Final Paper: Modeling Opinion Diffusion in Online Social Networks

Devin Gaffney

25 April 2015

Introduction

The mechanism behind how individuals coordinate and arrive at opinions has long been a source of academic interest across various fields of social science (Ryan & Gross, 1943; Katz, 1957). Whether attempting to explicate the process by which individuals adopt a new technology or shift their opinion on a given topic, the fundamental question underlying this strain of research concerns itself with how, for many individuals, their direct social environment may inform new opinions or practices.

Initially, some set of actors within a given social environment will adopt a new practice or receive new information. The interactions that these actors engage in within their social environment, and how their alters (and alters' alters, and so forth) adopt a new practice or interpret the received information is the diffusion process of interest. Beyond a simple transmission of information from one party to another, the diffusion of opinions and practices is fundamentally tied to the interpersonal relations between the transmitting parties. As Rogers observes of the Two-Step flow theory, while "the first step, from media sources to opinion leaders, is mainly a transfer of *information*, ... the second step, from opinion leaders to their followers, also involves the spread of interpersonal *influence*" (Rogers, 2010, 285). Similarly, Ryan & Gross showed that while salesmen were responsible for being the original source of information about a hybrid corn seed, neighbors were cited as being the most influential source, leading the authors to argue that "salesmen were credited with informing the majority of the operators but neighbors were credited with convincing them" (Ryan & Gross, 1943, 21).

While information transmission alone is an interesting topic of research, the second stage, in which information transmission is concomitant with interpersonal influence, is important for shifting the opinions of a much larger set of individuals. Indeed, the notion of "opinion leaders", a long-standing component of the two-step flow, as well as "early acceptors" in Ryan & Gross' work, assumes a smaller subset of individuals to be the initial set to whom information is relayed or amongst whom a practice is adopted (Katz & Lazarsfeld, 1970). Additionally, the interplay between social influence and information transmission is a rich point of methodological complication – to what degree does the information alone inform an opinion leader's alter? To what degree does the social tie instead inform the alter? Beyond dyadic interactions, how does an opinion leader shift the opinions of a larger network in which the opinion leader finds themselves? And in what way does an opinion leader's topographical position within a larger network of individuals shift the efficacy of an individual opinion leader? More fundamentally, what is an opinion leader (Dubois & Gaffney, 2014)?

In the intervening years since early observational research into the phenomenon of opinion diffusion, various research methods have opened new analytical avenues for exploring how individuals influence one another. Specifically, computational advances have allowed for both observational research at scale in the context in online social networks as well as computer-based simulations (Cha, Haddadi, Benevenuto, & Gummadi, 2010; Baldassarri & Bearman, 2007). While observational research provides rich examples of influence in action, operationalizing any complete theory of opinion diffusion into any particular dataset is difficult, and external validity to other networks and contexts is typically limited. Conversely, in simulation design researchers are free to develop theoretically sound operationalizations of a theory of opinion diffusion, and explore the results, though they may be divorced from the actual networks of opinion diffusion in the field.

This work seeks to design an simulation that explores how opinion diffusion may work in one particular online social network – Twitter. A simplified model of Twitter is proposed, where opinions diffuse from an initial actor in a network to the rest of the network. Several relevant variables are defined, and through a series of iterations, the way in which the model impacts the network's shift from an original "null" opinion distribution is considered. Various network topologies are employed in order to understand how the general structure of a social network informs how opinions diffuse through a network.

Ultimately, the question of interest is how the initial conditions of the network model shift a group of individuals from initially having no opinion about a particular topic towards the individuals co-creating an environment in which minorities of individuals become polarized as a result of their interactions. Polarization on Twitter is of particular interest in observational research, and constructing an agent-based model, or a particular simulation exemplified by it's emphasis on rules guiding individual interactions between a set of "agents", that explores this phenomenon may provide a powerful metaphor for how polarization may be constructed on this network (Yardi & Boyd, 2010; M. Conover et al., 2011).

First, a brief review of both observational and simulation analyses of opinion diffusion are considered. Limitations of both approaches are highlighted. Second, The methodological design of this simulation, along with a discussion of the relevant variables is provided. Through a series of trials against the simulation design, a statistical analysis is then provided, which allows for a larger discussion about how such research may help further unpack questions of opinion diffusion, and how a simulation design with observational components may aid in understanding opinion diffusion processes generally.

The Observational Approach

The observational approach to exploring questions of opinion change and polarization on Twitter has been well explored (Yardi & Boyd, 2010; M. Conover et al., 2011). Largely, the current literature falls into two broad categories – one explores how groups of individuals talk at a distance in existing structurally polarized networks (Weber, Garimella, & Batayneh, 2013), and another explores how groups of individuals negotiate the shift towards a new opinion within a larger political phenomenon (Lotan et al., 2011; Segerberg & Bennett, 2011). In both cases, researchers have attempted to unpack the relationship between structural positioning of actors within a given context and their ability to influence alters within that context. Ultimately, this research has stopped short of making strong causal claims about how structural topologies of either politically polarized networks or "hashtag" communities informs wide value shifts. This is largely due to several factors: first, identifying the causal relationship between structure and opinion change in observational data remains murky, second, each observation is highly contextual, and third, instrumentation is lacking in terms of exploring actual opinion change at a quantitative level.

One paper that is emblematic of these issues is the work done by Dubois and Gaffney, who explored the relationship between a wide array of network metrics as they relate to influential actors within the context of political actors. Specifically, the work aimed to identify opinion leaders via commonly used network metrics for two distinct and oppositional Canadian political hashtags, #CPC (denoting the Conservative Party of Canada) and #NDP (denoting the New Democratic Party of Canada). In their work, they argue that typical network metrics focused on the relative scale of direct degree-based influence such as indegree and eigenvector centrality identify "traditional" influential elites, such as politicians and establishment journalists. The authors introduce two other metrics, a "knowledge index", or the degree to which an individual employs context-relevant keywords in their tweets, and an "interaction index", or the number of unique interactions an individual has within the given hashtag community. In their work, they argue that these four distinct metrics allow for an exploration of Katz's two-step flow hypothesis – specifically, traditionally "popular" actors comprise the establishment, while the individuals that are highly interactive within the two hashtag communities and those who express high "knowledge" are the opinion leaders. While the work does indeed find that these four metrics are measuring similar but distinct facets of what one may consider in defining an individual as influential, the work does not attempt to address the issues raised earlier. The paper quantitatively examines the structure of the network, but does not attempt to unpack how the actors identified by these metrics are able to actually effect any sort of change within the hashtag communities. Additionally, the authors recognize that the results are specific only to the case under examination, though they encourage further use for this strategy. Third, lacking any instrumentation to measure adoption of a particular concept between these individuals, the authors do not explore how the "popular" actors or the identified "opinion leaders" actually fulfill the roles postulated in Katz.

