Are we bowling at all? An analysis of social capital in online networks

Javier Sajuria

University College London

Jennifer vanHeerde-Hudson

University College London

David Hudson

University College London

Niheer Dasandi

University College London

Yannis Theocharis

University of Mannheim

Corresponding author:

Javier Sajuria

j.sajuria@ucl.ac.uk

29/30 Tavistock Square, London WC1H 9QU, UK

+44 (20) 7679 4969

Abstract

Does social capital exist in online networks? This paper analyzes the formation of bridging and bonding social capital in online networks. It has been suggested by Putnam that online interactions are unable to foster social capital. We test this claim using Twitter data from three events: the Occupy movement in 2011, the IF Campaign in 2013, and the Chilean Presidential Election of the same year. Using Ronald Burt's concepts of closure and brokerage as indicators of bonding and bridging social capital, we observe the networks formed through online interactions and test them against several theoretical models. Our results support the claim that online ties are able to create social capital, but by distinguishing bridging and bonding types, and comparing across different issue networks, we tease out more specific findings. In the case of bonding social capital, online ties seem more effective to form close networks than theory predicts. Meanwhile, bridging social capital requires particular conditions for its creation, such as the presence of external events and professional brokers. Our results provide a confirmation that online networks are able to produce social capital, but it raises questions about the process of brokerage across highly connected groups.

Keywords: social capital, Twitter, network simulation, closure, brokerage

Are we bowling at all? An analysis of social capital in online networks

Introduction

Since the publication of Putnam's (2001) account of the decline of social capital in the United States, there has been a great deal of research testing his assertions in other settings (Colletta & Cullen, 2000; Bowles & Gintis, 2002; Hooghe & Stolle, 2003; Claridge, 2004; Pinchotti & Verwimp, 2007). The influence of Putnam's approach stems from the ability to provide a straightforward argument that social connections are vital for the sustainability and stability of a democratic society. Critically, then, Putnam advanced the field of social capital from an individual or group level of analysis (1994; 2001), to the understanding of how does social capital affect political institutions.

Nevertheless, the literature on social capital has not always been able to keep up with the pace of the development of information and communication technologies (ICTs). Or at least it is not always clear the extent to which existing theories — such as Putnam's — 'translate' to the online sphere. The explosion of new ways of communication, most of them Internet based, has brought new questions to the field. People do not interact the same way as they were used to, and this might affect the ways in which people create and maintain social connections. Since the emergence of ICTs, scholars have focused on understanding the role of social connections formed — or maintained — through the Internet (Bond et al., 2012; Ellison et al., 2006; Gibson et al., 2000; Kavanaugh & Patterson, 2001; Margetts et al., 2011; Shah et al., 2001; Wellman et al., 2001; Williams, 2006). It is worth emphasising that most of the research assessing the relationship between these new technologies and social capital assume that the ties formed through these platforms can carry similar quantity and quality of resources. The main purpose of this paper is to test this assumption and analyse the structure of online social networks as a way to understand how, and by which means, social capital might operate in the online world.

In this paper, we take up the question of the relationship between social media and social capital. More generally we are interested in the influence of the Internet – perhaps the key development since the publication of Putnam's seminal article – on social connectedness. In particular, we ask: is there evidence of social capital online? And if so, do we see evidence for both bonding and bridging varieties of social capital?

The distinction between bonding and bridging social capital, popularised by Putnam (2001), is a well-known and developed one, but one worth, briefly, rehearsing here. Bonding social capital exists in the strong ties that exist within, often homogeneous, groups — families, friendship circles, work teams, choirs, criminal gangs, and bowling clubs for example. Bridging social capital exist in the ties that link otherwise separate, often heterogeneous, groups — so for example, individuals with ties to other groups, messengers, or more generically the notion of brokers. Bonding social capital acts as a social glue, building trust and norms within groups, but also increasing intolerance and distrust of out-group members. Bridging social capital allows different groups to share and exchange information, resources, and help coordinate action across diverse interests. Putnam emphasises that these are not either / or categories, but that in well-functioning societies the two types or dimensions develop together.

This second question — the relative importance of bridging and bonding social capital — is of especial interest since one of the advantages of ICTs is to connect otherwise unconnected people, suggesting we might expect to see a different calibration between the two types of social capital than we see in the face-to-face world emphasised by Putnam. The other advantage of emphasising the two dimensions is that it allows us to develop a considerably more nuanced answer to the question of whether we see social capital or not online, i.e. we can ask to what extent is there both types of social capital online?

Similar to other studies (Coleman, 1988), we use Burt's (2005) structural notion of social capital and two associated metrics, closure and brokerage, as indicators of bonding and bridging social capital respectively. Closure refers to the level of connectedness between particular groups of members within a broader network and encourages the formation of trust and collaboration. Brokerage refers to the existence of structural holes within a network that are 'bridged' by a particular member of the network. Brokerage permits the transmission of information across the entire network. Social capital, then, is comprised of the combination of these two elements, which interact over time. We use the observed values for closure and brokerage over time, and compare it with different simulations based on theoretical network models. From this, we provide a thorough evaluation of the existence and formation of social capital in online networks.

We utilise Twitter data for three different events, the 2011 U.S. Occupy Movement, the UK-based NGO IF Campaign organised around the 2013 G8 meeting, and the 2013 Chilean Presidential Election, to analyse the networks created by the transmission of information around each one of these events in order to identify their structural features.

Our data show that, in general, online networks present higher levels of closure than what theory would expect, while the presence of professional brokers seems to be key in the existence of bridging social capital. Similar to traditional – offline – conditions, bridging social capital in online networks does not exist organically, and requires purposive efforts of certain members of the networks to connect across different groups. Furthermore, the interaction between closure and brokerage seems to go in the right direction, moving and growing together.

The paper proceeds as follows. In the first section we outline the theory of social capital and Putnam's scepticism towards the possibility of online social capital. We also outline the two key indicators of online social capital used in this paper — bridging and bonding capital — and a brief review of review of the literature on network approaches to social interactions. And finally we introduce our research hypotheses, derived from the theoretical discussion, and summarise the theoretical models that are used to test the research hypotheses; namely, Erdos-Renyi (random graphs), Barabasi-Albert, and the Watts-Strogatz. The second section describes the methodology used to collect and analyse the data. The third section documents our results and discusses the main findings. The conclusion brings together the paper and outlines fruitful directions for future research.

Theory and Hypotheses

Social Capital Online?

Social capital comprises of two parts: the networks of connections that exist between individuals and groups and the shared values and norms of behavior that facilitate cooperation. Crucially for Putnam (1994; 2001), the networks created by people, operating under norms of reciprocity and trust, are key for building and maintaining democracy. His work shows that communities where citizens interact with each other, primarily through participation in voluntary associations, are communities with higher levels of democratic performance and political stability. Moreover, in these communities, citizens show higher levels of political awareness, get involved more frequently and intensively in political processes, and consequently, overall government performance improves.

The conventional wisdom is that a key condition for the development of social capital is the connections and interactions between citizens take place face-to-face; social interactions that take place online do not foster social capital. According to Putnam (2001), four differences of computer-mediated communication make online interactions unsuitable for the formation of social capital. First, face-to-face interactions carry much more contextual information than online interactions due to the high degree of non-verbal communication that accompanies face-to-face verbal communication. Second, face-to-face interactions can bring diverse people together, whereas online interactions take place among like-minded people, something he calls 'cyberbalkanisation'. Third, online interactions do not foster social capital because of a digital divide in access to the Internet, which allows for the interaction of members of the elite and not the public in general. Fourth, the Internet has more potential to become a form of entertainment rather than communication. We take up each of these differences in turn, and set out why, a priori, online interactions may indeed foster the development of social capital.

