whole. Closeness centrality represents how quickly paths from all other actors can reach an actor, taking into account the direction of the links. Beta centrality is a weighted version of degree centrality, weighting by the degree centrality of the actors it is connected to. 2-step (2Step) centrality measures how many actors are within two steps of the focal actor, as where a focal actor A retweets actor B who has retweeted actor C. ARD centrality weights 2-step centrality by discounting steps that are further away. Betweenness centrality measures how many times an actor is on the shortest path between each other pair of actors. It indicates brokerage in the network, a position connecting otherwise disconnected parts of the network.

From Model 1 in Table 6.2 it can be seen that centrality among those *LeWeb'12* London attendees mentioning, replying to and retweeting (OutDegree, Outclosenes and Out2Step) is positively associated with participation in *LeWeb'13* Paris. Controlling for affiliated organisation type in Model 2 does not affect the results. To test the effect of participation in *LeWeb'13* Paris on tweeting by *LeWeb'12* London attendees, we reversed the regression relationship, with OutDegree as the dependent variable and participation in *LeWeb'13* Paris as the independent variable. As reported in table 6.3, participants had a 0.133 higher Outdegree than non-participants. That is, *LeWeb'12* London attendees who also participated in *LeWeb'13* Paris subsequently mentioned replied to or retweeted 13 per cent more to *LeWeb'12* London or *LeWeb'13* Paris participants in their subsequent tweets than those who did not attend.

**Table 6.3 Effect of participating in *LeWeb'13* Paris on tweeting by *LeWeb'12* London participants**

| **Dependent variable** | OutDegree | |
|---|---|---|
| **Participation in *LeWeb'13* Paris** | 0.133 | *** |
| **R-square** | 0.018 | |
| **Adj. R-square** | 0.014 | |
| **F** | 10.922 | |
| ***P*** | 0.001 | |
| **Observations** | 607 | |
| **Permutations** | 10000 | |

*** $p < 0.000$

So, while the overall pattern of tweeting among *LeWeb'12* London and *LeWeb'13* Paris participants did not differ structurally, participation by *LeWeb'12* London attendees in *LeWeb'13* Paris opened the door to increased tweeting interactions following the event.

To examine the dynamics of this process, we analysed changes to the tweeting network among participants in *LeWeb'12* London and *LeWeb'13* Paris with stochastic actor-oriented models (SAOM) for social networks, as discussed in section 3. The analysis was restricted to the subset of repeating tweeter pairs to provide sufficient density of interactions to observe sufficient variation in change and determinates; in the network of all tweets from this sample, most interactions were non-repeated bilateral relationships so the overwhelming change was termination, data not amenable to SAOM modelling.

In Table 6.4 we specify a SAOM model to estimate the changes to *Twitter* connections among participants most likely to have occurred for the pattern of *Twitter* interactions among *LeWeb'12* London attendees in the month before *LeWeb'13* Paris to transform to the pattern of interactions observed after the event. The values reported are odds ratios.

The dependent variable in these models is the decision by individual actors about whether to change their interactions with others in the network in the period between the two observation points, that is, before and after *LeWeb'13* Paris. The estimated rate parameter accounts for the frequency with which actors are assumed to consider new communication exchanges. The first parameter in the evaluation function (*outdegree)* controls for the overall density of the network. The parameter indicates the overall propensity of participants to engage in communication exchanges during the observation period. As very often in communication networks, the negative parameter indicates that overall participants are more likely not to initiate a communication tie when having the chance to do so. In other words, the decision to communicate is a rare event and communication exchanges happen only among a minority of actors.

**Table 6.4. SAOM models estimating determinates of change in the network between periods (among repeated tweeting pairs only)**

| | **Model 1** | | | **Model 2** | | |
|---|---|---|---|---|---|---|
| **N** = 249 | Coefficient | SE | | Coefficient | SE | |
| **Rate parameters:** | | | | | | |
| **0. Rate parameter** | 19.1108 | 2.1323 | | 20.3784 | 3.6309 | |
| **Evaluation parameters:** | | | | | | |
| **1. outdegree (density)** | -3.3221 | 0.0722 | ** | -4.7769 | 0.1949 | ** |
| **2. reciprocity** | 3.6387 | 0.1127 | ** | 4.0029 | 0.1587 | ** |
| **3. attendedP13 alter** | 0.3357 | 0.0769 | ** | 0.0137 | 0.1117 | |
| **4. attendedP13 ego** | 0.1546 | 0.1041 | | 0.3255 | 0.1486 | ** |
| **5. same attendedP13** | 0.0294 | 0.0572 | ** | 0.0506 | 0.0600 | |
| **6. indegree - popularity (sqrt)** | | | | 0.6974 | 0.0682 | ** |

** $P < 0.05$

In Model 1, the change in network ties between the two periods is accounted for by parameters 2 to 5; reciprocation of mentions, replies and retweets increased after *LeWeb'13* Paris (reciprocity). Attendees were engaged in significantly more ties, whether they attended and were recipients of mentions, replies or retweets (attended P13 alter), or both recipients and initiators attended (same attended P13). For example, by looking at the parameter estimate for the Reciprocity effect in Model 1 we can affirm that the odds of a participant to decide to communicate with another participant from whom communication was received as against to communicate with someone who did not communicate first are e^(3.6387)= 38.04.

Model 2 controls for indegree popularity. The effect tests whether participants already targeted by many others are progressively more likely to attract further participants. The statistically significant and positive parameter associated with the Indegree Popularity effect indicates that indeed an effect of progressively accumulated advantage – otherwise known as the Matthew or "rich-get-richer" effect – shapes the structure of the communication network by progressively increasing the likelihood of further engagements of already popular targets. On top of this overall effect of existing connections on the likelihood of future connections, Model 2 confirms that attendees in *LeWeb'13* Paris become more active tweeters after the event – they both attended and were initiators of these tweets (attended P13 ego) – and reciprocation was prevalent. However, the "alter effect" of attendees in *LeWeb'13* Paris is no longer significant in Model 2, thus leading to the conclusion that no extra popularity as recipients is gained by attendees once we control for the overall Mathew effect present in the network.

In conclusion, then, in terms of the effects of *LeWeb'13* Paris on retweeting decisions by *LeWeb'12* London attendees modelled as endogenous behaviour among this group participation increased the post-event tweeting activity by *LeWeb'12* London attendees. *LeWeb'12* London participants in *LeWeb'13* Paris mentioned, replied to or retweeted more participants in both events than *LeWeb'12* London attendees who did not attend *LeWeb'13* Paris. *LeWeb'12* London participants in *LeWeb'13* Paris were more central to the tweeting network as a whole and to close contacts than *LeWeb'12* London non-participants in *LeWeb'13* Paris. Reciprocal tweeting was an important part of the increased activity and increased centrality of these participants. Participation in the face-to-face event appears to have established reciprocated interactions after the event.

This is consistent with hypothesis 2, that participation in *LeWeb'13* Paris increased tweeting among participants. It also provides some support for hypothesis 3, that the extra connections gained at *LeWeb'13* Paris varied among distinct subgroups by organisation type, location and participation in multiple events. But it does not support hypothesis 4, that participation in *GitHub* projects was associated with different patterns of participation in *LeWeb* events and different patterns of *Twitter* follower, following and tweeting.

## 7. *GitHub* Collaboration

We identified a subset of *LeWeb* participants who had user accounts on *GitHub*, a major platform for open source software development. This opened a window to additional data on how these participants interacted outside the immediate context of *LeWeb*. It provided an opportunity to examine another dimension of the relationship between interaction in a face-to-face event and online collaboration.

We identified *GitHub* users among *LeWeb* participants through matches between data on all of the *LeWeb* participants' lists and the *GitHub* user directory. We accepted a match where:

i. the *Twitter* user name listed on *LeWeb* and the *GitHub* user name were the same ***and*** one of the three following conditions also applied:
ii. the last name was the same on the two lists; or
iii. the affiliated organisation name was the same on the two lists; or
iv. the affiliated website was the same on the two lists.

408 *LeWeb* participants were identified as *GitHub* users on these criteria. To provide a control group for our analysis we extracted a second random sample of *GitHub* users. This was the same size and had the same age distribution of user accounts as the group drawn from the *LeWeb* lists, using the numbering property of the site; *GitHub* ID numbers are generated sequentially, the lowest numbers being the oldest and largest the newest. Random numbers were generated in proportion to the number of IDs in the *Le Web* sample in the 100, 1000, 10000, 100000, and 1000000 ranges, to produce a sample stratified by age similar to the participant group. User details for these IDs downloaded. A comparison of means, medians and maximums of project characteristics indicates that both the *LeWeb* participant group and the control groups had similar levels of development activity on average.[20]

We collated details of software development repositories ("repos") maintained on *GitHub* by *LeWeb* attendants and the control group, including the number of collaborators and contributors on each project and performance indicators (watchers and forks). A collaborator is a user who has been provided read and write access throughout a repository by the owner. A contributor is a user who has provided software code suggestions that the owner may or may not decide to use. Watchers are users who have registered to be notified of any changes to a repository. Forks are the utilisation of code from one repository to develop a new software project elsewhere.