Indeed, a common theme across the literature is concessions made to the imperfect measurement abilities for unpacking how individuals change their opinions through interactions in a given context. Bekafigo and McBride examined which types of actors tweeted about the 2011 gubernatorial election cycle in the US, and then employed a survey to examine the values and attributes of those who engaged. The survey was conducted by eliciting respondents via tweets posted to any individual who mentioned gubernatorial candidates within the context of the four elections under review. Indeed, the design is a relatively useful one – the responses elicited confirmed a relationship with propensity to tweet and strongly held opinions about a particular candidate. Of course the immeasurable opinion change aspect is also present. Instead of exploring how these respondents' opinions related to the neighbors of these respondents', the authors largely avoid any serious discussion of how opinion change via interaction may occur, stating the largely deflating conclusion that "[t]hough this study did not attempt to explain who these Tweeters are reaching, if anyone, through their messages of support for the candidates or pleas for others to vote, it is probably safe to say that our highly involved partisans are paying attention to gubernatorial politics and how the candidates are presenting themselves on Twitter" (Bekafigo & McBride, 2013). Similar sentiments are echoed in the tempering review provided by Aday et al., which argues that "[w]ithout rigorous research designs or rich data, partisans of all viewpoints turn to anecdotal evidence and intuition" when trying to establish causal claims around many questions employing social media data.

Still other work does provide interesting mechanisms for analysis. Lotan et al. employs the use of cascade analysis to explore how different types of hand-coded actors are able to affect a series of downstream tweets within the context of the Arab Spring, though little is done to capture the values embedded within those tweets and the degree to which they change. Even better, Leskovec, Backstrom, and Kleinberg employs a mixture of observational research as well as a model attempting to emulate the observation. Still, even this work falls short – the authors do not attempt to predict the changing nature of the memes they model as the memes transmit through interactions, stating that such an endeavor is "challenging, however, since it requires reliable methods of labeling significant fractions of sources at this scale of data". While observational data can provide us with many of the qualitative contours of the general phenomenon of how interaction between individuals occurs in an online space like Twitter, and particularly, the general contexts which may generate opinion change, it is not yet at a point where measuring that change is tractable.

The Simulation Approach

Previous literature has focused on questions around simulated designs of opinion diffusion models. Specifically, the vein of research around the diffusion of innovations is of particular applicability to this research question. Much like opinion diffusion, the literature focuses on "the process by which a few members of a social system initially adopt an innovation, then over time more individuals adopt until all (or most) members adopt the new idea" (Valente, 1996). In short, it is assumed that innovation adoption is relatively fungible as an action as opinion acceptance – indeed, Valente explicitly uses the terms "innovation" and "idea" interchangeably. Still other research goes so far as to imply that they are fundamentally the same process (Iyengar, Van den Bulte, & Valente, 2011).

Opinion diffusion models can be tied to fairly far-reaching foundational theory. In establishing the grounds for a model, Valente cites Granovetter arguing that Granovetter's weak ties "[a]re necessary for diffusion to occur across subgroups within a system". In the introduction to this work, two seemingly distinct theoretical foundations, innovation diffusion as well as the two-step flow hypothesis, were raised as potential guiding frameworks for understanding opinion diffusion – Valente also raises these works as well. This trend is not exclusive to Valente's work (Aral, 2011; Deroıan, 2002)

If a diffusion model can be discussed in terms of the ability to get a job, to learn about the news of the day, or to start using a better corn seed, what operationally is being measured? At it's core, the fundamental similarity between these grounding theories is the degree to which interaction with social ties alters actions or opinions for the interactants. As a result, the vast majority of models in current literature necessarily involve some step in a process whereby actors "interact", and are changed as a result.

In some ways, this fundamental similarity could serve as the definitional feature of agent-based modeling, alternatively defined as a simulation that "show[s] how simple rules of interaction could explain macro-level phenomena such as spatial patterns and levels of cooperation" (Janssen & Ostrom, 2006). In other terms, whether the model aims to seek an explanation for opinion shifts, technological adoptions, or career advances, the simulation approach to this broad category of sociology is at it's core an alternative explanatory method for unpacking how individuals co-create a larger ecological phenomenon. Gilbert and Terna echo this sentiment, framing a narrative about agent-based models as a "third way" of analysis between overly terse analytical equations as found in economics and overly vague "verbal" form of scholarship defined as "representation of past events, abstracted and simplified to emphasise some events and some inter-relationships at the expense of others".

Beyond providing a "third way" to analyze some particular phenomena, in the case of opinion diffusion, and indeed, any research interested in unpacking the mechanics of influence, a simulated environment allows the researcher the ability to sidestep many of the observational issues previously discussed. Aral talks at length about the various issues surrounding the observational approach, namely that many data sources used "can confound assessments of peer influence and social contagion, including simultaneity (Godes & Mayzlin, 2004), unobserved heterogeneity (Van den Bulte & Lilien, 2001), homophily (Aral, Muchnik, & Sundararajan, 2009), time-varying factors (Bemmaor, 1994; Van den Bulte & Lilien, 2001), and other contextual and correlated effects (Manski, 1993)". In short, Aral highlights a myriad of issues affecting a persuasive measurement and distinction of how interaction drives a behavior or opinion change, and how the opinion change in turn drives further interactions. By employing a simulated approach, many of these issues can be directly controlled and accounted for in a rigid and precise model rather than being accepted as the product of nebulous processes. In other terms, instead of accepting the inherently nebulous process by which a particular observed network was originally generated (and how that may inform the phenomenon one attempts to measure), agent-based models allow us to control for the particular network topology *as well as* the mechanics driving interaction and resulting change in the network.

An agent-based model, or indeed any simulation, when designed properly, is an oversimplification of reality (Kiesling, Günther, Stummer, & Wakolbinger, 2012). To borrow Gilbert and Terna's metaphor of the third way, simulation approaches to sociological questions, while still more rich than a simple and reductive analytical equation, still lack the richness of a verbal narrative that captures many more nuances of any particular question, including the potential and necessary "inconsistencies between the various concepts and relationships" being explored. In this way, simulated analyses are at best defined as a metaphor for the reality that is of real interest. This notion of the agent's actions as a metaphorical ruleset roughly approximating the decision processes and interaction results involved in the phenomenon of interest is a commonly echoed sentiment through the literature (Contractor, 1998; Luck, McBurney, & Preist, 2003; Drogoul, Vanbergue, & Meurisse, 2003; Moretti, 2002). Just as metaphors lose their utility with every added contextual complexity to the represented phenomenon, the simulation as metaphor necessarily omits relevant facts.

Amongst the various criticisms of the agent-based model approach (O'Sullivan & Haklay, 2000; Louie & Carley, 2008), Louie and Carley constructively provide a clear set of advisements for proper design. In this work, they emphasize the importance of grounding parameters within the model with known empirical data, and that "simulation models are particularly useful in developing theory that is 'rough' and 'not yet logically precise and comprehensive"', which consistently echoes many of Aral's sentiments about the state of understanding networked influence (Louie & Carley, 2008; Davis, Eisenhardt, & Bingham, 2007).

Model Design

At it's core, the question of how opinions, or norms, diffuse across a network is a micro-level process of interaction, ultimately leading to a macro-level analysis of how the larger networked group of individuals collectively create a macro-level distribution of opinions, or norm values, for a particular topic. Empirical data, when available, should be incorporated into the model as often as possible, allowing of course for the necessary simplifications involved in the model in order to maintain a clear metaphor to be drawn. What follows is a short discussion of the parameters that will define our model of interaction.