Putnam argues that online interactions are unable to foster social capital due to the absence of non-verbal cues and information, which form a large part of inter-personal, communications. In the case of this first difference, we agree with Putnam: offline interactions lack this fundamental feature. However, to our knowledge, no study has empirically shown the extent to which non-verbal communication is necessary for the formation of social capital or social trust and cooperation that flows from it. Second, with respect to cyberbalkanisation, recent research has shown (Brundidge & Rice, 2009) that Facebook groups and profiles allow the emergence of political discussions among people who disagree, particularly through the connection of two persons who have a 'friend' in common. More generally, we argue the Internet has the potential to facilitate discussion amongst different groups. Online ties are not bound to their immediate communities creating the possibility of communication across traditional geographical boundaries. Moreover, online ties may facilitate communication amongst different individuals and groups because some of the initial barriers to communication in offline, face-to-face communication (gender, race/ethnicity, disability) are rendered less visible. While concerns of a digital divide persist, evidence shows a closing gap in access (Judge et al., 2006). Moreover, offline interactions do not provide any insurance for discussions outside elites. Other factors, such as geographical segregation, may be far more relevant for social integration than Internet access. Finally, while some scholars (Morozov, 2011) concur with Putnam's assessment of the Internet's greater potential for entertainment than communication, there is some evidence to show the Internet's communicative and mobilizing forces (Ward & Gibson, 2009). We also think this same assessment applies to offline organisations; joining organisations is not necessarily the same as interacting within those organizations.

Like Putnam, we also think there are differences between online and offline interactions, however, we argue that these differences do not prohibit the formation of social capital online. Ultimately social capital in online networks may differ from its offline counterpart. For example, online ties may be based more on the transmission of information than the personal characteristics of those interacting, such as geographical location, gender, ethnicity or even more important, who they know. Online ties may not be as stable or durable as those created face-to-face, because of the dynamic nature of the Internet. The level of engagement required to create a tie online might be lower than the engagement required offline, which might also have consequences in the type of resources they can mobilise.

Finally, the categorisation between weak and strong ties proposed by Granovetter (1973) might not operate in the same way. He argues that family connections and those based on close friendship are strong ties, while the others, such as work colleagues and acquaintances, are weak. However, it might be the case that the strength of a tie online should be measured by the quantity of interactions and the frequency and quality of the information it transmits, and not by the personal characteristics of those making the connection.

However, our task in this paper is not to identify the factors that contribute to those differences, but to test, drawing on two theoretical models, evidence of social capital online. Like the bowling leagues that Putnam used to illustrate social capital offline, we argue that Twitter and Facebook discussions can create social networks operating under norms of trust and reciprocity, that are able to mobilise resources and information. In the next section we examine the concepts of bonding and bridging social capital. Subsequently, we set out two theoretical models of social capital in online networks and drawing on these models, we identify three hypotheses relating to the formation and structure of online networks.

Observing Social Capital Online: Bridging and Bonding Social Capital

The concept of social capital has travelled a long way since its original inception by Hanifan (1920). That first reflection on the concept is related to the idea that the connections made in the classroom may affect the performance of the students. Since then, according to Webber (2008), there has been two streams of development of the concept: neo–capital and communitarian theories of social capital.

Neo–capitalists (e.g. Portes, 1998; Bourdieu, 1986; Burt, 2005) are concerned with the relative position of a person within a network, that is, how the position of a person might bring them benefits in relation to the rest of the members of the network. This approach allows us to determine how the relationships we form are able to mobilise resources or, as Bourdieu would prefer, how much 'capital' we can acquire through our social connections. In the case of communitarian approaches, as exemplified by Putnam, they look at the aggregate benefits of social connections. This approach is less concerned about the individual gains of participating in a network, but in the societal outcomes of them.

Within the communitarian approach, Putnam makes the distinction between bonding and bridging social capital. Bonding social capital exists in tight-knit networks that foster intragroup, strong ties. Putnam calls it a 'sociological superglue', and explains that it is useful to build trust between the members of the group and increases the levels of solidarity. Moreover, bonding social capital might be responsible for creating exclusion against those outside the group, which becomes the negative dimension of social capital. Bonding ties are the natural result of homophily (McPherson et al., 2001), where people who share similar relevant characteristics – such as geographical location, religion, ideology, among others – tend to group and work together. In a way, 'birds of a feather, flock together'.

The other dimension of social capital are bridging ties, that is, the connections that people form outside their circles. This is, in a way, what Granovetter (1973) called ‘weak ties’. Bridging social capital is responsible for coordinating action across different groups, and provides new information and resources to the more dense groups. Although both forms of social capital might be considered as competing with each other, Putnam explains that they are not "either-or categories". Instead, they operate in coordination, and become different dimensions at which we can measure social capital.

Meanwhile, Burt (2005, p.4) defines social capital as ‘the advantage created by a person’s location in a structure of relationships’. In that regard, he introduces two key indicators of what he considers to be social capital: closure and brokerage. The latter refers to the existence of a gap between two social groups, known as a structural hole. Brokerage takes place when two different groups are connected by a single node. Being a broker allows a person to have a better overview of the network and to become the only point of contact between two or more groups; hence, she can control the flux of information and resources through that network.

Social network structures consider the relationships built by people over time. These relationships can be dependent on contextual elements, such as work relations, or on a more personal level, such as friendship. Regardless of how we connect with others, the networks we build will have different shapes. Some networks will be denser, with everyone in that group interacting with all the other members (the basic definition of a cluster), while others will require someone to bridge different groups. The latter function of bridging is what we call ’brokerage’.

However, as Burt argues (2005), brokerage works in cooperation with closure (Coleman, 1988). That is, in order to broker something between two groups, each one has to host cohesive ties among their members, or some degree of closure. Conceptually, closure can mean different things depending on the network. In a group of friends, closure might mean trust, intimacy or frequency of contacts whereas in a group of colleagues, closure might mean that they share work on the same project or the same working space. In that sense, what we understand by closure may change depending on the type of social network we are observing. The important thing to consider is that closure allows a network to build trust among its members, by providing a safe environment for social relations. Hence, closure is essential for the creation of resources and information within a group, which in turn can be mobilised by a broker to another group.

A useful example of closure provided in the literature (Christakis & Fowler, 2011) are the dynamics of military companies. A company of 100 soldiers is usually composed by 10 groups of 10 soldiers each. It is really important for the efficiency of the whole company that each group of 10 becomes very close and that everyone in them knows each other. But within group closure is not enough for the emergence of social capital. It is also important that each group has ties with members of the other groups, i.e. what Granovetter (1973) would call 'weak ties', to transmit information and resources. Thus, it is the interplay of closure and brokerage that provides the company with an optimal level of social capital.

Neo-capital and communitarian approaches show us how levels of social capital may differ depending on whether the analysis is at the individual or group level. For example, at the group level we may have a community with low social capital, but there will be always some members of that community with more advantageous network positions than others. What makes social capital so important for a community is not only the relative position of any of its members, but the overall level of resources and information that are present and able to be mobilised.

As in the conjunction between closure and brokerage, the important element of social capital refers to a collective behaviour based on trust and reciprocity. Putnam claims that the benefits of participating in voluntary associations are not only individual, but also bring positive outcomes at a societal level. His distinction between bridging and bonding social capital takes the brokerage and closure discussion to an aggregate level by arguing that intragroup ties build trust and mobilise diverse resources.