We mapped the network of co-participation among the sampled users, their collaborators and contributors in these projects. Table 7.1 presents descriptive statistics for the repos maintained by the two samples. The number and distribution of repos maintained by *LeWeb* participants and the random sample are similar, means of 6.8 and 6.7 medians of 3.0 respectively. So the two samples have similar levels of hosting activity. But the repos maintained by *LeWeb* participants are considerably larger on average and the distribution of these in terms of size twice as large as the random sample. The *LeWeb* participants' repos have four times as many watchers and are

[20] There is no limit to the number of public repositories that can be maintained by a user on *GitHub*.

forked four times more than those of the random group. That means these are particularly successful as projects of interest to other *GitHub* users and as the basis for seeding new projects.[21]

**Table 7.1. Descriptive statistics – Repositories maintained by two samples of *GitHub* users, 2009-2013**

| | LeWeb Participants | | | | | Random Sample | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Mean** | **Std Dev** | **Med** | **Min** | **Max** | **Mean** | **Std Dev** | **Med** | **Min** | **Max** |
| **Reposit-ories** | 6.8 | 8.1 | 3.0 | 1 | 30 | 6.7 | 7.9 | 3.0 | 1 | 30 |
| **Size** | 56763 | 223462 | 4408 | 0 | 3015642 | 46105 | 104805 | 4112 | 0 | 964085 |
| **Watchers** | 29.2 | 129.0 | 1.0 | 0 | 1314 | 7.4 | 19 | 1.0 | 0 | 176 |
| **Forks** | 8.3 | 35.0 | 0.0 | 0 | 283 | 2 | 6.8 | 0.0 | 0 | 60 |

Figures 7.1 and 7.2 visualise the collaborators in and contributors to *GitHub* repositories maintained by each group. Blue squares represent repositories, red circles users. Green lines represent contributions to the repositories and black lines collaborations. *GitHub* contributors may submit issue reports, feature requests and offer code for use in a depository (a pull request). Collaborators have full read and write access to the repositories.

The *GitHub* network of *LeWeb* participants, illustrated in Figure 7.1, comprises a large number of isolated components around individual software projects (225). The largest component involves 154 users and projects, seven more contain more than 30 users and projects. As reported in Appendix Table A5, the collaborators are more dispersed than contributors, with a diameter of 3 steps as opposed to 1 and a slightly longer path distance; but the diameter of 1 indicating no connection between different projects.[22] But where collaboration occurs, it is concentrated with a higher average degree, higher degree centralization, betweenness centralization and transitivity. This reflects the weaker nature of contributive ties in comparison to collaborators. There is a much greater instance of reciprocal relationships and closed triads among collaborators than would be expected in a random graph.[23] This combination of intensive collaborative relationships, strong ties, with a wide range of weaker ties in the contributive

[21] While the two groups have distinctive collaborative patterns, this does not affect the robustness of the subsequent comparative analysis; this is simply a finding - a characteristic of the participants in *LeWeb* - with respect to GitHub users in general.

[22] The general form of these networks is two mode: Contributor A -> Project <- Contributor B. Simply counting the path distance from A to B without accounting for mode would equal 2. But since A and B are actually participating in the same activity, the Project, in one mode terms (contributor-contributor) the path is A-B, a distance of 1. Two-mode cohesion metrics account for this. A network diameter of 1 means the maximum path distance is no greater than Contributor A -> Project <- Contributor B, indicating that contributors make no connections between projects.

[23] Reciprocal relationships or mutual dyads among collaborators are formed where Collaborator A participates in a project of Collaborator B and Collaborator B participates part in a project of Collaborator A. Triadic closure or a complete subgraph occurs among collaborators where Collaborator A participates

**Figure 7.1. Collaborators and contributors to *GitHub* repositories 2009-2013 maintained by *LeWeb* participants 2009-2013.**

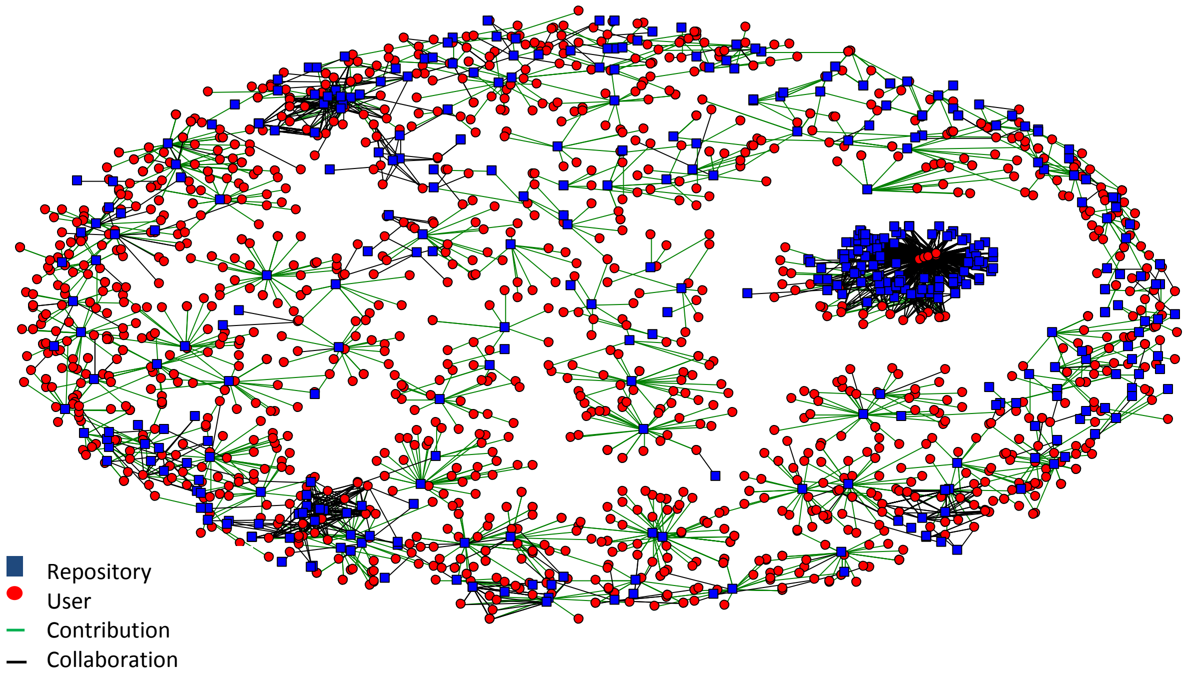
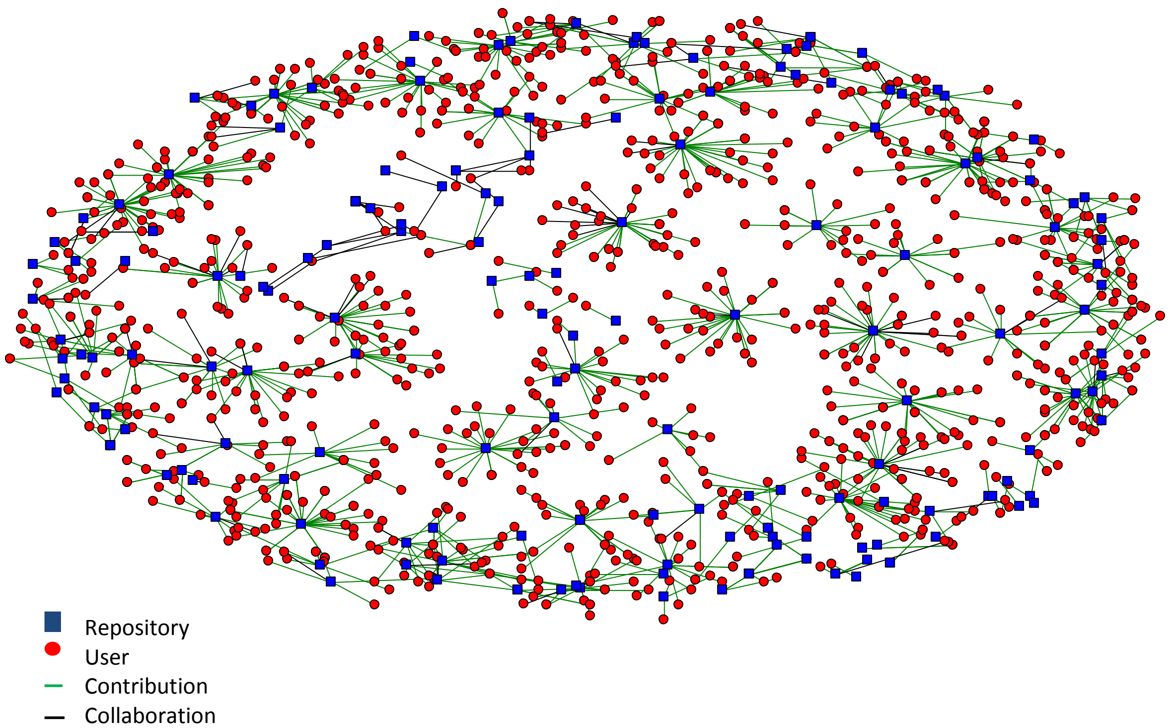



**Figure 7.2. Collaborators and contributors to *GitHub* repositories 2009-2013 maintained by randomly selected users.**




---

in a project of Collaborator B, Collaborator B participates in a project of Collaborator C and Collaborator C participates in a project of Collaborator A.

relationships is consistent with the exploitation/exploration combinations of dense regions of strong ties and sparse regions of weak ties very characteristic of high performing collaborations as discussed in the introduction.[24]

By contrast, the *GitHub* network of the random participants does not exhibit characteristics of a high performing collaborative network. This indicates that collaboration on *GitHub* is not uniformly high performing and in general is distinct from the collaboration patterns exhibited by *LeWeb* participants. As illustrated in Figure 7.2, the *GitHub* network of the random participants comprises 187 separate components, with only one containing more than 30 users and projects (31), a characteristic pattern of most networks; activity is much more intensely distributed among *LeWeb* participants than among *GitHub* participants in general. The random user network does not feature the large clusters of intense collaboration present in the networks of *LeWeb* participants. As reported in Appendix Table A5, the network actually appears more cohesive on most global indicators, a symptom of its smaller size, with lower levels of transitivity and clustering. These global metrics are not significantly different in conventional statistical terms, however. What is significant is a much lower incidence of reciprocated relationships and triadic closure in this network than in the one constituted by *LeWeb* participants and the presence of a small number of intensely collaborative components. Again, this indicates that collaboration in general on *GitHub* is not highly reciprocated and this and triadic closure is more evident in projects led by *LeWeb* participants.

So, the pattern of collaboration among *GitHub* developers who participate in *LeWeb* events is distinctive, with dense regions of strong ties combined with sparse regions of weak ties characteristic of high performing networks. There are eight highly interrelated clusters of users and projects not evident in the random sample of otherwise similarly active users.

