The Rise and Fall of FidoNet: The Geographic Growth and Decline of an ICT Community

Griffith Rees

St Cross College

University of Oxford

Doctor of Sociology

Trinity 2014

To my parents. Nothing in this world can ever recoup the morning drives to school, the late night blobs of peanut butter, the impeccable standards of diction, grammar and layout, and the exhausting drain of my boutique eduction. This is for you.

Acknowledgements

It takes a village; an indescribably generous, brilliant, indefatigable village guided me through this laborious swamp.

My oldest friend Ross inspired me with ***BBS: The Documentary*** *and the director—the generous Jason Sadofsky—gave excellent advice and crucial data. Phillip C. Dykstra, Ulrich Schroeter and the incredible IPUMS and NHGIS projects provided essential GIS data and support.*

Anne, Jane and Dorota gave pep-talks and help seasoned with wisdom, alongside the ground breaking ideas and reasources of the CABDyN group at the Saïd Business School. Michael's enthusiasm, encouragement and support started me on this journey I never thought I was capable of, and his analytic rigour and guidance ripened many parts of this haphazard thesis. Tess held my hand through the mazes my mind makes and Jacquie arranged those mazes into something readable while Richard made crazy coffee. Tom Snijders inspired, encouraged and gave excellent methodological suggestions and Omar saved me from a deep mistake I nearly made. And Christiaan's tough love got me through the end (ultimately as kindly as could be).

There is a greater debt I have to three people without whom this document would not exist. Andrew Elliott was enourmously helpful in developing the approach, the code and the math for chapter 5. When things became so difficult computationally, he parallelized running simulations on the server to save time. Nicole—who traveled thousands of miles, bought me jeans, read my drafts—held my hand in the final hours (as I write this now), wiped my tears and kept me human. Finally: Felix I will never know what my life would be without your incomparable generosity, your brilliance and your support, even when things were so difficult for you. I would not still be here without you, and this document certainly would not exist without you. Thank you for everything.

Funded by the Anthony Heath Scholarship and the European Commission: FP7 FET Open Project ITCeCollective (Contract 238597).

Abstract

This thesis studies the time evolution of a computer mediated community called FidoNet. The introduction explains what FidoNet is, briefly details its history and sociological significance and sets out research goals for the thesis as a whole. The second chapter—covering background and relevant literature—relates FidoNet to other social phenomena and reviews relevant sections of work on dynamic approaches to social systems. It also describes how FidoNet Nodelists were combined with geographic data on US telephone lines and 1990 Census data so that the geographic growth and decline of FidoNet could be mapped and analysed.

The third through fifth chapters are substantive starting with an empirical analysis on the spatial growth of FidoNet in the United States, covering a variety of different ways in which distance could have mediated the contagiousness of FidoNet as a system. The fourth considers decline as a similarly contagious process, demonstrating that FidoNet's most obvious competitor—the internet—may in fact have discouraged its decline while the short-term influence of leaving events on current sysops fits the data far better as an explanation. The other explanatory variable of this section—long-term social influence—exhibits an unexpected sign, suggesting that perhaps there were incentives to maintain FidoNet that were most prevalent in the long-term.

The fifth chapter attempts to tease out which different mechanisms may have been at work during decline, using an Agent-Based Modelling (ABM) approach and specifically considering individual rather than aggregated behaviour. The 6th and final chapter summarises the findings.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> *I had no idea it was going to become so amazingly large. Or long lived, that's the part that gets me. —Tom Jennings (BBS: The Documentary)*

In 1983 Tom Jennings, a computer hobbyist living in San Francisco, wrote a programme for sending messages to his friends via phone lines and early, primative, consumer-grade modems. He called the software FidoNet. Despite the severe limitations of the hardware and operating systems then available, FidoNet provided many of the features now expected of a modern day Information Communication Technology (ICT): leaving messages for for other people, like live or asynchronous computer games and chatting in real time. By today's standards the largely text-based interfaces would seem primitive, but it was highly innovative for its day, allowing people to interact socially across continents at relatively small costs, ten years before the internet became available to consumers.

In little over a decade, FidoNet grew from a small collection of friends into a community of millions, spanning all six inhabitable continents.[1] By 1995 there were nearly 40,000 sysops—volunteers who ran local connection points usually at their own expense—with at least a few million users in the US alone. However, from its peak that year, FidoNet began to decline, first in the US and then across the world. By 2005 the number of sysops was less than a twentieth of the peak, though it still covered a wide geographic area. Today it is a small but resilient

[1] There may have even been a node on all seven continents: Steve Freddette is listed as sysop of a node in Antarctica for the March 6, 1998 Nodelist.

and geographically diverse community, with membership in at least the thousands (though the current number of users is difficult to estimate).

Understanding the process by which this community grew and then declined is the subject of this thesis. The motivation for choosing this topic is fourfold. First it is a chance to model a community's growth and decline starting from its inception, and to compare the dynamics governing those different phases within one system. Second the dataset's quality, sheer size and temporal and geographic resolution is unusually high, allowing the system to be approached from a micro-mechanismal perspective rather than solely in the aggregate. Third, the dataset provides an opportunity to study key, micro-level components of spatial contagion, in particular the contagiousness of joining events and how that decays with distance (chapter 3) and the extent to which leaving events can also be contagious (chapters 4 and 5). Finally, the subject matter is an early example of an ICT that developed an extensive and diverse community, and such systems are only becoming significant in post-industrial societies.

