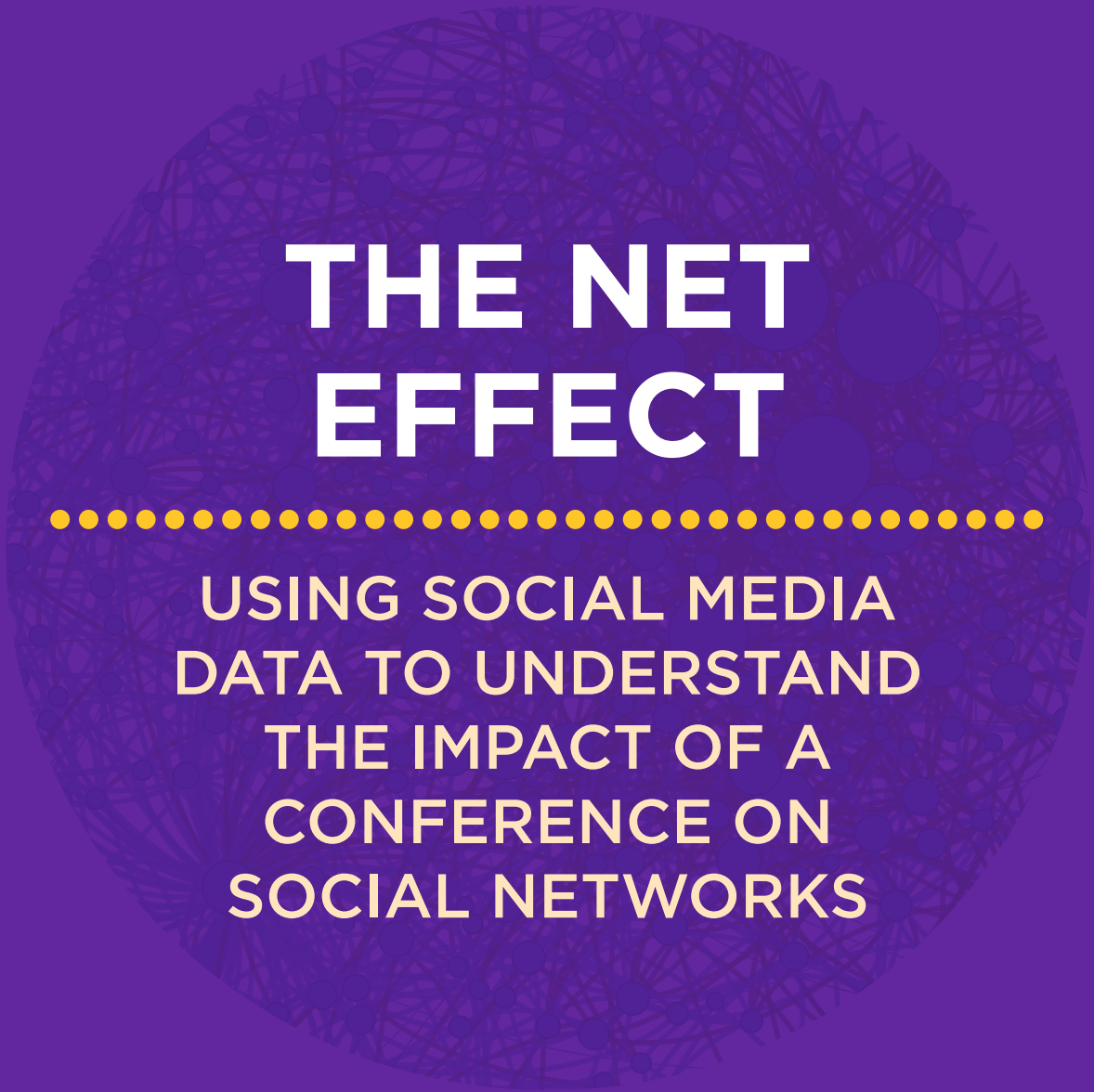
Nesta



THE NET EFFECT

USING SOCIAL MEDIA DATA TO UNDERSTAND THE IMPACT OF A CONFERENCE ON SOCIAL NETWORKS

Hasan Bakhshi, John Davies
and Juan Mateos-Garcia

June 2015

ACKNOWLEDGEMENTS

The authors would like to thank **Charles Armstrong** and **Craig McMillan** of Trampoline Systems for their help in collecting the Twitter data used in the analysis, and **Abla Kandalaft** for her research assistance.

Nesta...

Nesta is an innovation charity with a mission to help people and organisations bring great ideas to life.

We are dedicated to supporting ideas that can help improve all our lives, with activities ranging from early-stage investment to in-depth research and practical programmes.

Nesta is a registered charity in England and Wales with company number 7706036 and charity number 1144091.
Registered as a charity in Scotland number SCO42833. Registered office: 1 Plough Place, London, EC4A 1DE.

www.nesta.org.uk

THE NET EFFECT

USING SOCIAL MEDIA DATA TO UNDERSTAND THE IMPACT OF A CONFERENCE ON SOCIAL NETWORKS

CONTENTS

EXECUTIVE SUMMARY

It is widely accepted that innovation is often sparked by events where people meet to share ideas, and identify opportunities for collaboration and trade. Quantitative evidence about the impact of events on innovation is, however, hard to come by. Social media platforms which track the social networks of individuals over time can help address this.

In this paper, we explore the potential of using one particular social media platform – Twitter – to measure the outcomes of an innovation event, the LeWeb'12 tech conference held in London on 19–20 June 2012. Our analysis uses Twitter data for 702 participants in the event, including their 'follow networks' and public communications via Twitter, as well as additional data collected from websites like LinkedIn and TechCrunch. Our goal is to take a first step in developing a methodology for quantifying the impact of events on innovation, generating evidence that will be useful for:

- **Event organisers who want to understand and promote their impacts better, and identify good practices for running events.**
- **Individuals who want information about which events they should attend.**
- **Policymakers aiming to stimulate innovation through supporting networking events.**

Our key findings are:

- **Participants at LeWeb' 12 formed Twitter connections with each other at a greater rate than they did with non-participants (over four times faster in the six weeks after the event).** This is consistent with the event having stimulated increased networking between participants. In total, 1,520 new follows took place between participants at LeWeb'12, or within six weeks of the event. This is far more than in the following six weeks, so in this summary (and in the text, unless otherwise stated) we refer to these connections as having been generated 'at the event'.
- **A significant minority (30 per cent) of new Twitter following at the event involved the formation of a reciprocal connection**, indicating mutual awareness between participants. Specifically, the event was associated with the creation of 291 new reciprocal connections, 157 were completely new (A follows B and B follows A) and 134 were formed from consolidating an existing follow connection (A follows B who was following A before LeWeb'12).
- **The enhanced connectivity between people was also reflected in changes in the global structure of the network of LeWeb'12 participants**, with a decrease in the average distance between individuals in the network of 6 per cent.
- In order to gauge the additionality of LeWeb'12 (the extent to which it created connections that would not have happened anyway), we considered the distance between participants in the network that existed before the event. The idea being that participants who were further apart in pre-LeWeb'12 networks would have been less likely to connect without the event. **We find that a significant proportion (52 per cent), but not all, one-way follow connections at LeWeb'12 were between people who were previously just one step removed on Twitter**, perhaps indicating that these connections were more likely to occur at some point anyway even if LeWeb'12 had not taken place.

- **But there remain a significant proportion of new connections between individuals who were further apart in the network that existed before LeWeb'12; these people were arguably less likely to connect if they had not attended the event**. This was particularly the case for participants from different nationalities who connected at LeWeb'12, as we show these to have been further apart from each other on Twitter prior to the event.
- **The patterns of Twitter following between people in different roles and from different industries is consistent with the idea that participants were connecting to pursue economic opportunities** e.g. consultants linking up with potential clients, but not with each other.
- Although only a minority of LeWeb'12 participants connected by Twitter following connections communicated with each other on Twitter, we find that connections formed at the event were more likely to involve Twitter communication than those formed prior to the event. **Around half of the completely new reciprocal connections formed at LeWeb'12 involved communication between the people involved, in the majority of cases during or after the event**.
- We also mined the content of tweet communications between participants for keywords that might indicate face–to–face meetings bringing individuals together to discuss opportunities for collaboration. This analysis indicates that a number of the people had arranged to meet at or after the event.

Our analysis shows the potential for using social media data to attain a quantitative understanding of events and their impacts on networks. We would encourage event organisers to follow LeWeb'12's example, and publish participants' data in a way that enables the type of analysis we have undertaken in this paper. We also encourage policymakers to explore using our methodology to improve the evidence base on the impact of publicly supported events and networking initiatives on innovation.

Identifying the genuinely additional impact of events on attendees' networks is far from straightforward, because people who attend the same event share common unobserved interests and characteristics and so are more likely to connect in any case. As discussed above, we informally consider this by looking at the network distance between different participants prior to the event, although this is not the same as having a control group. A control group methodology is used in a forthcoming Nesta working paper by Greenwich Business School which analyses Twitter data along with participants' collaborations on the software platform GitHub.

1. MEASURING THE IMPACT OF AN EVENT

Ideas worth spreading and people worth meeting

Attending events, such as trade fairs, festivals and technology meet-ups, helps potential innovators solve some fundamental challenges: How can they identify the best opportunities? How can they find, and access, the skills and resources to get new projects underway? How can they meet potential collaborators?

Events address these information problems. They bring together people interested in relevant technologies and markets. They hold sessions where individuals and companies promote their offerings. They host thought-leaders discussing new developments, challenges and solutions. They offer dedicated networking, with face-to-face interaction enabling relationships to be built more quickly.[1]

Through this, events give participants an overview of what is going on, who is who, and who is doing what in their field, allowing them to network strategically and effectively. Events should therefore enhance economic activity and innovation: a speaker presents a new idea inspiring others to adopt it, a business meets a future client, investor or employee; a group of entrepreneurs decide to collaborate.

Yet, achieving this poses its own challenges:

1. **How can event organisers maximise the impact of their events, and demonstrate their benefits in a competitive market?** Without measures of performance, other than qualitative feedback from attendees and whether they choose to come again, event organisers may find it difficult to differentiate themselves from others, or to identify event practices that have greatest impact.
2. **How can time- and cash-pressed prospective attendees choose from the wide range of events on offer, and get the most out of their attendance?** Evidence about how an event impacts on participants' networks, and the quality of that connectivity e.g. what groups are getting connected at an event, would be valuable for participants considering going to future editions of an event, and planning how to get the most out of attending.
3. **Are there methodologies that would allow funders and innovation agencies to make better informed decisions on which events they should support?** Although it is plausible that events increase participants' connectivity and collaboration (potentially leading to innovation), we do not know of any studies quantifying these changes. That evidence would be invaluable for funders and innovation agencies aiming to improve innovation networks by supporting events. For example, the event studied in this paper, LeWeb'12, was sponsored by UK Trade and Investment (UKTI) funded by what was then London Tech City Investment Organisation, which covered between 10 per cent and 20 per cent of the budget.[2]

In order to help begin to address these challenges, rich quantitative information on events is required. This information would traditionally have been costly to obtain, but is increasingly available from the data generated by social media use.

Using social media data and social network analysis to measure the impact of an event

Most studies of the effects of events, festivals and conferences have been based on qualitative methods.[3] There have until now been few attempts to use social network analysis to explore events quantitatively. By measuring the connections created between participants attending the same event, and changes in the overall configuration of their network, social network analysis can improve our understanding of how events connect people, and whether these connections help disseminate valuable information and spur collaboration.

Part of the reason why social network analysis has rarely been used to study events is a lack of data: in contrast with other types of innovation network (e.g. scientists publishing papers together[4] or mergers between companies[5]) it has been difficult to measure relationships between individuals before and after an event because such connections have not traditionally generated a record (like academic papers, patents or press releases). This kind of constraint has been particularly pronounced at bigger events where the need for larger datasets increases the costs of data collection.

The widespread use of social media like Twitter or LinkedIn gives us a way to deal with this problem. The idea is that the ties between individuals in a social media platform – Alice follows Bob in Twitter, or Alice and Bob are 'contacts' in LinkedIn – are either valuable relationships in themselves or proxies for social relationships or connections outside of the social media platform: they can be used as a 'digital trail' for the connections that happen at the event. In other words, if Alice and Bob attended an event, and during it or shortly afterwards they connect on a social media platform, we might infer that Alice and Bob met or somehow interacted in the event, raising the possibility of measuring its impact by proxy.[6]

Ours is not the first paper to use Twitter analysis in research on conferences. Twitter communication has also previously been used in research by Ebner et al., (2010) which studied the EduCamp 2010 conference, and Letierce et al., (2010) which analysed the 2009 International Semantic Web Conference.[7, 8] Ebner et al., analysed the content of the tweets and the extent to which they related to the conference discussions. They found that it would have been hard for non–attendees to meaningfully follow the conference on the basis of the tweets generated. The Letierce et al. study found that participants' tweeting indicated that they were mainly trying to communicate with researchers in their academic field rather than with the wider public.

In this paper we study changes in the Twitter network of participants at London LeWeb'12, a digital technology conference held in June 2012 in Westminster. This brought together 1,500 digital entrepreneurs, investors, policymakers and others to explore the innovation opportunities afforded by real–time data. The event had the title theme of 'Faster than Real Time'.[9] There was some contemporaneous analysis of levels of tweeting activity at LeWeb London and how it related to LeWeb Paris.[10] This did not however examine the changes in follow connections between delegates at LeWeb London, or look at the structure of the event's network and communications between delegates on Twitter as a whole.

Using social media data to measure the network impacts of events like LeWeb'12 may therefore be an important component in estimating the effects of events. In this paper we present a set of metrics based on social network analysis that may be useful more generally for event organisers, funders and innovation agencies.

It must be acknowledged that Twitter data has some important limitations. People who use Twitter are a self–selecting subset of conference participants, and interaction on Twitter is only a small subset of the ways that conference participants may interact with one another.

Nevertheless, our sample of Twitter users is a significant proportion of those who attended the event (702 people out of 1,500 attendees), and the use of Twitter data allows us to analyse a rich set of information on networking between a significant number of delegates at low cost. This is information that might otherwise be hard to obtain without requiring attendees to carry a dedicated device (see the products developed by companies such as Blendology that allow interactions at events to be recorded) or using sensing equipment to monitor people's interactions in buildings.[11] For our purposes such devices would not capture information on attendees' networks from before the event, however.

Report structure

This paper is structured as followed:

- **Section 2 covers the research methodology; the data used in the analysis, and contains a short introduction to social network analysis.**
- **Section 3 describes the participants at LeWeb'12 London.**
- **Section 4 analyses how the Twitter follow connections between event participants changed as a result of the event.** It covers the strength of the connections formed and the effect on the overall network of connections between participants.
- **Section 5 analyses how likely people are to have connected due to the event itself. It does this by analysing how far apart the people were that connected at the event in the Twitter network before the event**. This being a proxy for whether people would likely have connected anyway if they had not attended the event.
- **Section 6 analyses the kinds of connections that were created between different groups of individuals at the event**. It does this by analysing the extent to which different groups connected within and between each other, relative to what one would have expected if the connections had been created randomly.
- **Section 7 analyses the communication flows that occurred between event participants on Twitter (before, during and after LeWeb'12)**. Communications on Twitter being a measure for people's engagement with one another due to the event. It examines the extent to which different kinds of connections formed were more, or less, likely to have information flowing over them. This also includes an analysis of tweet content on whether people met at the event.
- **Section 8 provides conclusions**.

2. RESEARCH METHODOLOGY AND DATA

Our analytical framework

In this paper, we use Twitter data to shed light on the different ways that events create connections between people that may in principle be conducive to innovation. They are as follows:

- **Event participants become aware of each other:** At events, speakers promote themselves, their projects, organisations, and ideas – some of which are useful to others. The same is true for other attendees who wear badges describing their affiliation, ask questions at sessions and exchange business cards. Other participants at events (including those not directly at the event, but tuning in via social media platforms) use this information to identify potentially valuable contacts. This process leaves an imprint on Twitter which we can analyse. The **formation of a Twitter 'follow' connection** between two participants indicates that, at the very least, one of them is aware of the other and considers them interesting enough to 'follow' on Twitter. **Reciprocal connections** between individuals (i.e. where two individuals follow each other) indicate they are aware of and interested in what each other has to say.
- **Event participants access new information:** An individual who follows another on Twitter starts receiving the other's tweets – in that sense, a 'follow' decision reveals interest in the information contained in those tweets. Previous research has shown that Twitter is used as a news medium as much as a social networking platform. In this sense, following someone on Twitter resembles 'subscribing' to their channel (Kwak, Lee, Park, and Moon, 2010).[12] A 'follow' could, however, also capture other things – interest in the originator of the tweets, an attempt at getting their attention, or simply a way of storing their contact details.
- **Event participants communicate with one another:** A useful measure of the interest in, or relevance of, the information that an event participant can access through new connections at an event are their **retweets** (when an individual republishes a tweet to their followers) from those new connections. In addition to retweeting, an individual can interact with other participants at the event by **replying** to their tweets, or **mentioning** them in tweets. These public communications may lead to more in-depth discussions, and, in principle, even collaboration. Researchers studying Twitter have argued that the intensity of communication between its members should be considered together with 'follow' connections in order to identify active social networks inside the platform (Huberman, Romero, and Wu, 2008).[13]
- **Event participants meet in person:** The interactions between individuals on Twitter are 'cheap' (only a click, or a few dozen words), so they might not tell us much about the significance of the relationships underpinning them. Direct Messages (DMs), which intuitively might signify closer relationships, cannot be observed by third parties, and we cannot, at least on Twitter, directly measure interactions taking place outside the platform. Face-to-face meetings represent a stronger commitment of time, potentially reflecting the perceived value that those involved see in each other, and are more likely to entail the exchange of 'private' (perhaps commercially valuable) information. Moreover, the literature suggests that resource-intensive types of interaction like face-to-face meetings

are more conducive to building trust and communication, leading to collaboration and innovation (Storper and Venables, 2004).[14] Opportunities for face-to-face interaction may be particularly important when attendees meet people from other industries and disciplines that are less likely to be in their existing networks (Boschma, 2005).[15] We try and informally capture the likelihood of face-to-face meetings by examining the tweet content for 'keywords' indicating if the two individuals met in person at the event.[16]

Additionality – what would have happened if the event had not occurred?

One important challenge that we face in measuring the effect of an event on participant networks is to find a suitable control group that allows us to consider what would have happened without the event (that is, what is the event's additional impact). The challenge here is that events often bring together individuals who are 'similar' to each other because they are interested in the same topic, work in the same industry, or have common socio-demographic characteristics and educational levels etc. These similarities mean that the participants are more likely to have connected with each other, regardless of whether they attended the event or not. A valid control group to identify the additional impact of the event on participants' networks would need to condition for all of these characteristics, many of which are unobserved.

In this paper, we do not formally identify a control group for the LeWeb'12 participants, but instead use their observed personal characteristics (country of origin, type of organisation they work for, their industry etc.), and their initial 'distance' in the network of participants (e.g. if they already had mutual connections or were 'further away' from each other), to assess how similar they were to each other, and get an informal sense of whether the connections they formed among themselves were likely to have happened anyway. In a forthcoming Nesta paper by the University of Greenwich, a control group methodology is used to help address the question of additionality.[17]

How Twitter is changing events: The distinction between networking on Twitter and face-to-face networking

In this research we have approached Twitter as a social media platform that can be used to measure networking activities at events. However, this networking need not necessarily be face-to-face. By tweeting about the event/each other and/or following its hashtag (tweets starting with a '#' that refer to an event specific label e.g. #LeWeb) participants are visible to one another and communicate without directly speaking. This means that Twitter is changing speaker sessions at events as it allows the audience (and others on Twitter) to publicly discuss talks as they are in progress. As Twitter users can, and do, follow people they do not know on Twitter, or refer to anyone in a public conversation, the activities on the platform do not necessarily have to follow the social conventions that may affect networking at some events.