From a conceptual point of view, Burt's concepts of closure and brokerage offer a useful way of bringing the neo-capital and communitarian approaches to social capital together. Burt provides a clear conceptual definition that fits most of the elements of Putnam's categories, but also provides a path for rationalising them. Closure operates in the same way as bonding social capital, favouring intragroup ties, fostering the formation of trust and building dense communities. On the other hand, brokerage provides a fresh flux of new information to the network, allows for the mobilisation of different resources, and uses the trust formed by closure to act as a tool for collective action. Our approach here has been to demonstrate the similarity of Putnam's bonding and bridging capital and Burt's closure and brokerage concepts. Thus, we employ Burt's measures as indicators of bonding and bridging capital at the aggregate level.

Finally, the decision of using these concepts (brokerage and closure) as measures for bonding and bridging social capital stems from the need to provide better indicators for them. Currently, measures of bonding and bridging social capital are analysed either using social network analysis, or survey instruments such as the name generator (McCallister & Fischer, 1978), the position generator (Lin, 2008) and, more recently, the resource generator (Van Der Gaag & Snijders, 2005). Some researchers (Ellison et al., 2011; Kwon et al., 2013) have also used survey instruments to assess the presence of bonding and bridging social capital in online platforms. In our opinion, this kind of exercise introduces two sources of bias. On the one hand, the use of self-reported data may lead to a misrepresentation of the actual networks. On the other hand, this type of data only allows for the analysis of ego-networks (i.e. the connections of a single node), and thus excludes the possibility of observing directly the interplay among different social groups. We argue that the use of observed networks provides an unbiased opportunity for analysing bonding and bridging social capital.

Theoretical Models and Hypotheses

We test the observed values we get from the online networks against different theoretical models that are commonly used to explain social networks formation: random graphs (Erdos-Renyi), Barabasi-Albert model, and Watts-Strogatz model. To construct the random graphs, we use the first variant of the Erdos-Renyi (ER) model, $G(n, M)$, which assumes that a graph is randomly selected from all the different possibilities of graphs with a fixed number of nodes n and vertices M. Each node in the graph, then, has the same probability of being connected with any other node from the same graph. We assigned the fixed number of nodes and edges according to the observed information.

The Barabasi-Albert (BA) model, on the other hand, is based in the idea of preferential attachment. That is, it starts an initial random graph and creates new nodes, one at a time. The main assumption is that nodes are more likely to connect with other nodes that are better connected. The aim of this model is to account for the level of influence of certain nodes in the network. Those who have more links, will attract more to connect with them. Formally, the model starts with a network with m_0 nodes. Each new node is connected to $m \leq m_0$ existing nodes with a probability that is proportional to the number of links that the existing nodes already have. The probability p_i that the new node is connected to node i is

$$p_i = \frac{k_i}{\sum_j k_j}, \tag{1}$$

where k_i is the degree of node i and the sum is made over all pre-existing nodes j. Heavily linked nodes tend to quickly accumulate even more links, while nodes with only a few links are unlikely to be chosen as the destination for a new link. The new nodes have a "preference" to attach themselves to the already heavily linked nodes.

Finally, the Watts-Strogatz (WS) model overcomes two main criticisms of the ER models. First, it accounts for the formation of triadic closure in a network – i.e. if we have three nodes A, B and C, where there are strong ties between A and C, and A and B, it is very likely that there will be a weak tie between B and C. Second, the degree distribution of ER models form a Poisson distribution, since it does not assume that highly connected nodes can link each other with higher likelihood. WS starts with a fix number of nodes N connected with degree K (which needs to be an integer), each one connected in a circular lattice with their neighbours. Then, the model rewires each one of the edges of a node i with another node k with a probability β that each node will be selected. No self-loops or duplicated edges are allowed. The main advantage of this model is that it accounts for the small-world effect (i.e. even if most nodes are not neighbours to each other, they can be easily connected from every other with a small number of steps) by producing higher levels of clustering coefficient than the BA model. The BA model, on the other hand, produces more realistic degree distributions.

All the models use the information from the observed networks – such as the number of edges and vertices, or the average degree – to build their own.

Drawing on the closure and brokerage concepts set out above, we test three hypotheses relating to the structural features of online networks and how they are related with the formation of social capital. We analyse the levels of closure and brokerage from a set of online networks and compare them with both random simulations and the most common theoretical models used to explain the formation of social networks. From that exercise, we use the outcome to test the following hypotheses:

Hypothesis 1 *The levels of bridging and bonding social capital formed through online interactions are significantly different than random.*

Hypothesis 2 *The networks formed through online interactions are, on average, less dense and weaker than the theoretical expectations.*

Hypothesis 3 *In online networks, bonding and bridging social capital operate in coordination, strengthening each other.*

For each model, we simulated a hundred different random iterations of the graphs and calculated their average values for closure and brokerage. We used the observed graphs as a reference for the number of nodes and edges required for the calculation of the models. From there, we used the values to test Hypotheses 1 and 2. For hypothesis 1 we run two-sample t-tests to compare the difference in means between the observed and the random (Erdos-Renyi) networks. For hypothesis 2, we compared the observed values against all the models. Our expectation is for the observed clustering coefficient to be lower and the network constraint to be higher than the theoretical models. Finally, in the case of hypothesis 3, we used the observed values for each event and calculate their correlation coefficient. The expected outcomes for each hypothesis are expressed in Table 1.

Data and Methods

The data for this paper is drawn from three Twitter datasets for three unique events: the Occupy Movement in the US (2011), the UK Enough Food for Everyone 'IF' campaign organised by UK-based NGOs to coincide with the UK G8 meeting (2013), and the Chilean presidential elections (2013). The three cases, outlined below, allow us to test our hypotheses across both spatial and temporal domains. Previous analyses of Twitter data have concentrated around events of the same nature: for example the use of Twitter for protests (González-Bailón et al., 2011); political campaigns (Vaccari et al., 2013); charitable campaigns (Clements, 2011) or using the entire population of tweets for a certain time period (Morstatter et al., 2013). Similar to the idea of voluntary associations and informal social interactions proposed by Putnam, our approach has been to select disparate events (cases) to analyse contexts in which users act and interact with others.

The Occupy movement started in October 2011, after a group of protesters decided to occupy Zucotti Park in New York. Their biggest aim was to demonstrate against the inequality and the monetary system holding it. From that initial occupation several occupations took place across the US and beyond. The data for Occupy was obtained through the Occupy Research project (`www.occupyresearch.net`), a collaborative network of researchers interested in the Occupy movement. The were gathered by R-Shief (`www.R-shief.org`) using the Twitter Streaming API for a period of 13 weeks, following the onset of the movement on October 2011. The data contain tweets using the different hashtags related to the movement, in particular those referring to cities where occupations took place. We focus on all the tweets using the 'official' hashtag of the movement (#ows; N= 4,352,071 tweets).

The IF campaign was a coalition of over 200 UK NGOs seeking to put pressure on the G8 governments meeting in the UK in the summer of 2013. The campaign's focus was on global hunger and sought to get the G8 leaders to make commitments to tackle four underlying drivers of malnutrition – insufficient aid and investment, the problem of land grabbing, the failure to tax multinational companies, and a lack of transparency around deals and investment. The data from the IF Campaign were gathered using DiscoverText (`www.discovertext.com`), from 23 January to 16 October, the official start and end dates of the campaign, using the live feed API. We pulled tweets that contained the official hashtags used by the campaign (e.g. #IF, #IFCampaign, #BigIF, #BigIFLondon, #BigIFBelfast). Because the main hashtag was somewhat generic our dataset had a very high number of non-campaign related tweets. As such the data were cleaned using DiscoverText's built in machine classifier (a naïve Bayesian classifier) resulting in a total of 101,842 units.