We next examined the effect of participating in *LeWeb* events on *GitHub* activity. As *LeWeb* is a prominent event within the Web-tech sector, the developers who participate in these events might gain status among the open-source development community and consequently attract more contributions to their development projects. We undertook a randomised permutation regression analysis of the relationship between participation in *LeWeb* events and subsequent contributions to their projects on *GitHub* (the number of commits to repos owned by the participating developers). Two sets of models were estimated, the first examining total contributions (technically termed ’commits’) and the second examining commits per repo i.e. controlling for a repository owner hosting a number of projects. We controlled for general prominence in *LeWeb* by including centrality measures from repeated participation in all *LeWeb events* (discussed in Section 4).[25]

---

[24] This can be distinguished from “small worlds” which combine local clustering with long path distance on a single dimension, in this case collaboration or contribution. The average path distance in the *GitHub* collaborations is low (see Table A5).
[25] Appendix Tables A6 and A7 demonstrate little correlation between participation in *LeWeb’12* London and the core centrality metrics from all LeWeb events but moderate correlation (0.6) between participation in *LeWeb’13* Paris and these metrics, indicating potential issues with multicollinearity. Because there are no assumptions about the distribution of observations, however, randomised

The general form of the regression was:

$$\textit{GitHub}\ \text{Commits}_{i(t)} = \alpha + \beta_1\ \textit{LeWeb}\text{Attendance}_{i(t-1)} + \beta_2\ \text{All}\textit{LeWeb}\ \text{Centrality} + \varepsilon \qquad (4)$$

Tables 7.2 and 7.3 present the results of the regressions. In both sets of regressions, the independent variables are dummies indicating whether the user participated in *LeWeb* during the previous year LeWebAttendance (t-1), a dummy variable indicating whether the user is one of the core of multiple participants analysed in Section 4 and five different measures of the centrality of the user within that core.

Also in both sets of regressions, the model fit as a whole (indicated by R-square and subsequent metrics) is low and statistically insignificant, indicating that there are other variables important to *GitHub* contribution behaviour that are not considered. These are likely to include demographic homophily, collaborative dynamics including leader-follower behaviour,

**Table 7.2 Relationship between *LeWeb* participation and contributions to participants' *GitHub* projects**[26]

| | Commits 2010 | | Commits 2011 | | Commits 2012 | | Commits 2013 | |
|---|---|---|---|---|---|---|---|---|
| **LeWeb (t-1) Paris** | -0.91 | | 3.10 | | -8.71 | | 25.81 | ** |
| **LeWeb (t-1) London** | | | | | | | 28.62 | ** |
| **Participant Core** | -23737746 | | -33932272 | | -62810460 | | -21222786 | |
| **Participant Core Degree**[27] | -13540.63 | | -19331.82 | | -35706.72 | | -12115.69 | |
| **Participant Core Betweenness** | 2.19 | ** | 2.81 | | 1.99 | | 2.60 | |
| **Participant Core Closeness** | 4046.95 | ** | 5784.98 | ** | 10708.245 | ** | 3618.35 | |
| **Participant Core Eigenvector** | -11934.05 | | -4409.92 | | -17032.12 | | -22984.65 | |
| **Participant Core 2-Local Eigenvector** | 0.19 | ** | 0.07 | | 0.19 | | 0.36 | |
| **R-square** | 0.025 | | 0.012 | | 0.030 | | 0.031 | |
| **Adj. R-square** | -0.014 | | -0.028 | | -0.009 | | -0.013 | |
| **F** | 0.721 | | 0.330 | | 0.849 | | 0.777 | |
| ***p*** | 0.626 | | 0.861 | | 0.608 | | 0.676 | |
| **Observations** | 232 | | 232 | | 232 | | 232 | |
| **Permutations** | 10000 | | 10000 | | 10000 | | 10000 | |

** $p < 0.05$

---

permutation-based regression techniques are well-isolated from problems of multicollinearity and autocorrelation (Dekker, Krackhardt & Snijders, 2007)

[26] Unstandardised coefficients. In Model 4, an increase in *LeWeb'12* Paris of 1, that is participation in the event, is associated with a 26-fold increase in commits in 2013. While there is some mulitcollinearity among the independent variables, permutation-based regression techniques are robust in the face of this (Dekker, Krackhardt & Snijders, 2007).

[27] Degree centrality is used rather than the absolute number of attendances at *LeWeb* events as this metric is normalised for the differing number of participants each year.

reciprocity, core-periphery structures (see Conaldi & Lomi 2013). However, among the variables that are considered, there are some statistically significant effects, indicating an association between participation in *LeWeb* and *GitHub* activity, all else being equal.

In Table 7.2, the dependent variable in each of the models is the total number of software code contributions (commits) to a user's projects (repos) in four separate years. In all four years, there is a significant positive association between participation in *LeWeb* at some point and contributions to the user's *GitHub* projects, but this does not necessarily mean causation. In the first three models, the relationship is not with participation in specific *LeWeb* events but with centrality in the network of people who had been in multiple past *LeWeb* events. In particular, those users with the highest closeness centrality throughout the 2009-13 period had more software code contributions to their projects than others. 2010 contributions were also associated with betweenness and local eigenvector centrality of users participating in multiple events. In the most recent year the relationship was more specific; *GitHub* users who attended *LeWeb'12* London and Paris, regardless of their centrality in multiple events, had more software contributions to their projects than those who didn't.

The centrality variables are status indicators. Closeness and betweenness centrality reflect reach across the whole network of multiple participants across the period. Local eigenvector centrality reflects interactions with highly connected participants. So these users are prominent within the *LeWeb* community. Contributions of code to their projects is then related to their status and prominence within the web-tech community that *LeWeb* provides a focus for.

The specific relationship between participation in *LeWeb* and subsequent code contributions in 2013 is likely to be a product of immediacy overwhelming longer term prominence (indicated by the significant effect of centrality). The projects and ideas of developers involved in the most recent *LeWeb* events will be prominent among observers of the events, they will be talked about by other participants and so will attract attention of code contributors. And participation in the events is likely to stimulate new ideas and generation of novel projects that will also attract more contributions. But a year later, the next event will capture the immediate attention of participants and observers and the immediate effects of the previous year's events will be subsumed by longer term status and prominence. An effect of Paris 2009 attendance on 2010 thus won't be found as it gets swamped by the person attending the 2010 event.

A user with multiple projects, however, is likely to gain prominence within the *GitHub* network simply because there are more opportunities to contribute. To examine the impact of user prominence on code contributions, we undertook a second set of regressions, controlling for the number of projects maintained by a user. In Table 7.3, the dependent variable in each of the models is the total number of software code contributions (commits) to a user's projects (repos), divided by the number of projects each year.[28]

The results are broadly similar to those for total contributions. In all four years, there is a significant positive association between participation in *LeWeb* and contributions to the user's *GitHub* projects. In the first three years, those users closest to all others in multiple *LeWeb*

[28] An alternative approach would be to include the number of projects in the initial regression.

events received more code contributions than others, regardless of the number of projects they maintained. Betweenness centrality is significant in the first year as with total projects.[29] But, when controlling for the number of projects, eigenvector centrality is more prominently associated with code contributions in three of the four years, including the fourth. Finally, when controlling for the number of projects, participation in *LeWeb'13* Paris was no longer significant in subsequent code contributions, though participation in *LeWeb'12* London was, suggesting this particular event was of particular interest to larger developers. The particularity of *LeWeb'13* Paris was diluted by simply being a multiple participant (core and core degree centrality); these centrality metrics will be picking up effects of attendance before and after any particular year.[30]

**Table 7.3 Relationship between *LeWeb* participation and contributions per project to participants' *GitHub* projects[31]**

| | Commits / Repo 2010 | | Commits / Repo 2011 | | Commits / Repo 2012 | | Commits / Repo 2013 | |
|---|---|---|---|---|---|---|---|---|
| **LeWeb (t-1) Paris** | -0.68 | | -1.36 | | -1.86 | | 3.04 | |
| **LeWeb (t-1) London** | | | | | | | 1.12 | * |
| **Participant Core** | -3046279 | | -2134895 | | -8563614 | | 12775060 | ** |
| **Participant Core Degree** | -1739.29 | | -1216.73 | | -4875.91 | | 7280.34 | ** |
| **Participant Core Betweenness** | 0.36 | ** | 0.16 | | 0.36 | | -1.02 | |
| **Participant Core Closeness** | 519.34 | ** | 363.96 | | 1459.96 | ** | -2177.95 | |
| **Participant Core Eigenvector** | -1054.77 | | 294.734 | | -7386.90 | | 3610.99 | ** |
| **Participant Core 2-Local Eigenvector** | 0.02 | ** | -0.00 | | 0.118 | ** | -0.06 | |
| **R-square** | 0.056 | | 0.013 | | 0.031 | | 0.065 | |
| **Adj. R-square** | 0.017 | | -0.029 | | -0.010 | | 0.021 | |
| **F** | 1.600 | | 0.347 | | 0.842 | | 1.640 | |
| ***p*** | 0.395 | | 0.878 | | 0.618 | | 0.368 | |
| **Observations** | 232 | | 232 | | 232 | | 232 | |
| **Permutations** | 10000 | | 10000 | | 10000 | | 10000 | |

* $p < 0.10$ ** $p < 0.05$

[29] It is not unusual for Betweenness in particular to be prominent one year and not another in a sequence of cross-sectional network analyses. Betweenness centrality is particularly sensitive to the structure of the network as a whole and modest changes in network structure from year to year can have a large effect on the metric.

[30] It is possible that the appearance of the LeWeb'12 London effect in 2013 arises because the cumulative 3-core centrality measures are then in the past and so have less impact. A test for this would be to introduce a smaller moving window for the cumulative centrality measures.

[31] Unstandardised coefficients. In Model 4, an increase in *LeWeb'12* London of 1, that is participation in the event, is associated with a 12 per cent increase in commits in 2013.

Adjusting by the number of projects maintained by a user reinforces the status findings of the first set of regressions. Users who were central participants in multiple *LeWeb* events attracted more code contributions to their projects. This was generally both in terms of closeness, most similar patterns of participation to all other multiple participants, and eigenvector centrality, similar participation to the most connected participants. In the final year, participation in the latest Paris event, particularly by multiple participants, had a specific effect on the most recent year of code contributions.