Twitter is therefore not just a passive medium that records networking at events, but is changing the dynamics of events itself. We do not, however, attempt to distinguish between whether the networking we observe was enabled by the combination of Twitter and the event, as distinct from just the event itself. This would require data on face-to-face networking and access to a similar event where participants did not use Twitter, which we do not have.

As Twitter use at many events, particularly tech events, is routine, it is therefore reasonable to assume that their impact will often include the impact of Twitter networking as well as face-to-face networking. This does though raise the question as to how events organisers can best facilitate the use of Twitter to encourage networking.

Our Twitter dataset

LeWeb London 2012 (which we refer to as LeWeb'12) took place on 19–20 June 2012. Ahead of the conference, the organisers published a **list** of all registered participants with their country of origin, company and Twitter ID. Out of 1,281 names on that list, 831 included a Twitter ID. After merging those observations which shared a Twitter ID (usually individuals working in the same company who had registered its Twitter ID), Trampoline Systems used 789 IDs to extract additional social media data through Twitter's open Applications Programming Interface (API). They also matched these IDs to the **speaker** list in order to identify those who had spoken at a LeWeb session. The data Trampoline extracted included:

- Three snapshots of the Twitter network of 789 Twitter users, before the event ('pre-LeWeb'12 period'), six weeks after the event ('post LeWeb'12 period'), and 12 weeks after the event ('post LeWeb'12 period 2').
- All tweets published by participants between June 2011 (one year before LeWeb) and September 2012 (1.17 million tweets).

The data had to be cleaned. A limitation of the Twitter API is that it sets a restriction on the number of past tweets for any given individual that can be retrieved.[18] This means that our measures of tweet activity for individuals who posted more tweets than that limit over the considered period were skewed towards most recent times (since we could only collect their latest tweets, instead of all). We opted to treat these individuals as outliers and remove them from the database.[19] This left us with 703 individuals. Additionally, we identified a single individual in the database who had been following 556 event participants before the event (almost 80 per cent of the participants for which we had data), and unfollowed 485 of them afterwards. This individual was, singlehandedly, having a significant impact on some of our results – for example when we considered the 'net' number of follow connections created at LeWeb'12. We therefore decided to exclude this individual from our analysis as their behaviour was unrepresentative of those attending the event. These two situations illustrate some of the challenges of working with social media data – as well as the need for caution in collecting and analysing it.

Secondary data

In addition to the data obtained from Twitter, we also manually collected information about participants from other websites like the professional social network LinkedIn, and CrunchBase, a startup directory. A full list of the data collected is given in Appendix 1.

The user metadata obtained from Twitter alone does not contain information on key factors that are likely to influence the propensity to connect on Twitter e.g. country of origin (as compared with the location on their Twitter profile, a field captured in our initial list of participants), occupational role (entrepreneurs, investors, corporates, etc.), or industry of work. We also wanted to use this secondary data to triangulate the measures of 'distance' in the social network of participants before LeWeb'12 that we used to help assess how robust our measures of distance are.

Brief guide to network terminology

As social network analysis is central to what follows and has a specialist terminology, below is a short introduction to some of the terms used in the text.

A network/graph

A network (or graph as it is sometimes known) consists of a set of nodes (also referred to as vertices) connected by edges (also referred to as connections). In the case of the event networks studied in this paper, the participants are the nodes/vertices in the network and the edges/connections are the Twitter follow connections or the tweeting between them.

Directed vs undirected connections

A directed connection is a link between two nodes in a network where the connection exists in one direction but not another. On Twitter, for example, connections are directed, i.e. if person A follows person B on Twitter, then that does not necessarily imply that person B follows person A. On the social network LinkedIn, if A is connected with B, then B will also be connected to A; the connection is not directed.

Degree

This is the number of connections that a node has in the network. The interpretation of the degree of a node in the network can depend on whether the connections in the network are directed or not. For example, if person A follows five people on Twitter, but nobody follows them, then although their degree would be five and their out-degree (number of outward bound connections) would be five, their in-degree (number of inward bound connections) would be zero.

Path length

This is how far apart two people are in a network. There are typically a number of different of paths that travel between two people through the web of network connections. The average distance of the paths between two people, or what the shortest path is, are often measures that are of interest.

Centrality

This is a measure of how important/central a person is within a network. One measure is the in-degree that a person has (in the case of Twitter, this would be the number of people that follow them), but there are also other measures for this such as betweenness centrality (which measures for the specified individual how many of the shortest paths between the members of the network pass through that person) and eigenvector centrality (which takes into account how well connected are the people that a person is themselves connected to within the network).

Components

Sometimes a network may be split into a number of unconnected smaller networks, called components. An example of this would be an event where there are two groups of people who know everyone in their own group, but do not know anyone in the other group. In this case the event network would split into two unconnected network components.

Network density

This is the total number of connections in a network divided by the total number of possible ties that could exist between the people in the network.

3. WHO ATTENDED LeWeb'12?

We refer from this point on to the 702 individuals in our sample as the 'participants'. Table 3.1 shows that they came from a range of different countries (separate nationalities are shown for those with more than ten participants). UK country of origin was, unsurprisingly, the most common, but at 23 per cent only a minority of participants. LeWeb'12 was an international conference, with France and the US accounting for almost 30 per cent of participants between them.[20] The 'Other countries' category includes nationalities from a range of different countries e.g. Denmark, Turkey, Bulgaria, Brazil, Lebanon and Russia. Among the 702 participants are 49 speakers. UK and US speakers accounted for the majority of the conference programme of keynote speeches and panel sessions, and US participants alone accounted for just over 40 per cent of the speakers in the sample. In what follows we refer to participants who were not speakers at LeWeb'12 as 'attendees'.

Table 3.1: **Nationalities and status of participants**

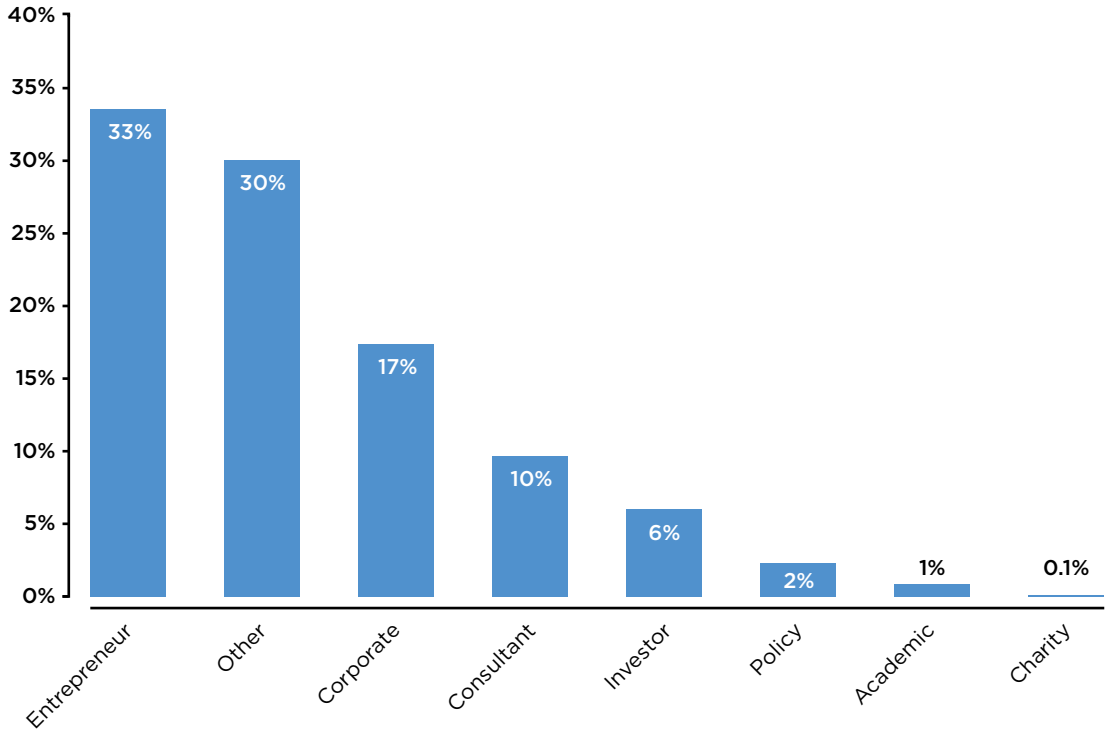
Nation	Participants	Per cent (Inclusive of speakers)	Speakers	Per cent of speakers
UK	161	23%	8	16%
France	118	17%	1	2%
US	82	12%	20	41%
Germany	25	4%	1	2%
Belgium	14	2%	0	0%
Spain	13	2%	1	2%
Italy	13	2%	1	2%
Netherlands	12	2%	0	0%
Canada	11	2%	0	0%
Other	253	36%	17	35%
Totals	702	100%	49	100%

Note: The Other category includes 100 individuals where it was not possible to identify a country of origin, either because it was not possible to match the Twitter account with further online information or because there was no clear answer from the information obtained.

The gender split between male and female delegates was 81 per cent male and 19 per cent female. The largest age group among attendees was that aged between 26–35 years (44 per cent of attendees were from this group), followed by the 36–45 age group (35 per cent).[21] Those older and younger than these age groups counted for approximately 10 per cent each of the remaining delegates.

Figures 3.1, 3.2 and 3.3 plot the distribution of participants by their occupational role, industry, and the discipline of their highest academic qualification. They show that the largest group at LeWeb'12 were entrepreneurs (a third of all participants), although there were also significant numbers of participants in other roles, like corporates at 17 per cent (we use this term to refer to people working in large companies and brands), consultants (10 per cent) and investors (6 per cent).

Figure 3.1: **The largest group of participants at LeWeb'12 were entrepreneurs**



Figures 3.2 and 3.3 illustrate the range of industries and disciplines present at LeWeb'12 – creative industries like software, media, internet and advertising companies predominated (Figure 3.2).[22] Business and Finance was the academic discipline that the largest number of participants had their highest qualification in (and almost three in ten of those with Business and Finance qualifications had MBAs), but there were also large numbers of participants with degrees in computer science, other Science, Technology, Engineering and Maths (STEM) subjects (e.g. physics or mathematics), Marketing and the Arts and Humanities. These statistics suggest that there was a great deal of potential for both multidisciplinary and cross-sectoral connections to be made at LeWeb'12. Section 6 explores how much of this potential may have been realised.

Figure 3.2: **There were a wide range of industries represented at LeWeb'12**

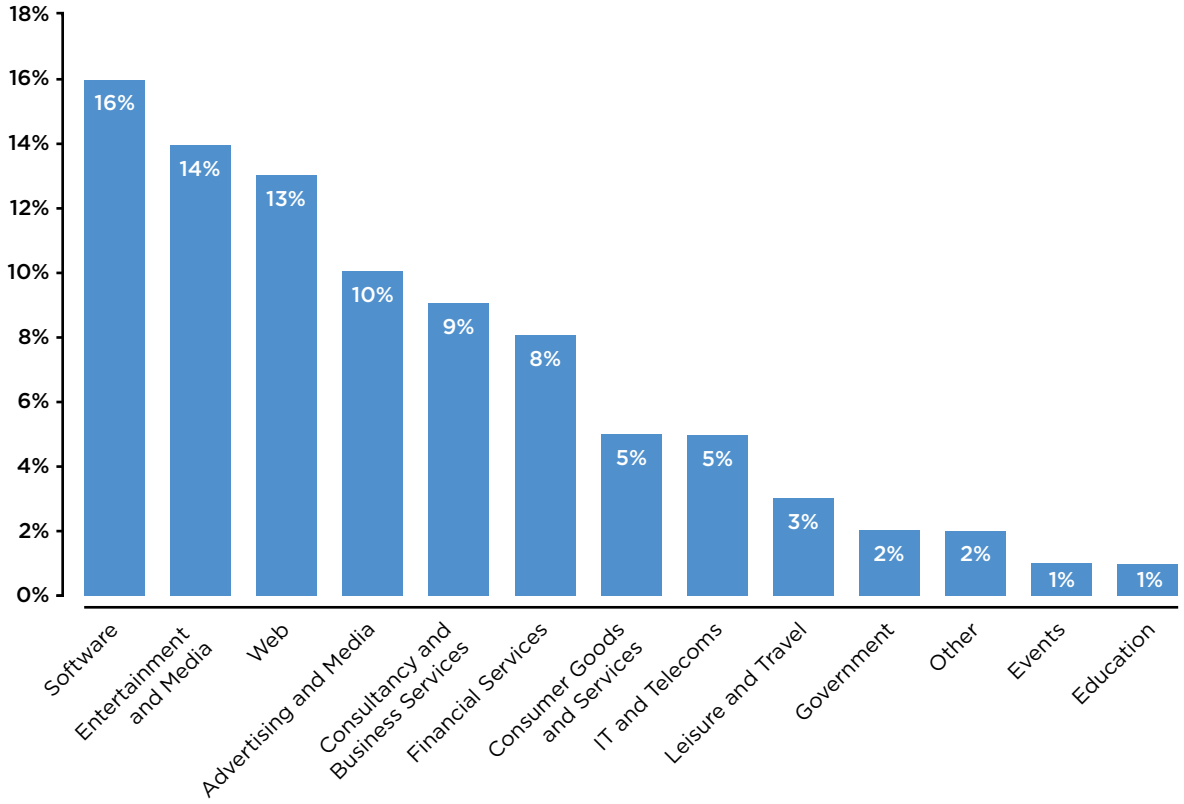
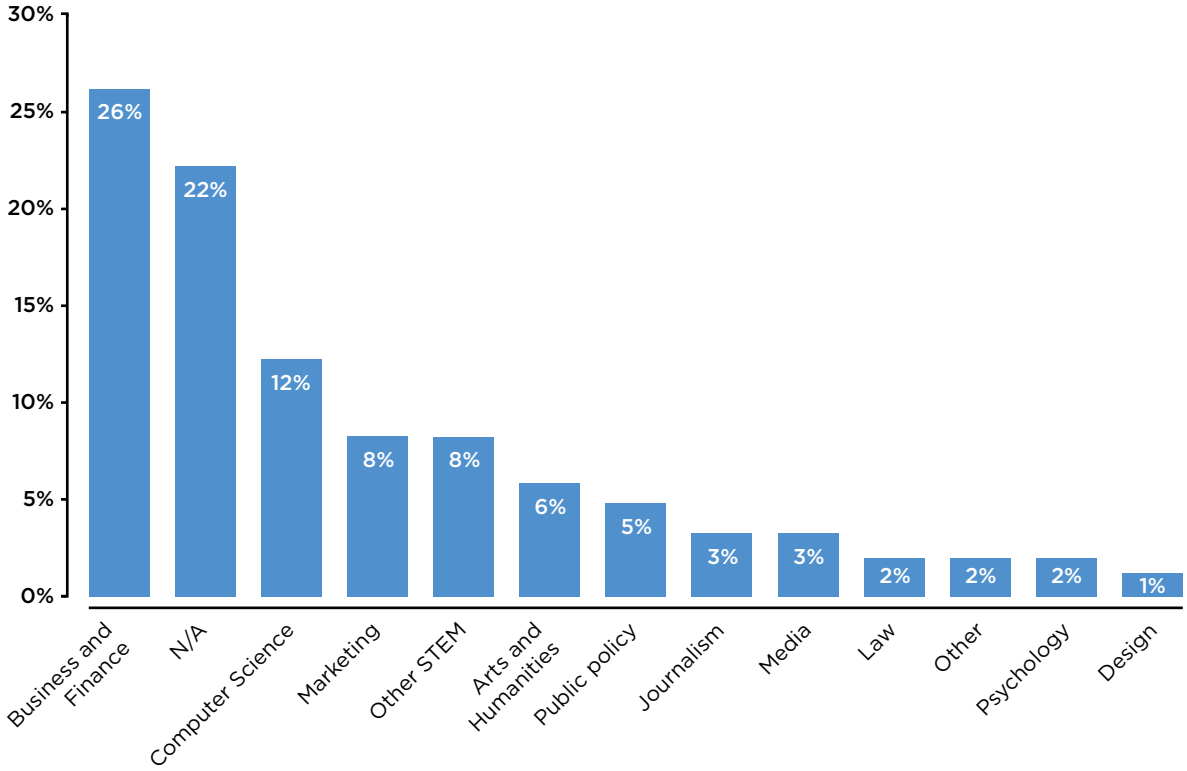


Figure 3.3: **Business and Finance (including MBAs) was the most common discipline that LeWeb'12 participants had studied**



4. HOW DID LeWeb'12 CHANGE CONNECTEDNESS BETWEEN PARTICIPANTS?

What we did:

- We measured changes in the Twitter 'follow network' of participants at LeWeb'12 before and after the event, distinguishing between one-way connections and reciprocated ones, and considered how this connectivity drove changes in the structure of the network.

What we found:

- There was a jump in connectivity (as measured by Twitter following) among LeWeb'12 participants during the event, with participants connecting at a faster rate to other LeWeb'12 participants than with other Twitter users outside the event.
- Around three out ten follows formed at LeWeb'12 involved the formation of two-way relationships between participants (i.e. were reciprocal). The rest were one-way (a participant followed another, but they weren't followed back or the follow was not reciprocating a pre-existing connection).
- This enhanced connectivity was reflected in changes in the global structure of the network of LeWeb'12 participants, which became denser and less fragmented.

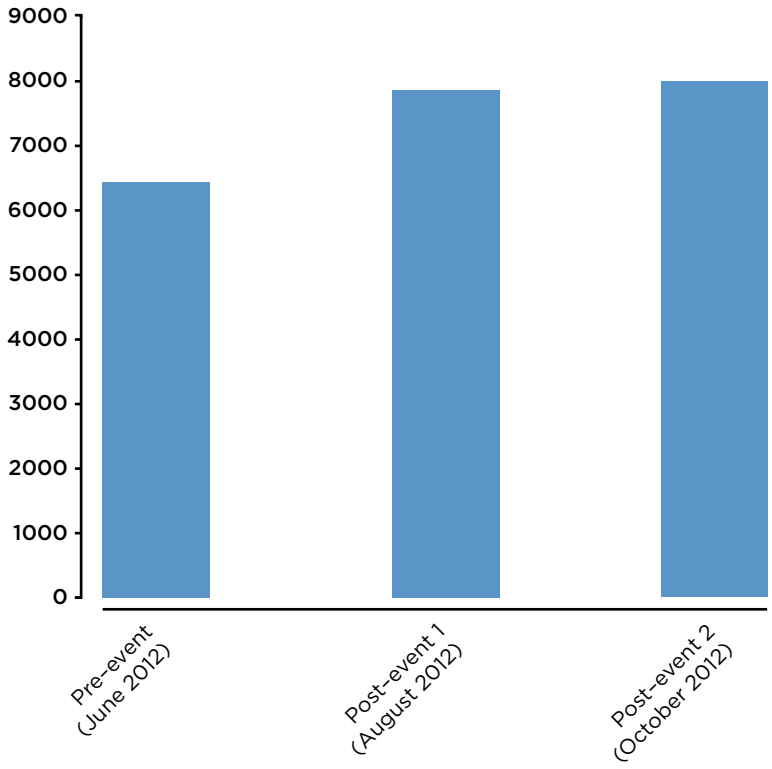
The formation of Twitter follow connections between participants at an event like LeWeb'12 indicates that they perceived value in each other, at least as sources of potentially valuable information. In this section we analyse the Twitter follow connections created at the event.[23]

Aggregate changes in connectivity between LeWeb'12 participants

Figure 4.1 shows the total number of Twitter follow connections within our group before the event (June 2012), six weeks afterwards (August 2012), and 12 weeks after in October 2012. It suggests that participation in LeWeb'12 spurred increased connectivity between those at the conference.