The data for the Chilean election were obtained through the Analitic platform (`www.analitic.cl`), which uses the Twitter Gardenhose API. We collected the tweets related to the two main candidates for this election, Michelle Bachelet and Evelyn Matthei. The tweets were selected based on the use of the name of the candidates, either as a mention, in hashtags containing the names, or their names without an "@" at the beginning. The time period spanned from 7 weeks before the run-off election until December 17 2013, which covered the entire legal campaign period for both rounds (N= 1,556,109 tweets).

The datasets[1] were filtered, leaving the username of the sender, the date of the tweet and any corresponding text. Each dataset was then divided into weekly static networks, creating a list of all usernames contained within the text of the tweets. An edge list was created using the username of the sender, and assigning a directed edge to any other usernames mentioned in their tweets. In order to account for more stable relationships among users, we filtered out any edges (ties) with a degree less than two. Descriptive statistics for each dataset is presented in Table 2.

Measures. To assess the level of closure for each network, we used the average local clustering coefficient metric. This value, for each weekly network, was calculated using an algorithm (Watts & Strogatz, 1998) that determines how close a node and its neighbours are to becoming a clique (a graph of fully connected nodes). Any graph $G = (V, E)$ formally consists of a set of vertices V and a set of edges E between them. An edge e_{ij} connects vertex v_i with vertex v_j. The neighbourhood N_i for a vertex v_i is defined as its immediately connected neighbours as follows:

$$N_i = \{v_j : e_{ij} \in E \wedge e_{ji} \in E\}. \tag{2}$$

We define k_i as the number of vertices, $|N_i|$, in the neighbourhood, N_i, of a vertex. The local clustering coefficient C_i for a vertex v_i is then given by the proportion of links between the vertices within its neighbourhood divided by the number of links that could

[1]Each dataset contains the text of the tweet, date and time, the user who sent it (username and user identification number), and relevant metadata, such as location and the profile image of the sender.

possibly exist between them. For a directed graph, e_{ij} is distinct from e_{ji}, and therefore for each neighbourhood N_i there are $k_i(k_i - 1)$ links that could exist among the vertices within the neighbourhood (k_i is the number of neighbours of a vertex). Thus, the local clustering coefficient for directed graphs is given as

$$C_i = \frac{|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)}. \tag{3}$$

From this, we can calculate the average local clustering coefficient for each node n:

$$\bar{C} = \frac{1}{n}\sum_{i=1}^{n} C_i. \tag{4}$$

To measure brokerage, we used an average of Burt's Network Constraint Index (2005) for each network. This metric observes the lack of structural holes within a network. As Burt explains,

> Constraint is a function of network size, density, and hierarchy that measures the extent to which relations are directly or indirectly concentrated in a single contact (see Burt 2009, p.50 for a detailed discussion). Contact-specific constraint, the extent to which manager (node) $i's$ network is concentrated in the relationship with contact j is defined as follows:
>
> $$C_{ij} = (P_{ij} + \Sigma_p p_{iq} p_{ij})^2, i \neq q \neq j, \tag{5}$$
>
> where $p_{ij}$is the proportion of $i's$ relations invested in contact
>
> $$j(0 < p_{ij} < 1, \Sigma_j p_{ij} = 1) \tag{6}$$
>
> Measuring indirect connections, the sum $\Sigma_p p_i q p_i j$ is the portion of $i's$ relations invested in contacts q who are in turn invested in contact j. In the parenthetical expression, the sum of the two terms is the proportion of $i's$

relations that are directly or indirectly invested in the connection with contact j. Summing across contacts, $\Sigma_j c_i$, yields a constraint index C measuring the concentration of a manager's (node's) direct and indirect relations in one contact - more constraint, less social capital. (Burt, 1997, pp. 367-368)

Both metrics are good indicators of closure and brokerage. In summary, a higher value in clustering coefficient means that the level of closure is higher. Instead, lower Network Constraint values mean higher levels of brokerage. Previous findings (Burt, 2000, 2005) show that both measurements are associated with higher levels of individual social capital.

Results and Discussion

To test our first hypothesis, that the three observed networks are different from random, we compared the mean scores for closure and brokerage for the random simulations against each network. The results, shown in Tables A1-A3, present strong evidence in support of H1. In nearly every instance, the means for closure and brokerage are statistically different between the observed networks and the models ($p<.01$). However, in a few cases, the statistical tests do not allow us to reject the null hypothesis. The calculation of closure for weeks 10-13 for the IF campaign using the BA algorithm are not statistically different. However, this time interval coincides with the period in which the number of tweets is the smallest for the whole series, and consequently, the size of the networks is also much smaller. Since BA models are calculated based on the count of vertices from the observed models, this may well explain the lack of significant differences between the theoretical model and observed networks. In substantive terms, these results show that the Barabasi-Albert and Watts-Strogatz theoretical models, in the way we simulate them, are not able to replicate the same levels of brokerage and closure of our observed networks. Furthermore, the particular networks created by the Twitter conversations differ significantly from the random models simulated for this study.

Moving to the actual difference between the models and the observed data, Figures 1 and 2 show the development of closure and brokerage over time for each network. Figure 1 shows the progress of closure, week by week, in comparison with the different theoretical models. The data shows that the levels of closure are higher (slightly) for the observed networks than for any of the models, in each of the three datasets. That is, given the number of edges, vertices, and the average degree of the networks, none of the simulated models is able to create higher levels of closure. This finding supports H2, by showing that online networks seem to be more efficient in forming small, denser communities than what theory would expect. This suggests that online networks are able to produce bonding social capital and their levels of closure are not explained simply by random allocation of nodes and ties.

In the case of network constraint (Figure 1), the support for H2 is only partial. None of the observed networks is able to produce higher levels of brokerage than the theoretical models. Moreover, in the case of Occupy, the levels of brokerage are even lower than the random graphs. In the case of the IF campaign and the Chilean election, brokerage was consistently above the random models, which shows that the connections across structural holes present in these networks is higher than what we would expect on any random network.

Two points warrant further consideration. First, the presence of brokerage opportunities is lower in online networks than the theoretical expectations, and the ability of members' of the networks to connect groups across structural holes is less efficient that what we would expect. Second, the difference between the OWS movement and the other cases raises questions about the nature of the events and whether differences in the type of event may explain the differential findings with respect to brokerage. One plausible hypothesis of this difference is that the Occupy case is less constrained in two particular aspects: geography and scope of issues. As has been described by the literature (Conover et al., 2013), the Occupy Movement reached places beyond the USA, but was highly concentrated on local events in each city. Moreover, the issues raised by the demonstrators ranged from the – rather vague – claim for more equality, to more concrete topics (e.g. the change in the financial system) depending on the place of the occupation (Chomsky, 2012; Castells, 2012). For these reasons, we performed a second set of analyses on the Occupy case.

Using the data from two cities in the US – Oakland and Boston – we calculated the levels of brokerage for each network and compared it with the simulated networks (using hashtags #OccupyOakland and #OccupyBoston respectively). The aim of this analysis is to establish whether the trend of low brokerage is something inherent to the Occupy movement, or was simply less evident in the wider, national network given its diffuse set of issue concerns and sizeable geographic constituency. Our test of the Oakland and Boston chapters of Occupy may show higher levels of brokerage (in relative terms) than the broad-based Occupy/#OWS.