In this chapter I discuss how FidoNet resembles aspects of a technological innovation and a community, as well as components of a voluntary organisation, a social movement and a public good. I then provide an overview on the diffusion literature that is methodologically closest to the chapters that follow, and conclude by outlining the remainder of the thesis, as well as summarising its structure, and motivating the research questions addressed.

1.1 A Technology and Community

It is perhaps most sensible to think of FidoNet as both a technological innovation and as a community. While the technical details of how FidoNet grew and functioned are covered in chapter 2, I will briefly summarise its features and what made it innovative for its time. Next I will highlight how, like many ICTs, FidoNet stitched together people into what I argue can be thought of as a community, and indeed thinking of FidoNet as a community is helpful when trying to understand its dynamics, particularly in decline (see chapters 4 and 5).

I close this section by comparing FidoNet to other sociological phenomena—voluntary organisations, social movements and public goods—and how the models and theories of their respective literatures are useful for the chapters to come. In turn, their similarity to FidoNet implies that my results could be useful in modelling their dynamics as well.

1.1.1 A Technological Innovation

FidoNet was an innovation that built on a pre-existing technology called Bulletin Board Systems (BBSs). BBSs were an early ICT which resembled modern online discussion forums. A person called a sysop would leave their computer and modem on and phone line free so that users—people who used BBSs—could call and connect via their computers to post messages and read replies.[2]

A limitation for BBSs was the cost of phone calls. While US local calls tend to be unlimited with a normal monthly phone bill, calls outside a local region incur extra charges. With the very slow speed of early modems, BBSs were forced to remain spatially isolated communities because connecting long-distance was too expensive.

FidoNet was developed to solve this problem. Each FidoNet BBS was set to forward all the messages it received that day to its nearest neighbour at midnight when long-distance calls were cheapest. That meant that each BBS sent messages that could eventually reach much further, allowing social ties to be maintained over great distances and a greater diversity of people to be involved, especially once FidoNet spread beyond North America. At its peak FidoNet was the largest BBS community in the world, and despite contracting significantly since 1995, arguably remains among the most resilient of the pre-internet ICTs, continuing to connect people from Southeast Asia to Russia to South America.

Like most technologies FidoNet has had competitors. Initially other BBSs, Advanced Research Projects Agency NETwork (ARPANET) and Joint Academic

[2] See section 2.1 for a more detailed description of the development and use of BBSs.

NETwork (JANET) (government projects that would become the internet)—and the newsgroup network Usenet that operated on those projects—provided similar services. Eventually the internet itself, through early Internet Service Providers (ISPs) like CompuServ and America Online (AOL) became popular in the mid 1990s, offering far more features, and as these technologies spread FidoNet declined. Modern ICTs like Internet Relay Chat (IRC), internet forums, Facebook and mobile phone applications like WhatsApp provide many of the features that once made FidoNet unique and highly attractive. These platforms are more advanced and convenient—and in many cases profitable companies—but bear profound resemblance to FidoNet.

It would seem, therefore, that FidoNet can be understood as a technological innovation that improved upon an existing technology, spread rapidly and eventually was largely replaced by newer competitors. In modelling its dynamics, we can prudently draw upon the extensive innovation diffusion literature described below in section 1.2.

However, unlike many of its contemporaries, FidoNet remains in operation three decades after its release, and while small in membership, it continues to connect people across much of the world. It could be the case that it is simply an old and outdated technology which has a particularly stubborn set of 'laggards'—those who are last to change over to a new technology in Rogers's (2003) categorisation.

Alternatively, this may be because the internet was not a direct competitor. FidoNet was connected to Usenet by 1985 and the internet as early as 1993, so calling them direct competitors is perhaps inappropriate. It is also highly likely that FidoNet's current persistence in part stems from its internet connectivity, which further reduces the costs of long-distance connections (replacing the midnight transfer system) and provides access for anyone with an internet connection, anywhere in the world. Thus in that sense FidoNet has taken advantage of new innovations to survive.

However, the results of chapter 4 suggest that approaching FidoNet exclusively as an innovation—especially with the internet as a pure competitor—only explains part of the results observed, and its social side, which I argue fits best with the community and social capital understanding of some social systems, provides some insight into other significant results from FidoNet's decline, and in part motivates the mechanisms included in the simulation of chapter 5.

1.1.2 Community

What separates FidoNet from the internet, and connects it conceptually to ICTs like IRC is its standing as a community created and run by its users and sysops. This classification is complementary to its standing as a technology in as much as what makes FidoNet attractive to potential new users may be interest in socialising with current users. Equally: what induces one to leave or stop using FidoNet may be a function of social ties leaving, just as much as it may be a preference for a new supplanting technology.

In the following sections I review some of the theoretical literature on community and argue that it is important to consider FidoNet as a community, both in terms of analytical approach and the applicability of the results of this thesis to community dynamics in general.