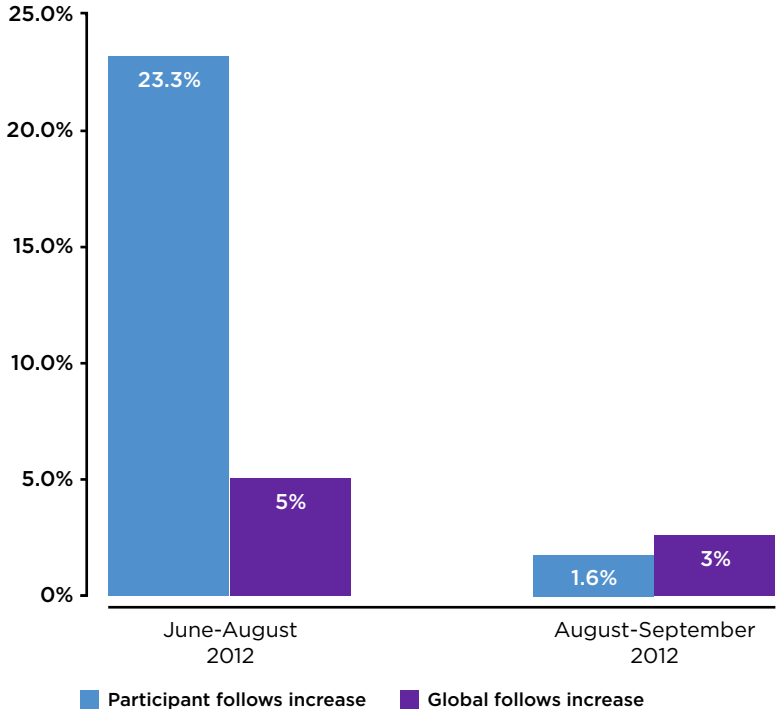
Between June 2012 and August 2012, the total number of follow connections between the 702 participants increased by 1,416 (a 22 per cent rise in connections), 449 LeWeb'12 participants (63 per cent of our sample) had created at least one new connection with another participant (either by following someone, or being followed by someone at LeWeb'12) within six weeks of the event. Networking between August and October 2012 was much slower, with the total number of connections increasing by 127. These figures include the effects of the, by comparison, small number of connections that were 'destroyed' in the aftermath of LeWeb'12 (i.e. where people unfollowed each other).

Figure 4.1: Twitter follow connections between participants grew fastest during the period when LeWeb'12 happened



The formation rate of new Twitter connections between those who participated at the event greatly exceeded that of new connections between participants and other people on Twitter, i.e. Twitter users who did not attend LeWeb'12 (Figure 4.2). Specifically, the rate at which participants' connections with others at LeWeb'12 increased was 4.6 times faster than the formation rate of new global follow connections.

Figure 4.2: **The growth in connectivity within LeWeb'12 was much faster than outside**



Appendix 4 contains further information on the average changes in follow activity for individuals in different groups (occupations, industries, and nationalities) participating in LeWeb'12.

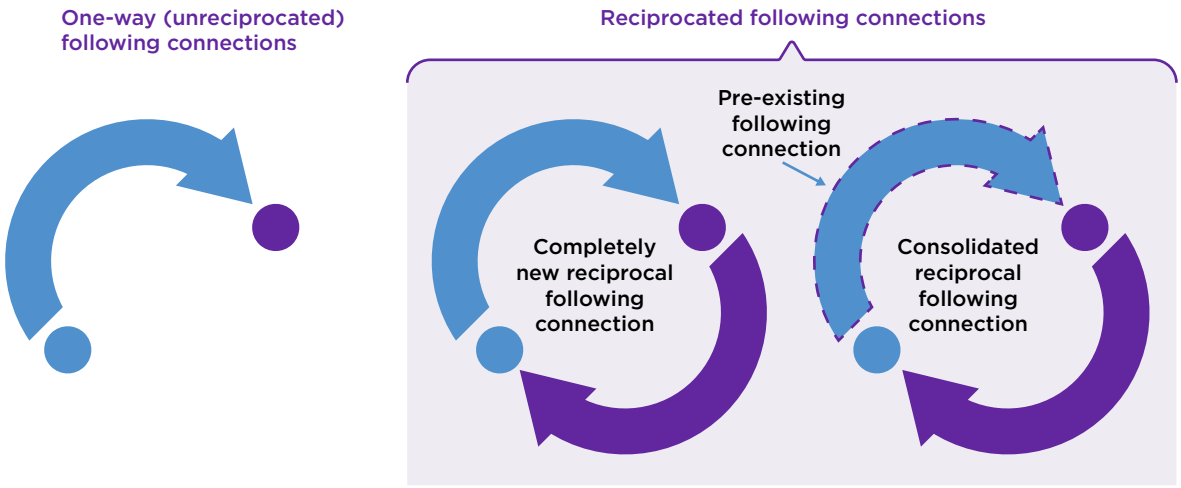
Not all connections created at LeWeb'12 were equal

As discussed in Section 2, reciprocity in the formation of a Twitter connection indicates mutual awareness between individuals (including the possibility that they have met face-to-face or connected online), potentially linked to future communications and collaborations. This contrasts with 'one-way' follow connections, which just indicate that a participant has subscribed to another's Twitter updates.

To take into account these differences we distinguish between the following types of connections at the event (Figure 4.3):

1. **One-way (unreciprocated) follow connections**: at the event, Person A followed Person B (who did not follow them), but Person B did not follow Person A back.
2. **Reciprocal follow connections**, which are either:
   - A completely new reciprocal connection i.e. Person A followed Person B at the event, and vice versa.
   - Consolidation of an existing one-way connection i.e. Person A was already following Person B before the event, and Person B began to follow Person A at the event.

Figure 4.3: **Describing connections on Twitter**



1520 new follows were created between participants at LeWeb'12, this is greater than the total increase in follow numbers due to 104 unfollows among pre-event connections. When we categorise these into the three types of relationship, we find that 70 per cent were one-way follow connections and 30 per cent resulted in the formation of reciprocal relationships (of which just under a third consolidated a one-way connection that existed prior to the event):

1. Seventy per cent (1,072) were one-way follow connections
2. Nine per cent (134) created 134 reciprocal connections by consolidating an existing one-way follow connection
3. Twenty-one per cent (314) created 157 completely new reciprocal connections

The global picture

We have used this connections data to plot two Twitter networks of LeWeb'12 participants (see Figures 4.4 and 4.5). This has been done using the software package, Gephi.[24]

Figure 4.4 shows the network of the 1,520 new Twitter follow connections that were created between participants during LeWeb'12 (it excludes participants that did not form any new follow connections). The purple nodes are speakers and the blue nodes are attendees. The size of the nodes represents the in-degree of each node (how many new follows it gained). It shows that in general – and as one would have expected – speakers tended to gain more followers than attendees. The colour of the edges shows whether a follow connection was made between speakers (purple colour), between attendees (blue colour) or between speakers and attendees (brown colour).

Figure 4.4 **All new follow connections**

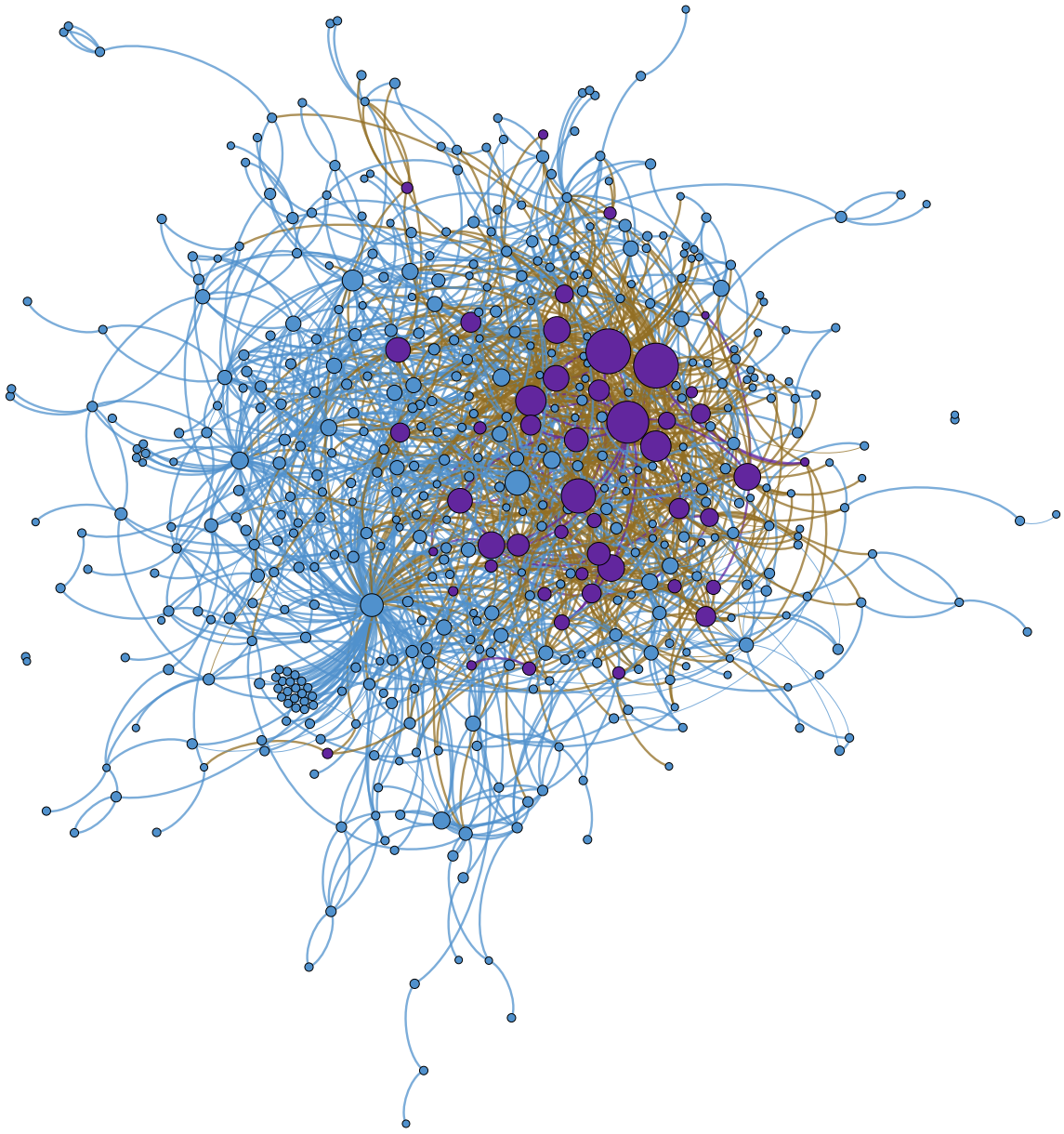
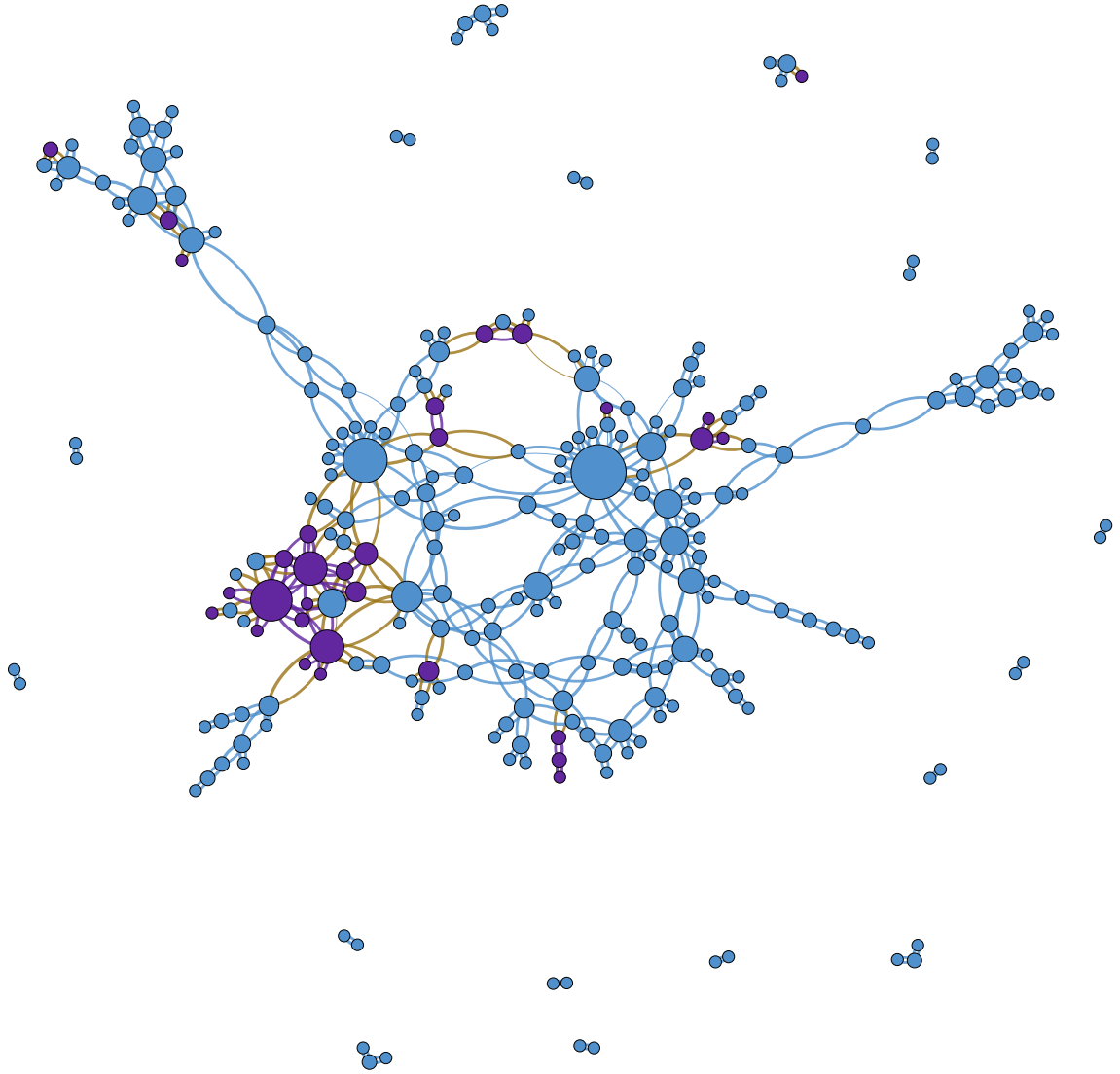


Figure 4.5 below shows the smaller number of the 291 reciprocal connections created at LeWeb'12 (again, it does not show participants who did not gain any new reciprocal connections). It uses the same colour coding as before: purple nodes are speakers, blue nodes are attendees, purple edges are connections between speakers, blue edges are connections between attendees, and brown edges are connections between speakers and attendees. The size of nodes indicates their degree (how many new reciprocal follow connections they achieved at LeWeb'12). It reveals some hubs of reciprocal connectivity at LeWeb'12, including a cluster of speakers who reciprocally connected at the event.

Figure 4.5 **All new reciprocal connections**



The increase in connectivity between participants at LeWeb'12 also led to changes in the overall structure of the networks between participants on Twitter. This is shown in Table 4.1 for the Twitter network involving all follow connections and the network consisting of reciprocal connections only:

- **Efficiency** (second row): the average path length decreased slightly in the all follows network – in other words, there was a reduction in the number of people that, on average, a message would need to pass through to get between any two participants at LeWeb'12, implying that the all follows network became more efficient in the transmission of information. By contrast the average path length went up in the reciprocal network.[25] The likely explanation for this is that during LeWeb'12, a component (an inter–connected group of people) connected with another inter–connected group of people that was previously separate. Parts of the reciprocal network that had been until then disconnected became inter–linked, increasing the size of the network of reciprocal connections and hence increasing the average distance in the network.
- **Density** (third row): Logically, given the increase in connections, the percentage of connections that were realised as a proportion of all possible connections went up in the all follows and the reciprocal networks.

- **Cohesiveness** (fourth row): A clustering coefficient that captures, on average, how many of the connections a person has are themselves connected. Intuitively, unconnected individuals who have a connection in common are more likely to connect with each other than unconnected individuals who do not have a common connection (e.g. they are more likely to be introduced by that third party). Table 4.1 shows that clustering increased in both the network of all follow connections and in the network consisting exclusively of reciprocal connections.
- **Isolation** (fifth row): As would be expected, the number of participants in LeWeb'12 who were totally disconnected from the all follows network (i.e. they did not follow or were followed by anyone there), or from the reciprocated network (i.e. they did not have a reciprocated follow relationship with anyone) went down.
- **Fragmentation** (sixth row): The number of completely disconnected components in the all follows network and in the reciprocated network after LeWeb'12 fell; this means that connections were established between previously fragmented groups.

Table 4.1: **Changes in the global features of the LeWeb'12 network: all follow connections and reciprocal connections only**

Metric	Definition	Pre-event All follow connections	Post-event All follow connections	Pre-event Reciprocal connections only	Post-event Reciprocal connections only
1. Average degree	Average connections per participant	18.1	22.1	3.19	3.98
2. Average path length (efficiency)	Average path length between participants in the network (measured in terms of edges between people)	3.54	3.34	3.91	3.94
	% change in average path length		-6%		1%
3. Density	Percentage of all pairs of participants with an edge between them	1.30%	1.57%	0.45%	0.57%
	% change in density		21%		27%
4. Average clustering coefficient (cohesiveness)	Average % connections that a participant has who are themselves connected	28.9%	30.5%	13.6%	14.8%
	% change in clustering		6%		9%
5. Isolates	Number of participants disconnected from the network	52	37	225	170
	Proportion of isolates in the network overall	7%	5%	32%	24%
6. Number of components (fragmentation)	Number of disconnected sub-networks	56	40	249	184

Note: reciprocal connections (which consist of two follow connections) are counted as one undirected connection in this analysis.

5. WERE THE CONNECTIONS FORMED AT LeWeb'12 LIKELY TO HAVE OCCURRED IF THE EVENT HAD NOT HAPPENED?

What we did:

- We considered the distance between LeWeb'12 participants in the pre-event network as an indication of whether connections at the event were more likely to have happened if the event had not in fact occurred (i.e. whether the connections only happened as a result of the event). We also checked the robustness of our measures of distance by checking if people who were identifiably different from each other in other ways (e.g. in terms of their nationality, role etc.) were further apart in the pre-LeWeb'12 network.

What we found:

- Our analysis reveals that a substantial number of people who connected at LeWeb'12 had a mutual connection on Twitter in the pre-LeWeb'12 network. There were, however, a significant fraction of new connections between people who were further apart in the pre-LeWeb'12 network.
- Connections between individuals from different nationalities and industries were more likely to have been between people who were further away from each other in the pre-event Twitter network.

Distance matters

A central question for our research is whether the connections at LeWeb'12 would have formed anyway if the individuals involved had not participated in the event. Or in other words, the extent to which the connections made were 'additional'. In this section, we use the distance between individuals in the network that existed prior to LeWeb'12 to explore this issue.[26] Intuitively, proximity between people in the Twitter network before the event might be related to people's proximity in other social (and professional) networks, and therefore, their probability of connecting in the absence of their attendance at LeWeb'12. Conversely, connections between individuals located further apart in the pre-event network were arguably less likely to happen without the proximity and opportunities for serendipitous interaction generated by LeWeb'12.

There are two reasons for this:

1. We know from the social networks literature that individuals who have mutual friends (that is, are close to each other within the network) are more likely to meet.[27]
2. Individuals closer to each other in a network are more likely to be similar in terms of their demographics, the industry where they work, their interests and so forth – and for that same reason, more likely to connect with each other.[28]

Both mechanisms are reinforced in online social networks by platforms like Twitter, allowing us to see our wider social network, and recommending people for us to connect with on the basis of this. These social media platforms also lower the transaction cost of connecting.