Figure 3 shows the results for both networks. In the case of Boston, the trend was exactly the same as in the OWS networks: brokerage was lower than any of the theoretical models, including the random simulations. The difference is statistically significant, and is consistent with the results from the general Occupy movement. The case of Oakland, on the other hand, shows more disparate results. The results remain different at a 0.05 level, which means that the observed values differ significantly from the simulations, however, the results show no clear trend over time. The observed networks show, at points, even higher levels of brokerage than most of the models (with the exception of Watts-Strogatz), and during other weeks the brokerage is lower than the simulations. Looking at the results more closely, the weeks where brokerage is lower are those where the number of edges is higher. This is consistent with the idea that more ties within a limited network will eventually work against the existence of structural holes. Nevertheless, this does not answer the question of why the levels of brokerage are consistently lower in the other Occupy datasets, but not in this one[2]. We consider the differences in the results for brokerage across the three cases in the discussion of our findings below.

In summary, we find only partial support for H2 with respect to closure: online networks are able to foster the creation of tight, small groups within the network and do so better than what would be predicted if random. With respect to brokerage, the story is twofold. On the one hand, the IF campaign and the Chilean election networks show similar results (as in closure), whereas the OWS networks do not any more brokerage than what we might expect if random. In the case of the Occupy, this result was tested with smaller groups within the Occupy movement, but with disparate results.

For H3, the results are consistent with our expectations. In all three events analysed, the correlation between brokerage and closure is positive[3]. In the case of the IF campaign, the Pearson coefficient is 0.48 (n=22), and 0.80 (n=7) in the case of the Chilean election. The OWS dataset shows a significantly lower degree of correlation (r= 0.09, n=13), however this is to be expected given the results from H2[4].

[2] As a plausible explanation, we could argue that Occupy movements radicalised in smaller, not mainstream cities, might benefit from more local, offline organisation. Hence, the levels of brokerage might look more dynamic and higher

[3] As explained above, the way in which Network Constraint is measured is such that higher levels of brokerage is expressed in lower levels of Network Constraint. For that reason, the original results from the correlation coefficients are negative.

[4] On a related note, the difference in the results for the OWS networks also provide an interesting test for the overall validity of our findings. One of the most common criticisms of network analysis is that the metrics used to observe the networks seem to account for the same phenomena from different angles. As such, high levels of correlation are not only expected, but would also provide evidence in support of that theory. Contrary to that expectation, the levels of brokerage and closure in the case of the OWS seem not to be correlated at all, which defies the notion that the metrics are not providing new information. This,

The findings from the OWS, the IF campaign and the Chilean election provide a compelling overview of the formation of social capital online. The three cases show patterns of behaviour that cannot be explained fully by the most widely used theoretical models nor respond to mere random allocation of nodes and ties. In sum, the data suggest evidence of social capital formation online.

Our result showing differences in brokerage between OWS and the other two cases warrants further consideration. Beyond the more technical inferences about the differing results, we argue that that OWS may differ substantively from the other two cases. Both the Chilean election and the IF campaign are highly organised, well-funded and tightly focused events. Given that the main aim of campaign communications, Twitter or otherwise, is to influence attitudes, preferences and ultimately vote choice, we would expect to see a higher number of 'professional brokers', i.e. people whose main job it is to connect the different supporters of a given candidate, transmit information from the campaigns, and engage potential supporters. Moreover, the election itself narrow in focus with two main events: the first round and the run-off election. This means that the professional brokers not only had a goal, but also a deadline, to focus their resources and efforts. Similarly, IF was a coordinated campaign focusing on two key events and four key issues. Each of the participating organisations, though varied in their level of resources, may have served as professional brokers whose primary aim was engaging the sector and the broader public, by transmitting relevant information across them.

On the other hand, the OWS movement was more organic in its origins. The demonstrators themselves tried to foster the idea of a 'leaderless revolution' and aimed to keep momentum for a long period of time. There were few singular events that served to focus their resources and activities and the way in which they organised, both locally and

in turn, supports to the idea that closure and brokerage, yet related, are different theoretical and empirical concepts.

globally, was explicitly designed to foster egalitarian and horizontal interactions. Analysed at a more local scale, the results from the Occupy show different patterns. While in some cases the trend was similar to the aggregate movement, in other cases, local networks show higher levels of coordination and intergroup interaction. Future research could delve deeper into this difference to get more information about this macro-to-micro phenomenon.

Conclusion

Going back to Putnam's conceptions of bonding and bridging social capital, our results present an interesting puzzle for those studying online networks. In terms of the formation of social capital, online interactions seem able to bring together like-minded people, and create small, dense groups among them. That is, the potential of ICTs to create bonding social capital is better than any of the theoretical models. On the positive side, this means that online networks might have more potential than we expected to foster the creation of trust and reciprocity, based on the idea of intragroup ties. On a more worrying note, this might also lead to what Putnam calls "cyberbalkanisation", keeping like-minded people together, and not allowing the members of the groups to be exposed to more diverse information, while excluding those outside of them.

In terms of the bridging social capital, the results are diverse. It seems that the presence of professional brokers in the networks allows for bridging across structural holes. That is, the formation of bridging social capital seems possible by the presence of people whose aim is to produce those ties. The connection between small groups does not occur randomly or organically. Although more in-depth research of the in-group dynamics is needed to confirm this conclusion, the differences in brokerage between the OWS networks and the IF and Chilean election datasets seem a good starting point for that hypothesis. In essence, this is not much different that what we would expect in traditional, offline social connections. The alleged horizontal and spontaneous nature of online interactions might not be enough to produce, without intention, bridging social capital.

Putnam also claims that healthy societies foster the formation of both bonding and bridging social capital in coordination. One is required for the presence and operation of the other, and as such, the interplay between them creates trust, appreciation for diversity, and communication among different social groups. Our results show that online interactions are able to produce the same positive interplay. Furthermore, the evidence presented also provides support to the idea that this positive interplay requires intentionality. Online social capital seems to be in the right direction, allowing and fostering the coordination between bridging and bonding social capital. However, this is also present in events where part of the ethos of the network is the communication across people from different groups.

This paper has attempted to provide a preliminary approach to the formation of social capital in online contexts, by analysing three different Twitter datasets. Our main goal was to determine whether social cpaital exists in online networks. Our findings suggest that the current theoretical expectations of how social connections are created and maintained are not able to explain the network structure of online social interactions. Furthermore, on the question of the existence of social capital in online settings, we fall on the side of caution. Online connections seem able to easily create bonding social capital, but they require a concentrated effort to create bridges across those groups. The ideal setting presented by Putnam, where bonding and bridging social capital operate in conjunction, requires intention and effort.

References

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., et al. (2012, September). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295–298. Available from `http://www.nature.com/nature/journal/v489/n7415/full/nature11421.html`

Bourdieu, P. (1986). The forms of capital. In *Handbook of theory and research for the sociology of education* (pp. 241–58). New York: Greenwood Press.

Bowles, S., & Gintis, H. (2002, November). Social capital and community governance. *The Economic Journal*, *112*(483), F419–F436. Available from `http://www.jstor.org/stable/798455` (ArticleType: research-article / Issue Title: Features / Full publication date: Nov., 2002 / Copyright © 2002 Royal Economic Society)

Brundidge, J., & Rice, R. (2009). Political engagement online. do information rich get richer and the like-minded more similar. In *Routledge handbook of internet politics* (pp. 144–156).