Parameter Definitions

Network Model

First, what is the structural network topography upon which actors will interact? While one may employ a sampled network from Twitter, this paper instead eschews the contextual particularities of any given Twitter graph, as was highlighted by Aral, and opts for a wide diversity of established and well-measured network models. This approach is well explored in existing literature, and also allows for a greater degree of replicability and modification of this research design (Centola & Baronchelli, 2015; Deffuant, 2006; Van Eck, Jager, & Leeflang, 2011; Kiesling et al., 2012). Of course, choosing wildly differing graphs can only elucidate so much - for this reason, the graphs considered all share many attributes, though the topological construction differs. In this case, four graphs are considered that provide for a range of network topologies within which to explore. Three models, the Barabasi/Albert, Erdos-Renyi, and Lattice models, are relatively prevalent according to Kiesling et al.'s review of agent-based model literature. One, the "Barabasi/Albert Barbell", is distinct to this study, but draws many parallels to observational research, which finds the presence of strongly polarized networks in contexts such as political debates (M. D. Conover, Gonçalves, Ratkiewicz, Flammini, & Menczer, 2011; Yardi & Boyd, 2010; M. Conover et al., 2011; Gruzd & Roy, 2014; Adamic & Glance, 2005). A visual review of these networks is provided in figure 1. For each trial of the model, the chosen network is simply referred to as the network.

Norm Values

In the case of Twitter, much empirical and qualitative information is available. In this simulation, a topic upon which no one has any particular opinion about is represented as the initial "norm value" for individuals on the platform. In the context of Twitter, this initial state is not particularly unbelievable – as an example, Justine Sacco, a hitherto unknown person, was the subject of a vast negative swarm of opinion about a particularly (and intentionally but apparently unnoticeably sarcastic) abrasive Tweet about her upcoming trip to South Africa as it related to contracting the AIDS virus (Bercovici, 2013). Additionally, by presenting the network with a uniform and "neutral" initial norm value, the model doesn't have to be further complicated by deciding the precise contours of a non-neutral initial norm value, which would necessarily engage with the issue of value homophily, which is precisely the intended research variable (McPherson, Smith-Lovin, & Cook, 2001). This parameter, "norm value", is a value held by a particular agent in the network, and as such is a function of that agent and is denoted $nv(i)$.

Initial Deviant Opinion

Qualitatively, one could imagine that few individuals had a strong opinion about this particular person prior to the incident – one likely (and to be clear, other potential causes exist) pathway towards this swarm of negative opinions is that a single actor or small subset of actors signalled to their respective audiences that this tweet was a transgression of some nature. In terms of the Sacco case, these initial actors were signalling a negative opinion or norm towards Justine Sacco at an audience which previously had no opinion on the topic of Justine Sacco. This line of reasoning is largely consistent both with research on understanding how individuals post content on Twitter assuming an imagined audience as well as literature surrounding persuasion of this audience on the platform (Marwick & boyd, 2011; Chang, 2010). This parameter, or the "initial deviance", is a randomly set value denoted as DE within $DE \sim U([0, 1])$ for each simulation.

Concession Salience

The next important parameter deals with the nature of issue salience for individuals. Specifically, this paper adopts the relatively reasonable notion that "[t]he level of perceived personal importance of an issue, therefore, should predict the extent to which information about an issue is transmitted and accepted across individuals" – namely, that in cases of high topical importance, we would expect less constant fluctuation and shifting of opinions (Cullum & Harton, 2007). For this parameter, this model specifies a "concession" score – when the value is low, the network's agents hold strong importance to the issue, and so the effect of each interaction has less impact on the change of opinion. When the value is high, the agents are more willing to shift their norm values. Formally, "concession salience" is a randomly set value denoted as CS within $CS \sim U([0, 1])$ for each simulation.

Social Pressure

How does social influence factor into opinion changes? In this model, the relatively simple and widely employed Jaccard index is used as a proxy for social strength between agents (Liben-Nowell & Kleinberg, 2007; Wang, Pedreschi, Song, Giannotti, & Barabasi, 2011; Szell, Lambiotte, & Thurner, 2010). When two agents have a high Jaccard index, their social strength moderates the degree to which they are willing to concede to their alters – stronger social strength implies that they will more willingly accept an alter's opinion, or at least be sympathetic to it. In the context of this model, however, a Jaccard index of 0 implies that a tie with no mutual friends would not affect an agents opinion to the same degree as if the person were a complete stranger. In context, two agents interact if they share a social tie. For this reason, the Jaccard index is modified to be:

$$J(i,j) = \frac{\sum A_{ik} * A_{kj} + 1}{A_i * A_j^T + 1},$$

Where A represents a network in adjacency matrix notation, and i, j are a pair of tied agents. By adding 1 to the numerator and denominator, we allow for cases which still impress some social pressure on agents when no shared alters are present – this is done in order to simulate the degree to which having *any* tie is itself a source of social pressure, regardless if no alters are present.

Tweet Frequency

The last parameter in the model concerns itself with the frequency of tweets per user. By default, users who post tweets typically do so within the context of either a web-based or mobile-based Twitter interface, both of which display one's personal timeline, or a reverse-chronologically sorted listing of tweets posted by users the individual actor follows (Schmidt, 2014). In this model, it is assumed that some users may tweet more often than others. In order to ground this notion in empirical data, which is readily available, a random 10% sample of all tweets occurring on February 28th, 2015 were collected – while the collection strategy may be problematic due to it's short time window, it is assumed that the distribution of tweet frequencies per user, at least from the large ecological view of 10% of all users, randomly selected, ensures that this day, as well as any particular day in general, is more or less representative of tweet frequencies. Figure 2 displays this frequency distribution as a log-binned probability against the number of tweets occurring at that probability. Formally, this is denoted as $T(i)$, where any value of $T(i)$ is a randomly selected frequency of tweets based on the probability distribution found in figure 2.

Finally, what is the dynamic of change in this network? This model assumes a dynamic of "compliance gaining", or the notion that "an individual accepts influence because [they hope] to achieve a favorable reaction from another person or group" (Kelman, 1958). Specifically, the inducement strategy that agents employ is to overstate their "norm value" as a method for gaining compliance from their alters. When an issue is particularly salient, they will overstate their actual norm value in a more extreme way relative to the individuals they are tied to as an inducement to normalize the alters towards their actual norm value. When they are placed between extremes, this mechanism will serve to balance between those extremes – in short, the dynamic for actors is to attempt to place themselves in the equilibrium between all interactional partners in order to attempt to draw all alters closer to their actual norm value at any given iteration of the model. This parameter is the most important (and most complicated) aspect of the model, because the compliance gaining mechanism dynamic is a function of all other parameters, and occurs in a temporal context. Formally, for any time t in the model, agent i's norm value $nv(i)$ is changed through the compliance gaining dynamic through the function:

$$nv(i,t) = nv(i,t-1) + \frac{[\sum (nv(i,t-1) - nv(j,t-1)) * J(i,j) * T(j) * C]}{\sum_{j=1}^{n} A_{ij}},$$

$$nv(i,t) = nv(i,t-1) + \frac{[\sum(nv(i,t-1) - nv(j,t-1) * J(i,j) * T(j) * C)]}{\sum_{j=1}^{n} A_{ij}},$$

$$nv(i,t) = nv(i,t-1) + \frac{[\Sigma(nv(i,t-1) - nv(j,t-1)) * J(i,j) * T(j) * C(i)]}{\Sigma_{j=1}^{n} A_{ij}},$$

$$nv(i,t) = nv(i,t-1) + \frac{[\Sigma(nv(i,t-1) - nv(j,t-1) * J(i,j) * T(j) * C(i))]}{\sum_{j=1}^{n} A_{ij}},$$

Where $\sum_{j=1}^{n} A_{ij}$ is the number of total neighbors for a given agent.