The Idea of Community

Perhaps the most influential work on community within sociology is Tönnies' (1887) *Community and Society*. Tönnies combined locality, strongly reciprocal social ties, a sense of belonging and a pastoralist narrative into what he called *Gemeinschaft*. *Gemeinschaft* was contrasted with a threatening, atomistic, modern form of social association he called *Gesellschaft*, which he associated with the rise of urban centres and the de-coupling of occupations and upbringing.

Gesellschaft derives from the social contract ideas of Rousseau (1762) and others, where an individual's relationship with their state is codified in the rule of law, and that formal agreement generates a sense of common identity. Tönnies believed *Gesellschaft* was becoming dominant with the advent of industrialisation, but astutely noted that these were merely two endpoints on a continuum: no community embodied a pure form of either extreme. Nevertheless, he made clear his preference for *Gemeinschaft* over *Gesellschaft*, and worried about our descent into the latter.

Tönnies' ideas were soon recast by Durkheim and Halls (1893) into mechanical and organic solidarity. Mechanical solidarity resembles *Gemeinschaft* in that it is associated with rural, close-knit, often highly religious villages, homogeneous in world-view and life-trajectories. Similarly, organic solidarity has roots in *Gesellschaft*: complex interrelated divisions of labour in densely populated urban areas that break with old traditions. However, Durkheim upends Tönnies' normative framework, suggesting that organic solidarity is the positive force, freeing us from artificial historical social bonds and forging new, dense, organic ties while pushing society forward.

Both Tönnies and Durkheim present a continuum between an informal, indigenous, local source of community, and formal, institutional bonds unfettered by origin. This dichotomy persists within the literature (and Tönnie's terminology seems more prevalent), but much of the work since has fragmented with regards to more specific and stringent definitions, and the advent of ICTs has only exacerbated the complexity of the debate.

Recent Conceptual Strands

Bell and Newby (1974) provide an influential overview of many of the developments in community research since Tönnies. Of the 29 chapters, six devote large sections to the difficulties and disagreements surrounding the idea of community, lamenting Hillery's (1955) oft-cited list of 94 distinct definitions of community and König's (1968) rejection of community as a useful concept for comparison altogether. Indeed Stacey's (1974) chapter *The Myth of Community Studies* ends with her own list of 30 'tentative propositions about local social systems'. In a separate work, Plant (1974) suggests that defining community is as problematic as defining 'democracy' or 'game' (following Wittgenstein) and that it should be approached as a nexus of related ideas and not as a well-defined conceptual whole.

Blackshaw's (2009) more recent overview of developments that began in the early 1980s suggests that the situation has become even more complex.[3] Blackshaw mentions three key developments. First is Anderson's (1983) 'imagined community': a theory of nationalism that suggests communities larger than small villages require some imagined sense of belonging. Cohen (1985) studies the 'symbolic construction' of community and the extent to which community identification, customs and symbols are a *part* of community rather than a consequence of it. More recent is Wegner's (2002) concept of 'imaginary community', which stems from utopian idealised[4] images of community which help define and express a culture's values and aspirations.[5]

To these I add two further developments: first the aggregative (macro) branch of social capital research most prominently pursued by Putnam (2001) and Wellman (1999), and second the rise of ICTs and their sudden and deep embedding in social intercourse.

Work on individual social capital focuses on the advantage an individual accrues from their *position* in a network of social relationships (Lin, 2002). Aggregative social capital considers the effect of the *structure* of social relations on the *group as a whole*, and in turn the benefit (or detriment) to members from their association (Putnam, 2001, 19). Social capital research primarily sees communities as networks of relationships unbounded by their spatial proximity, a view Keller (2003, 293–298) criticises sharply.

Putnam (2001) also highlights two useful concepts related to space—bridging and bonding ties.[6] Bonding ties relate back to the intimate, spatial homogeneity of *Gemeinschaft*, 'creating strong in-group loyalty' (Putnam, 2001, 23). Bridging ties are weaker, unbounded by distance and more diverse, providing a 'sociological WD-40' (Putnam, 2001, 23) which seems to resemble the kind of association suggested by

[3]Strangely, Blackshaw suggests that before 1980 'broadly speaking, everyone concurred', specifically citing Bell and Newby (1974) as evidence.

[4]Tönnies' *Gemeinschaft* is a canonical example of a utopian sense of community.

[5]Keller (2003) calls her opening chapter *The Passionate Quest*.

[6]These concepts are quite similar to Granovetter's (1973) strong and weak ties from an explicitly macro-social perspective.

Gesellschaft: less personal but still some sense of common identity and association.