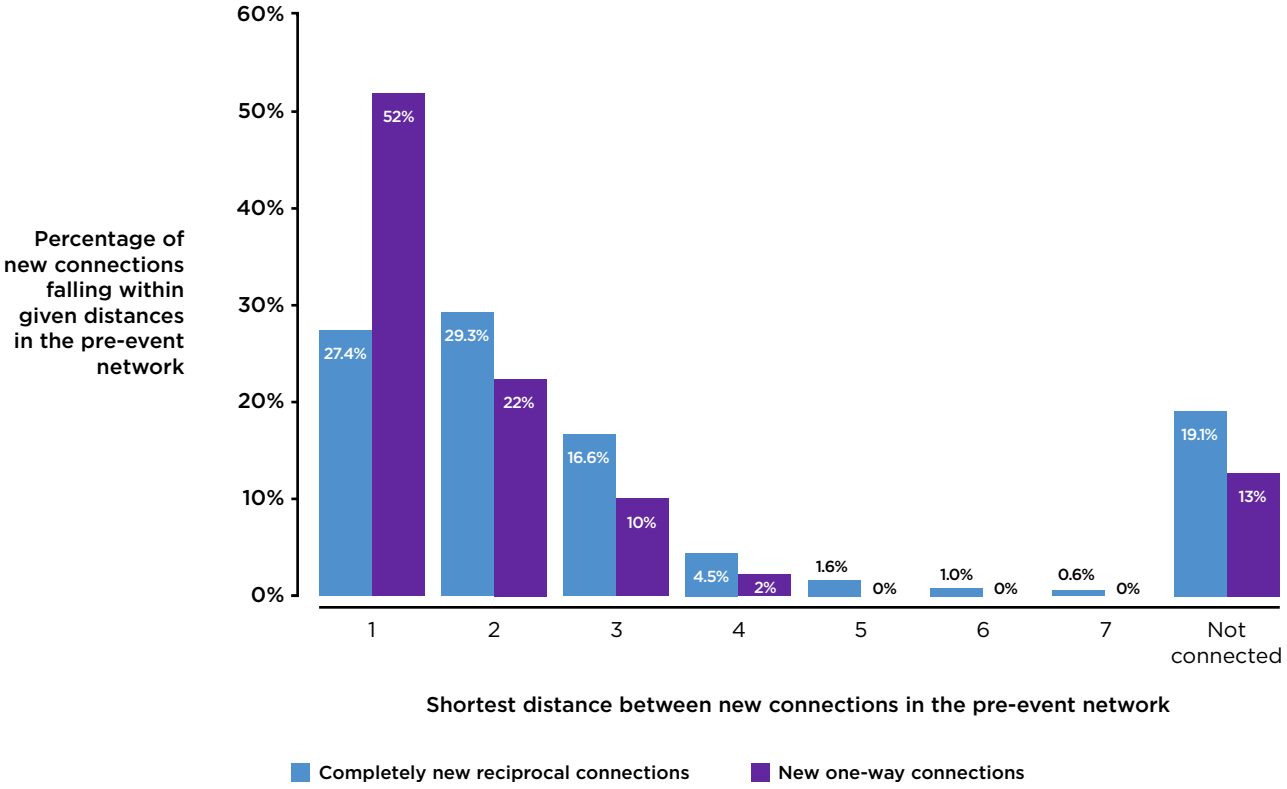
In this section, we introduce the idea of distance in our analysis by examining: how far apart people who connected on Twitter at LeWeb'12 were in the pre–event Twitter network, and whether the distance between people who connected at LeWeb'12 was related to their personal characteristics.

In addition to this, we assess the robustness of our distance measure. Our results are consistent with the idea that these distances capture genuine differences between individuals, and therefore tell us something indirectly about the extent to which LeWeb'12 had additional effects on participant connectivity.[29]

How far apart were participants who connected at LeWeb'12 in the Twitter network prior to the event?

We measure the initial distance between two individuals who connected at LeWeb'12 by calculating 'the shortest path' between them in the Twitter network that existed before the event.[30] Figure 5.1 plots the distribution of initial distance for the one–way (unreciprocated) and completely new reciprocal connections formed at the event.[31] Distance in this Figure, and others that appear in this section, is defined as the number of people that the shortest path between people in the network had to pass through e.g. a distance of one means that the people connecting at the event had one contact in common before the event.

Figure 5.1: How far apart were individuals who connected at LeWeb'12 in the network prior to the event?



It shows that, in the case of one-way follows, individuals were more likely to follow people who were already closer to them in the network – for example, just over half of the one-way follow connections at LeWeb'12 formed between participants who already had a mutual connection.

Completely new reciprocal connections were slightly more likely than one-way connections to be with people who were, before LeWeb'12, further away in the network (both following connections that contributed to a new reciprocal connection are counted in the analysis).[32] While only 13 per cent of the one-way follow connections were between individuals who had been in totally disconnected parts of the network before the event, this was the case for 19 per cent of the following connections that formed complely new reciprocal connections at LeWeb'12.[33, 34] Compared with one-way connections, reciprocal connections were also more likely to involve people who had been at least two steps removed from each other in the network before LeWeb'12.

One possible interpretation of this finding is that LeWeb'12 participants tended to follow/be followed by others in their neighbourhood (that is, individuals who are close to them), but on average form reciprocal relationships with people further apart in the network. This is consistent with the idea that they were more likely to see mutual value in relationships with others with whom their existing social networks overlapped less – perhaps because those relationships were seen as more likely to help them access new information.[35] Another interpretation is that one-way follows were likely to be of people with a higher profile, such

as speakers, who were themselves very well connected (and so a shorter distance away from anyone in the network).[36] By contrast reciprocal connecting was arguably more likely to be between people of similar status.

If we make the strong assumption that new connections with one degree of separation prior to the event would have happened anyway and those with two or more degrees of separation would not, then this suggests an indicative upper bound on the number of new follow connections likely to have been additional. It turns out that around 50 per cent were one step removed for new one-way follow connections and around 30 per cent were one step removed for completely new reciprocated follow connections. On this basis perhaps as many as half of the unreciprocated follow connections created at the event were additional and around 70 per cent of the following connections that formed completely reciprocal connections.

Were participants from different nationalities who connected at LeWeb'12 further apart in the pre-LeWeb'12 network?

If our working assumption is that people who were further apart in the Twitter network were less likely to have connected in the absence of the event, we should find that individuals who connected to others with observable common characteristics (e.g. their country of origin, or industry) were closer in pre-LeWeb'12 networks than those who connected with people who had identifiably different characteristics.

As a robustness check we estimated the impact of personal differences between individuals on the probability that they were connected (or at least shared a mutual connection) in the pre-LeWeb'12 network.[37] The intuition being that if our measures of pre-event distance captured the fact that individuals had different backgrounds and were in different social circles, then we would have expected a connection between differences in their nationality, industry, role, age, and their likelihood to have been directly or indirectly connected before LeWeb'12. We report the findings of this analysis in Appendix 2. We find significant links between several indicators of personal difference and network connectivity before LeWeb'12.

When we analysed the pre-event distance between participants who connected at LeWeb'12 we found that it was correlated with personal differences along a number of dimensions.[38] For example, participants from different academic backgrounds who created a one-way follow connection were more likely than others to have been further apart in the network before LeWeb'12, and something similar happened with participants in different occupational roles (e.g. investors and entrepreneurs): participants who created reciprocal connections with others in different roles from them were less likely to have had mutual connections (be close to each other) in the pre-LeWeb'12 network.

The association between distance in the pre-LeWeb'12 networks and personal differences was, however, clearest when we considered differences in nationalities. Figure 5.2 shows that participants who created a follow connection with others with the same country of origin were just under 25 per cent more likely to have shared a mutual connection before LeWeb'12. Figure 5.3, where we consider only completely new reciprocal connections and distances in the pre-LeWeb'12 network, shows that individuals who connected with those with the same country of origin as them were more likely to have had a mutual connections, i.e. they were both connected by an intermediate connection.

We interpret this as suggesting that those new connections that happened between individuals from the same country were more likely to have happened anyway because those people were closer in the pre–LeWeb'12 network. By contrast, there was a substantial proportion of reciprocal relationships that were established between people from different nationalities who were further away in the reciprocal network. Specifically, there were 88 follow connections (forming part of a completely new reciprocal connection) between participants from different nationalities who had been removed from each other by at least two people, indicating that they would have been less likely to happen without LeWeb'12. Perhaps LeWeb'12 generated its main additional impacts on networks by bringing together entrepreneurs and innovators from different nationalities?

Figure 5.2: Participants from the same country who formed one–way connections were more likely to have had a mutual connection prior to LeWeb'12

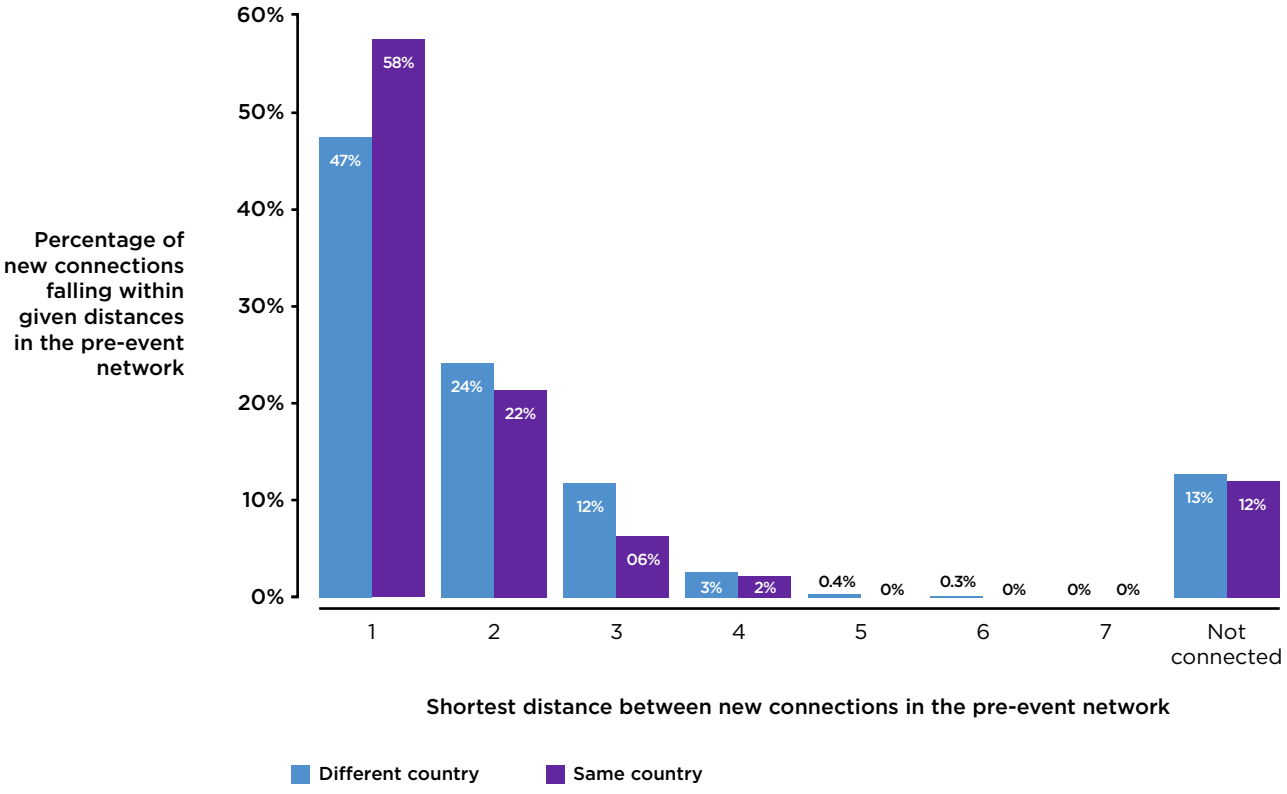
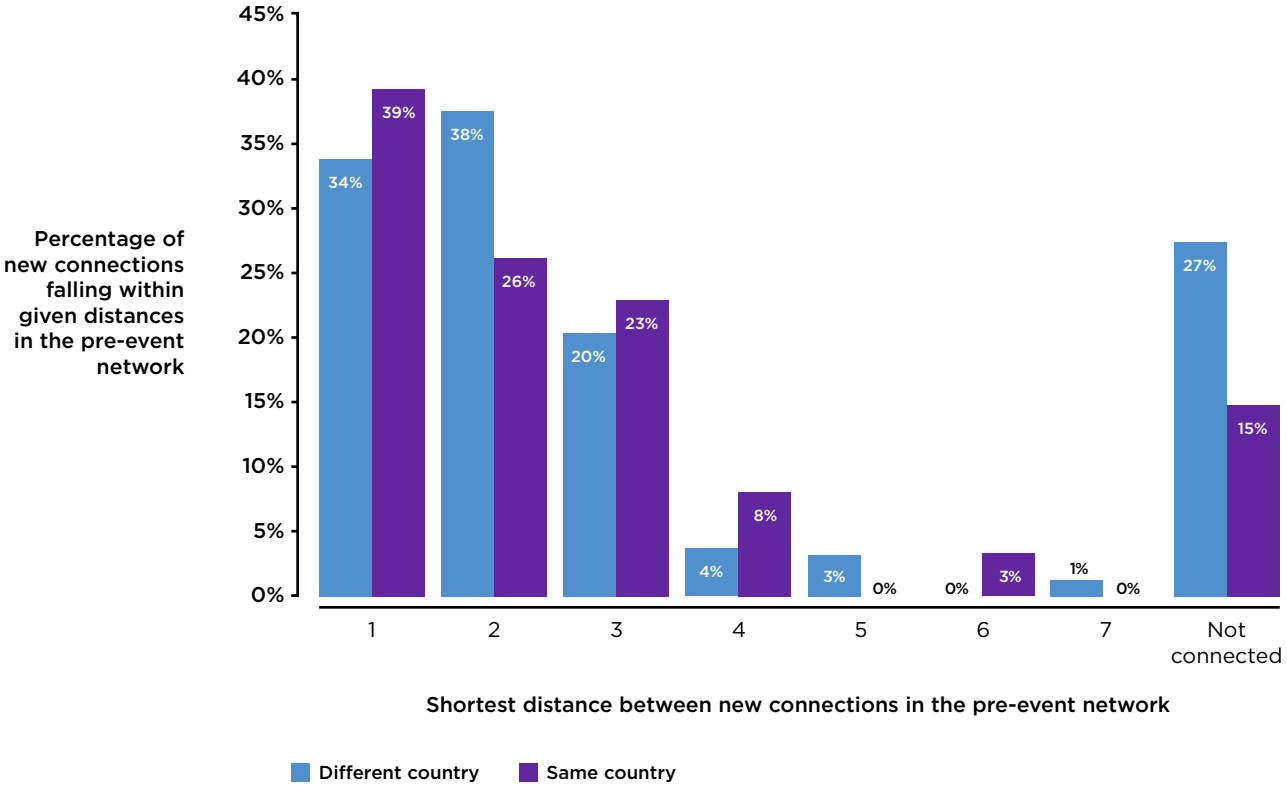


Figure 5.3: Participants from different nationalities who created completely new reciprocal connections created connections with people further away in the network



6. BRIDGING THE GAP: DID LeWeb'12 PARTICIPANTS CONNECT INSIDE THEIR GROUPS, OR OUTSIDE?

What we did:

- We looked at the patterns of completely new reciprocal follow connections formed within and between different groups on Twitter at LeWeb'12, considering how the levels of connectivity observed contrasted with a benchmark of connections generated at random.

What we found:

- Participants from the main countries represented connected intensively with those of the same nationality, although there were also international connections which seemed, in part, to relate to language or geographic proximity (e.g. connections between UK and US delegates, and between Dutch and Belgian delegates). As indicated from the previous section, these groups were likely to be more closely connected in the Twitter network and therefore arguably more likely to have connected anyway.
- When we look at connectivity between people in different roles, and different industries, we see some patterns in Twitter following that are consistent with the idea that participants were connecting with economic opportunities in mind – consultants linking up with potential clients, but not with each other, for example. We also see patterns in the network of reciprocal connectivity between industries, which might also reflect commercial motivations.

In this section, we examine the connections within and between groups of participants at LeWeb'12. Systematic patterns of connectivity between groups conceivably indicate situations where participants are establishing relationships with economic potential: an entrepreneur connecting with an investor for example, or a consultant connecting with a potential client.

We would also expect to see connectivity happening within groups as events, like LeWeb'12, act as fora where industry participants can exchange information and ideas with each other, look for like-minded collaborators, or seek new job opportunities. From the previous section we also saw that people with similar characteristics were likely to be closer together on Twitter, and were therefore more likely to have connected anyway.

Measuring crossover

In this section, we focus on the completely new reciprocal connections formed at LeWeb'12 (i.e. reciprocal connections formed at the event between people who were previously unconnected on Twitter) across three categories for which we have data – participants' country, occupational role and industry. We are interested in reciprocal connections because they are intuitively more substantive than unreciprocated ones. Furthermore, completely new reciprocal connections are arguably more likely to have been generated by the event itself

than those formed from someone reciprocating a following connection from before the event.

We generate a simple reciprocal connectivity indicator C_{ij} between any two groups i and j (or within a group e.g. C_{ii}) which measures the ratio of the proportion of connections between two groups (or within a group) observed at LeWeb'12, and the proportion of connections we would have expected to have seen if connections had been generated randomly.[39]

A C_{ij} score above 1 indicates that group i connected with group j at a higher rate than we would have expected if individuals had been connecting at random (considering the overall presences of both groups at LeWeb'12). A score below 1 indicates the opposite.

We present our findings in Tables 6.1–6.3, and in Figures 6.1–6.3. The tables display reciprocal connectivities, C_{ij} inside/across the groups in the three categories we consider. We give cells a darker shade the higher the score. The diagonal of every table captures connections within a group.[40]

In Figures 6.1–6.3 we plot these tables as network graphs where the nodes represent categories in each of the groups (e.g. a country, or an industry), and edges are plotted between groups where the connectivity indicator is greater than one (i.e. levels of connectivity are higher than the random benchmark). In most cases a 0 corresponds to no connections between/within groups. Groups that have higher levels of connecting between their members than the random benchmark are given a darker shade.

Which countries were connected at the event?

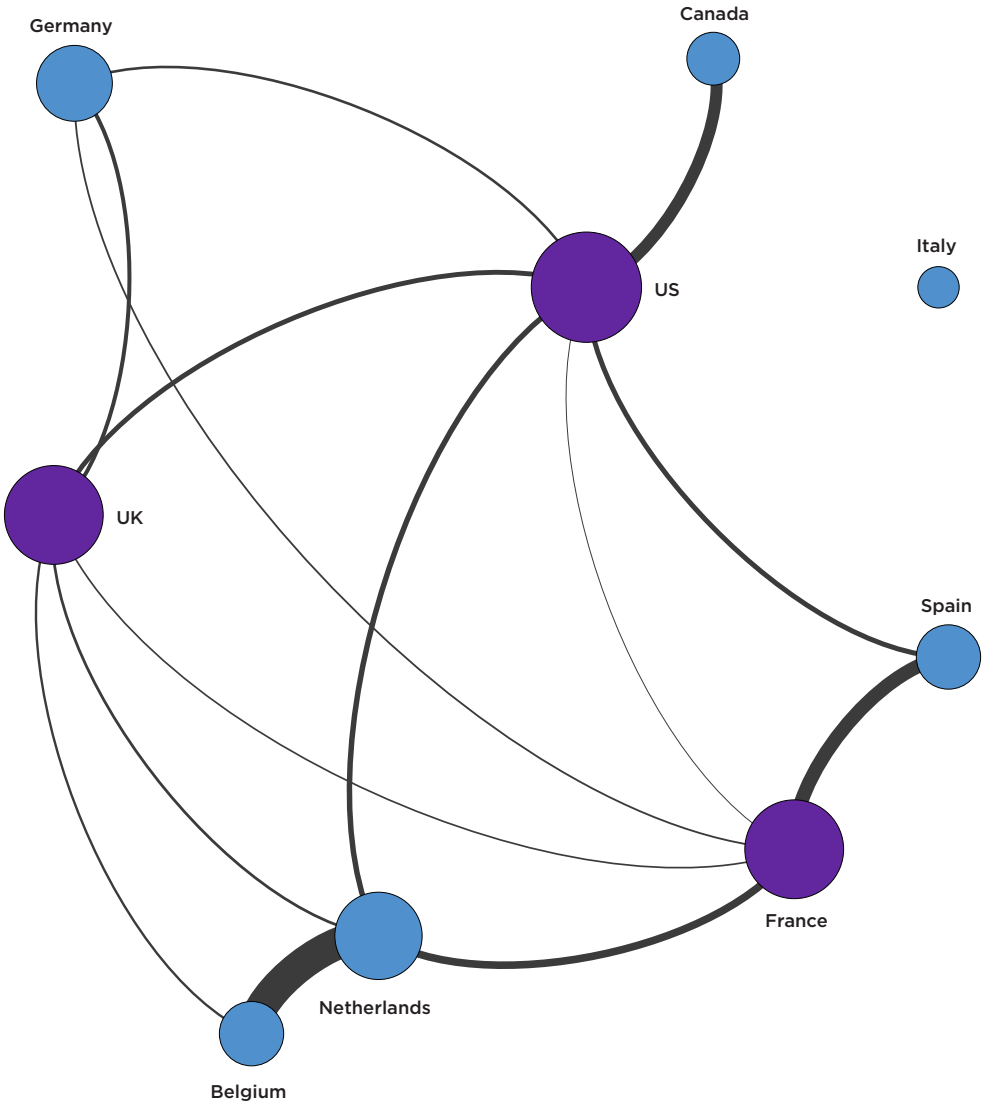
Table 6.1 and Figure 6.1 show that participants from the three countries with the most participants (UK, US and France) connected intensively with others from their own country. The thickness of the edges represents the extent to which people from different countries connected with each other more than we would have expected if they had been connecting at random. The colour of the nodes shows whether people from the same country connected with each other more than expected (in which case the node is purple) or less (in which case it is light-blue). The size of the nodes is according to the number of degrees each group has. Larger nodes were connected to more groups (i.e. more central).