Burt, R. (1997, October). A note on social capital and network content. *Social Networks*, *19*(4), 355–373. Available from `http://www.sciencedirect.com/science/article/pii/S0378873397000038`

Burt, R. (2000). The network structure of social capital. *Research in organizational behavior*, *22*, 345–423.

Burt, R. (2005). *Brokerage and closure: An introduction to social capital.* Oxford University Press, USA.

Castells, M. (2012). *Networks of outrage and hope: Social movements in the internet age.* Polity Press.

Chomsky, N. (2012). *Occupy.* Penguin.

Christakis, N. A., & Fowler, J. H. (2011). *Connected: The surprising power of our social networks and how they shape our lives – how your friends' friends' friends affect everything you feel, think, and do* (Reprint ed.). Back Bay Books.

Claridge, T. (2004). *Social capital and natural resource management.* Unpublished doctoral dissertation, University of Queensland, Brisbane, Australia. Available from http://www.socialcapitalresearch.com/

Clements, K. (2011). Social media–the twilight years of fundraising. *BMJ Supportive & Palliative Care*, *1*(2), 252–252. Available from http://spcare.bmj.com/content/1/2/252.2.abstract

Coleman, J. S. (1988). Social capital in the creation of human capital. *The American Journal of Sociology*, *94*, S95–S120. Available from http://www.jstor.org.libproxy.ucl.ac.uk/stable/2780243 (ArticleType: research-article / Issue Title: Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure / Full publication date: 1988 / Copyright © 1988 The University of Chicago Press)

Colletta, N. J., & Cullen, M. L. (2000). *Violent conflict and the transformation of social capital: lessons from cambodia, rwanda, guatemala, and somalia.* World Bank Publications.

Conover, M. D., Davis, C., Ferrara, E., McKelvey, K., Menczer, F., & Flammini, A. (2013, March). The geospatial characteristics of a social movement communication network.

PLoS ONE, *8*(3), e55957. Available from http://www.plosone.org/article/citationList.action;jsessionid=AF3940510EE045619B369947BD869F42?articleURI=info%3Adoi%2F10.1371%2Fjournal.pone.0055957

Ellison, N., Steinfield, C., & Lampe, C. (2006). The benefits of facebook "Friends": social capital and college students' use of online social network sites. *Journal of Computer Mediated Communication*, *12*(1), 1143–1168.

Ellison, N., Steinfield, C., & Lampe, C. (2011, January). Connection strategies: Social capital implications of facebook-enabled communication practices. *New Media Society*. Available from http://nms.sagepub.com/content/early/2011/01/26/1461444810385389

Gibson, R. K., Howard, P. N., & Ward, S. J. (2000). Social capital, internet connectedness and political participation: A four-country study.

González-Bailón, S., Borge-Holthoefer, J., Rivero, A., & Moreno, Y. (2011, December). The dynamics of protest recruitment through an online network. *Scientific Reports*, *1*. Available from http://www.nature.com/srep/2011/111215/srep00197/full/srep00197.html#/supplementary-information

Granovetter, M. S. (1973, May). The strength of weak ties. *The American Journal of Sociology*, *78*(6), 1360–1380. Available from http://www.jstor.org.libproxy.ucl.ac.uk/stable/2776392 (ArticleType: research-article / Full publication date: May, 1973 / Copyright © 1973 The University of Chicago Press)

Hanifan, L. J. (1920). *The community center*. Nabu Press.

Hooghe, M., & Stolle, D. (2003). *Generating social capital: civil society and institutions in comparative perspective*. Palgrave Macmillan.

Judge, S., Puckett, K., & Bell, S. M. (2006). Closing the digital divide: Update from the early childhood longitudinal study. *The Journal of Educational Research*, *100*(1), 52-60. Available from http://www.tandfonline.com/doi/abs/10.3200/JOER.100.1.52-60

Kavanaugh, A. L., & Patterson, S. J. (2001). The impact of community computer networks on social capital and community involvement. *American Behavioral Scientist*, *45*, 469–509. (3)

Kwon, M.-W., D'Angelo, J., & McLeod, D. M. (2013, February). Facebook use and social capital to bond, to bridge, or to escape. *Bulletin of Science Technology & Society*, *33*(1-2), 35–43. Available from http://bst.sagepub.com/content/33/1-2/35

Lin, N. (2008, May). A network theory of social capital. In *The handbook of social capital.* Oxford University Press, USA.

Margetts, H., John, P., Escher, T., & Reissfelder, S. (2011). Social information and political participation on the internet: An experiment. *European Political Science Review*, *3*(03), 321–344.

McCallister, L., & Fischer, C. S. (1978, November). A procedure for surveying personal networks. *Sociological Methods & Research*, *7*(2), 131–148. Available from http://smr.sagepub.com/content/7/2/131

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, 415–444. Available from http://www.jstor.org/stable/10.2307/2678628

Morozov, E. (2011). *The net delusion: How not to liberate the world.* ePenguin.

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the sample good enough?

comparing data from twitter's streaming api with twitter's firehose. *Proceedings of ICWSM*.

Pinchotti, S., & Verwimp, P. (2007). *SOCIAL CAPITAL and the RWANDAN GENOCIDE a micro-level analysis.* Brighton.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology*, *24*, 1–24. Available from http://www.jstor.org.libproxy.ucl.ac.uk/stable/223472 (ArticleType: research-article / Full publication date: 1998 / Copyright © 1998 Annual Reviews)

Putnam, R. (1994). *Making democracy work: Civic traditions in modern italy.* Princeton: Princeton University Press.

Putnam, R. (2001). *Bowling alone: The collapse and revival of american community.* Simon & Schuster Ltd.

Shah, D. V., Kwak, N., & Holbert, L. (2001, April). "Connecting" and "Disconnecting" with civic life: Patterns of internet use and the production of social capital. *Political Communication*, *18*(2), 141–162.

Vaccari, C., Valeriani, A., Barberá, P., Bonneau, R., Jost, J. T., Nagler, J., et al. (2013). Social media and political communication. *RIVISTA ITALIANA DI SCIENZA POLITICA*, *43*(3). Available from https://files.nyu.edu/jat7/public/2013_IPSR.pdf

Van Der Gaag, M., & Snijders, T. (2005). The resource generator: social capital quantification with concrete items. *Social networks*, *27*(1), 1–29.

Ward, S., & Gibson, R. (2009). European political organizations and the internet:

mobilization, participation, and change. In *Routledge handbook of internet politics* (pp. 25–39).

Watts, D., & Strogatz, S. (1998). Collective dynamics of a "small world" networks. *Nature*, *393*, 440–442.