ICTs and Community

These ancient questions of scale, boundedness, *Gemeinschaft* and *Gesellschaft* have exploded over the advent of ICTs. Internet and/or computer based ICTs have diffused through western societies much faster than older ICTs such as the telephone and television (Putnam, 2001, 169). Their unprecedented capacity for social intercourse has given rise to 'online communities', where the internet provides a major (if not the primary) means of interaction between groups of individuals.[7]

The capacity for new forms of ICT interaction to cross social and physical boundaries has polarised research into those supportive of online community (Rheingold, 2000; Castells, 1996; Wellman, 1999) and those who find it 'lacks 90 per-cent of the defining criteria' (Keller, 2003, 293). Indeed, as Wellman (1999, 331) observes, this debate can be seen as *presentist*: treating current worries about potential changes in community as entirely new, ignoring the history of similar debates over similar worries. In many ways online communities are the modern *Gesellschaft* conjuring for some a looming spectre of anomie:

> *Virtual communities, as indicated by their name, are mirages. The absence of full sensory contact creates shadow communities at best ... individuals increasingly isolated and cyberville as an escape from the world. The illusion of community may mask that isolation and eliminate the awareness of human interdependence. Whatever cyberville does offer—and there is as yet no consensus on what that is—it does not demand the investment in time, effort, and commitment to shared goals or the give and take of real communities. (Keller, 2003, 298)*

I find Wellman (1999, 355) an appropriate rebuttal for this claim, despite being published four years prior:

> *Pundits worry that virtual community may not truly be community. These worriers are confusing the pastoralist myth of community for the reality.*

[7]While telephones have long provided long-distance means of social intercourse, they are a primarily dyadic technology while the internet is well suited to group discussion. Arguably, it is this distinction that has led to internet communities without comparable telephone communities.

> *Community ties are already geographically dispersed, sparsely-knit, connected heavily by telecommunications (phone and fax), and specialized in content. There is so little community life in most neighborhoods in western cities that it is more useful to think of each person as having a personal community: an individual's social network of informal interpersonal ties, ranging from a half-dozen intimates to hundreds of weaker ties. Just as the Net supports neighborhood-like group communities of densely-knit ties, it also supports personal communities, wherever in social or geographical space these ties are located and however sparsely-knit they might be.*

Throughout his work, Wellman argues for a more nuanced view which avoids Keller's (2003) reification of *Gemeinschaft*-like villages and blends offline and online social interaction. He argues that in reality complex, overlapping contexts of social interaction have supplanted the old idea of *polis*. Rather than approaching the social fabric of western life as a set of localised communities, we should approach it as a set of ego-networks, where each individual builds their own individual community through the diverse ties that bridge the many contexts of their lives: work and home, school and play, offline and online.

FidoNet and Glocalisation

FidoNet occupies a strange intersection within these debates. On the one hand it is clearly an ICT: interaction flows through phonelines, modems and computers via message lists on a variety of topics. It was explicitly conceived to transcend the limitations of distance, and its eventual global reach meant that Russians might have discussed politics with Thais, who in turn argued about sexuality and programming with Swedes (provided they knew a common language). And in all likelihood these individuals never would meet, their only affiliation being their use of FidoNet. This certainly coheres closely with *Gesellschaft*, though with at most a virtual notion of an underlying social contract.

However, FidoNet was composed of BBSs, which were by necessity *local* communities. Certainly in the early days of FidoNet, the fact that each BBS would forward its daily messages on provided a 'bridging' connection between otherwise tightly 'bonded', localised communities, traditionally based around computer hobbyists. As mentioned above, there is evidence to suggest that BBS users and their sysops would meet offline, and because of a lack of media coverage it is quite likely FidoNet propagated via face-to-face interaction through pre-existing social ties.

Robertson (1995) calls this combination of local and global 'glocalisation', a concept Wellman uses frequently (1999; 2002). As it pertains to community, glocalisation seems to unify the *Gemeinschaft*/*Gesellschaft* distinction within ego-networks and the modern tendency for people to have ties that are dense and local and ties that are distant physically and emotionally. FidoNet seems to embody this tendency as an organisation through its local BBS sysop communities and means of propagation coupled with the potential for long-range relationships emerging through participation in FidoNet message boards.

Calling FidoNet a community is controversial, but that controversy stems more from disagreements about the word 'community', than from ICT communities like FidoNet. Through its local component, FidoNet fits old ideas about *Gemeinschaft* far better than other ICT communities, such as IRC and Usenet. Even Keller (2003, 294) acknowledges some potential for locally embedded ICT communities such as San Francisco's famous WELL, Charlotte's Web (from Charlotte, North Carolina) and a community based in Blacksburg, Virginia. Whether we prefer imagined communities based around a common purpose, or Wellman's globalised ego-nets, aspects of FidoNet are relevant and substantive.

Treating FidoNet as a voluntary community organisation in the US (where the network was largest and densest), is especially interesting against the background of Putnam (2001). Putnam studies what he believes to be the decline of American social capital, which he attributes to decreased membership in voluntary, civic, and religious organisations. By the early 1990s FidoNet was a registered 503c non-profit organisation in the US, and membership—which I define as the number of users and sysops—surpassed estimates of 2 million worldwide (Bush, 1993). Over this same period, the membership of the organisations Putnam tracked was steadily declining. FidoNet and the internet in general may provide a new form of social involvement that is replacing traditional community organisations, though I do not have the space to explore this hypothesis here. As a comparison, FidoNet was certainly among the largest American voluntary organisations at its peak, comparable to the Rotary Club and the Boy Scouts of America.[8]

FidoNet combines aspects of traditional voluntary organisations, such as the Rotary Club and community sports; open content communities like Wikipedia, Open Source Software, and file sharing communities; and general discussion and social networking sites like Usenet, Facebook, and Twitter. In particular it closely resembles a free-rider public good (Olson, 1971), where users (who connected to the network for free) free-rode on the efforts of sysops (who ran the BBSs).