Table 6.1: **Reciprocal connectivity within/between countries**

	Belgium	Canada	France	Germany	Italy	Netherlands	Spain	UK	US
Belgium	0.0								
Canada	0.0	0.0							
France	0.0	0.0	3.5						
Germany	0.0	0.0	0.5	0.0					
Italy	0.0	0.0	0.0	0.0	0.0				
Netherlands	9.3	0.0	2.2	0.0	0.0	0.0			
Spain	0.0	0.0	4.1	0.0	0.0	0.0	0.0		
UK	0.7	0.0	0.5	1.2	0.0	0.8	0.0	1.6	
US	0.0	3.5	0.2	0.8	0.0	1.6	1.5	1.4	1.9

Note: Table excludes the numbers from countries with fewer participants

Figure 6.1: Reciprocal connectivity between main countries represented at LeWeb'12



New connections between people from the same country were particularly visible for French participants, who connected with each other at more than three times the level expected if they were connecting at random. This tendency is also visible within UK and US participants, albeit less marked. Reciprocal connectivity also tended to be more intense along language lines – so, connectivity between UK and US participants was higher, while the opposite was true when considering their reciprocal connectivity with French participants. There were also some high levels of connectivity, which may relate to geographic proximity (France and Spain; Germany, France etc.), although the numbers of connections involved were small.

Which occupations connected at the event?

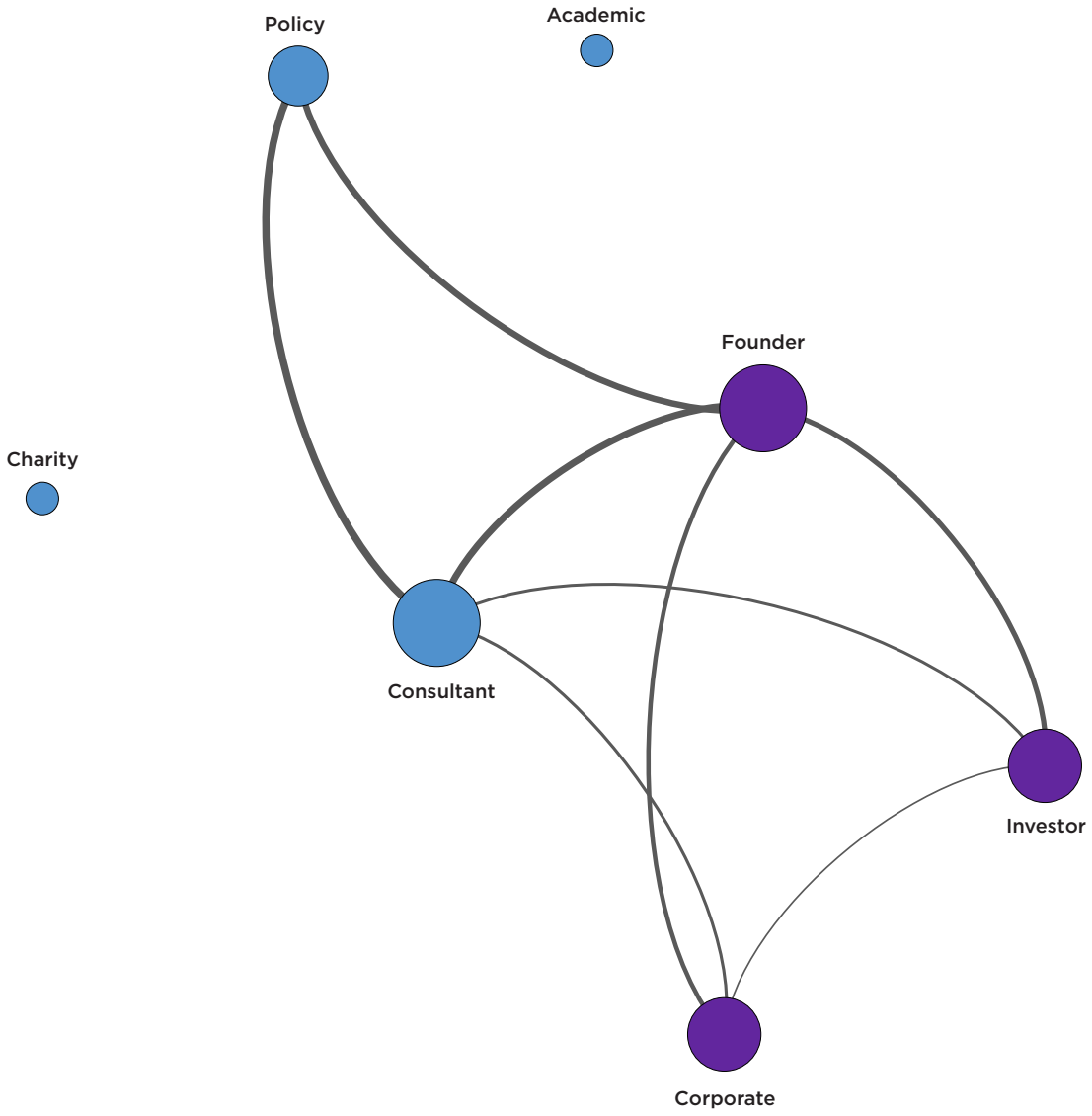
Table 6.2 and Figure 6.2 show reciprocal connectivity within and between occupational roles at LeWeb'12 (the interpretation of edges, colour coding and node size is the same as above).

We see divergence between occupational groups in their propensity to connect internally (with each other) and externally (with other occupational groups). While investors, corporates and entrepreneurs tended to connect within their groups with more intensity than might have been expected based on their attendance numbers in the sample alone, the opposite was true for consultants. This is consistent with the intuition that consultants use events like LeWeb'12 to establish relationships with potential clients, and not with each other. It is interesting to compare their behaviour with investors, who connected with each other at close to twice the expected rate. We know from the literature on venture capital that investors often collaborate formally (through syndicates) and informally (by passing each other investment opportunities), and this appears to have occurred at LeWeb'12 in the relatively intense connectivity within this group that we observe in the data.[41]

Table 6.2: **Reciprocal connectivity within/between roles**

	Academic	Charity	Consultant	Corporate	Founder	Investor	Policy
Academic	0.0						
Charity	0.0	0.0					
Consultant	0.0	0.0	0.7				
Corporate	0.0	0.0	0.6	1.1			
Founder	0.0	0.0	1.3	0.8	1.4		
Investor	0.0	0.0	0.5	0.3	1	1.8	
Policy	0.0	0.0	1.4	0.0	1.2	0.0	0.0

Figure 6.2: **Reciprocal connectivity across roles in LeWeb'12**



Which industries connected at the event?

Table 6.3 and Figure 6.3 show reciprocal connectivity within/between the industries represented by the participants in LeWeb'12 (with the same interpretation as above). As with occupational roles, the propensity of sectors to network internally or externally may have been informed by commercial considerations.

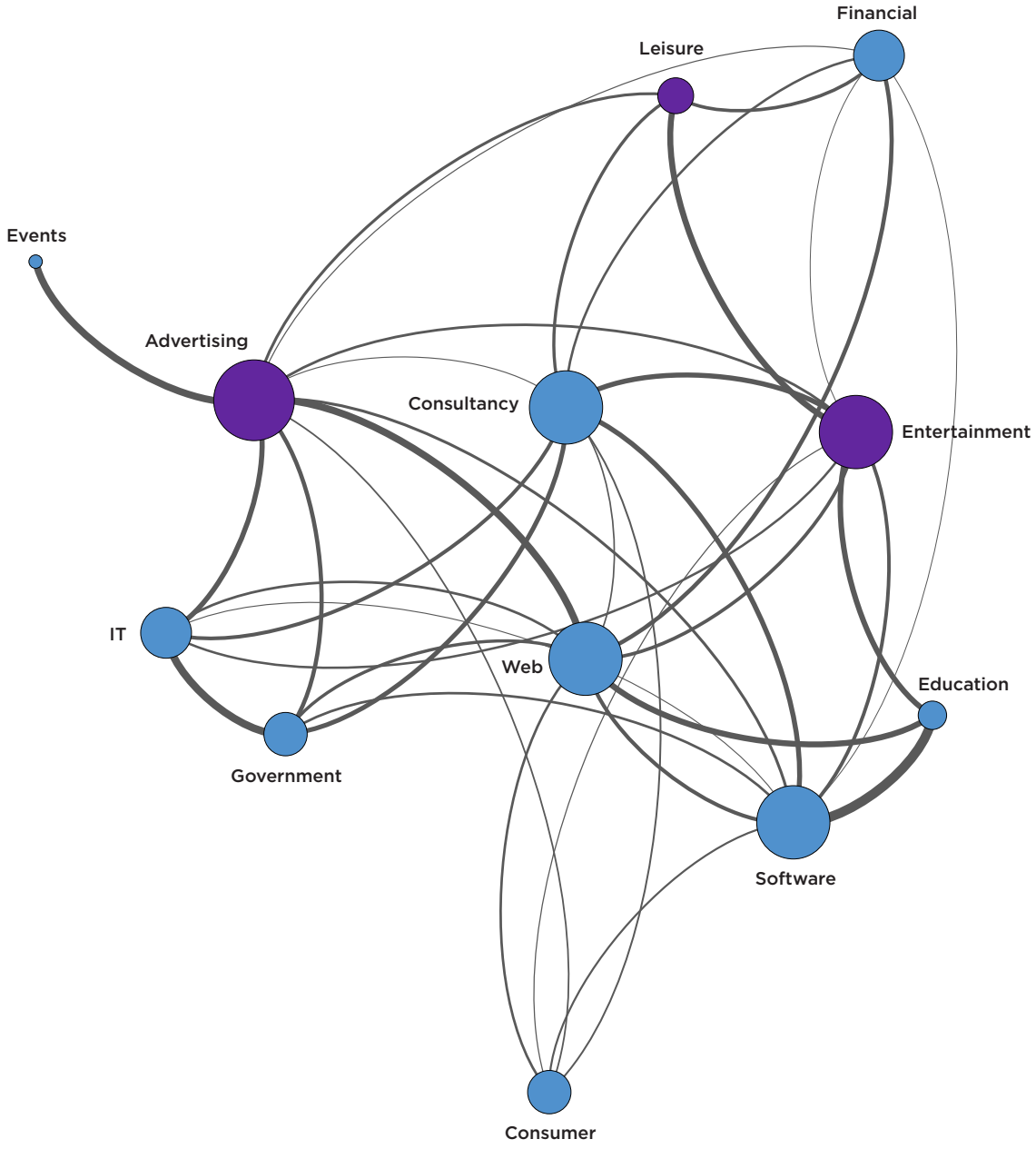
Individuals from consultancy and business services, web companies and software companies were less likely at LeWeb'12 to network with their industry competitors, perhaps because they preferred to use the opportunity to look for clients or collaborators in other sectors. By contrast, advertising and marketing companies, entertainment and media, and leisure and travel all engaged in high levels of networking within their own industries.

When we consider connectivity between groups (perhaps easier to see in Figure 6.3), our data reveals a 'hub' of inter–sector connectivity involving digital sectors (like software and web), and application sectors (entertainment and media, advertising and marketing). This is particularly the case between advertising and marketing and web businesses, and between entertainment and media and software companies. Again, these patterns of connectivity are consistent with business and technological considerations, such as opportunities merging from the convergence of advertising and web (social media) platforms, and the digitisation of the entertainment and media sector.

Table 6.3: **Reciprocal connectivity within/between industries**

	Advertising & Marketing	Consultancy & Business Services	Consumer goods & services	Education	Entertainment and Media	Events	Financial services	Government	IT & Telecoms	Leisure and travel	Software	Web
Advertising & Marketing	2.8											
Consultancy & Business Services	0.4	0.9										
Consumer goods & services	0.6	0.7	0.0									
Education	0.0	0.0	0.0	0.0								
Entertainment and Media	1	1.9	0.4	2	2.1							
Events	2.6	0.0	0.0	0.0	0.0	0.0						
Financial services	0.4	0.9	0.0	0.0	0.3	0.0	1					
Government	1.6	1.7	0.0	0.0	0.0	0.0	0.0	0.0				
IT & Telecoms	1.9	1.4	0.0	0.0	0.9	0.0	0.0	2.8	0.0			
Leisure and travel	1.1	1.2	0.0	0.0	2.2	0.0	1.3	0.0	0.0	7.1		
Software	1.1	1.9	0.7	3.5	1.3	0.0	0.2	0.9	0.4	0.0	0.8	
Web	2.6	0.6	0.9	2.2	1.3	0.0	1.6	1.2	0.9	0.0	1.4	0.8

Figure 6.3: **Reciprocal connectivity between industrial sectors**



7. WHAT WERE THE COMMUNICATION FLOWS BETWEEN PARTICIPANTS AT LeWeb'12?

What we did:

- Using the tweets that directly related to event participants (tweets that mentioned, replied or were retweets of event participants by other participants), we reconstructed the levels of communication on Twitter and the structure of communication flows on Twitter between participants.

What we found:

- Twitter communication activity was strongest among participants during the event itself, and during event sessions in particular, when there was generally less scope for offline interaction. Communication activity was higher on the first day of the event than on the second.
- A significant minority (around one-third) of new one-way following connections formed between people at LeWeb'12 had some form of Twitter communication interaction since the start of the event. This was a higher level of interaction than that for pre-existing connections, though this could also reflect older connections on Twitter in general being less active, rather than an impact of the event.
- Completely new reciprocal connections had the highest proportion of reciprocal communication (27 per cent of connections were associated with reciprocal communications), more than connections that existed prior to the event or, new one-way follow connections (7 per cent).
- The first Twitter communication between those who followed one another at the event was most likely to have been at the event (in the case of new reciprocal connections or one-way follow connections.) The first communication between reciprocal follow connections formed from a consolidation of an existing follow connection was likely to have been before the event.
- The analysis of tweet content suggests that a number of the people may have arranged to meet as a result of attending the event.

So far, we have looked at the creation of Twitter follow connections between individuals as a proxy for connections made more generally at LeWeb'12. Here, we consider the communications between these individuals. We are interested in this for at least two reasons:

1. Follow relationships are not a complete picture of the strength of relationships between individuals, or whether they endure. By contrast, communication between individuals on Twitter perhaps indicates a deeper level of engagement. The extent to which there is two-way communication may also tell us something about the strength of the relationship.

2. Communication flows across newly created connections vary in content, and may be more interesting if, for example, they contain information on people meeting up, new technology, business opportunities or market developments.

We collected data on three different kinds of Twitter communication interaction between LeWeb'12 participants:

1. **Mentions**: Where one participant included another person's Twitter user name in their tweet e.g. 'Hello @LeWebparticipant'.
2. **Replies**: A form of mention, where the tweet began with the user name of the person being replied to e.g. '@LeWebparticipant, have you read this new report from Nesta?'
3. **Retweets**: Where a participant retweeted a tweet from another participant i.e. where a tweet from one participant was rebroadcast by another to their own followers.

The data on these was collected from 19 December 2011 to 23 September 2012 (LeWeb'12 having taken place on 19–20 June 2012). This resulted in a dataset of:

- **13,138 mentions**[42]
- **4,746 replies**
- **2,571 retweets**

It is interesting that there are many more mentions than retweets. Perhaps this reflects the personal interactivity aspects of LeWeb, including attendees' interest in raising their visibility with each other in a more direct and targeted way by mentions (where the sender has to type out the mentioned individual's Twitter handle) than a retweet (which just requires a button press).

Table 7.1 shows the tweeting activity in the months before, during, and after LeWeb'12. For all forms of communication there was a peak in activity in June which, as can be seen from the table, related in large part to the Twitter activity over the two days of the event itself.