In summary then, the *GitHub* projects of *LeWeb* participants exhibited higher performance than those of the random control group in terms of numbers of watchers and forks (the seeding of new projects). The pattern of collaboration in *GitHub* around the *LeWeb* participants' projects with dense regions of strong ties combined with sparse regions of weak ties was characteristic of high performing networks in other contexts. Central participation in multiple *LeWeb* events appears to be associated with the status of developers among the *GitHub* community as these users receive more code contributions to their projects than others. And there appears to be a more specific effect with prominent developers participating in the latest event gaining additional code contributions.

These results are consistent with hypothesis 4, that different patterns of collaboration in *GitHub* projects were associated with different patterns of participation in *LeWeb* events. They are consistent with hypothesis 6, with respect to the pattern of participation in *LeWeb* events, that these particular *GitHub* collaborative patterns were associated with higher project performance than other collaborative patterns. They are also consistent with hypothesis 7, that these *GitHub* collaborative patterns involved network structures typical of high-performance collaboration in other contexts. However, the results of section 5 do not support hypothesis 4, that participants in *LeWeb'13* Paris involved in *GitHub* gained more Twitter friends and followers than those involved in other patterns of collaboration.

In general, activity and performance in *GitHub* is associated with particular patterns of collaboration in *LeWeb*. But *Twitter* is not a significant channel in this relationship. Rather, we surmise that these results may be a symptom of exposure in *LeWeb*, in that these developers and their projects are more visible through this prominent activity in the web-tech sector. It may also be that participation in the *LeWeb* events is a source of novel projects that are intrinsically attractive to code contributors. And it may indicate a selection effect, that high status developers within the *GitHub* community participate more extensively in *LeWeb*.

# 8. Conclusions

This project extends earlier Nesta research on the effects of conferences and events. It confirms aspects of the earlier research in an important event for technological innovation in Europe and adds further insight into the complex interactions between in-person events and online collaboration. However, the restriction of the dataset to lists of participants in *LeWeb*, *GitHub* and their *Twitter*-based interactions limits the scope of the study. Much of the interaction in these forums, particularly the LeWeb events will be driven by live interpersonal dynamics not captured in this way. While data mining approaches such as those employed in this study can provide some insight, a comprehensive analysis of the interaction between online and offline dynamics around these events would involve interviews with the people involved.

Our follow-up study of the impact of the *LeWeb'13* Paris conference on online collaboration and the examination of the series of *LeWeb* conferences from 2009-13, adds some support to earlier findings that face-to-face events provide a catalyst for new and deeper connections that contribute to online collaboration. A core of *LeWeb* participants, around eight per cent of the total, attended multiple events. The central participants are predominantly males, located in France and associated with media organisations, corporations, and venture capitalists.

To investigate *Twitter*-based collaboration in a rigorous manner, we distinguished two comparator groups of *LeWeb* participants with similar characteristics, from those that had attended *LeWeb'12* London. The first, the participant group, were those *LeWeb'12* London attendees who went on to participate in *LeWeb'13* Paris. The second, the control group, were those *LeWeb'12* London attendees who did not participate in *LeWeb'13* Paris. This allowed us to consider the impact of the *LeWeb'13* Paris event in relation to online activity by similar people over the same period of time. It also provided continuity with the earlier Nesta study of the same group of *LeWeb'12* London attendees.

With respect to following activity around *LeWeb* events, we found no statistically significant difference among the participant group and control group in changes to the total number of followers or following on Twitter after *LeWeb'13* Paris; participating in *LeWeb'13* Paris did not affect the number of followers or number of Twitter users that attendees followed after the event. A more detailed examination of following among the specific participants of LeWeb'13 Paris also found no change associated with the event, however the sample size of this analysis was small (fewer than 20 people). This may result from the prominence of *LeWeb* participants themselves, with large numbers of friends and followers in a pattern typical of thought leaders; there may have been an effect but it would be too small to isolate relative to the median 1400-2400 contacts per user. It may also relate to the character of the participants, particularly prominent or multiple participants who may have established their following within the *LeWeb* community at an earlier point.

*LeWeb'12* London attendees classified as entrepreneurs gained significantly more *Twitter* followers during the period than any other identified subgroup. And *LeWeb'12* London attendees classified as developers increased the number of *Twitter* users during the period they followed more than any other identified subgroup. We surmised that entrepreneurs are likely to use *LeWeb* as one of many channels for publicising their activities in a general search for weak ties (followers) that may be beneficial in some combination in the future. By contrast, those participants classified distinctly as developers, distinct from those involved in large corporations, are typically working for smaller start-ups, are likely to be more engaged in specific projects and are more purposeful in their search, seeking stronger ties and add users to their following lists that may be more immediately beneficial.

Participation in *LeWeb'13* Paris did influence tweeting behaviour, however. Participants in *LeWeb'13* Paris *mentioned*, replied to, or retweeted *LeWeb'13* Paris attendees in tweets 13 per cent more often in the subsequent six weeks than the control group itself did. They were more likely to mention, reply to or retweet those at the centre of the whole tweeting network, those within 2 steps of themselves (2-step centrality) and those who had mentioned, replied to or retweeted them (reciprocation).

Our study of collaborative relationships among software developers involved in *LeWeb* and *GitHub* highlighted distinctive patterns of collaboration, with dense regions of strong ties combined with sparse regions of weak ties characteristic of high performing networks, compared to a randomly selected control group of individuals on *GitHub*. Their projects had four times as many watchers and were used to seed new projects by other users four times more often than those of the control group.

*GitHub* developers who were central participants in multiple *LeWeb* events consistently attracted much greater code contributions to their projects than *GitHub* users in the control group. In the final year of code contributions examined, 2013, contributions were 112 per cent greater to the projects of those developers who had participated in the *LeWeb'12* London event than to those in the *GitHub* control group, all else being equal

Thus, activity and performance in *GitHub* is associated with particular patterns of collaboration in *LeWeb*. But *Twitter* is not a significant channel in this relationship; *GitHub* users participating in *LeWeb'13* Paris did not gain more followers or tweet more than other participants. We surmise that the positive relationship between *LeWeb* participation and distinctive collaboration, project performance and greater open source contributions on *GitHub* may be a symptom of exposure, in that these developers and their projects are more visible through this prominent activity in the web-tech sector. It may also be that participation in the *LeWeb* events is a source of novel projects that are intrinsically attractive to code contributors. And this may indicate a selection effect, that high status developers within the *GitHub* community participate more extensively in *LeWeb* and commercially successful developers are more likely to be able to afford the relatively high registration fees.

With regard to the specific hypotheses proposed at the start of the study, as summarised in Table 8.1, we found support for our expectations in terms of tweeting associated with event participation, distinct subgroup behaviours but not following activity. Participation in the *LeWeb* face-to-face events had a demonstrable impact on subsequent online communications and collaboration. This was particularly evident for those participants who were more central to the tweeting network as a whole and among close contacts. Participation in the face-to-face event established the basis for reciprocated interactions after the event.

Participants in multiple *LeWeb* events who we observed on *GitHub* also had distinctive forms of online collaboration. Those most central in the network of people who attended *LeWeb* were engaged in network structures conducive to high performance and exhibited higher levels of performance and greater contributions from other *GitHub* users than developers not engaged in *LeWeb*.

**Table 8.1 Summary of findings with regard to hypotheses**

| Hypothesis | Findings |
|---|---|
| *Hyp#1 – Participants in LeWeb'13 Paris gained more Twitter followers and increased the number of users followed more than non-participants with otherwise similar characteristics.* | Not supported (sections 5 & Appendix B). |
| *Hyp#2 – Participation in LeWeb'13 Paris increased tweeting among participants.* | Supported (section 6). |
| *Hyp#3 – The extra connections gained at LeWeb'13 Paris varied among distinct subgroups by organisation type, nationality and participation in multiple events.* | Some support (sections 5 & 6). |
| *Hyp#4 – Different patterns of collaboration in GitHub projects were associated with different patterns of participation in LeWeb events.* | Supported with respect to LeWeb participation (section 7). |
| *Hyp#5 – Participants in LeWeb'13 Paris involved in GitHub gained more Twitter friends and followers than those involved in other patterns of collaboration.* | Not supported (section 5 and section 6). |
| *Hyp#6 – These particular GitHub collaborative patterns were associated with higher project performance than other collaborative patterns.* | Supported (section 7). |
| *Hyp#7 – These GitHub collaborative patterns involved network structures typical of high-performance collaboration in other contexts.* | Supported (section 7). |

Participation in the *LeWeb* events provided opportunities for exposure to new ideas and the basis for novel projects as well as opportunities to prominently highlight existing projects, which may have stimulated online collaboration. There may also be selection effects, with more successful online developers more likely to engage in the face-to-face events. But the specific association between participation in *LeWeb'13* Paris and the collaboration indicators subsequently points to the stimulus of the face-to-face event to online collaboration.

The broad relationship between participation in face-to-face events and subsequent online communication and collaboration identified in this study of *LeWeb* is likely to be found in conference events in general. Conference events normally bring together a variety of participants and interests around a broadly common theme. They thus provide a source of novelty in personal interaction and ideas. With increasing use of social media, it is likely that the interactions forged in the face-to-face event will be translated into interactions in social media as well as other forms of collaboration, while some event interactions may occur solely through social media. The particular relationships identified in this study, increased *Twitter* interaction, particularly involving core participants, close contacts and reciprocation are likely to be found in other conference events.

More comprehensive open online collaboration is currently largely restricted to software development and, to a limited extent, *Wikipedia* but has parallels in corporate use of online collaborative systems such as *Sharepoint*, *Jive*, *Lithium* and *Yammer.* So there are likely to be similar relationships between face-to-face events and online collaboration within corporate settings as found here. There may be parallels between the relationships found here and semi-open online collaborative platforms such as *Facebook*, *LinkedIn* and *Google+*. But since these platforms are generally more diffuse and less task-focussed than the software development and corporate contexts, any relationships are likely to be much weaker.