FidoNet resembles classic voluntary organisations because of its free-rider nature and because such organisations require physical infrastructure, be they BBS servers in the case of FidoNet or local chapters and offices in the case of offline organisations. Like an organisation that collects dues for membership, FidoNet BBSs sometimes imposed time limits on use, requiring users to upload a file or message in order to extend their connection quota. FidoNet's geographical dispersion faced similar challenges of expanding and supporting a growing membership, and work on Rotary clubs (Huang and Gould, 1974; Sugiura, 1986) and community sports (Liljeros, 2001), as well as unions (Hedström, 1994; Biggs, 2003) provide analogues.

Open content communities and open source software provide other similarities. FidoNet, like Usenet, BitTorrent, and Rapidshare, was a community which shared files, including illegally copied content. It also, like an open source project, was centred around freely available code that was improved by the community to accommodate its changing needs. These communities often benefit from network externalities:[9] as the number of users increases, so does the quantity and quality[10] of

[8]Since the American part of FidoNet accounted for about half of all the BBSs, a reasonable estimate would be greater than a million members (half of the 1993 estimate of 2 million, two years before the network membership peaked). Total Boy Scouts of America membership in 2007 is nearly 3 million (Boy Scouts of America, 2007).

[9]Network externalities capture the tendency for certain systems to benefit from increases in size. eBay, for example, is useless without sellers, and the more sellers eBay attracts, the more valuable it is as a marketplace.

[10]The quality of the code should increase because more people will be improving and checking it.

content or code. Also, the speed of transfer in content communities tends to increase due to the distributed nature of how files are provided. It is unclear whether FidoNet became more or less efficient over time, net of exogenous improvements such as faster modems and computers. However, it is quite plausible that the network became more robust, diverse and produced more content as the number of BBSs and users increased.

The free-rider framework applies here as well, where users 'leech' content (download files without uploading, thereby using the network without contributing) or use software at little to no cost while contributors upload files or contribute bug reports and code. Also like these communities, FidoNet solved its own problems. The community radically restructured the network in 1986, solving an approaching crisis in the code which created a hard growth limit at 256 BBSs. Later as the network reached parts of Asia, the community developed a system for representing non-Latin character sets, just as the internet community eventually did with Unicode. All this was facilitated by the community itself, and made possible by free access to the code the network ran on.

Finally, FidoNet resembles the online forums and social networking sites that are becoming increasingly popular and important in everyday life.[11] Like other ICTs, FidoNet was a means of social interaction, providing an alternative forum for maintaining social ties and creating new ones. ICT communities also have network externalities, since the more people in each community, the more potential social connections per person. Like other ICT communities, anonymity was important in FidoNet since each user controlled how much other users knew about them, and the variety of discussions rivalled the diversity that now pervades the internet, with Echolists (FidoNet's topic-driven message threads like Usenet Newsgroups) devoted to politics, philosophy, television, art, and sexual-deviance.

FidoNet is unique in spanning these different sociological categories, but exhibits key features of all of them. As such I hope whatever results may emerge from this

[11] At least within developed nations.

research, and methodologies used, will have significant bearing on research into these other areas.

1.2 Diffusion

I now turn to the modelling literature that inspires much of the quantitative analysis that follows.

'Diffusion refers to the spread of something within a social system.' (Strang and Soule, 1998). More formally it is the dynamics of adoption behaviour[12] in a *susceptible population*, meaning the set of potential *adopters*. Generally, adoption is a binary activity (you have either adopted or not), and diffusion studies examine the increase in adopters over time, often with mathematical models. The size of the susceptible population is important because the dynamics of adoption tend to be sensitive to the proportion of the population that has adopted, rather than the overall level. This causes asymptotic behaviour at the extremes of very low and very high proportions of adoption.[13]

1.2.1 Diffusion Processes

The most studied diffusion mechanisms within sociology are broadcast, contagion and threshold effects.

A broadcast effect is an *exogenous* signal applied to the susceptible population (Mahajan and Muller, 1979). This often corresponds to an advertising campaign or media coverage that urges individuals to adopt. In most models such effects are assumed to be constant throughout the adoption process and to apply equally throughout the population (for an exception see Andrews and Biggs, 2006). FidoNet is unusual because while it spread rapidly and widely, there is little evidence that

[12] I will use adoption as the verb in this context, in the sense that you can adopt an innovation or a new behaviour. I may occasionally use *join* as well, in the sense that a person joins a social movement, or joins FidoNet (becoming a user or sysop).

[13] Proportion of adoption is the number of adopters divided by the susceptible population.

traditional broadcast effects were involved.[14] Further, the members of the network made no centralised attempts to increase the number of users.[15] However, I will incorporate a kind of exogenous effect into the model of local growth: the total number of BBSs in the network.

Contagion is the effect of an interaction event between an adopted individual A and a potential adopter B. If A has adopted an iPhone and B sees A using it, B's likelihood of adopting increases. Each interaction event presents an opportunity for B to *contract* (following an epidemiology metaphor) the innovation or behaviour A has adopted. As more people adopt, the effect on the remaining non-adopters intensifies because there are more adopters to interact with, and therefore a greater number of potentially contagious interaction events per person.