Table 7.1: **Number of mentions, replies and retweets, before during and after LeWeb'12** (the bracketed numbers are the figures for the two days of LeWeb – 19–20 June)

	Dec	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Total
Retweets	40	206	187	234	222	288	629 (302)	307	223	235	2571
Replies	135	341	317	455	407	491	1420 (628)	455	383	342	4746
Mentions	268	930	875	1249	1060	1328	4130 (2100)	1277	999	1022	13138

Figure 7.1: **Levels of monthly mentions, replies and retweets communication on Twitter**

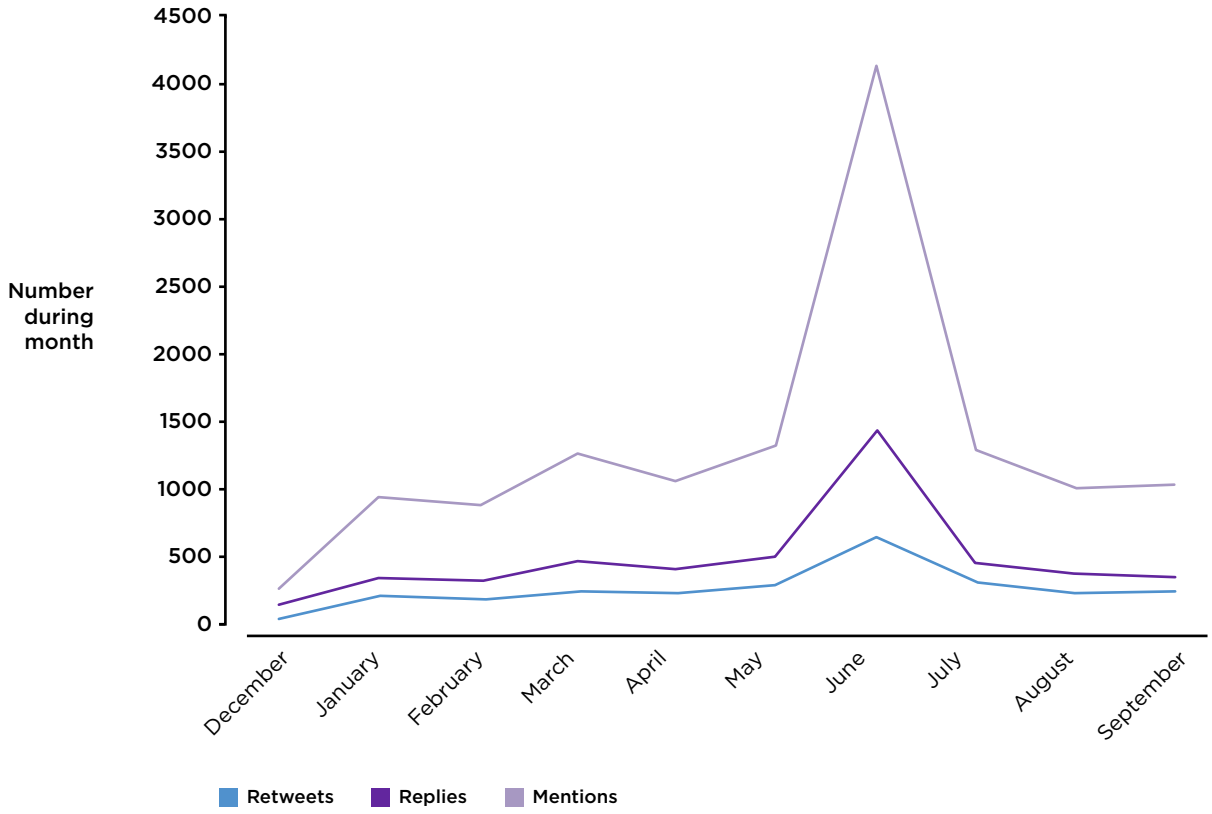
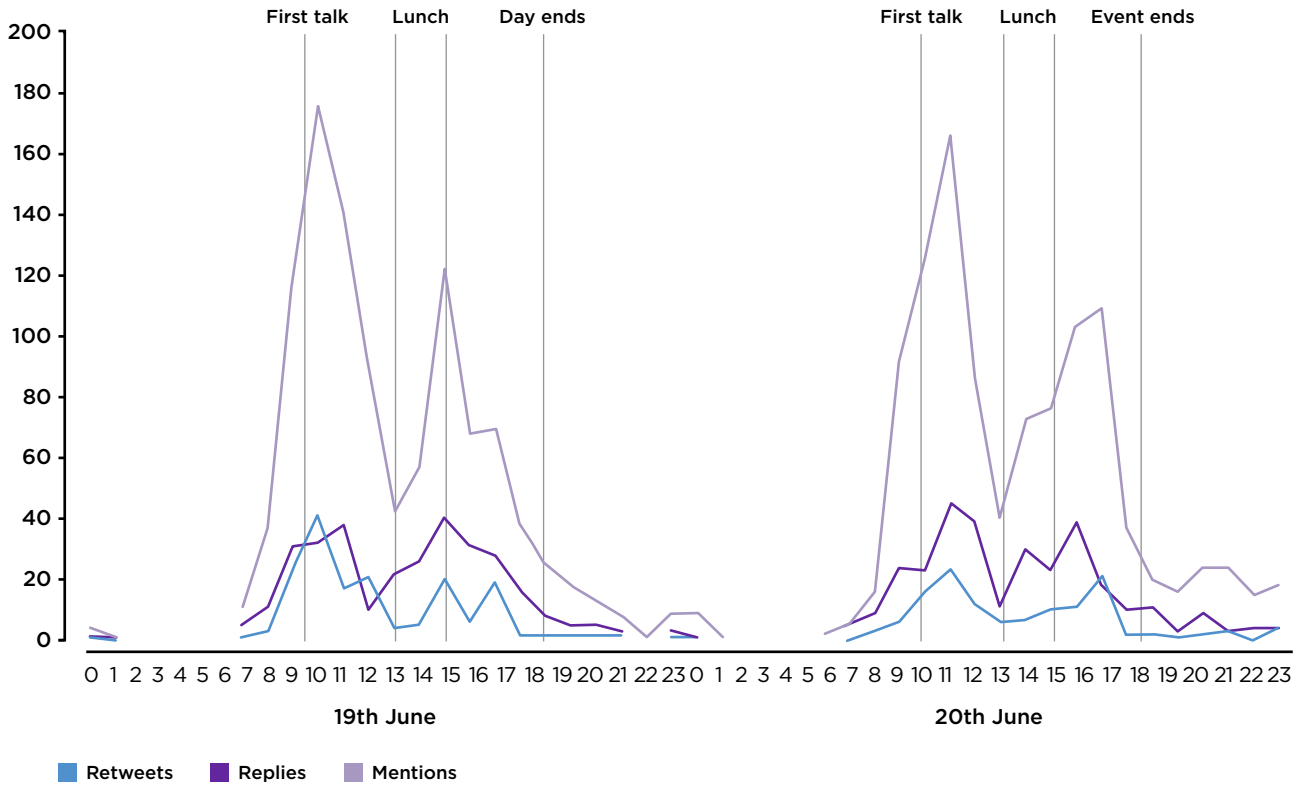


Figure 7.2 displays the levels of Twitter mentions activity over the two days of the event. It shows that activity peaked during the first sessions of the day and fell during lunchtime and after the event had ended. There was also slightly more activity on the first day of the conference than on the second.

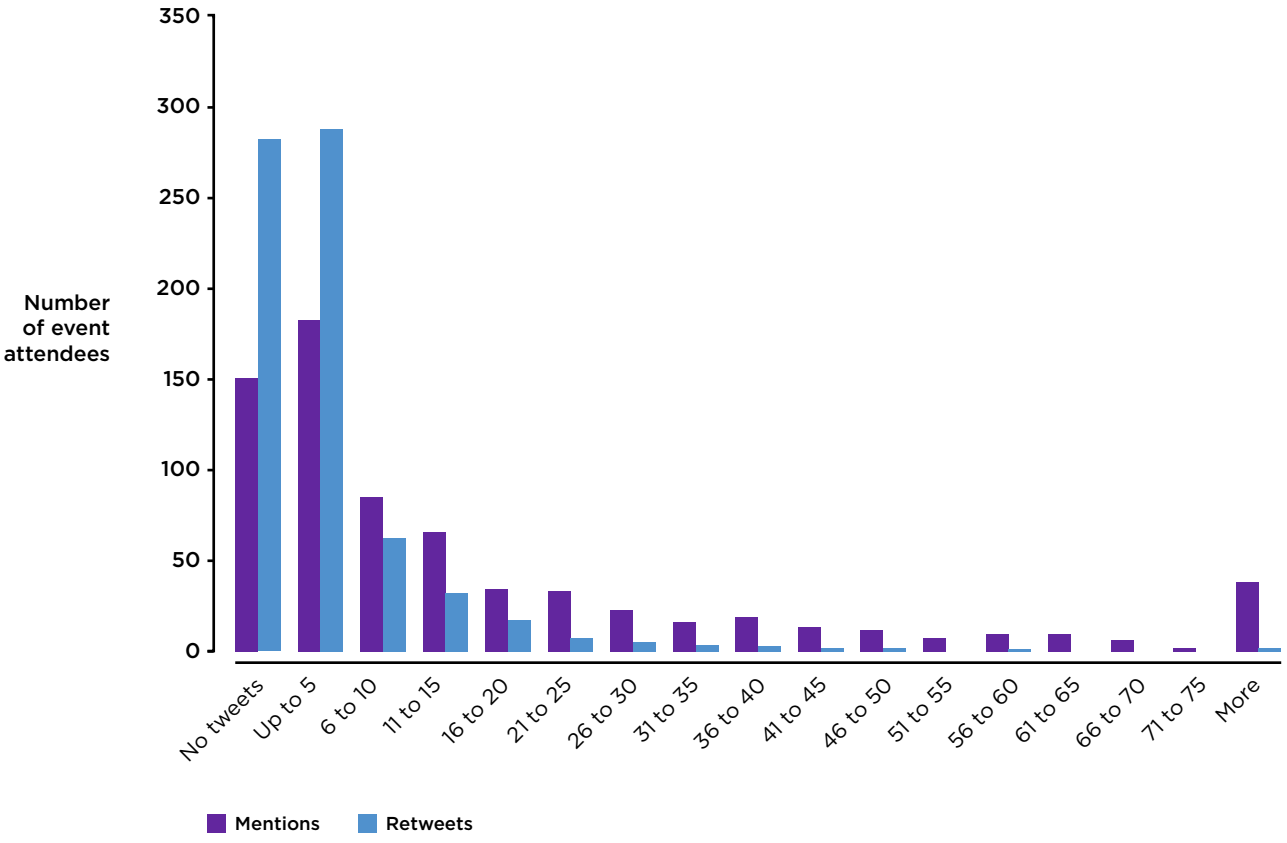
Figure 7.2: **Levels of tweeting activity during LeWeb'12**



Distribution of mentions and retweets across participants

Figure 7.3 shows the distribution of retweets and mentions across LeWeb'12 participants on Twitter over the period of data collection. Large numbers of participants did not actively use Twitter to interact with fellow participants at all, or only did so with a small number of tweets. There was, however, a long tail of people involved in high levels of mentioning and retweeting of other participants.

Figure 7.3: **Distribution of mentions and retweets across LeWeb'12 participants on Twitter**

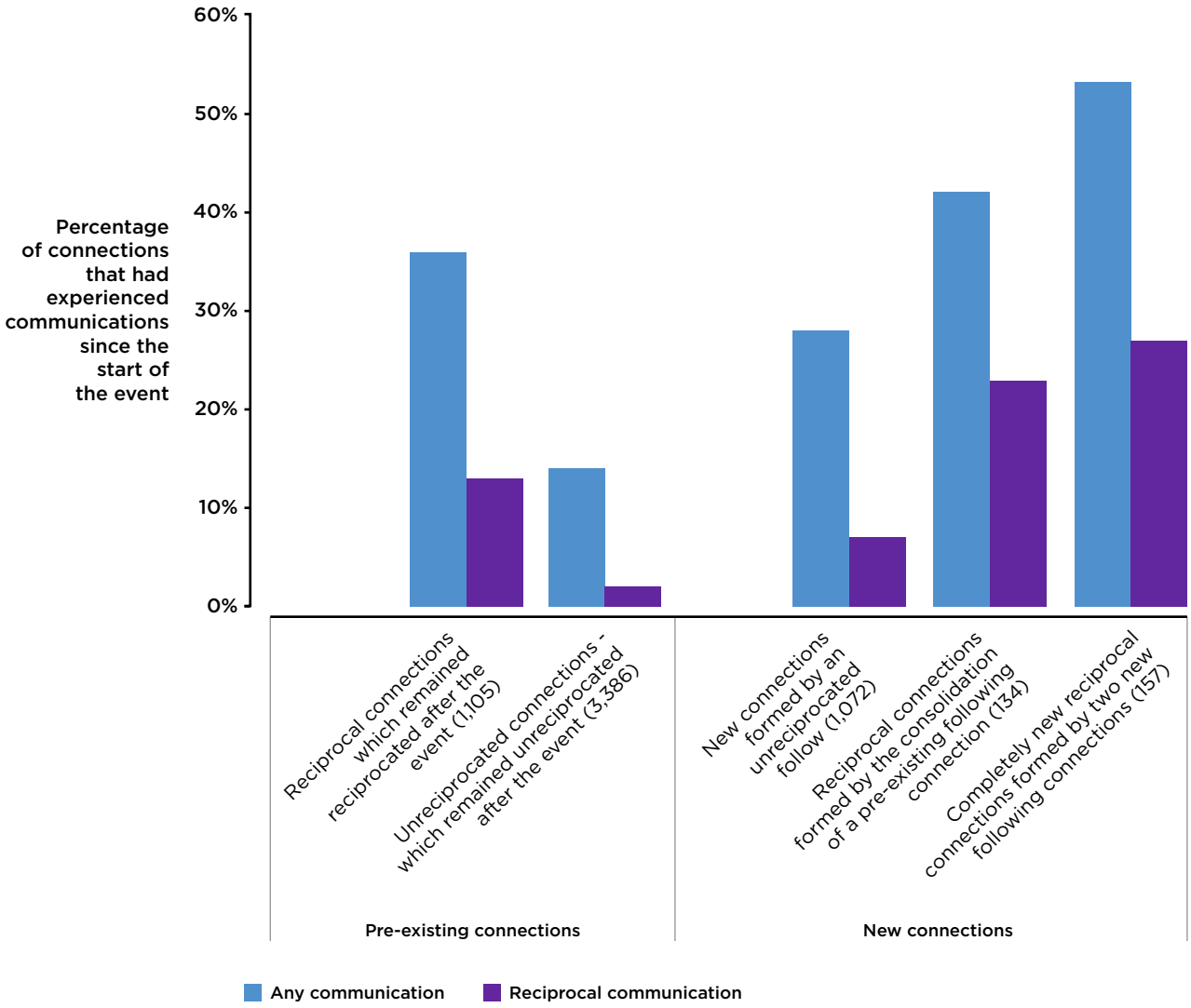


The relationship between tweeting behaviour and new follow connections created at the event

To examine the extent to which the new follow connections generated at the event also involved active communications, Figure 7.4 shows communication flows involving the different kinds of follow connections that existed before LeWeb'12 and those that were created at the event (communication being measured in terms of the three types of Twitter communication – retweets, replies and mentions – that had occurred since the start of LeWeb'12). 'Any form of communication' is defined as having occurred if at least one of these three kinds of communication was made between one person and another. 'Reciprocal communication' occurred where communication passed in both directions one or more times. Separate results for each of the three different kinds of Twitter communication are set out in Appendix 3.

A follow connection is counted as having existed between two people if there was at least one follow connection between them (be it reciprocal between the two people or one-way) – see the classification of new connections on page 18.

Figure 7.4: Communication levels between participants by follow connection type since start of LeWeb'12



Note: In this Figure all follow connections are undirected i.e. reciprocal connections and unreciprocated connections both count as one connection.

Figure 7.4 confirms that the majority of the follow connections in the pre–LeWeb'12 follow network, and most of the new connections created after the event, did not experience any communication over them. The proportion of connections that experienced communication flows between them after the event was higher for new connections than it was for pre–existing ones. It was also higher the more substantive the connection. New reciprocal connections – either completely new or formed from consolidating an existing connection – were more likely to have had some communication interaction than unreciprocated follows.

The new reciprocal connections created at the event (including completely new reciprocated connections and those created from the consolidation of an existing follow relationship) were also more likely to involve reciprocal communications than connections that pre–dated the event or the unreciprocated connections formed at the event.

In the pre-LeWeb'12 network, there were 1,105 reciprocal connections between participants (consisting of 2,210 individual follow connections). Of these 1,105, 143 (13 per cent) had experienced some reciprocal communication since the start of the event. This is consistent with the idea that new connections had enabled the mutual sharing of useful information. However, the lower levels of communication relative to the reciprocal connections formed at the event may also reflect the possibility that communications between connected individuals generally declines with the age of a connection.

In terms of the number of people that engaged in reciprocal communication for the entire period for which we have Twitter data, around one-third of the sample engaged in no reciprocal communication at all with other event participants, and another third engaged in reciprocal communication with one to two people.

The relationship between Twitter following and first communication interaction

As we do not know the precise time that the follow relationships were created, it is possible that the communication preceded the formation of connections, and in some instances the causality will run from communication leading to a follow on Twitter. Table 7.3 shows for different kinds of connections the timing of first communication (as measured by the date of the first retweet, reply or mention between people).

Table 7.3 Timing of first communication on Twitter

First communication	Before event (Dec to June 18)	During event (19 to 20 June)	After event (June 21 to 23 Sept)
Completely new reciprocal follow connections: 157	9%	50%	41%
Reciprocal connections formed by consolidation of a pre-existing follow connections: 134	61%	16%	23%
Unreciprocated one-way follow connections: 1,072	9%	57%	34%
Pre-existing follow connections: 4,491	76%	10%	14%

It shows that the completely new reciprocal connections and the new unreciprocated follow connections were those where the first communication between the two people involved was most likely to be at or after LeWeb'12 (only a small percentage of these connections had communication before the event). Intuitively, reciprocal connections formed by consolidation of a pre-existing follow relationship, or connections that existed before the event are more likely to be associated with communications before the event.

Offline networking

In addition to examining if participants at LeWeb'12 communicated with each other, we also extracted the content of @mentions between participants to examine whether they had used keywords suggesting that they may have met and/or were planning to meet face–to–face (such as 'meet', 'meeting' or 'coffee'). Some examples include:[43]

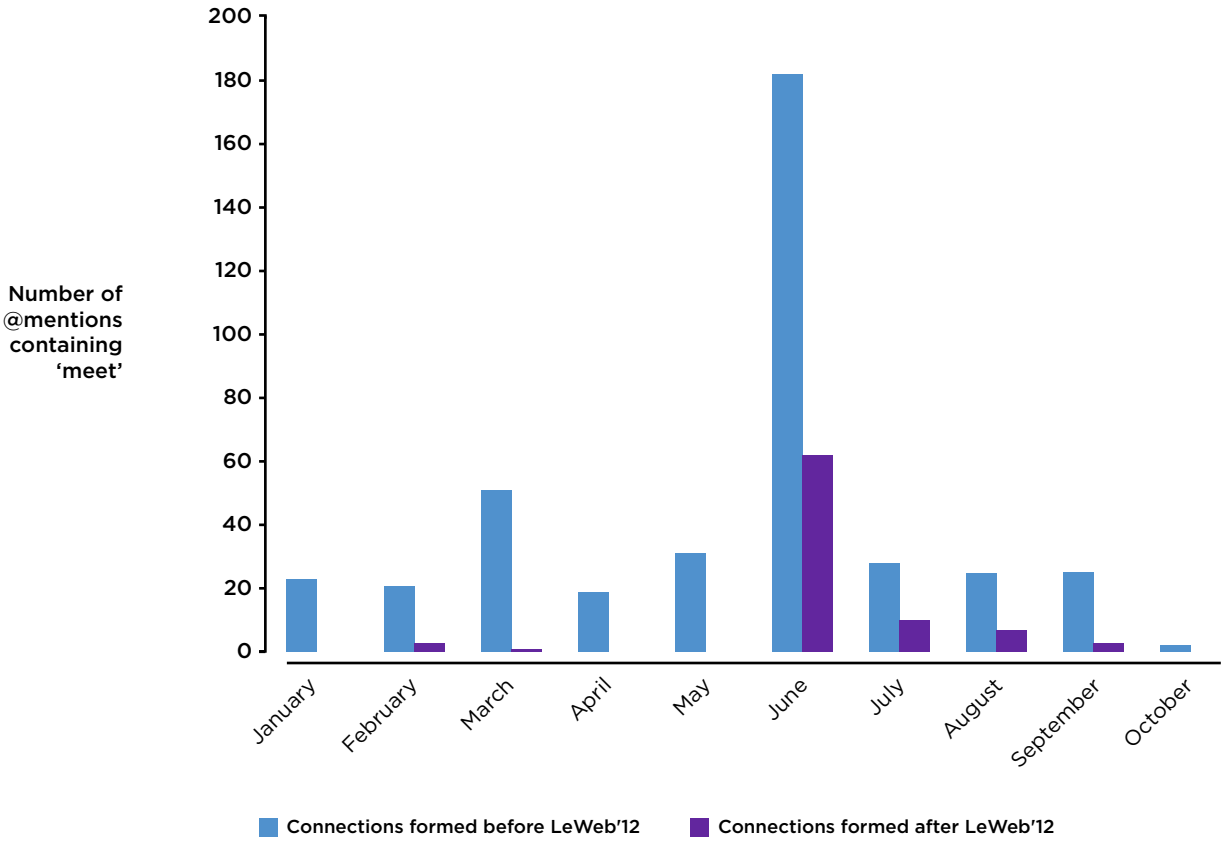
@user_y Nice meeting you at dinner last night! Hope to catch you on your SF LA visit :)

@user_z hey X, just to confirm, we are meeting in 10 min at box networking zone?

As discussed in Section 2, decisions to take a conversation 'offline' suggest that those involved saw some value in spending time with each other (the opportunity cost of having a face–to–face meeting is higher than following someone, or communicating with them on Twitter).

Figure 7.5 plots the number of occurrences of such terms over the period for which we have data, distinguishing between communications involving those previously connected in the follow network, and individuals that connected at the event.