The particular context of *LeWeb* is likely to reduce the ability to generalise the findings of this study to smaller events. *LeWeb* is a prominent annual meeting of major figures in the European web-tech community and the participants are often thought leaders or community stars, as evidenced by the high average *Twitter* follower/followed ratios. This is likely to give central participants greater reach among both event attendees and their broader collaborative communities. As a meeting place of participants from different types of organisations (corporations, media, venture capitalists, entrepreneurs, developers etc.) there is a brokering dimension to *LeWeb* events, evidenced by the differing follower and following behaviour of entrepreneurs and developers, than might be found in a conference of a particular profession, for example.

This study adds to the general understanding of online/offline interactions pioneered by Haythornthwaite (2005). It demonstrates the value of the live continuous streams of data available from sources such as *Twitter* and *GitHub*. The challenges of managing these large datastreams are considerable, however. Comprehensive, freely available *Twitter* data can only be accessed via live daily download and aggregation; historic data is only available on a limited basis at high cost. This demands experimentation with research design. While the amount of data available is unprecedented for such studies, scale is a major issue; both *Twitter* and *GitHub* impose download limits, which require parallel processing during data collection and analysis and careful planning to navigate and manage. Data cleaning for the scale of data collected for this project involves extensive programming. For the more complex social network analytic techniques used in this study, scale also provides computational challenges as many of the analytical techniques require exponentially greater computation as the number of people included in the analysis increases.

The current Nesta research programme would certainly benefit from further replication and extension of the methods to future *LeWeb* events. With a longer timescale before the event there would be greater scope for sensitivity testing of sampling periods and a greater range of micro-sociological interactions could be considered with the SAOM modelling. It would be useful to sample following and follower behaviour on a daily basis rather than at six week intervals, to analyse the smaller-scale dynamics undetectable in aggregate. There would be value in developing further randomised permutation-test regression models to account for the evident skewed nature of the data and the event-history context. As there seems to be considerable seasonality in the data, data collection over more extended periods and comparison with *Twitter* behaviour in other contexts would be valuable to determine longer baseline trends with which to compare short term changes in activity. And of course the methods used here would be usefully employed in other event contexts to provide further comparative insight into the relationship between events and innovation.

# 9. References


Ahuja, G. (2000). Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly*, *45*(3), 425–455. doi:10.2307/2667105

Bakhshi, H., Davies, J. & Mateos-Garcia, J. (2013). Joining the Dots: Can events grow networks of innovators? http://www.nesta.org.uk/blogs/Economic%20growth/joining_the_dots_can_events_grow_networks_of_innovators

Borgatti, S.P., Everett, M.G. & Freeman, L.C. (2002). UCINET for Windows: Software for social network analysis. Harvard, MA: Analytic Technologies.

Borgatti, S. P., Everett, M. G. & Johnson, J. C. (2013). *Analysing social networks*. London: Sage.

Burt, R. S. (2005). *Brokerage and closure*. New York: Oxford University Press.

Conaldi, G., & De Vita, R. (2011). At the core of innovation: network reconfiguration during radical and incremental innovation episodes in an open source software project. *International Journal of Information and Decision Sciences*, *3*(1), 5–25. doi:10.1504/IJIDS.2011.038838

Conaldi, G., & Lomi, A. (2013). The dual network structure of organizational problem solving: a case study on open source software development. *Social Networks*, *35*(2), 237–250. doi:10.1016/j.socnet.2012.12.003

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*(Supplement), S95-S120.

Costenbader, E. & Valente, T. (2003). The stability of centrality measures when networks are sampled. *Social Networks*, 25(2), 283–307

Cronin, B. (2007). Managing collaborative networks. In S. Al-Hawamdeh (Ed.), *Creating Collaborative Advantage Through Knowledge and Innovation* (pp. 135–153). Singapore: World Scientific.

Dekker, D., Krackhardt, D., & Snijders, T. A. B. (2007). Sensitivity of MRQAP Tests to Collinearity and Autocorrelation Conditions. *Psychometrika*, 74(2), 563–581.

Ebner, M., Mühlburger, H., Schaffert, S., Schiefner, M., Reinhardt, W. & Wheeler, S. (2010). Getting granular on Twitter: Tweets from a conference and their limited usefulness for non-participants. In N. Reynolds & M. Turcsányi-Szabó (Eds.), *Key Competencies in the Knowledge Society* (pp. 102–113). Heidelberg: Springer Berlin. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-15378-5_10

Fleming, L., King, C. & Juda, A. I. (2007). Small worlds and regional innovation. *Organization Science*, *18*(6), 938–954. doi:10.1287/orsc.1070.0289

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology, 78*(6), 1360-1380.

Haythornthwaite, C. (2005). Social networks and Internet connectivity effects. *Information, Communication & Society*, *8*(2), 125–147. doi:10.1080/13691180500146185

Honey, C. & Herring, S. C. (2009). Beyond microblogging: conversation and collaboration via Twitter. In *42nd Hawaii International Conference on System Sciences, 2009. HICSS '09* (pp. 1–10). doi:10.1109/HICSS.2009.89

Java, A., Song, X., Finin, T. & Tseng, B. (2007). Why we Twitter: Understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis* (pp. 56–65). New York, NY, USA: ACM. doi:10.1145/1348549.1348556

Kieslinger, B., Ebner, M. & Wiesenhofer, H. (2011). Microblogging practices of scientists in e-learning: A qualitative approach. *International Journal of Emerging Technologies in Learning*, *6*(4), 31–39.

Letierce, J., Passant, A., Breslin, J. & Decker, S. (2010a). Understanding how Twitter is used to spread scientific messages. Presented at the WebSci10: Extending the Frontiers of Society On-Line, Raleigh, NC: US. Retrieved from http://journal.webscience.org/314/

Letierce, J., Passant, A., Breslin, J. G. & Decker, S. (2010b). Using Twitter during an academic conference: The# iswc2009 use-case. In *ICWSM*. Retrieved from http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewPDFInterstitial/1523/1877

McFadden, D. L. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105-142). New York: Academic Press.

McKelvey, M. (2001). The economic dynamics of software: Three competing business models exemplified through Microsoft, Netscape And Linux. *Economics of Innovation and New Technology*, *10*(2-3), 199–236. doi:10.1080/10438590100000009

Reinhardt, W., Ebner, M., Beham, G. & Costa, C. (2009). How people are using Twitter during conferences. In V. Hornung-Prähauser & M. Luckmann (Eds.), *Creativity and innovation competencies on the Web* (pp. 145–156). BoD – Books on Demand.

Sopan, A., Rey, P. J., Butler, B. & Shneiderman, B. (2012). Monitoring academic conferences: Real-time visualization and retrospective analysis of backchannel conversations. In *2012 International Conference on Social Informatics (Social Informatics)* (pp. 62–69). doi:10.1109/SocialInformatics.2012.20

Snijders, T.A.B., 2001. The statistical evaluation of social network dynamics. *Sociological Methodology, 31*(1): 361–395.

Snijders, T.A.B., van de Bunt, G.G. & Steglich, C.E.G. (2010) Introduction to stochastic actor-based models for network dynamics, *Social Networks*, *32*(1): 44-60.

Twitter, Inc. (2013). *Annual Report, 2013*. Twitter, Inc. San Francisco, CA.

Uzzi, B. & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology*, *111*(2), 447–504. doi:10.1086/432782

White, D. & Harary, F. (2001). The cohesiveness of blocks in social networks: Node connectivity and conditional density.  *Sociological Methodology, 31* (1): 305–359

Zhao, D. & Rosson, M. B. (2009). How and why people Twitter: The role that micro-blogging plays in informal communication at work. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work* (pp. 243–252). New York,NY: ACM. doi:10.1145/1531674.1531710

## Appendix A. Supplementary Tables and Figures

**Table A1. Structural characteristics of tweeting networks**[32]

| | London 2012 Period 1 | London 2012 Period 2 | Paris 2013 Period 1 | Paris 2013 Period 2 |
|---|---|---|---|---|
| Vertices (N) | 302 | 355 | 378 | 490 |
| Arcs | 908 | 3205 | 1113 | 4034 |
| Loops | 323 | 1379 | 202 | 1135 |
| Multiple lines | 435 | 2212 | 489 | 2591 |
| Density (loops allowed) | 0.010 | 0.025 | 0.008 | 0.017 |
| Average Degree | 6.013 | 18.056 | 5.889 | 16.465 |
| Arcs after Loops, Multiples | 329 | 754 | 517 | 1261 |
| Lines value 1 | 206 | 401 | 315 | 680 |
| Lines value >1 | 123 | 353 | 202 | 581 |
| Density1 [loops allowed] | 0.004 | 0.006 | 0.004 | 0.005 |
| Density2 [no loops allowed | 0.004 | 0.006 | 0.004 | 0.005 |
| Average Degree | 2.179 | 4.248 | 2.735 | 5.147 |
| All Degree Centralization | 8.467 | 31.554 | 14.207 | 23.523 |
| Clustering Coefficient | 0.080 | 0.167 | 0.101 | 0.137 |
| Transitivity: | 0.048 | 0.088 | 0.079 | 0.072 |
| Number of unreachable pairs | 84529 | 85844 | 119211 | 125123 |
| Average distance among reachable pairs | 5.206 | 4.010 | 6.824 | 4.505 |
| Max distance | 13 | 9 | 20 | 11 |
| Number of components | 28 | 16 | 40 | 17 |
| Vertices in largest component N | 159 | 245 | 259 | 425 |
| Vertices in largest component % | 0.526 | 0.690 | 0.685 | 0.867 |
| Louvain Clusters N | 106 | 102 | 90 | 65 |
| Louvain Clusters Modularity | 0.833 | 0.663 | 0.854 | 0.698 |
| n ln n | 1724.55 | 2084.60 | 2243.39 | 3035.26 |
| Total adjacency index m | 329 | 754 | 517 | 1261 |
| Zagreb group index M1: | 3716 | 22646 | 6042 | 35996 |
| Zagreb group index M2 | 9621 | 168548 | 17582 | 257696 |
| Randic index Xr | 99.093 | 112.673 | 152.215 | 180.822 |
| Platt index F | 3058 | 21138 | 5008 | 33474 |
| T-test independence of variables[33] | | | .118 | .118 |

[32] Calculated with Pajek 3.12.
[33] 2-tailed paired t-test of statistical independence of items Average degree to Platt index.