Broadcast and contagion effects are often combined (under the labels internal and external effects respectively) to form the Bass Model (for a review see Mahajan et al., 1990). These labels capture the fact that contagion is an *endogenous* process—meaning the dynamics of the effect are internal to the model and sensitive to the time-evolution of the process as a whole—and a broadcast effect is exogenous—where the dynamics are externally defined and independent of the state of the system.

While models based around contagion generally assume population homogeneity with random interactions providing the driving force behind adoption, a threshold process takes individual heterogeneity as its starting point. Early models suggested by Granovetter (1978); Schelling (1969) have thresholds based on how many people in the population have already adopted. The critical number of others required for A to adopt is A's threshold. Thresholds can be absolute numbers (A will only adopt if x people have adopted) or proportions (A adopts if x/p, where p is a population, have adopted). Young (2009) calls this a social influence process,

[14]Searching for FidoNet on ProQuest yields 32 results, four of which prior to 1990, and 20 between 1990 and 1995 (the peak). I take this to indicate that the media *reacted* to the growth of FidoNet, rather than drove it.

[15]The original system architects never expected the system would grow as large as it did, and two years after FidoNet was first released, Jennings and others had to re-write the software from scratch to cope.

focusing on the extent to which it is the collective influence exerted by a group on an individual, rather than random dyadic interactions, that are the driving force.

Thresholds may be many things, including differing levels of risk aversion or uncertainty, political alignment, or some proxy for an individual's tendency to conform to social norms. Rogers (2003) suggests that thresholds in innovation adoption correspond to different levels of technological interest and risk aversion, where first adopters are willing to try any new innovation (and take on potentially high risks), then early adopters adopt if the first adopters approve, and then mainstream adopters follow early adopters and so on.

One final type of diffusion from the economics literature is a social learning model of the sort described in Young (2009). He surveys a broad range of work[16] to develop a model that captures the basic theoretical features that apply across that literature. Since his formulation is new and less explored than contagion and threshold models I will not cover it in depth. Briefly: potential adopters are treated as rational agents collecting information on the innovation by observing adoption events in the population. Each has an information threshold, past which they will consider adopting. Agents accrue information as others in the population adopt and use the innovation, and the adoption curve is similar to a threshold model.

1.2.2 Application of Diffusion Models

While initially diffusion was primarily studied with respect to innovations and individuals (Ryan and Gross, 1943), similar models have been applied to business practices (Abrahamson and Rosenkopf, 1993), geographic regions (Hägerstrand, 1967; Hedström, 2005) and cities (Andrews and Biggs, 2006). I will be studying both locational diffusion (using cities) and person-to-person diffusion.

Empirically, diffusion research traditionally requires time-series data on adoption and the total susceptible population. This could quantify *How many people*

[16] He cites Bikhchandani et al. (1992); Banerjee (1992); Kirman (1993); Ellison and Fudenberg (1993, 1995); Kapur (1995); Bala and Goyal (1998); Smith and Sorensen (2000); Chatterjee and Xu (2004); Banerjee and Fudenberg (2004); Manski (2004); Golub and Jackson (2008) among others.

from susceptible population have adopted at t? or *Which cities have adopted at t?*. A mathematical model is then proposed on theoretical grounds and estimated with respect to the observed adoption data. As it happens the majority of available diffusion datasets of individual adoption exhibit a sigmoidal or 'S-curve' shape, with slow initial growth, then marked acceleration, followed by deceleration as adoption asymptotically approaches saturation (total adoption within the susceptible population). Those datasets that do not exhibit the deceleration phase tend to have some truncation effect which intervenes before the process can approach saturation (for example Biggs, 2003).[17]

The S-curve shape relates to the three mechanisms discussed above in the following ways. First a constant broadcast effect will not by itself generate an S-curve,[18] and will only do so when coupled with a contagion or threshold process (Figure 1.1). A contagion process generates an S-curve because its growth is exponential until it approaches full adoption, at which point it decelerates as it asymptotically approaches saturation. With thresholds, an S-curve merely requires a distribution of thresholds that is single-peaked (similar to a normal distribution) or an appropriate network structure.[19] Most population thresholds are assumed to be normal or truncated normal (Rogers, 2003; Young, 2009), and the cumulative normal distribution is sigmoidal.[20]

Most of the mathematical models of diffusion processes take inspiration from epidemiology, consisting of a system of differential equations (for an overview see Daley and Gani, 2001) which estimate the rate of adoption (rather than infection) over time. Like most epidemiological research, most social science diffusion studies

[17]Spatial diffusion research has not focused on whether locations adopt in an S-curve fashion, so it is possible the spatial data may contradict the S-curve regularity (though I think it is unlikely).

[18]A dynamic broadcast effect could produce an S-curve, but these are rarely considered. For an exception see Van den Bulte and Lilien (2001).

[19]Valente (1995); Centola and Macy (2007) study thresholds in a network context, where the threshold relates to a potential adopter's *ego network*: those they are socially tied to. It is possible for everyone in a population to have the same adoption threshold (say, 4 people in their ego network) but have an S-curve adoption pattern due to the network structure.

[20]Social learning models also feature the cumulative normal effect, via a distribution of information thresholds.