Simulation Execution

Random values were selected for DE and CS, and the number of time steps per simulation was held constant at 100. Simulations were run in parallel until a sufficiently representative number of simulation executions were performed. After an initial set of iterations, CS was converted into a randomly selected value from a normal distribution per node, denoted as $CS \sim (\mu, \sigma^2)$, where the mean was set to 0.5 and the standard deviation was set to 0.33, echoing the settings specified in Baldassarri and Bearman, in order to capture how differing saliences for a given individual and topic may affect the outcome of the model.

Results

A general review of the model reveals surprising results. Figure 3 shows that, across all values of DE and CS, a surprising trend occurs regardless of the underlying network model that agents interact upon. Specifically, across many trials, a clear trend emerges – after the 100 timesteps, regardless of the underlying network model, there are clear cases of polarization. This implies that the parameters specified above are sufficient for generating a Twitter-like network that is capable of generating polarization when the initial number of actors asserting a new norm value is only a single agent.

Two values in the simulations were randomly assigned due to, more or less, a lack of clear empirical data – the initial deviance of an actor is highly contextual and difficult to measure in the context of a Twitter network, and the degree to which an individual on Twitter is willing to concede to an alter's norm without regard to social influences is similarly intractable.

Understanding the degree to which these randomly assigned parameters may affect the network, a simple scatter plot is evidence enough to understand the degree to which DE and CS impact the ultimate distribution of changed norm values within the networks.

As is clear in figure 3, a clear set of agents do not deviate widely from the initial null norm value of 0.5. Additionally, the graph shows that relatively equivalent numbers of agents are pulled to either polar extremes after 100 trials, regardless of the underlying network topology. As a result, examining the average change of an agent's norm change per simulation would likely lead to little insight. The standard deviation of these changes, however, allows for an interesting avenue of discussion. In short, an underlying assumption of the model is that, regardless of DE or CS, agents will attempt to re-balance their local network by shifting their norm value to an optimal value between the sum of their alters, which implies a necessary and inverse reaction to alters in order attempt to sway alters to the norm value of the agent.

The standard deviation of the net changes for agents, however, is able to elucidate the relative impacts of DE and CS. While the model deterministically forces agents to further polarize, the degree to which they polarize is captured in the standard deviation of changes per agent. In other terms, while the mean changes per agent per simulation per network model will be relatively constant, the standard deviation is likely either a function of DE or CS. As figure 4 and figure 5 show, no matter the initial deviance, the CS parameter almost completely determines the ultimate standard deviation of net changes per agent to a point where the regressions holding standard deviation as dependent and either DE or CS as independent show that CS is almost entirely the same underlying construct as the standard deviation of net changes per simulation.

As was alluded to earlier in the paper, CS being a uniformly set parameter is likely unrealistic. In other terms, for a given topic, agents likely assign a different concession salience – some agents will care more about conceding to their interactants' norm values, while others may not care as much. In general, however, it would be reasonable to expect that most actors care a certain amount. This qualitative description of norm values per agent roughly approximates a normal distribution – while some may have extreme preferences as to the degree to which the opinions of others matter, most will take the opinions of others into account in a relatively similar way.

Again, this notion is found in the model that Baldassarri and Bearman employ, as well as seeming to be a generally qualitatively innocuous assumption.

Interestingly, when CS is set as a normal distribution with a mean of 0.5 and a standard deviation of 0.33, figure 6 shows a surprisingly reduced amount of random results. While the r-squard values remain insignificant, the degree to which the scatter plots, sorted by the network topology, express a purely random result diminishes towards more conclusive relationships. Of these charts, the most compelling chart is how the average standard deviation of the change of norm values per node relates to DE in the context of an Erdos-Renyi Graph. In particular, while this oservation remains qualitative, it appears that as the DE value deviates from the null norm value, the degree to which the standard deviation changes increases. Substantively, one would conclude that in the case of Erdos-Renyi graphs, when CS is normally distributed, the effects of DE become clear, and the effect is that as an initial deviation becomes more extreme, it may be responsible for a larger degree of norm polarization within the network. More work is needed, particular in statistical tests and transformation of these parameters in order to assert this. Namely, creating a statistical variable that converts values for DE as absolute value deviating from the null value would prove to be particularly useful for exploring this relationship. Indeed, figure 7 shows this conversion and indeed supports the notion that there is a substantive relationship when CS is normally distributed.

Finally, one more aspect of this model is worth discussion. Figure 3 implied that the net change after 100 trials per simulation was not dissimilar across various network topologies. By examining a different metric of dynamics, however, it becomes clear that network structure plays a qualitatively distinct role across the simulations. Specifically, by treating the sum of node trial differentials (i.e. the sum of norm value differences per node between trial 1 and 2, 2 and 3, and so forth) as the x-axis and the magnitude of these changes as the y-axis, it is possible to see how different network topologies play distinct roles. As figure 8 shows, Barabasi-based graphs, regardless of if it is a Barbell model or not, experiences stronger magnitudes of net change per node in earlier trials than Lattice or Erdos-Renyi graphs, holding DE and CS as random values. When shifting to CS as a normal distribution per node, the trend is much noisier, but the basic qualitative assessment remains identical, as figure 9 shows. Ultimately, while the results shown in figure 3 show relatively similar processes, the rate at which these processes affect agents within the various network topologies are distinct – this is an important point, as it draws the implication that as we insert agents into networks that more closely approximate real-world networks, we see an attenuated rate of change of norm values within this model, which may provide insight into the dynamics of online polarization.

Discussion

This agent based model seeks to elucidate the mechanisms underlying how a practical network may produce polarization via a simple set of parameters that are grounded either in theoretical conceptions or empirical data. As has been shown, DE and CS are likely the fundamental parameters driving ultimate norm deviation in this model, though they are likely modified by the contextual parameters involved in any observational case. As opposed to observational research, however, this agent-based model allows for a relatively accurate model of the phenomenon to be approximated while still affording the ability to explore counterfactual cases as a metaphor for the ways in which polarization may occur on Twitter. Notably, and admittedly, this model is, almost inexorably, a model that predictively moves only towards polarization. At the very least, the R-squared values presented in figure 5 should give a reason for pause. While this agent-based model shows a relatively grounded and potentially plausible model for polarization, it does not yet account for cases in which polarization does not occur. One possible modification of this model that may more accurately capture a concession towards the mean is one where $nv(i, t)$ is defined as:

$$nv(i,t) = nv(i,t-1) + \frac{[\Sigma(nv(i,t-1)*J(i,j)*T(j)*C - nv(j,t-1)*J(i,j)*T(j)*C)]}{\Sigma_{j=1}^{n} A_{ij}},$$

Or in other terms, the magnitude of change for an agent is not a product of the difference of norms multiplied by the various factors, but is instead the norm that the agent bears multiplied by the strengths the alters exert minus the norms alters exert multiplied by the various factors. Substantively, this alternative model would allow for both concession as well as polarization – if the norm as well as the strength of relationship, tweet frequency, and/or concession salience was more extreme than the agent's opinion, the agent would act in a similar way as presented in the current model. If the opinion of the agent, factored against the other variables, outweighed or was significantly more deviant than the alters, they may still move towards a polarizing network. Ultimately however, the model currently presented displays a potentially powerful grounded metaphor concerning the dynamics of polarization within the Twitter network.