**Table A2. Triadic census of tweeting networks**[34]

| | London 2012 Period 1 | | London 2012 Period 2 | | Paris 2013 Period 1 | | Paris 2013 Period 2 | |
|---|---|---|---|---|---|---|---|---|
| **Balance:** | o | (o-e)/e | o | (o-e)/e | o | (o-e)/e | o | (o-e)/e |
| **3 - 102** | 26192 | 147.780 | 74348 | 94.380 | 49898 | 142.580 | 171830 | 107.380 |
| **16 - 300** | 1 | 97886273 | 35 | 101478627 | 5 | 245559024 | 44 | 106271859 |
| **Clusterability:** | | | | | | | | |
| **1 - 003** | 4473791 | 0 | 7210396 | 0.010 | 8787828 | 0 | 19058494 | 0 |
| **Ranked Clusters:** | | | | | | | | |
| **4 - 021D** | 63 | -0.640 | 779.470 | -0.610 | 117 | -0.660 | 803 | -0.490 |
| **5 - 021U** | 93 | -0.470 | 779.470 | -0.460 | 98 | -0.720 | 610 | -0.620 |
| **9 - 030T** | 6 | 3.690 | 9.410 | 0.590 | 7 | 1.770 | 36 | 1.150 |
| **12 - 120D** | 0 | -1 | 0.030 | 1055.360 | 1 | 216.040 | 30 | 675.030 |
| **13 - 120U** | 4 | 1721.090 | 0.030 | 597.610 | 21 | 4556.800 | 34 | 765.170 |
| **Transitivity: 2 - 012** | 44443 | -0.540 | 104417 | -0.600 | 91608 | -0.520 | 250740 | -0.580 |
| **Hierarchical Clusters:** | | | | | | | | |
| **14 - 120C** | 2 | 429.520 | 0.060 | 439.150 | 0 | -1 | 23 | 258.150 |
| **15 - 210** | 4 | 237042 | 58 | 169174 | 7 | 208625 | 76 | 161855 |
| **No Model:** | | | | | | | | |
| **6 - 021C** | 121 | -0.660 | 1558.95 | -0.580 | 147 | -0.790 | 1068 | -0.660 |
| **7 - 111D** | 137 | 106.120 | 9.410 | 107.290 | 242 | 94.620 | 1339 | 78.820 |
| **8 - 111U** | 132 | 102.210 | 9.410 | 103.570 | 223 | 87.110 | 1720 | 101.530 |
| **10 - 030C** | 0 | -1 | 3.140 | -0.040 | 1 | 0.190 | 3 | -0.460 |
| **11 - 201** | 111 | 47786 | 861 | 30317 | 173 | 37546 | 1430 | 32223 |
| **T-test of London / Paris independence** | | | | | | 0.333 | | 0.334 |

[34] Sample size as for Table A1. Calculated with Pajek 3.12. See Table A8 for codes.

**Table A3. *LeWeb'12* London participants repeated tweeting network - cohesiveness**[35]

| | Before Paris 2013 | After Paris 2013 | Change |
|---|---|---|---|
| **Average Degree** | 0.955823 | 2.032129 | 113% |
| **H-Index** | 6 | 9 | 50% |
| **Density** | 0.003854 | 0.008194 | 113% |
| **Components** | 181 | 124 | -31% |
| **Component Ratio** | 0.725806 | 0.495968 | -32% |
| **Connectedness** | 0.06654 | 0.372425 | 460% |
| **Fragmentation** | 0.93346 | 0.627575 | -33% |
| **Closure** | 0.051756 | 0.077816 | 50% |
| **Average Distance** | 4.631054 | 3.83442 | -17% |
| **Std. Dev Distance** | 2.117771 | 1.341788 | -37% |
| **Diameter** | 11 | 10 | -9% |
| **Breadth** | 0.980536 | 0.887049 | -10% |
| **Compactness** | 0.019464 | 0.112951 | 480% |
| **Betweenness Centralization** | .0305 | 0.1433 | 370% |
| **Overall clustering coefficient** | 0.108 | 0.161 | 49% |
| **Weighted overall clustering coefficient** | 0.074 | 0.065 | -12% |

[35] Calculated within components with UCINET 6.474.

Table A4. *LeWeb'13* Paris participants repeated tweeting network - cohesiveness[36]

| | Before Paris 2013 | After Paris 2013 | Change | Change +/- London 2012 Participants |
|---|---|---|---|---|
| **Average Degree** | 0.923214 | 2.251786 | 144% | 31% |
| **H-Index** | 6 | 13 | 117% | 67% |
| **Density** | 0.001652 | 0.004028 | 144% | 31% |
| **Components** | 421 | 278 | -34% | -2% |
| **Component Ratio** | 0.751342 | 0.495528 | -34% | -2% |
| **Connectedness** | 0.074415 | 0.365726 | 391% | -68% |
| **Fragmentation** | 0.925585 | 0.634274 | -31% | 1% |
| **Closure** | 0.095327 | 0.087085 | -9% | -59% |
| **Average Distance** | 6.823825 | 4.505245 | -34% | -17% |
| **St. Dev Distance** | 3.111072 | 1.458426 | -53% | -16% |
| **Diameter** | 20 | 11 | -45% | -36% |
| **Breadth** | 0.985396 | 0.907455 | -8% | 2% |
| **Compactness** | 0.014604 | 0.092545 | 534% | 53% |
| **Betweenness Centralization** | 0.0394 | 0.0622 | 58% | -312% |
| **Overall clustering coefficient** | 0.101 | 0.137 | 36% | -13% |
| **Weighted overall clustering coefficient** | 0.079 | 0.072 | -9% | 3% |
| **T-test of independence from *LeWeb'12* London (Table A3)** | 0.309 | 0.314 | | |

[36] Calculated within components with UCINET 6.474.

Table A5. *GitHub* repo collaborators and contributors network - cohesiveness[37]

| | LeWeb Participants - collaborators | LeWeb Participants - contributors | Random Sample - collaborators | Random Sample - contributors |
|---|---|---|---|---|
| **N** | 408 | 408 | 408 | 408 |
| **Average Degree** | 13.857 | 13.359 | 14.315 | 13.213 |
| **H-Index** | 29.000 | 29.000 | 29.000 | 29.000 |
| **Density** | 0.009 | 0.009 | 0.011 | 0.011 |
| **Components** | 295 | 346 | 234 | 279 |
| **Component Ratio** | 0.193 | 0.226 | 0.187 | 0.223 |
| **Connectedness** | 0.009 | 0.009 | 0.011 | 0.011 |
| **Fragmentation** | 0.991 | 0.991 | 0.989 | 0.989 |
| **Closure** | 0.990 | 1.000 | 1.000 | 1.000 |
| **Average Distance** | 1.037 | 1.000 | 1.000 | 1.000 |
| **Std. Dev Distance** | 0.228 | 0.000 | 0.000 | 0.000 |
| **Diameter** | 3.000 | 1.000 | 1.000 | 1.000 |
| **Breadth** | 0.991 | 0.991 | 0.989 | 0.989 |
| **Compactness** | 0.009 | 0.009 | 0.011 | 0.011 |
| **Degree Centralization %** | 0.215 | 0.146 | 0.871 | 0.133 |
| **Betweenness Centralization %** | 0.010 | 0.000 | 0.000 | 0.000 |
| **Overall clustering coefficient** | 6.852 | 0.000 | 1.449 | 0.000 |
| **Weighted overall clustering coefficient (transitivity)** | 3.500 | 1.0000 | 1.254 | 0.000 |
| **Triadic Census (ni-ei)/ei** | | | | |
| **Balance** | | | | |
| **3 -102 (mutual dyad)** | 102.59 | 52.33 | 0.86 | 2.00 |
| **16 – 300 (complete subgraph)** | 141231320 | 8514207 | | |
| **Clusterability** | | | | |
| **1 – 003 (unconnected)** | 0.03 | 0.06 | 0.17 | 0.17 |
| **Other** | | | | |
| **11 – 201 (2 intransitive mutual dyads)** | 143.73 | | 11.95 | 33.95 |
| **T-test of independence – cohesion metrics** | | | 0.27 | 0.32 |
| **T-test of independence – triadic census** | | | 0.19 | 0.50 |

[37] 2-mode cohesion metrics calculated in UCINET 6.474. Lengths of geodesics computed within components. Triad census of 1 mode data (actors) calculated in Pajek 3.12. See Table A8 for codes.