Webber, M. P. (2008). *Access to social capital and the course of depression: A prospective study.* PhD thesis, King's College London, London. Available from `http://martinwebber.net/wp-content/uploads/2012/02/webber-phd-thesis.pdf`

Wellman, B., Haase, a. Q., Witte, J., & Hampton, K. (2001, November). Does the internet increase, decrease, or supplement social capital?: Social networks, participation, and community commitment. *American Behavioral Scientist*, *45*(3), 436–455. Available from `http://abs.sagepub.com/cgi/doi/10.1177/00027640121957286`

Williams, D. (2006). On and off the 'Net: scales for social capital in an online era. *JCMC*, *11*(2). Available from `http://jcmc.indiana.edu/vol11/issue2/williams.html`

Table 1

Hypothesis and expected outcomes

Hypothesis	**Indicator**	**Expected outcome**
H1. Observed networks are different than random.	Average local clustering coefficient and network constraint (t-tests)	$\neq$
H2. Observed brokerage and closure are lower than the theoretical models	Average local clustering coefficient and network constraint	$<$ clustering coefficient, $>$ network constraint
H3. Closure and brokerage work in cooperation	Correlation coefficient (Pearson)	$+$

Table 2

Descriptive Statistics

	IF Campaign		OWS		Chilean Election	
Week	*Vertices*	*Edges*	*Vertices*	*Edges*	*Vertices*	*Edges*
1	3334	478	40223	28480	94768	30682
2	3333	478	69799	86308	45156	9606
3	1660	220	42747	23483	87220	16445
4	1514	266	47067	36721	83333	13607
5	1221	162	60323	71216	34261	6372
6	1363	118	42168	28564	37450	9287
7	2637	284	30793	16289	68499	18115
8	3617	711	45118	35314		
9	2176	239	63185	86258		
10	380	31	53687	46380		
11	932	70	47361	36027		
12	932	70	41153	31683		
13	1028	124	25874	11585		
14	1946	111				
15	1053	116				
16	2469	255				
17	1677	523				
18	1504	190				
19	4146	728				
20	12532	3481				
21	4813	1135				
22	347	7				

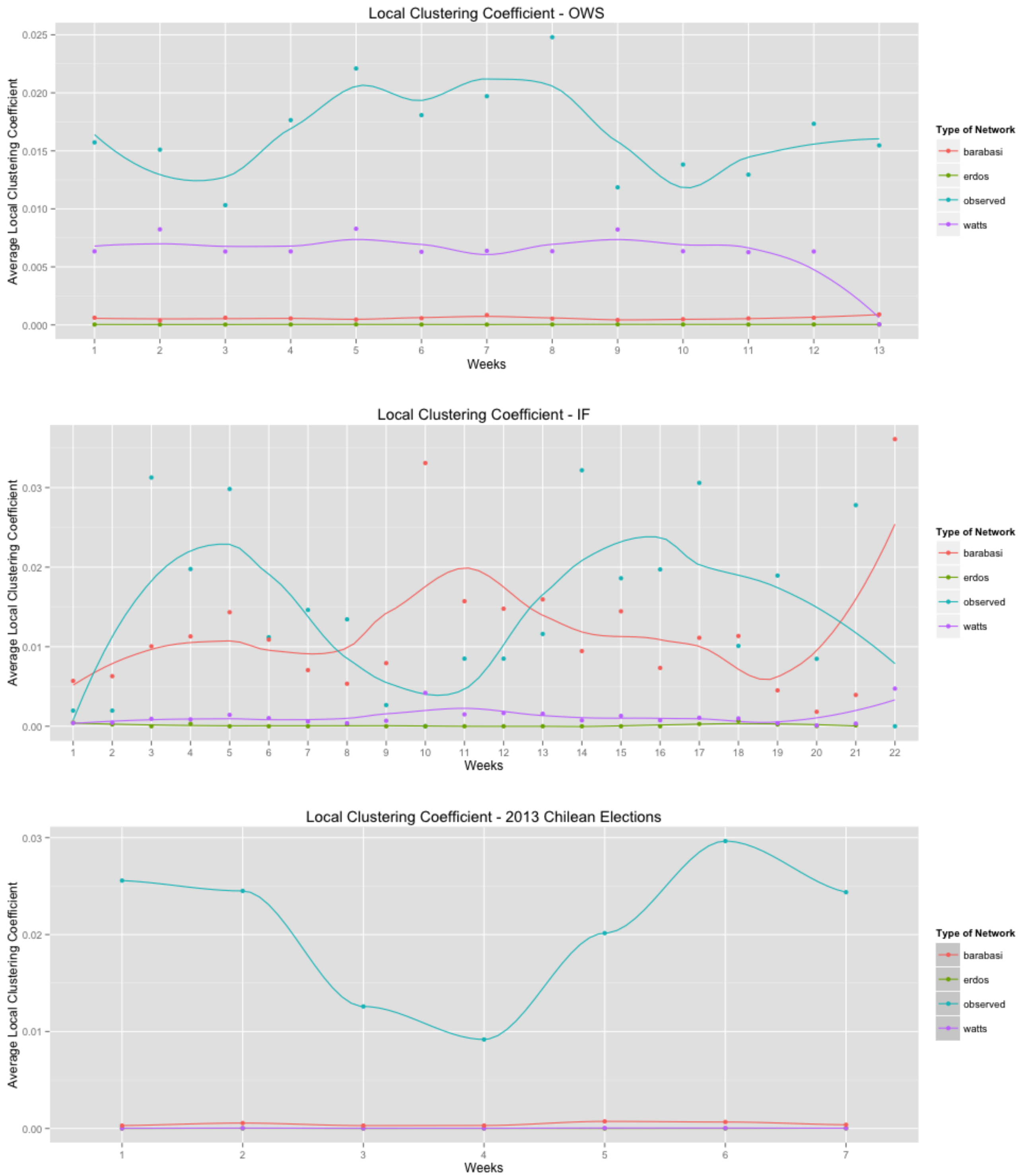

Figure 1. Closure for the three networks (Higher Clustering Coefficient means higher closure)