At their best, agent-based models are only metaphors that serve as guideposts for further research. By their nature, they can rarely be entirely validated via empirical methods (Louie & Carley, 2008). Yet, it precisely because of this reason that agent-based models provide substantively useful guideposts for research in cases where causality is observationally difficult (Aral, 2011). Arguably, Schelling's "agent-based model" contributed greatly to earning the author among the highest of intellectual awards offered, not to mention it's seminal role in literature (Nobelprize.org, 2015). Schelling's intellectual contribution was not to provide an accurate description of the precise mechanics of segregation in his work – instead, the author presented a model in which agents with relatively conventional motivations collectively generate an unconventional result. In this work, the same intellectual tradition is being sought. By defining a few simple empirically or theoretically grounded parameters, the aim of this model is to highlight a predisposition of agents on Twitter's network to trend toward a state of affairs that each individual is decidedly attempting to mitigate. At the local level, each actor is specifically attempting to approach a mediating norm value – by weighing the social ties, stated norm values, and frequency of appearance for each alter, the agent attempts to reach a middle ground representing the sum of their connections against their own norm value. Across many trials, however, regardless of whether the model modulates DE randomly, CS randomly, or indeed, entirely substitutes CS with a different approach, the march towards polarization continues.

As was stated above, there are more modifications that may be made to the model which may substantively alter the outcome of this model. Indeed, other complications may be added into the model – for example, while alters may tweet often, what is the probability that agents actually see and react to those tweets? Nevertheless, this simplified model provides at least one plausible avenue of exploration into how polarization occurs. In the case of the phenomenon being explored, observational data remains suggestive at best – by creating and exploring these models, researchers have, at the very least, the ability to explore possible explanations.

References

Adamic, L. A., & Glance, N. (2005). The political blogosphere and the 2004 US election: divided they blog. In *Proceedings of the 3rd international workshop on Link discovery* (pp. 36–43).

Aday, S., Farrell, H., Lynch, M., Sides, J., Kelly, J., & Zuckerman, E. (2010). Blogs and bullets: New media in contentious politics.

Aral, S. (2011). Commentary-identifying social influence: A comment on opinion leadership and social contagion in new product diffusion. *Marketing Science*, *30*(2), 217–223.

Aral, S., Muchnik, L., & Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, *106*(51), 21544–21549.

Baldassarri, D., & Bearman, P. (2007). Dynamics of Political Polarization. *American Sociological Review*, *72*(5), 784–811.

Bekafigo, M. A., & McBride, A. (2013). Who tweets about politics? Political participation of Twitter users during the 2011 gubernatorial elections. *Social Science Computer Review*, 0894439313490405.

Bemmaor, A. C. (1994). Modeling the diffusion of new durable goods: Word-of-mouth effect versus consumer heterogeneity. In *Research traditions in marketing* (pp. 201–229). Springer.

Bercovici, J. (2013, 23). *Justine Sacco And The Self-Inflicted Perils Of Twitter.* Retrieved from `http://www.forbes.com/sites/jeffbercovici/2013/12/23/justine-sacco-and-the-self-inflicted-perils-of-twitter/`

Centola, D., & Baronchelli, A. (2015). The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, *112*(7), 1989–1994.

Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, P. K. (2010). Measuring User Influence in Twitter: The Million Follower Fallacy. *ICWSM*, *10*, 10–17.

Chang, H.-C. (2010). A new perspective on Twitter hashtag use: Diffusion of innovation theory. *Proceedings of the American Society for Information Science and Technology*, *47*(1), 1–4.

Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. In *ICWSM.*

Conover, M. D., Gonçalves, B., Ratkiewicz, J., Flammini, A., & Menczer, F. (2011). Predicting the political alignment of twitter users. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on* (pp. 192–199).

Contractor, N. S. (1998). Self-organizing systems research in the social sciences: Reconciling the metaphors and the models. *Urbana*, *51*, 61801.

Cullum, J., & Harton, H. C. (2007). Cultural evolution: Interpersonal influence, issue importance, and the development of shared attitudes in college residence halls. *Personality and social psychology bulletin.*

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2007). Developing theory through simulation methods. *Academy of Management Review*, *32*(2), 480–499.

Deffuant, G. (2006). Comparing extremism propagation patterns in continuous opinion models. *Journal of Artificial Societies and Social Simulation*, *9*(3).

Deroıan, F. (2002). Formation of social networks and diffusion of innovations. *Research policy*, *31*(5), 835–846.

Drogoul, A., Vanbergue, D., & Meurisse, T. (2003). Multi-agent based simulation: Where are the agents? In *Multi-agent-based simulation II* (pp. 1–15). Springer.

Dubois, E., & Gaffney, D. (2014). The Multiple Facets of Influence Identifying Political Influentials and Opinion Leaders on Twitter. *American Behavioral Scientist*, *58*(10), 1260–1277.

Gilbert, N., & Terna, P. (2000). How to build and use agent-based models in social science. *Mind & Society*, *1*(1), 57–72.

Godes, D., & Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing science*, *23*(4), 545–560.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 1360–1380.

Gruzd, A., & Roy, J. (2014). Investigating political polarization on Twitter: A Canadian perspective. *Policy & Internet*, *6*(1), 28–45.

Iyengar, R., Van den Bulte, C., & Valente, T. W. (2011). Opinion leadership and social contagion in new product diffusion. *Marketing Science*, *30*(2), 195–212.

Janssen, M. A., & Ostrom, E. (2006). Empirically based, agent-based models. *Ecology and Society*, *11*(2), 37.

Katz, E. (1957). The two-step flow of communication: An up-to-date report on an hypothesis. *Public Opinion Quarterly*, *21*(1), 61–78.

Katz, E., & Lazarsfeld, P. F. (1970). *Personal influence, the part played by people in the flow of mass communications*. Transaction Publishers.

Kelman, H. C. (1958). Compliance, identification, and internalization: Three processes of attitude change. *Journal of conflict resolution*, 51–60.

Kiesling, E., Günther, M., Stummer, C., & Wakolbinger, L. M. (2012). Agent-based simulation of innovation diffusion: a review. *Central European Journal of Operations Research*, *20*(2), 183–230.

Leskovec, J., Backstrom, L., & Kleinberg, J. (2009). Meme-tracking and the dynamics of the news cycle. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 497–506).

Liben-Nowell, D., & Kleinberg, J. (2007). The link-prediction problem for social networks.

Journal of the American society for information science and technology, *58*(7), 1019–1031.

Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., & boyd, d. (2011). The revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. *International Journal of Communication*, *5*, 31.

Louie, M. A., & Carley, K. M. (2008). Balancing the criticisms: Validating multi-agent models of social systems. *Simulation Modelling Practice and Theory*, *16*(2), 242–256.

Luck, M., McBurney, P., & Preist, C. (2003). *Agent technology: Enabling next generation computing (a roadmap for agent based computing).* AgentLink/University of Southampton.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, *60*(3), 531–542.

Marwick, A. E., & boyd, d. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society*, *13*(1), 114–133.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 415–444.

Moretti, S. (2002). Computer Simulation in Sociology: What Contribution? *Social Science Computer Review*, *20*(1), 43–57.