**Table A6. Correlation matrix – *LeWeb'12* London participation and core centrality metrics**

| | Participant 3-core Degree | Participant 3-core Betweenness | Participant 3-core Closeness | Participant 3-core Harmonic Closeness | Participant 3-core Eigenvector | LeWeb 2012 London |
|---|---|---|---|---|---|---|
| **Participant 3core Degree** | 1.000 | 0.918 | 0.947 | 0.958 | 0.992 | 0.097 |
| **Participant 3core Betweenness** | 0.918 | 1.000 | 0.749 | 0.773 | 0.897 | 0.132 |
| **Participant 3core Closeness** | 0.947 | 0.749 | 1.000 | 0.997 | 0.948 | 0.063 |
| **Participant 3core HarmonicCloseness** | 0.958 | 0.773 | 0.997 | 1.000 | 0.963 | 0.057 |
| **Participant 3core Eigenvector** | 0.992 | 0.897 | 0.948 | 0.963 | 1.000 | 0.057 |
| **LeWeb 2012 London** | 0.097 | 0.132 | 0.063 | 0.057 | 0.057 | 1.000 |

**Table A7. Correlation matrix – *LeWeb'13* Paris participation and core centrality metrics**

| | Participant 3-core Degree | Participant 3-core Betweenness | Participant 3-core Closeness | Participant 3-core Harmonic Closeness | Participant 3-core Eigenvector | LeWeb Paris 2013 |
|---|---|---|---|---|---|---|
| **Participant 3core Degree** | 1.000 | 0.918 | 0.947 | 0.958 | 0.992 | 0.596 |
| **Participant 3core Betweenness** | 0.918 | 1.000 | 0.749 | 0.773 | 0.897 | 0.504 |
| **Participant 3core Closeness** | 0.947 | 0.749 | 1.000 | 0.997 | 0.948 | 0.598 |
| **Participant 3core HarmonicCloseness** | 0.958 | 0.773 | 0.997 | 1.000 | 0.963 | 0.607 |
| **Participant 3core Eigenvector** | 0.992 | 0.897 | 0.948 | 0.963 | 1.000 | 0.612 |
| **LeWeb Paris 2013** | 0.596 | 0.504 | 0.598 | 0.607 | 0.612 | 1.000 |

**Table A8. Coding of triadic census**

Triads are named as (number of pairs that are mutually tied) (number of pairs that are one-way tied) (number of non-tied pairs) in the triple. There are 16 possible combinations:

| Number | Configuration | Notes |
|---|---|---|
| **1** | 003 | The empty triad |
| **2** | 012 | |
| **3** | 102 | |
| **4** | 021D | "Down": the directed edges point away |
| **5** | 021U | "Up": the directed edges meet |
| **6** | 021C | "Circle": one in, one out |
| **7** | 111D | "Down": 021D but one edge is mutual |
| **8** | 111U | "Up": 021U but one edge is mutual |
| **9** | 030T | "Transitive": two point to the same vertex |
| **10** | 030C | "Circle": A->B->C->A |
| **11** | 201 | |
| **12** | 120D | "Down": 021D but the third edge is mutual |
| **13** | 120U | "Up": 021U but the third edge is mutual |
| **14** | 120C | "Circle": 021C but the third edge is mutual |
| **15** | 210 | |
| **16** | 300 | The complete |

**Figure A1. Distribution of tweets by day 1/11/13 – 31/1/14 – *LeWeb'12* London and *LeWeb'13* Paris participants – and sampling periods.**

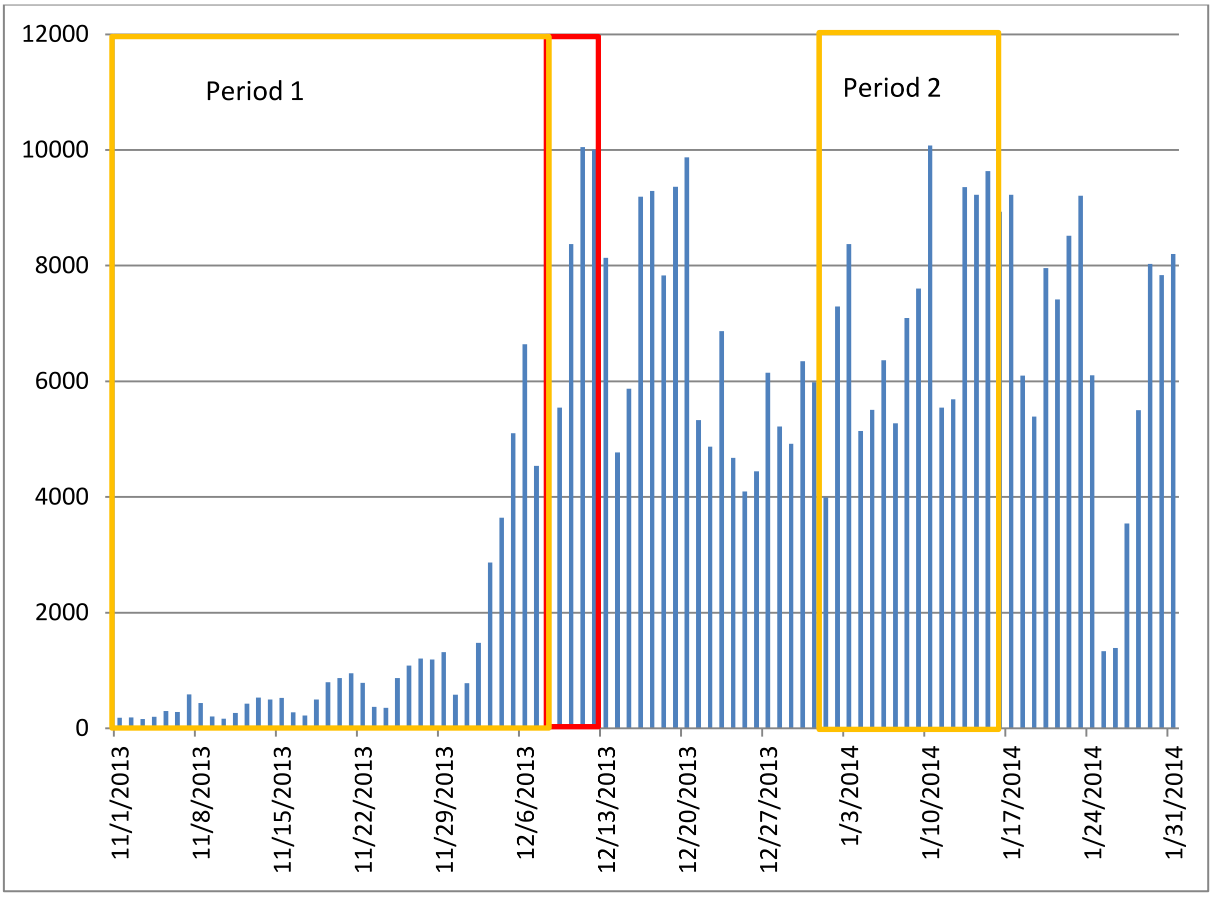
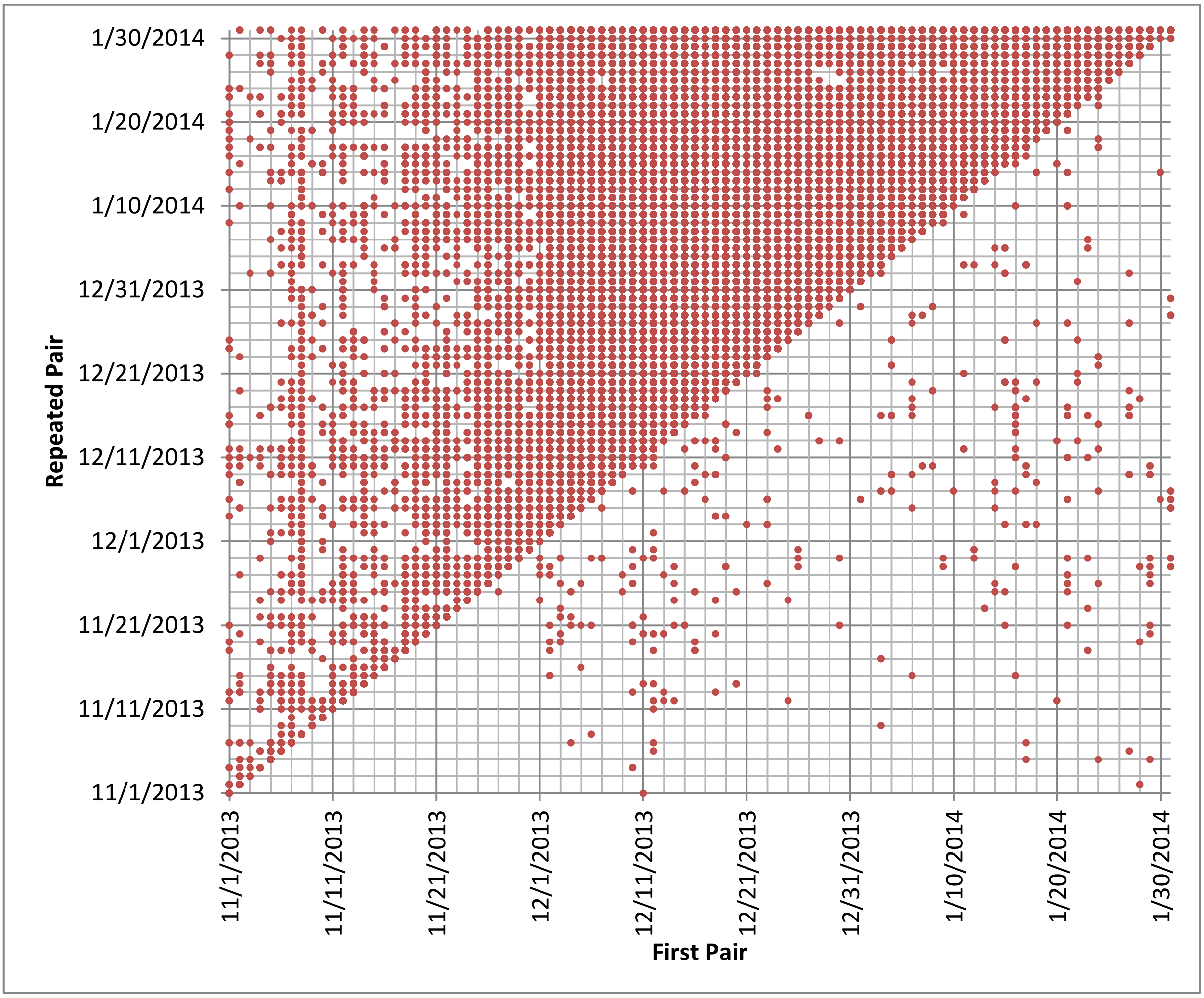


**Figure A2. Distribution of repeated tweeting pairs by day**




The x-axis records the first occasion during the observations that the first member of a repeating pair first mentions, replies to or retweets the other member of the pair. The y-axis records each subsequent occasion that the first member of the pair mentions, replies to or retweets the other. Data points below the diagonal are where the second member of the pair mentions, replies to or retweets the first, after the first has done so.

# Appendix B. *Twitter* Following among Event Participants

A central hypothesis of the study (hypothesis 1) is that participation in in-person events is likely to increase following activity on *Twitter* because participants are likely to encounter new contacts and ideas and seek to maintain contact with these subsequently, following *Twitter* users being a low-cost mechanism for this.