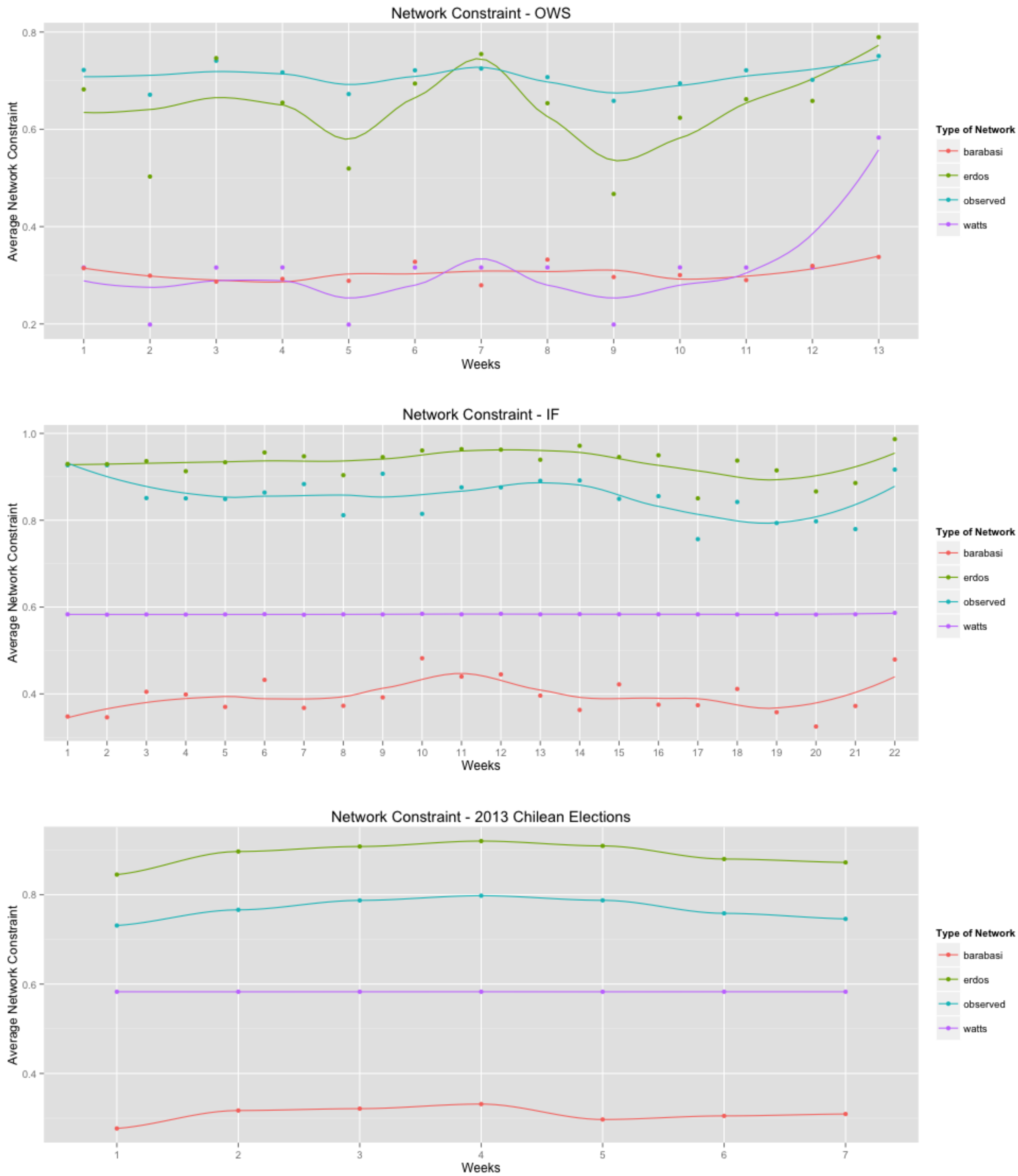

Figure 2. Brokerage for the three networks (Lower Network Constraint means higher brokerage)

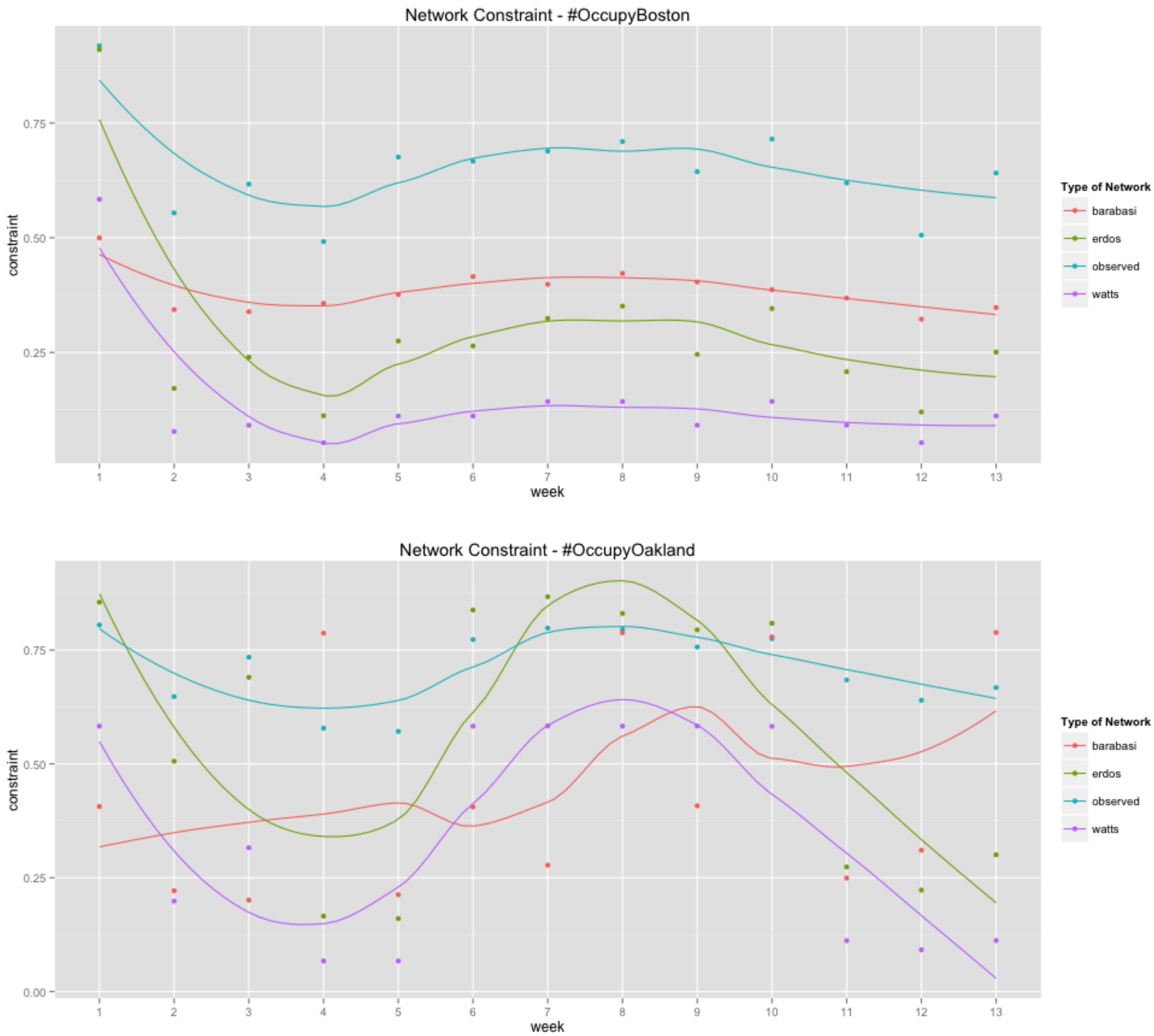

Figure 3. Brokerage for Oakland and Boston (Lower Clustering Coefficient means higher brokerage)

Appendix

T-tests

Table A1

P-values from t-tests using observed values against models - OWS

	Brokerage			Closure		
	Network Constraint			Avg. Clustering Coefficient		
Week	Barabasi	Random	Watts	Barabasi	Random	Watts
1	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000
11	0.000	0.000	0.000	0.000	0.000	0.000
12	0.000	0.010	0.000	0.000	0.000	0.000
13	0.000	0.000	0.000	0.000	0.000	0.000

Table A2

P-values from t-tests using observed values against models - IF Campaign

	Brokerage			Closure		
	Network Constraint			Avg. Clustering Coefficient		
Week	Barabasi	Random	Watts	Barabasi	Random	Watts
1	0.000	0.750	0.000	0.000	0.000	0.000
2	0.000	0.790	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000
9	0.000	0.010	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.640	0.000	0.000
11	0.000	0.010	0.000	0.860	0.000	0.000
12	0.000	0.010	0.000	0.950	0.000	0.000
13	0.000	0.030	0.000	0.120	0.000	0.000
14	0.000	0.000	0.000	0.030	0.000	0.000
15	0.000	0.000	0.000	0.000	0.000	0.000
16	0.000	0.000	0.000	0.000	0.000	0.000
17	0.000	0.000	0.000	0.000	0.000	0.000
18	0.000	0.000	0.000	0.000	0.000	0.000
19	0.000	0.000	0.000	0.000	0.000	0.000
20	0.000	0.000	0.000	0.000	0.000	0.000
21	0.000	0.000	0.000	0.000	0.000	0.000
22	0.000	0.240	0.000	0.000	0.160	0.800

Table A3

P-values from t-tests using observed values against models - Chilean Election

	Brokerage			Closure		
	Network Constraint			Avg. Clustering Coefficient		
Week	Barabasi	Random	Watts	Barabasi	Random	Watts
1	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000