Figure 7.5: **Number of @mentions including keywords related to 'meeting'**



Nine per cent of the mentions between newly connected people at LeWeb'12 included meeting–related keywords (this drops to 3.5 per cent if we exclude from the calculation mentions in the month of June in which LeWeb'12 took place). In comparison, a smaller 4 per cent of the mentions between those who had been connected before LeWeb'12 included meeting–related terms (falling to 2 per cent excluding June).

This supports the idea that people who connected with each other at, or shortly after, the event on Twitter, were more likely to meet than people who were already connected on Twitter. However, an alternative interpretation is that people who were already connected on Twitter before the event were less likely to use Twitter as a way of arranging a meeting because the relationship was of longer standing and therefore more likely to have other communication channels, such as email or phone.

8. CONCLUSIONS AND NEXT STEPS

In this paper we have shown how the data that is freely accessible from a social media platform like Twitter can provide rich insights into the impact of events on participants' social networks. These insights may be an important element in evaluating the impact of events like LeWeb'12 on collaborations and innovation, and as such should be of interest to event organisers, prospective attendees and innovation policymakers.

By collecting Twitter data from before and after the event (which was possible in this instance as people were asked to provide their Twitter handle when registering their attendance) we can assess the event's impact on social networks. We are able to gauge attendees' awareness of each other and whether this awareness is mutual by studying the connections and communications on Twitter, with a particular focus on reciprocal following and communication. We can assess the likely additional impact of the event (whether people were likely to have met each other in any case) by examining the distance between them in the pre-event network. By matching the Twitter data with data from other sources we can also analyse the patterns of network formation between different groups (such as occupations and industries) at the event. We can track the extent to which connections formed at the event led to communications (one-way or reciprocal) between participants and can examine the content of these communications.

This type of analysis illustrates how the use of information from social media platforms such as Twitter holds the promise of rigorous methodologies for tracking the impact of events and other networking initiatives. As well as potentially enabling the systematic benchmarking of events, allowing event organisers to differentiate their events from others, this will be of value to funders and agencies charged with supporting innovation by stimulating knowledge sharing and through developing collaborative networks. Ultimately, this type of event analytics information should also be valuable to participants choosing between events. It also illustrates the value that can be created from public access to the data generated by platforms like Twitter.

We are unable in this study to distinguish networking that was facilitated by the use of Twitter at the event from networking that was entirely due to the event itself (and would have happened without Twitter). This is a potentially important distinction, as we know that Twitter is changing the dynamics of events, enabling as it does public conversations among the audience during talks, and making it easier for people to connect with those they do not know. How event organisers can best structure their events to facilitate networking on Twitter (and other social networks), and hence networking at an event overall, is important.

Overall, we have shown that LeWeb'12 was associated with the formation of a large number of connections between participants on Twitter. This improved the network of connectivity between delegates. There is evidence that a significant proportion of these connections would have happened anyway in the longer term, as the people forming them were not far apart in the Twitter network before the event. Nevertheless, many were further away and so more likely to have formed as a result of the event itself. This is particularly the case for connections between different nationalities. Reciprocal Twitter connections were found to be associated with higher levels of communication and in terms of maximising event impact it may therefore be beneficial for organisers to focus on encouraging such connections at events.

Limitations and next steps

The research in this paper has uncovered a rich picture of the networking and communication flows that took place at LeWeb'12 London. Attributing network impacts to any one event, however, presents significant data-related and analytical challenges, and the limitations of our analysis must be borne in mind. Aside from the inherent constraints in using Twitter data for this purpose – namely, that Twitter is not necessarily representative of people's social networks – we would highlight the following methodological limitations:

1. There is no control group in this study – that is, we only track the Twitter networks of LeWeb' 12 participants. To rigorously assess the additional impact of the event on attendees' social networks would require tracking the Twitter network of a group of non-attendees with otherwise identical characteristics to event attendees, insofar as these relate to the propensity to connect with each other.
2. We do not model individual behaviour and network dynamics. Although a number of measures of network change are analysed, these are not based on any explicit analysis of individuals' decisions and how they are affected by the behaviour of others in the network. Completely new reciprocal connections are, for example, not in general created spontaneously between people – one person will typically follow the other first.
3. There is no measure of innovation outcomes. Although we can observe the impact of LeWeb'12 on Twitter connections, and communications between people on Twitter, we cannot directly observe innovation outcomes that may have arisen from connections made at the event.

We attempt to address these issues in a forthcoming Nesta working paper by Greenwich Business School researchers.[44] In particular, this paper evaluates the impact of the LeWeb Paris 2014 conference and:

1. **Has a control group**: The paper divides LeWeb'12 London attendees into those who did and did not go on to attend the LeWeb Paris 2014 conference. London 2012 participants who did not attend LeWeb Paris'14 are treated as a control group for the effect of attending LeWeb Paris 2014.
2. **Models individual behaviour**: The paper statistically models individual agent behaviour using Stochastic Actor-Orientated modelling (SAOM) to understand how the event affected the interactions between people on Twitter at the event.[45]
3. **Has a measure of innovation outcomes**: The paper links attendance at LeWeb'14 Paris to whether participants collaborated on the online code sharing platform GitHub. This allows whether attendance at LeWeb Paris led to greater collaboration to be explicitly analysed, relative to a random control group of GitHub participants. The performance of the collaborations is also measured by how many times the code that was created had gone on to be reused.

BIBLIOGRAPHY

Amalaya, O. and Montgomery, K. (2008) Using field-configuring events for sense-making: a cognitive network approach. 'Journal of Management Studies.'

Aral, S. and Walker, D. (2011) Identifying Social Influence in Networks Using Randomized Experiments. 'IEEE Intelligent Systems.' Vol.6, Issue 5.

Aral, S., Muchnika, L. and Sundararajana, A. (2009) 'Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks.' Proceedings of the National Academy of Sciences of the United States of America (PNAS).

Bastian, M., Heymann, S. and Jacomy M. (2009) 'Gephi: an open source software for exploring and manipulating networks.' International AAAI Conference on Weblogs and Social Media.

Boschma, R. (2005) Proximity and innovation: a critical assessment. 'Regional Studies. 39:1, 61–74.

Brown, C., Efstratiou, C., Leontiadis, I., Quercia, D., Mascolo, C., Scott, J. and Key, P. (2014) 'The architecture of innovation.' Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing – UbiComp '14 Adjunct. ACM – Association for Computing Machinery pp. 811–822.

Chon, B. S., Choi, J. H., Barnett, G. A., Danowski, J. A. and Joo, S. H. (2003) A structural analysis of media convergence: Cross-industry mergers and acquisitions in the information industries. 'The Journal of Media Economics.' Vol. 16, Issue 3.

Cramers, M. (2012) 'LeWeb Paris vs LeWeb London: The Social Media Scorecard.'

Cronin, B., De Vita, R. and Conaldi, G. (2015) 'Joining up the dots: Using social data to measure the effects of events on innovation.' London: Nesta. Working paper.

Csardi, G. and Nepusz, T. (2006) The igraph software package for complex network research. 'InterJournal.' Complex Systems 1695.

Easley, D. and Kleinberg, J. (2010) 'Networks, Crowds, and Markets Reasoning About a Highly Connected World.' Cambridge: Cambridge University Press.

Ebner, M., Mühlburger, H., Schaffert, S., Schiefner, M., Reinhardt, W. and Wheeler, S. (2010) 'Getting Granular on Twitter: Tweets from a Conference and Their Limited Usefulness for Non-participant.' In Reynolds, N. and Turcsányi-Szabó, M. (Eds.) 'Key Competencies in the Knowledge Society.'

Evers, N. and Knight, J. (2008) Role of international trade shows in small firm internationalization: a network perspective. 'International Marketing Review.' Vol. 25.

Granovetter, M. (1973) The strength of weak ties. 'American Journal of Sociology.' Vol. 78, Issue 6.

Hidalgo, C. (2011) 'The value in the links: Networks and the evolution of organizations.' In Allen, P., Maguire, S. and McKelvey, B. (Eds.) 'SAGE Handbook of Complexity and Management.' London: SAGE Publications Ltd.

Hochberg, A., Liungqviust, A. and Lu, Y. (2007) Whom You Know Matters: Venture Capital Networks and Investment Performance. 'The Journal of Finance.' Vol. 62, No.1.

Huberman, B., Romero, D. and Wu, F. (2013) 'Social networks that matter: Twitter under the microscope.' arXiv preprint arXiv:0812.1045.

Kossinets, K. and Watts, D. (2006) Empirical analysis of an evolving social network. 'Science.' Vol.311, No. 5757.

Kwak, H., Lee, C., Park, H. and Moon, S. (2010) 'What is Twitter, a social network or a news media?' Proceedings of the 19th International Conference on World Wide Web.

Letierce, J., Passant, A., Breslin, J. G. and Decker, S. (2010) 'Using Twitter During an Academic Conference: The #iswc2009 Use–Case'.

Leskovec, J., Backstrom, L., Kumar, R. and Tonkins, A. (2008) 'Microscopic evolution of Social Networks.' Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Maskella, P., Bathelt, H. and Malmberg, A. (2006) 'Building global knowledge pipelines: The role of temporary clusters.' DRUID Working Paper No.05–20.

Newman, M. (2001) 'The structure of scientific collaboration networks.' Proceedings of the National Academy of Sciences of the United States of America (PNAS).

Shalizi, C. and Thomas, A. (2011) Homophily and Contagion Are Generically Confounded in Observational Social Network Studies. 'Sociological Methods & Research.' 40 (2), 211–239.

Snijders, T. (2001) The Statistical Evaluation of Social Network Dynamics. 'Sociological Methodology.'

Storper, M. and Venables, T. (2004) Buzz: Face to Face contact and the urban economy. 'Journal of Economic Geography.' Oxford University Press, Vol. 4.

Vega–Redondo, F. (2007) 'Complex Social Networks.' Vol. 44 of Econometric Society Monographs. Cambridge: Cambridge University Press.

Weller, S. (2008) Beyond 'Global Production Networks': Australian Fashion Week's Trans–Sectoral Strategies. 'Growth and Change.' Vol. 39(1): 104–122.

APPENDIX 1: INFORMATION COLLECTED IN SECONDARY DATA COLLECTION

Table A1: **Secondary participant data**

Field	Question	Source/Observations
Twitter active	Has this person published a tweet in the last month?	Twitter
Tweets	How many tweets has this person published?	Twitter
Twitter friends	How many people does this person follow in Twitter?	Twitter
Twitter followers	How many people follow this person in Twitter?	Twitter
Age	How old is this person?	From LinkedIn. This was estimated as: 2013 – Year of graduation from university + 21 (we used sizebands rather than continuous estimates)
Gender	What is this person's gender?	LinkedIn
Country of Origin	Where was this person born?	From LinkedIn. Estimate based on location of university education/mother tongue in LinkedIn/country in LeWeb information.
Highest level	What is the highest academic qualification that this person has?	LinkedIn
Year of qualification	When did this person obtain their most recent qualification?	LinkedIn
Institution	In what institution did this person obtain their most recent qualification (name, country)?	LinkedIn
Subject	In what subject did this person obtain their most recent qualification?	LinkedIn
Position	What is this person's current position or job title?	LinkedIn
Employer	What organisation currently employs this person?	LinkedIn
Location of work	Where is this person based?	LinkedIn
Role	What is the role of the respondent in the 'innovation ecosystem'?	LinkedIn/CrunchBase
Change in job	Has this person changed jobs since LeWeb'12?	LinkedIn
Industry	In what sector does the company operate?	LinkedIn/CrunchBase
Company age	When was the company established?	LinkedIn/CrunchBase
Company size	How large is the company?	LinkedIn/CrunchBase

APPENDIX 2: TESTING THE ROBUSTNESS OF OUR MEASURE OF DISTANCE

We have explored the robustness of our measure of distance by estimating the impact of personal differences between individuals on the probability that they were connected before LeWeb'12, and on the probability that they were indirectly connected (i.e. that they at least shared a mutual connection before LeWeb'12), considering 245,350 possible pairs of individuals that could have been connected (directly or indirectly) before LeWeb'12.[46] The intuition here is that if our measures of pre–event distance are capturing the fact that individuals have different backgrounds and are in different social circles, then we would expect to find a connection between personal differences in their country of origin, industry, role, age, and their likelihood to have been directly or indirectly connected before LeWeb'12 etc.

If, however, the distance measures were not related to personal characteristics in some way, then this would cast doubt on whether the Twitter distance before LeWeb'12 is an informative measure of people's proximity in social networks.

Table A2.1 shows the outputs of a logistic regression. Our outcome variable is whether individuals were directly or indirectly connected before LeWeb'12 through a reciprocal connection (columns one and two), or through a one–way connections (columns three and four).[47] Our predictors are binary variables indicating whether those two individuals were different in their personal characteristics (see Appendix 1, Table A1 for information).[48] We find significant links between several indicators of personal difference and follow/reciprocal connections before LeWeb'12.

Table A2.1 There is a relationship between network distances and several measures of personal difference

	Dependent variable: Pre-LeWeb'12 connectivity			
	Rec-dir (1)	Rec-ind (2)	One-dir (3)	One-ind (4)
Different age	-0.166 (0.118)	-0.040 (0.050)	-0.017 (0.078)	0.036 (0.034)
Different gender	0.434*** (0.133)	0.340*** (0.082)	0.051 (0.115)	-0.047 (0.065)
Different industry	-0.459*** (0.141)	-0.035 (0.069)	-0.264*** (0.078)	0.131*** (0.042)
Different country of origin	-1.457*** (0.140)	-0.895*** (0.092)	-0.900*** (0.101)	-0.433*** (0.061)
Different org. size	-0.035 (0.119)	-0.034 (0.080)	-0.211** (0.090)	-0.136*** (0.048)
Different role	-0.123 (0.127)	0.025 (0.073)	0.140* (0.083)	-0.016 (0.044)
Different discipline	0.119 (0.134)	0.094 (0.089)	0.158* (0.090)	0.017 (0.051)
Constant	-3.974*** (0.144)	-2.580*** (0.110)	-2.997*** (0.107)	-0.267*** (0.090)
1=Different country of origin				
Observations	107,343	107,343	107,343	107,343
R^2	0.048	0.023	0.020	0.010
chi^2 (df = 7)	295.572***	718.384***	404.249***	768.103***

Note: * (p<0.1), ** (p<0.05),*** (p<0.01)

- Differences in gender between individuals meant that their odds of being directly reciprocally connected in the pre-LeWeb'12 network were around 50 per cent higher than if they had been from the same gender. This is probably linked to the disparities between genders at LeWeb'12 we indicated in Section 3.
- Differences in nationalities between a pair of individuals mean that their odds of being reciprocally connected in the pre-LeWeb'12 network were a quarter less than if they had the same country of origin.[49] The odds that individuals from different nationalities would be at least indirectly connected in the pre-LeWeb'12 network were 40 per cent less than if they had the same nationalities.
- The other variable that appears to be systematically connected to direct and indirect connectivity in pre-LeWeb'12 networks is whether individuals were in the same industry or not. In this case, individuals working in different industries had odds of less than half to be directly connected in the pre-LeWeb'12 reciprocal network than if they had been in the same industry.[50] Interestingly, when we consider the one-way network, although individuals

from different industries were less likely to be directly connected, they were more likely to be indirectly connected (i.e. to share a connection) – one potential explanation for this is the presence of boundary–spanning participants at the event, such as consultants and journalists, acting as bridges between industries.

Going beyond this, working for a larger company was negatively associated with being connected before LeWeb'12, but only when we consider one–way follow connections – that is to say, individuals were significantly less likely to follow those in companies of a different size, but not less likely to be reciprocally connected with them. It is interesting that differences in roles and discipline of study were both positively associated to proximity in the one–way network – the odds for an individual to follow or be followed by another were 15 per cent higher when they belonged to different types of organisations (e.g. investors and corporates) – this suggests that participants might have been using Twitter to obtain information flows from outside their area of activity before LeWeb'12, or perhaps it could indicate an aspiration to connect with those other groups. This is what one might expect in the case of entrepreneurs and investors, or of entrepreneurs and corporate clients. We have explored some of these issues further in Section 6.

APPENDIX 3

Table A3.1: The number of connections that experienced communication since the start of LeWeb'12

	Old connections 4991		New connection formed 1363		
Percentage of connections with communication in brackets	Pre-existing reciprocal connection from before event that existed after the event: 1105	Pre-existing unreciprocated connection from before event that existed after the event: 3386 (exc of those consolidated at event)	1. New connections formed by an unreciprocated follow connection: 1072	2. Reciprocal connections formed by the consolidation of a pre-existing follow connection: 134	3.Completely new reciprocal connections formed by two new follow connections: 157
Mentions	326 (30%)	415 (12%)	297 (28%)	52 (39%)	82 (52%)
Replies	171 (15%)	158 (5%)	126 (12%)	28 (21%)	56 (36%)
Retweets	184 (17%)	165 (5%)	80 (7%)	27 (20%)	29 (18%)
Any communication	395 (36%)	470 (14%)	300 (28%)	56 (42%)	83 (53%)

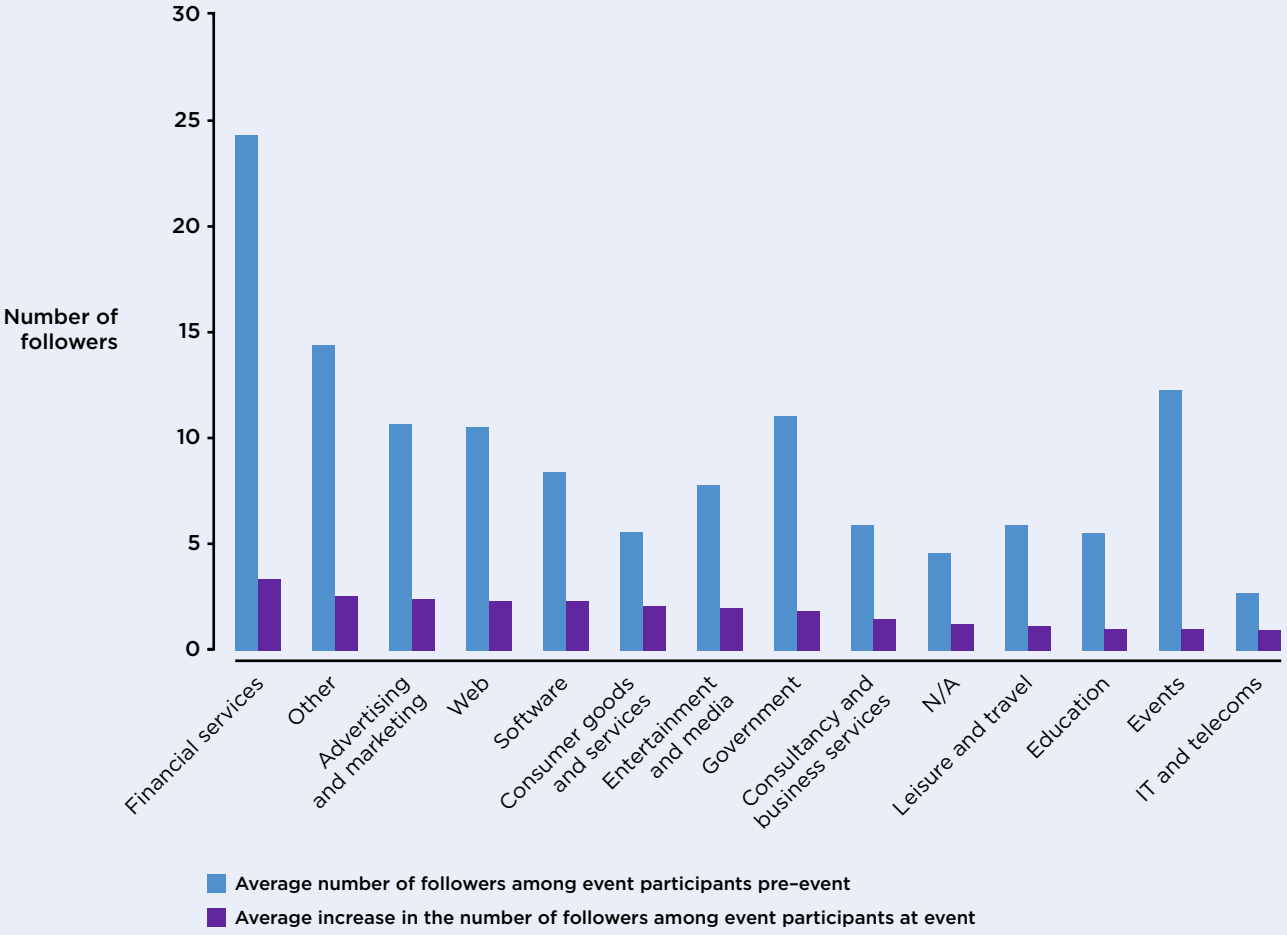
Table A3.2: The number of connections that experienced reciprocal communication since the start of LeWeb'12

	Old connections		New connection formed		
Percentage of connections with communication in brackets	Pre-existing reciprocal connection from before event that existed after the event: 1105	Pre-existing unreciprocated connection from before event that existed after the event: 3386 (exc of those consolidated at event)	1. New connections formed by an unreciprocated follow connection: 1072	2. Reciprocal connections formed by the consolidation of a pre-existing follow connection: 134	3.Completely new reciprocal connections formed by two new follow connections: 157
Mentions	79 (7%)	54(1%)	69 (6%)	27 (20%)	41 (26%)
Replies	58 (5%)	36(1%)	38 (4%)	14 (10%)	32 (20%)
Retweets	22 (2%)	4 (1%)	1 (0.1%)	5 (4%)	5 (3%)
Any communication	143 (13%)	68(2%)	71 (7%)	31 (23%)	43 (27%)

Note: In these tables all follow connections are undirected i.e. reciprocal and unreciprocated connections both count as one connection.