Nobelprize.org. (2015, 15). *Thomas C. Schelling - Facts.* Retrieved from `http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2005/schelling-facts.html`

O'Sullivan, D., & Haklay, M. (2000). Agent-based models and individualism: is the world agent-based? *Environ Plann A*, *32*(8), 1409–1425.

Rogers, E. M. (2010). *Diffusion of innovations.* Simon and Schuster.

Ryan, B., & Gross, N. C. (1943). The diffusion of hybrid seed corn in two Iowa communities. *Rural sociology*, *8*(1), 15–24.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology*, *1*(2), 143–186.

Schmidt, J.-H. (2014). Twitter and the rise of personal publics. *Twitter and Society, edited by Karin Weller, Axel Bruns, Jean Burgess, Merja Mahrt, and Cornelius Puschmann*, 3–14.

Segerberg, A., & Bennett, W. L. (2011). Social media and the organization of collective action:

Using Twitter to explore the ecologies of two climate change protests. *The Communication Review*, *14*(3), 197–215.

Szell, M., Lambiotte, R., & Thurner, S. (2010). Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences*, *107*(31), 13636–13641.

Valente, T. W. (1996). Social network thresholds in the diffusion of innovations. *Social networks*, *18*(1), 69–89.

Van den Bulte, C., & Lilien, G. L. (2001). Medical innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology*, *106*(5), 1409–1435.

Van Eck, P. S., Jager, W., & Leeflang, P. S. (2011). Opinion leaders' role in innovation diffusion: A simulation study. *Journal of Product Innovation Management*, *28*(2), 187–203.

Wang, D., Pedreschi, D., Song, C., Giannotti, F., & Barabasi, A.-L. (2011). Human mobility, social ties, and link prediction. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1100–1108).

Weber, I., Garimella, V. R. K., & Batayneh, A. (2013). Secular vs. islamist polarization in egypt on twitter. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (pp. 290–297).

Yardi, S., & Boyd, D. (2010). Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of Science, Technology & Society*, *30*(5), 316–327.

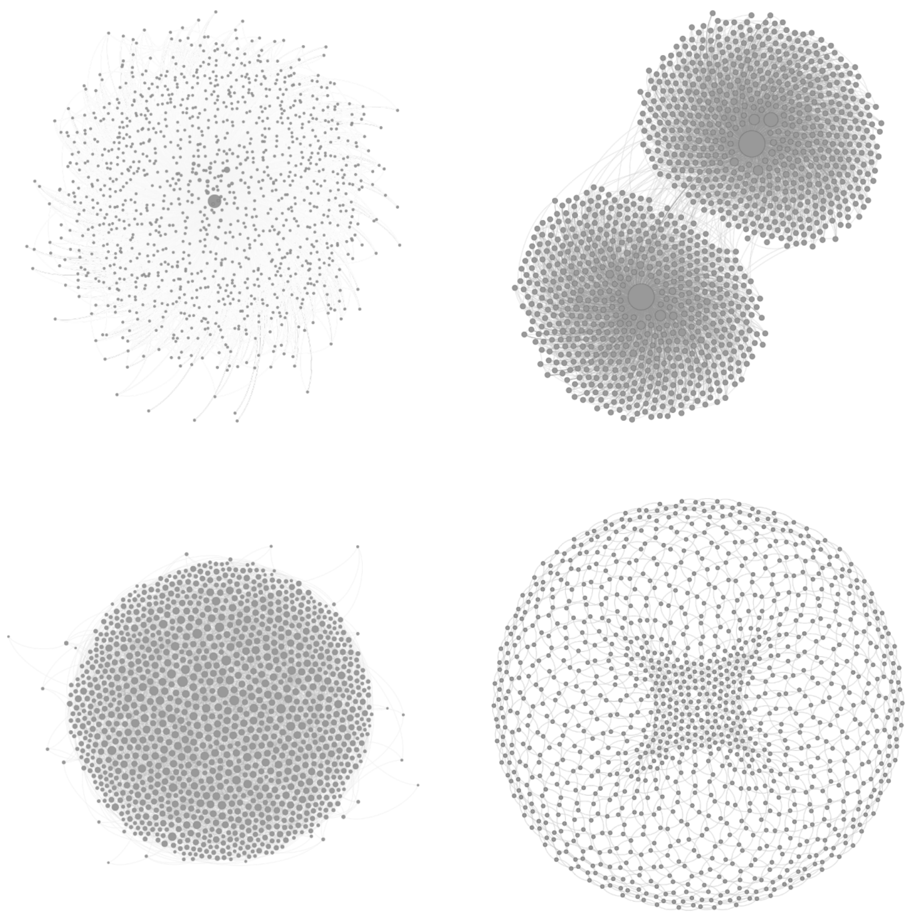

Figure 1: The four network models utilized in this model. Clockwise, from top right, they are 1. A Barabasi/Albert network with 1,200 nodes and $< k >= 4$, B. Two Barabasi/Albert networks with 600 nodes each, both approximating $< k >= 4$, joined by 100 random nodes between each "pole" (defined as a "Barabasi/Albert Barbell"), C. An Erdos-Renyi graph with 1200 nodes and $< k >= 4$, and 4. A lattice network with roughly 1200 nodes and $< k >= 4$.

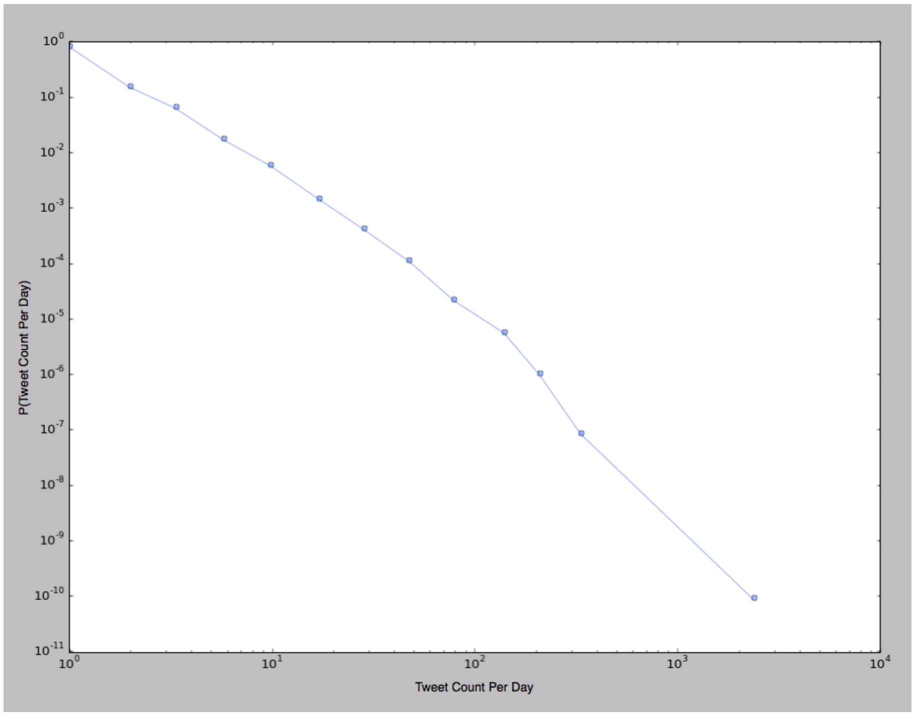

Figure 2: A log-binned probabilistic distribution of tweets per user in a given day on Twitter. February 28, 2015 Tweets per user selected from a 10% sample, including 31,162,421 tweets and 11,363,984 users