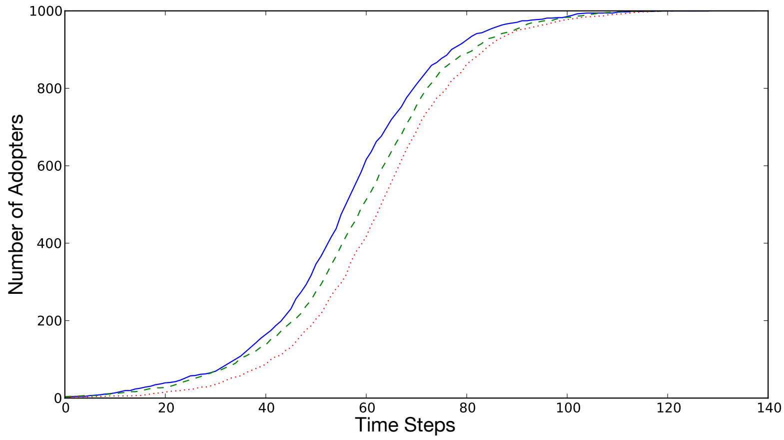


Figure 1.1: *Sigmoidal, S-shaped curves generated by 3 realisations of the same stochastic contagion simulation.*

also hinge on a homogeneous mixing assumption: every individual in the susceptible population is equally likely to interact with any other. Contagion models further assume that individuals are homogeneous, meaning they all have equal propensity to adopt.

As one might imagine, the more complex models attempt to relax these assumptions, often through the use of an event-history framework (Strang, 1991; Strang and Tuma, 1993) or social network analysis (Valente, 1995; Gould, 1993) (or some combination thereof). Event history models are well suited to incorporating population heterogeneity because they allow each individual to have an arbitrary set of properties which affect their probability of adopting. Social network analysis incorporates the tendency for interaction to be heterogeneous in both strength and probability. We tend to interact with and be influenced by our close friends and family more than our passing acquaintances.

While many of these extensions have yielded important, non-trivial results (Gould, 1993; Valente, 1995; Liljeros, 2001; Centola and Macy, 2007), thus far spatial extensions, which use distance as a proxy for interaction probability, have provided the most fruitful, substantive extension to classic diffusion research (Strang and Tuma, 1993; Hedström, 1994; Hedström et al., 2000; Andrews and Biggs, 2006), in great part due to the availability of high quality spatial time-series, and the relative paucity of data on network time-series and population heterogeneity in general.[21] The advantage of spatial studies which treat locations such as cities or regions as the unit of analysis is that much data exists on relevant properties of cities, allowing further incorporation of population heterogeneity.

One final assumption which pervades the social science literature on diffusion is that of permanent adoption.[22] Most diffusion models assume that individuals do not 'de-adopt' or leave the social movement. This means the process either continues until saturation or some exogenous event intervenes to slow or reverse adoption. This is in part due to the nature of many of these phenomena. For example, adopting an iPhone is a discrete event. 'De-adopting' an iPhone (discontinuing use) is a separate (though related) process.

However, any attempt to model the growth of FidoNet must address the possibility that growth slowed in part due to attrition. In fact, attrition occurred throughout the growth phase of FidoNet, as one would expect of any social organisation, and if we are to model social systems with dynamic membership (like social movements) we must take care to address this likely significant effect. I shall call this a *net effect*, where the total level observed is in fact net of de-adoption. Fortunately, I can address this issue because my data allows me to identify which BBSs are new.

1.3 Research Goals

The relationships between the aggregate or macro-level dynamics of a social system like FidoNet and the underlying micro-level mechanisms generating the observed behaviour is complex and constitutes a serious and interesting research challenge. Typically, this challenge has been addressed by focusing on settings and time windows in which certain mechanisms or processes can be assumed to dominate overall behaviour. For instance, much of the empirical literature that models innovation diffusion restricts itself to those known—*post hoc*—to have been successful, during

[21] See Valente (1995) for available datasets on network diffusion.
[22] See Young (2009); Hedström (2005) for exceptions.

a period of rapid growth where the rate of adoption is high and de-adoption events can be neglected.

The data used below covers an initial period of sustained growth, followed by a period of sustained decline. These periods correspond to two distinct phases in terms of the aggregate or macro-level behaviour of the system. Moreover, in each distinctive phase multiple mechanisms act at the micro-level. Hence, although joining events outnumber leaving events during the growth phase, individual sysops (not to mention geographic regions) do also leave FidoNet. Similarly, during the macro-level decline phase individual sysops are conversely joining FidoNet, although they are outnumbered by the leavers. Axtell et al.'s (2002) empirically calibrated simulation studies of the rise and fall of the Anasazi civilisation, and Mercken et al.'s (2010) work on the smoking adoption and cessation show that such considerations are not limited to a context like FidoNet.

The goal of the research reported here is to identify micro-level processes and mechanisms corresponding to the distinctive macro-level growth and decline phases, and to understand the relationship between micro-level processes and mechanisms that act in these two phases. From a macro-perspective, we could simply think of growth and decline as the outcome of the same joining and leaving mechanisms, with different proportions of joiners and leavers. However, from the micro-perspective, the key question is whether influence effects that cause sysops or users to join are structured in a fundamentally different way from influence effects that cause sysops or users to leave FidoNet.