APPENDIX 4: AVERAGE NUMBER OF FOLLOWERS BY GROUP

Figure A4.1: **Average number of followers pre-event and their increase by industry group**



Note: Averages in this appendix are calculated inclusive of people that did not have followers among participants before the event, and who did not gain followers at the event.

Figure A4.2: Average number of followers pre-event and their increase by occupation group

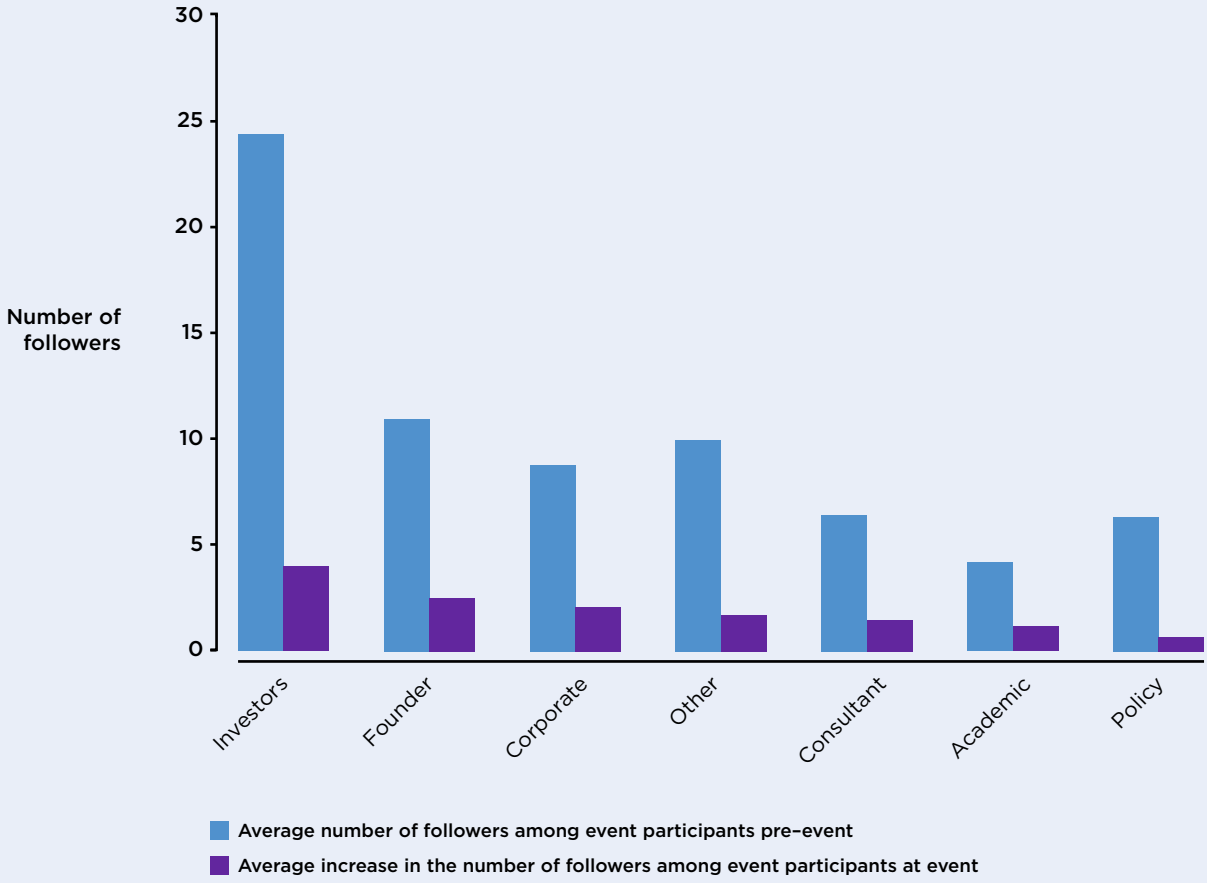
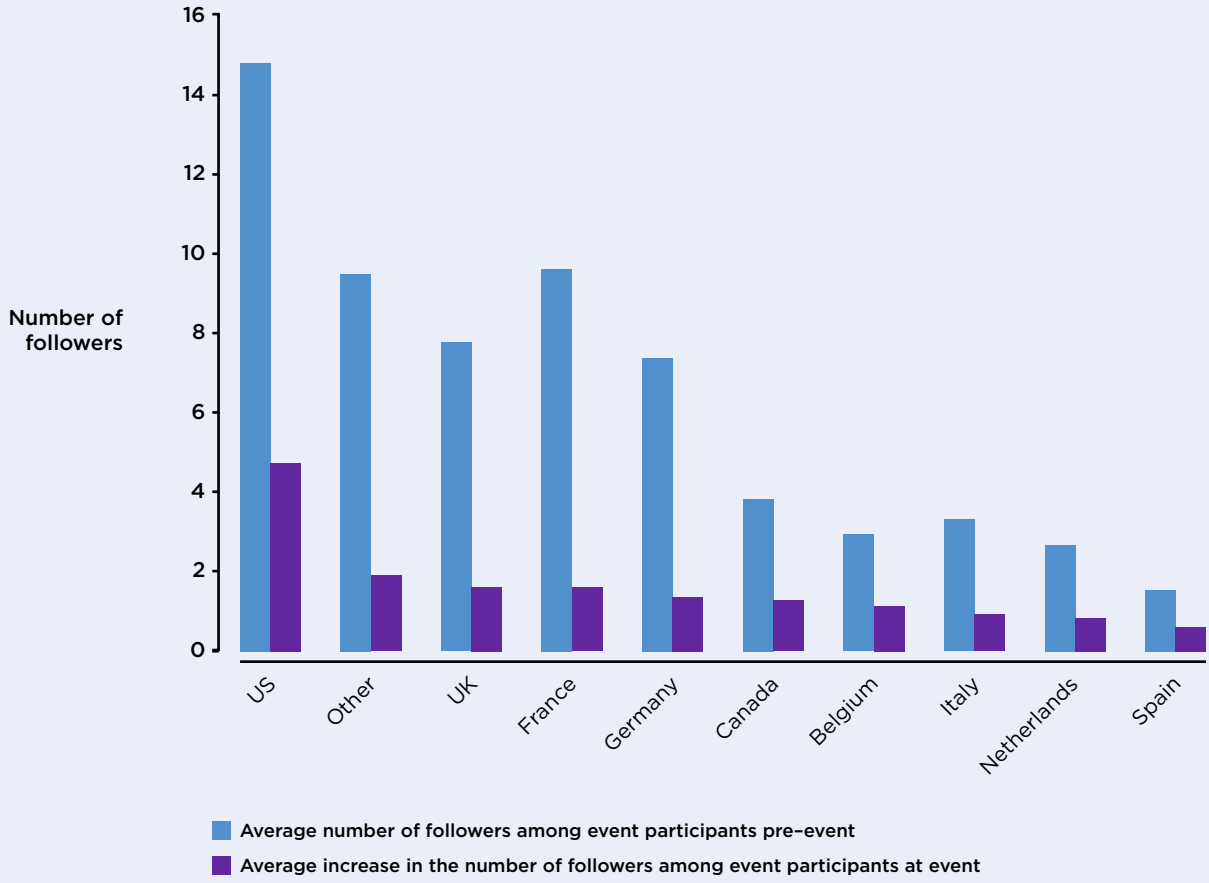


Figure A4.3: Average number of followers pre-event and their increase by country group



ENDNOTES

1. Boschma, R. (2005) Proximity and innovation: a critical assessment. 'Regional Studies.' Vol. 39. Issue 1.
2. http://gigaom.com/j2012/05/29/leweb-london-almost-happened-in-san-francisco
3. Evers, N. and Knight, J. (2008) Role of international trade shows in small firm internationalization: a network perspective. 'International Marketing Review.' Vol. 25. Issue 5. Maskella, P., Bathelt, H. and Malmberg, A. (2006) Building global knowledge pipelines: The role of temporary clusters. 'European Planning Studies.' Vol.14, Issue 8. Weller, S. (2008) Beyond 'Global Production Networks': Australian Fashion Week's Trans-Sectoral Strategies. 'Growth and Change.' Vol. 39(1): 104-122.
4. Newman, M. (2001) The structure of scientific collaboration networks, 'Proceedings of the National Academy of Sciences of the United States of America (PNAS)'.
5. Chon, B. S., Choi, J. H., Barnett, G. A., Danowski, J. A. and Joo, S. H. (2003) A structural analysis of media convergence: Cross-industry mergers and acquisitions in the information industries. 'The Journal of Media Economics.' Vol. 16, Issue 3.
6. Of course, an important question is whether Alice and Bob would have connected regardless of the event, say, because they are part of the same social group that attended – that is, whether the connection made was 'additional'. We will have more to say about this important issue later in the paper.
7. Ebner, M., Mühlburger, H., Schaffert, S., Schiefner, M., Reinhardt, W. and Wheeler, S. (2010) 'Getting Granular on Twitter: Tweets from a Conference and Their Limited Usefulness for Non-participants.' In Reynolds, N. and Turcsányi-Szabó, M. (Eds.) 'Key Competencies in the Knowledge Society'.
8. Letierce, J., Passant, A., Breslin, J. G. and Decker, S. (2010) 'Using Twitter During an Academic Conference: The #iswc2009 Use-Case.' In ICWSM.
9. LeWeb conferences happen yearly in Paris, and occurred in London in 2012 and 2013. For the rest of the paper, we use LeWeb'12 to refer to the London 2012 event.
10. Cramers, M. (2012) 'LeWeb Paris vs LeWeb London: The Social Media Scorecard.'
11. Brown, C., Efstratiou, C., Leontiadis, I., Quercia, D., Mascolo, C., Scott, J. and Key, P.(2014) 'The architecture of innovation.' Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing.
12. Kwak, H., Lee, C., Park, H. and Moon, S. (2010) 'What is Twitter, a social network or a news media?' Proceedings of the 19th International Conference on World Wide Web.
13. Huberman, B., Romero, D. and Wu, F. (2013) 'Social networks that matter: Twitter under the microscope.' arXiv preprint arXiv:0812.1045, 200.
14. Storper, M. and Venables, T. (2004) Buzz: Face–to–Face contact and the urban economy. 'Journal of Economic Geography.' Oxford University Press, Vol. 4.
15. Boschma, R. (2005) Proximity and innovation: a critical assessment. 'Regional Studies.' 39.
16. Of course, this approach only allows us to find the subset of all those individuals who met physically at the event who also talked about this with each other afterwards.
17. Cronin, B., De Vita, R. and Conaldi, G. (2015) 'Joining up the dots: Using social data to measure the effects of events on innovation.' London: Nesta. Working paper.
18. At the time that the data was collected, this was 3,200 tweets.
19. This means that if, for example, these individuals were more likely than others to increase their Twitter activity having attended the event, we would be underestimating the actual impact of the event on participants' networks.
20. This pattern is not explained by variation in the propensity to use Twitter in different countries. The same group of countries we mentioned above comprise 86 per cent of all individuals registered for the event when we include Twitter users and non-Twitter users.
21. The age information was inferred from the graduation year of participants' Linkedin profiles.
22. Web in Figure 3.2 refers to companies that provide online services, generally to consumers. This includes search, social networks and social media, and e-commerce.
23. We assume that the connections that happened between the period before the event and six weeks after the event (when we took our second 'snapshot') were made predominantly at LeWeb'12 itself, and refer to them as connections made at the event. The reduction in the rate of change in the number of follows between the first, second and third snapshot supports this.
24. Bastian, M., Heymann, S., Jacomy, M. (2009) 'Gephi: an open source software for exploring and manipulating networks.' International AAAI Conference on Weblogs and Social Media.
25. It is, though, also true that information on Twitter does not necessarily have to flow through the network of follow connections. In particular, one person on Twitter can look at the profile of another person on Twitter (who they do not follow) and retweet their tweets, or Twitter may show the tweet to the user as one they might potentially be interested in.
26. Another way one could assess additionality is by randomly assigning people to the thing that is being evaluated (which in this case would be attending LeWeb). This avoids the self-selection that can bias the evaluation of effects i.e. the fact that people who attend LeWeb are likely to be systematically different from those who do not attend in ways that increase their probability of connecting with other people who attend. This is not possible in this instance as delegates choose whether or not to attend LeWeb and we do not have any information on drop-outs (which could, in principle, have been random). For an example of the use of randomisation in a social network context see: Aral, S. and Walker, D. (2011) 'Identifying Social Influence in Networks Using Randomized Experiments.' IEEE Intelligent Systems.
27. Social network analysts refer to this phenomenon as 'triadic closure'. http://en.wikipedia.org/wiki/Triadic_closure
28. The technical term for the propensity of individuals to connect with others who are similar to them is known as 'homophily'. Disentangling these two effects is a difficult problem see Shalizi, C. and Thomas, A. (2011) Homophily and Contagion Are Generically Confounded in Observational Social Network Studies. 'Sociological Methods & Research.' 40 (2), 211-239.

29. The reason why this is important is that, for all we know, people who look far apart in the pre-event network could be more closely connected with each other (on Twitter, or via personal relationships) through people who did not attend LeWeb'12, and for whom we therefore do not have data (or connected via unobserved personal relationships among participants) – that is, we could be overestimating their distance before LeWeb'12. A significant association between our measures of distance and personal differences between individuals provides some reassurance that our network distance measures can be interpreted in the way we are doing.
30. This network includes all follow connections.
31. Figure 5.1 only includes completely new reciprocal connections. Reciprocal connections that were formed at the event from consolidating existing unreciprocated following connections already involved a direct connection between participants i.e., in undirected distance terms they will all be the same distance (one-step removed) in the existing network.
32. Note distances are calculated for each of the following relationships in the reciprocal connection i.e. Person A follows person B at the event, and Person B follows person A. Then in the chart above the length of the shortest path from Person B to Person A is calculated in the pre-event network and vice versa. These two distances may not necessarily be the same. The distance is calculated in terms of the number of people that a Tweet passing from the newly followed would have to travel to get to the follower when passing through the shortest path of the pre–event network.
33. We have repeated this distance analysis removing speakers (who as we showed in section Three account for a disproportionate amount of the one-way following connections) – the pattern that we have reported above (where one-way connections tend to happen between people who were closer to each other in the pre-LeWeb'12 network) remains.
34. Implicit in the analysis is that this finding would not be substantially changed if we had access to the global Twitter network. It cannot be known for certain, but if the entire network of users on Twitter was examined then it is much more likely that there is a path through the network that connects people, but one might hypothesise that those who are disconnected in the event's Twitter network would be connected, but still further away from each other than those that are directly connected in the global network.
35. This is the 'Strength of Weak Ties' identified by Granovetter. Granovetter, M. (1973) The strength of weak ties. 'American Journal of Sociology.' Vol. 78, Issue 6.
36. The average number of followers before the event (counting only followers who participated at the event) was 44 for speakers and 6 for the other attendees.
37. When doing this, we have considered the one-way follow Twitter networks as undirected – this saves us having to deal with the fact that distances between individuals in directed networks can change depending on the direction they are calculated e.g. the distance travelling from A to B may not be the same distance as travelling B to A as between the two there may be a links that exist in one direction (e.g. in the intermediate link C follows D) but not the other (D does not follow C). A limitation of this is that it may suggest proximity between individuals when in fact there are no direct ways for this proximity to translate into actual mutual awareness. For example, where two people are both following a third high-profile individual (who does not follow them back). This proximity can, however, still indicate hard to measure similarities and interests, and this is why we consider it.
38. We have established this with a χ^2 test of independence which considers whether the distribution of distances between both groups ('connected inside the group' and 'connected outside the group') are as one would expect if distance was independent from group.
39. More formally, $C_{ij} = LW_{ij}/B_{ij}$ is a ratio comparing the observed proportion of connections between (and within) groups (LW_{ij}) with the expected proportion of connections between group i and j if connections were generated at random between people (B_{ij}). The way the benchmark is calculated is now described. The number of possible reciprocal connections (with each reciprocal connection counting as one) inside a group of n people is the number of combinations of pairs that can be selected from the group (Labelled as C_2^n – this being the mathematical notation for the number of combinations of 2 objects, (in our case pairs of people), that can be selected from a group of n objects). Event participants can be divided into different groups, be they countries, occupations or industries. The set of all possible connections is partitioned into two kinds of connections: 1. Connections within a group (assuming connections are generated at random the probability of a connection falling within a group of size n_1 at an event with n participants is $C_2^{n_1}/C_2^n$), 2. Connections between two groups (Assuming connections are generated at random, and the group sizes are size n_1 and n_2, then the probability of a connection being between these two groups is $(C_2^{n_1+n_2} - C_2^{n_1} - C_2^{n_2})/C_2^n$) i.e. the number of connections that would fall within a group if the two groups were merged minus the connections that fall within the two groups in their own right, divided by the total number of possible connections among all participants. As some reciprocal connections already exist between groups prior to the event, the benchmark is calculated conditional on those connections that already exist.
40. This is a measure of homophily.
41. Hochberg, A., Liungqviust, A. and Lu, Y. (2007) Whom You Know Matters: Venture Capital Networks and Investment Performance. 'The Journal of Finance.' Vol. 62, No.1.
42. Some of these tweets will in addition to referring to people at the event refer to some people who did not attend LeWeb'12 as more than one person can be mentioned in a single tweet.
43. Individual names have been removed.
44. Cronin, B., De Vita, R. and Conaldi, G. (2015) 'Joining up the dots: Using social data to measure the effects of events on innovation.' London: Nesta. (Forthcoming.)
45. Snijders, T. (2001) The Statistical Evaluation of Social Network Dynamics. 'Sociological Methodology.' Vol. 31, Issue 1.
46. One participant was omitted from the analysis taking the sample size to 701.
47. Note that the measure of indirect connections nests the measure of direct connections – a score of one in Rec-ind would indicate that two individuals are either directly or indirectly connected in the pre-LeWeb'12 reciprocal network.
48. Our estimation uses a logit with clustered standard errors by participant to account for systematic differences in the connectivity patterns of those participants.
49. exp(-1.457) = 0.23
50. exp(-0.459) = 0.43

Nesta...

Nesta

1 Plough Place
London EC4A 1DE

research@nesta.org.uk
@nesta_uk
www.facebook.com/nesta.uk

www.nesta.org.uk