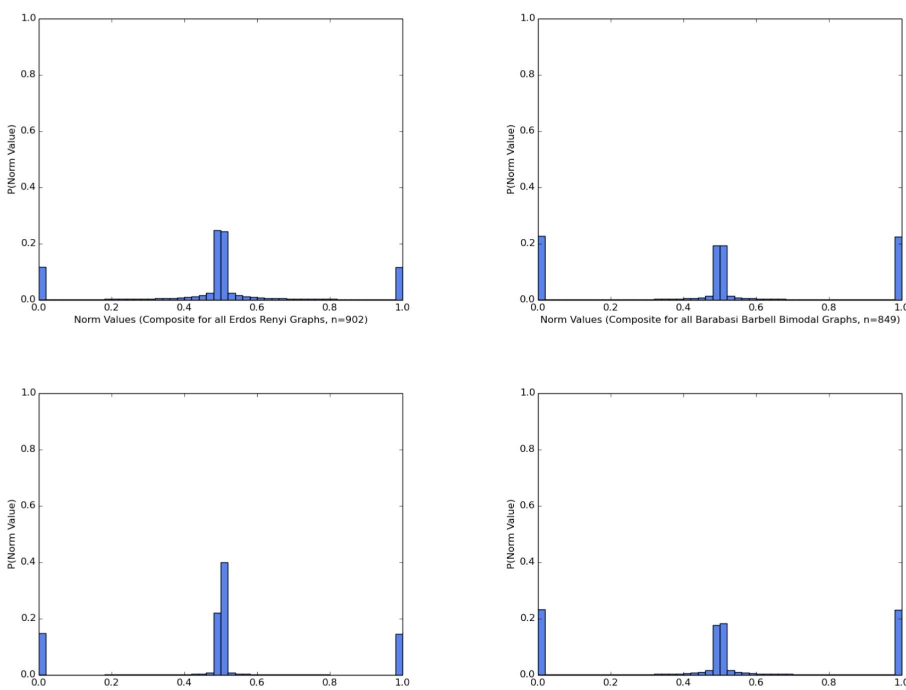

Figure 3: Norm value distributions for agents across all trials of simulations at timestep 100 per network model – note that in each case, there are clear cases of polarization for groups, while a subset of individuals remain relatively close to the default null norm value of 0.5

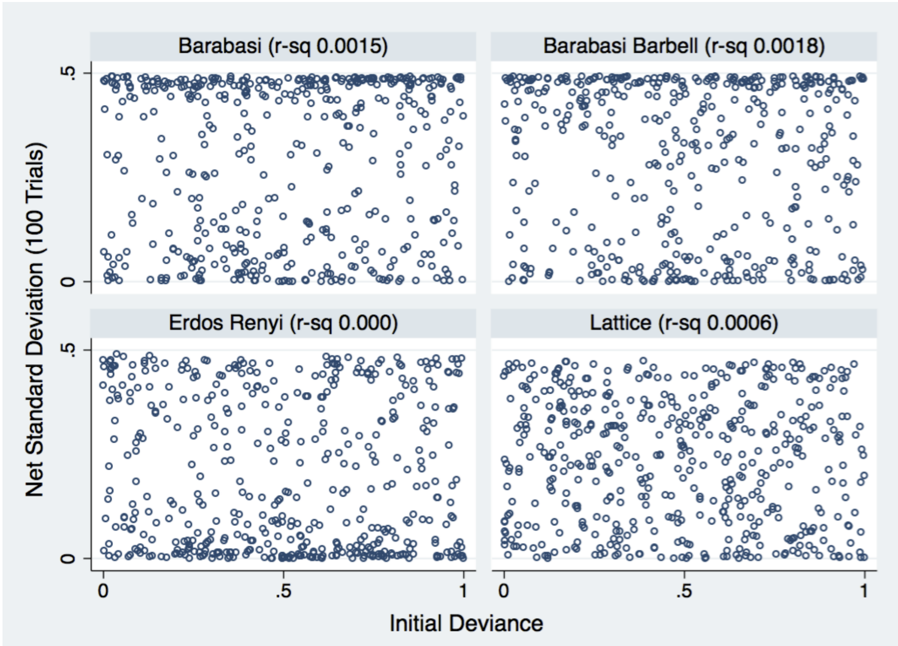

Figure 4: Standard Deviation per simulation given a DE value. Each point represents a simulation, where a value of 0.5 along the y-axis represents total polarization and a value of 0.0 represents no polarization given a particular value of DE between 0 and 1. R-squared is reported for each scatter plot, none of which are significant at p ¡ 0.05

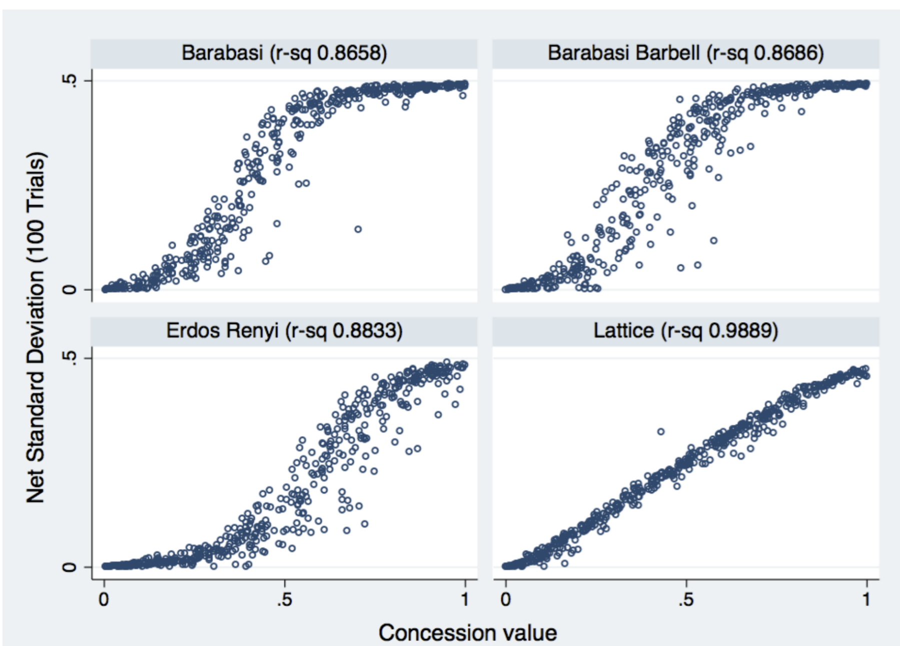

Figure 5: Standard Deviation per simulation given a CS value. Each point represents a simulation, where a value of 0.5 along the y-axis represents total polarization and a value of 0.0 represents no polarization given a particular value of CS between 0 and 1. R-squared is reported for each scatter plot, all of which are significant at p ¡ 0.05.