To test this hypothesis, we collected follower lists from participants registered to attend *LeWeb'13* Paris before and after the event. Follower lists from registrants were downloaded on 5/12/13, three days before the event started, and again on 25/1/14, seven weeks after the event. The structures of the follower/followed networks on these two dates were compared to determine the extent to which the structures had changed subsequent to the in-person event.

We had planned to analyse the follower networks of all participants in *LeWeb'13* Paris and *LeWeb'12* London but we encountered a variety of technical difficulties collecting the data before the event due to set restrictions on the Twitter API. In the end, we collected complete follower lists for 18 *LeWeb'13* Paris registrants on 5/12/13 and follower lists for the same 18 registrants on 25/1/14 but none from *LeWeb'12* London. The selection of *LeWeb'13* Paris registrants was determined by various rate limits in the Twitter API, but those successfully downloaded were approximately every 50$^{th}$ user searched for so the sample approximates a random selection. However, as indicated by Table B1, the sample tended to comprise more multiple *LeWeb* participants than *LeWeb'13* Paris participants in general and were disproportionately drawn from Media, Marketing and Government.

**Table B1. Sample characteristics**

| | All Paris 2013 Twitter IDs | Sample |
|---|---|---|
| **n** | 1876 | 18 |
| **Mean *LeWeb* Events Attended** | 2.00 | 4.12 |
| **SD *LeWeb* Events Attended** | 1.34 | 2.03 |
| **Media** | 15% | 22% |
| **Corporation** | 20% | 22% |
| **Venture Capitalist** | 5% | 6% |
| **Developer** | 16% | 12% |
| **Government** | 2% | 6% |
| **Marketing** | 5% | 12% |

The Twitter users on the follower list for each of *the LeWeb'13* Paris participants in the sample as of 5/12/13 were compared to the *LeWeb'13* Paris participants list and matches were identified. The resulting network of followers among *LeWeb'13* Paris participants was mapped. This was repeated for 25/1/14.

The follower/following network as of 5/12/13 is presented in Figures B1 and B2. The two figures are identical other than the size of nodes, the arrows representing a user following a user pointed at. In Figure B1 node size represents the number of followers; the network is

**Figure B1. Follower/following network among *LeWeb'13* Paris participants 5/12/13 – Node size scaled by indegree**[38]

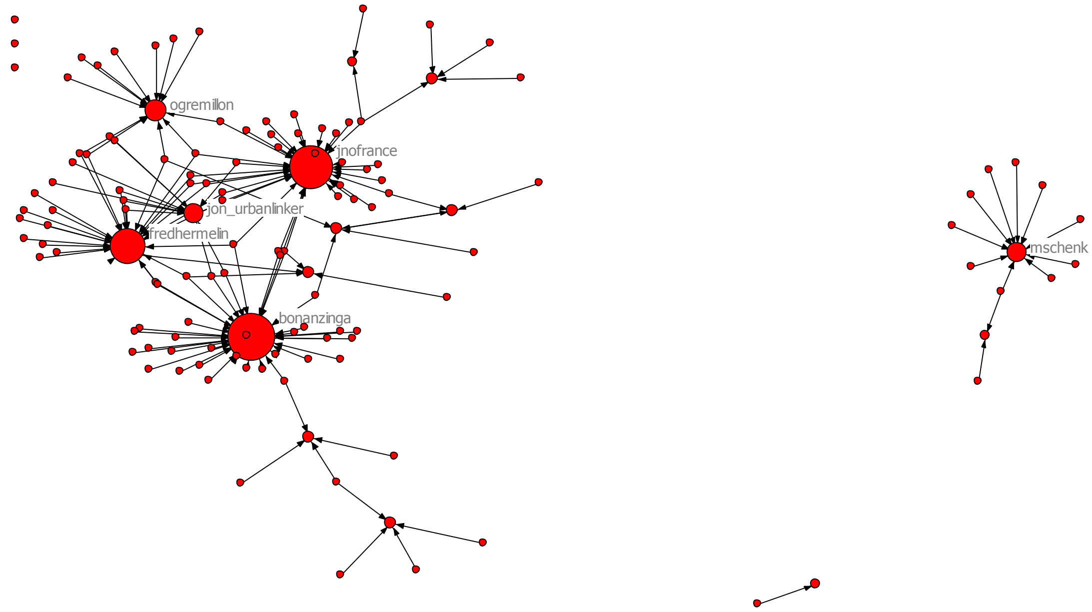
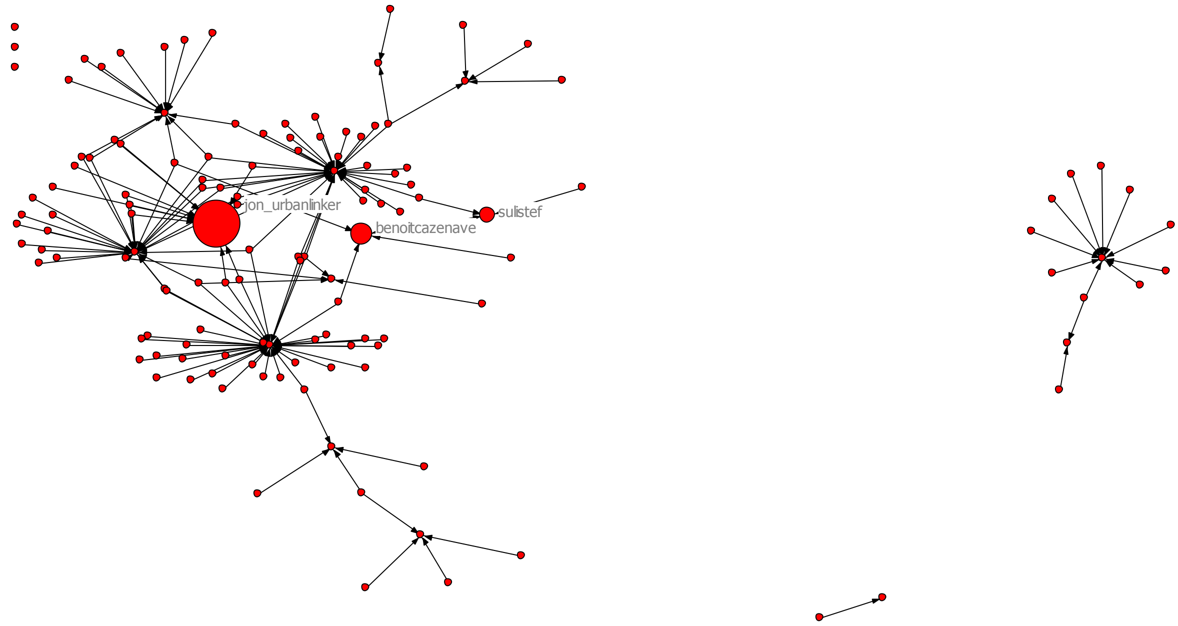



**Figure B2. Follower/following network among *LeWeb'13* Paris participants 5/12/13 – Node size scaled by betweenness**[11]




[38] Visualised in NetDraw 2.131 using a spring embedded algorithm, starting with a Gower Scaling.

dominated by six *Twitter* users, each with a large number of followers among other *LeWeb'13* Paris Twitter users. Three users have no followers among *LeWeb'13* Paris Twitter users. In Figure B2 node size represents the betweenness of each user, the number of times a node appears on the shortest path connecting each pair of nodes. Three nodes have particularly high betweenness, potentially influential brokering positions within the network.

Mappings of the follower/following network in 25/1/14 are not presented as they are identical to Figures B1 and B2; there was no change to the network after *LeWeb'13* Paris. Table B2 presents the structural properties of the network as at 5/12/13 and 25/12/14. The structure of

**Table B2. Structural properties of following networks among *LeWeb'13* Paris participants**[39]

| | 05/12/13 | 25/01/14 |
|---|---|---|
| **n** | 134 | 134 |
| **Ties** | 155 | 155 |
| **Avg Degree** | 1.157 | 1.157 |
| **H-Index** | 6.000 | 6.000 |
| **Density** | 0.009 | 0.009 |
| **Components** | 6.000 | 6.000 |
| **% nodes in main component** | 0.866 | 0.866 |
| **Component Ratio** | 0.992 | 0.992 |
| **Connectedness** | 0.009 | 0.009 |
| **Fragmentation** | 0.991 | 0.991 |
| **Closure** | 0.067 | 0.067 |
| **Avg Distance** | 1.083 | 1.083 |
| **SD Distance** | 0.276 | 0.276 |
| **Diameter** | 2.000 | 2.000 |
| **Breadth** | 0.991 | 0.991 |
| **Compactness** | 0.009 | 0.009 |
| **Markov clusters** | 14 | 14 |
| **Triadic Census:** | | |
| **1 - 003** | 373481 | 373481 |
| **2 - 012** | 16752 | 16752 |
| **3 - 102** | 127 | 127 |
| **4 - 021D** | 41 | 41 |
| **5 - 021U** | 1668 | 1668 |
| **6 - 021C** | 9 | 9 |
| **7 - 111D** | 5 | 5 |
| **9 - 030T** | 1 | 1 |
| **Network Centralization (Outdegree)** | 1.41% | 1.41% |
| **Network Centralization (Indegree)** | 26.60% | 26.60% |

[39] Calculated with UCINET 6.474. See Table A8 for Triadic Census codes.

both networks is identical. As there is no change to the network we were unable to apply Stochastic Actor-Oriented Models techniques for further analysis.

So, contrary to our first hypothesis, there was no measurable effect of participation in *LeWeb'13* Paris on Twitter followers among the sample examined. This may be an effect of the timing of the measurement; it may take more than six weeks after an event for users to add to their Following lists or intending participants may add to their following lists in anticipation of encountering them. Even so, some extra following would be expected during the in-person event if there was to be any at all but in this case there was none. More likely, the result is related to the characteristics of the sample. As indicated in Table B1, these were established *LeWeb* participants, having attended multiple events, and are likely to be relatively well known among other participants. It is likely, then, that their followers were established at earlier events and have stabilised and little change in this would be expected from a single event. Alternatively, following activity may have been concentrated among newcomers outside those sampled.