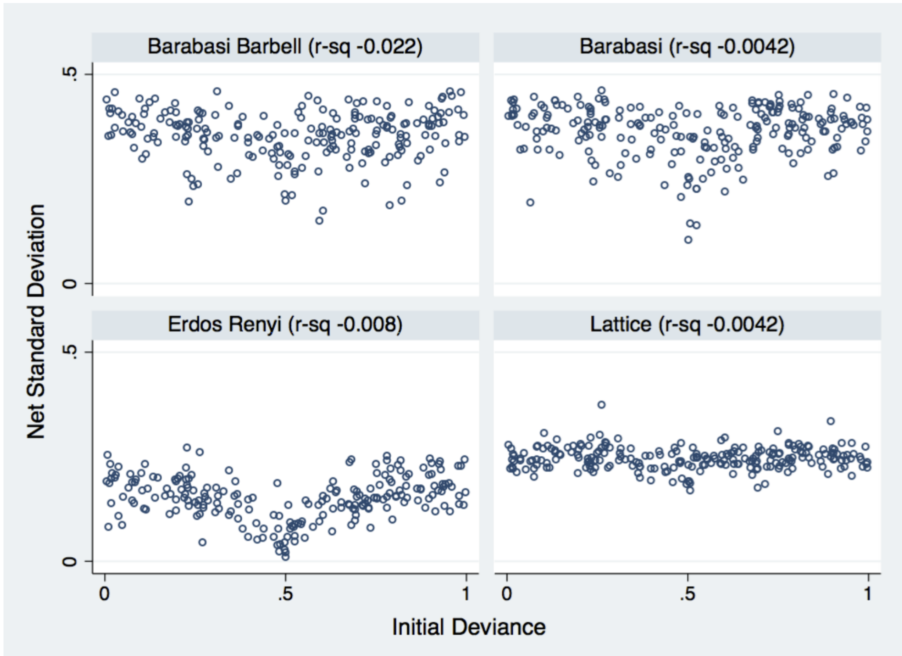

Figure 6: Standard Deviation per simulation given a DE value. Each point represents a simulation, where a value of 0.5 along the y-axis represents total polarization and a value of 0.0 represents no polarization given a particular value of DE between 0 and 1. R-squared is reported for each scatter plot, none of which are significant at p ¡ 0.05. Note, however, that the results are significantly less randomly distributed than in the initial mode in which CS was uniformly assigned across agents.

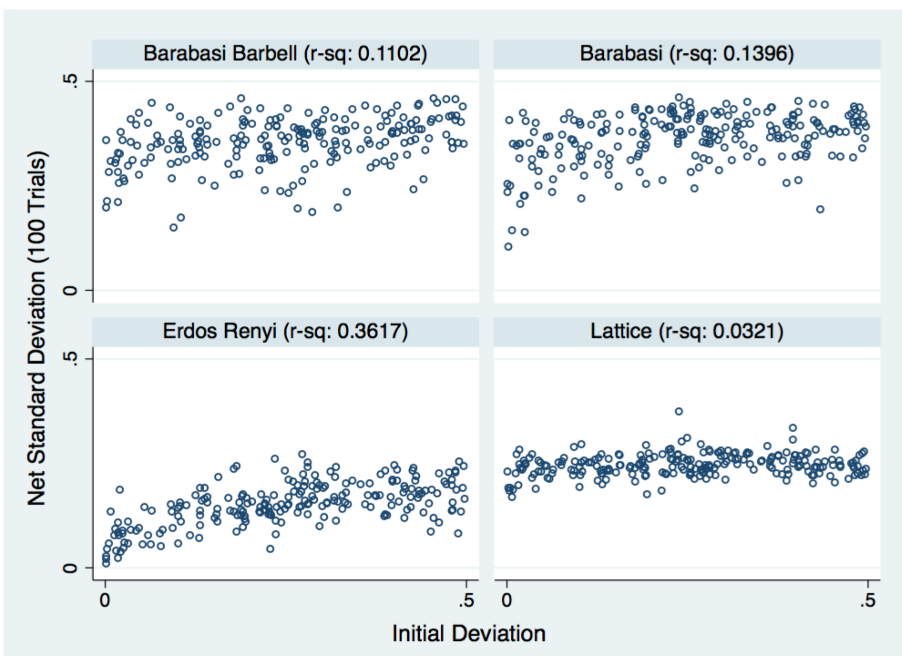

Figure 7: Standard Deviation per simulation given a DE value. Each point represents a simulation, where a value of 0.5 along the y-axis represents total polarization and a value of 0.0 represents no polarization given a particular value of DE between 0 and 1. R-squared is reported for each scatter plot, all of which are significant at p ¡ 0.05. Note that when cast as an absolute deviance from 0.5, the relationship becomes clear.

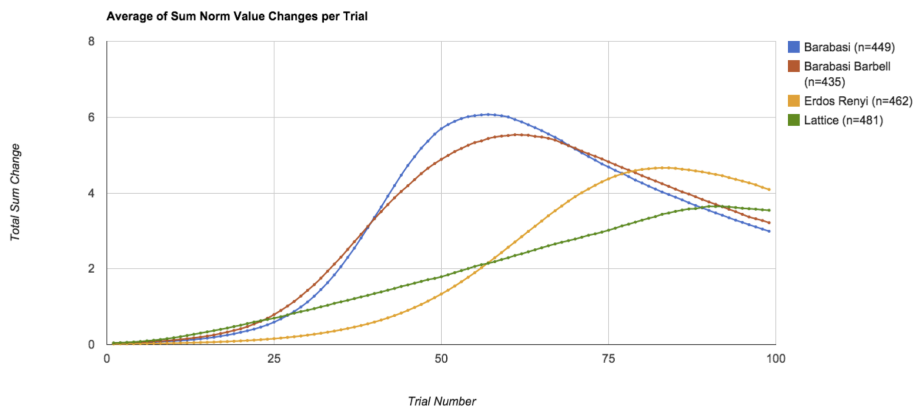

Figure 8: Average net norm value shifts of nodes per trial per network topology in a model where both DE and CS are held as constants across agents. Note that Barabasi-style graphs, which bear much closer resemblance to real world networks, experience earlier and larger shifts in norm change per agent relative to Erdos-Renyi and Lattice networks.

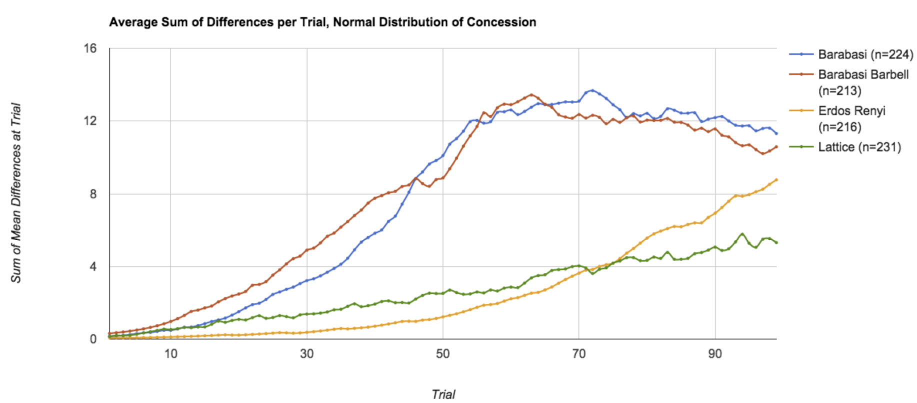

Figure 9: Average net norm value shifts of nodes per trial per network topology in a model where only DE is static and CS is a normal distributed value. Note that Barabasi-style graphs, which bear much closer resemblance to real world networks, experience earlier and larger shifts in norm change per agent relative to Erdos-Renyi and Lattice networks. In this graph, results are subject to much more variation than 8, yet the general qualitative pattern remains clear.