Concretely: if our analysis reveals a spatial contagion effect with respect to joining events, should we then also expect a spatial contagion effect (with a similar functional dependence on distance) for leaving events? If social influence can be thought of as relying on social imitation, then joining and leaving processes may indeed mirror each other. If, on the other hand, social influence should be seen as a form of social learning, then joining and leaving processes could differ significantly. This is not a question that can be addressed by arguing from first principles, so the research presented in this thesis seeks to provide an empirical answer.

In Chapter 3 I provide an empirical analysis of the growth of FidoNet in the US, with the objective of determining whether and to what extent the pattern of joining events is consistent with a spatial contagion model. My analysis also reveals the functional relationship between the probability of joining and the Vincenty distance to regions of active sysops and users exchange areas. In Chapter 4 I examine to what extent leaving events also exhibit spatial contagion effects, with an additional focus on whether specific spatial configurations reduce the probability of leaving, while controlling for the exogenous effect of internet adoption.

While the hazard rate models used in Chapters 3 and 4 allow me to demonstrate that spatial contagion acts both during growth and decline, I can neither uniquely identify and choose between specific mechanisms at the micro-level nor systematically explore the consequences of different combinations of distinct mechanisms. The agent-based model simulations presented in Chapter 5 represent an important, first step in this direction. Simply looking at the combination of short-term and long-term contagion effects demonstrates that the resulting behaviour of the system cannot in most circumstances be understood just as a linear combination of short-term and long-term effects.

Hence: if our objective (in the spirit of analytic sociology) is to isolate and identify the micro-level mechanisms that give rise to observed behaviour, the non-linearity of interaction effects between different mechanisms suggests that agent-based simulations will need to augment the more traditional methodological repertoire. The ability to model complex interaction effects separately during growth and decline represents a significant advance beyond the current state of the art, and is an absolutely necessary step if we wish to develop integrated, endogenous models that can explain both growth and decline in a social system.

The data is unique in a various ways, that shape both the opportunities and challenges. Namely: the quality and time resolution of the data, the size of the community being studied and our ability to map the locations of those involved geographically while controlling for demographic factors, as well as the fact that the observation window covers both sustained growth and decline. It is worth noting that on this count alone the research reported here is able to bring a new level of empirical clarity to our understanding of social dynamics. While other papers consider dynamics at annual time resolution (Hedström, 1994) time resolution, or sporadic events over a few years (Andrews and Biggs, 2006; Vasi and Strang, 2009), the FidoNet data covers three decades[23] and is weekly for 23 of those years.

Of course, it could be argued that FidoNet corresponds to an idiosyncratic setting, or that its membership constitutes an especially unusual sub-population, so that the findings reported here reflect unusually complete and precise data but cannot be generalised. I hope that the remainder of this thesis will dispel such objections on two levels. First: FidoNet was a social system with global reach, connecting millions of people across nations and elements of those nations' respective cultures. This suggests the mechanisms are applicable to a wide variety of individuals, arguably more diverse and international than many of the studies cited here. While the analysis (purely because of feasibility) only considers the continental US, understanding the dynamics of a major component[24] of a system of this magnitude has implications for understanding other geographically growing and declining social systems, be they ICTs or (as we show in covering the pre-internet growth of FidoNet) communities that spread through non-digital means.

Second, the findings in Chapter 3 very much resemble those for spatial contagion mechanisms in other social systems that have been studied, such as trade unions. Hence, it is my contention that FidoNet provides an opportunity to study social dynamics of broader relevance at a level of detail precluded by the data used in previous studies.

[23]Though the analysis at present ends after 26 years
[24]The US had the largest population of sysops throughout the time-series studied in this thesis.

Chapter 2

FidoNet

Before delving into the analysis of the coming chapters, it is helpful to have a basic understanding of what FidoNet is, how it came to be, how people are[1] involved and what elements are relevant to considering its growth and decline. This chapter seeks to provide that background, such that the terminology and salient features of this complex and novel social system are clear and correctly frame the work to come.

In addition I will cover the construction of the dataset that underpins all the subsequent analysis. I will not go into too much technical detail,[2] but understanding the basic structure, features and limitations of the data will also help with understanding the analysis to come, in particular Exchange Areas: the basic geographic unit of analysis used throughout this thesis. Covering these issues here allows me to focus on the specifics of each chapter individually without repeatedly covering the information that pertains to all.

We begin with the early history of BBSs, how they work and are used and the crucial distinction between people who connected to and socialised via BBSs—users—and people who, in addition to socialising, created and maintained them—SYstem OPerators or sysops. I then summarise the history of FidoNet, how it began, what an innovation it was over BBSs at the time, how it grew and declined, but still persists in the present.

[1] I use the present tense because FidoNet is still used at the time of this writing, even though its membership is now quite small.

[2] Some of the woolier issues and code are covered in greater detail in section 6.2.4

After I will explain the dataset preparation. More than a quarter century of archived records of US telephone numbers of sysops—the volunteers who provided FidoNet's grassroots infrastructure—were combined with the geographic coordinates of US telephone exchanges to approximate their location. The resultant geographic time-series of involvement was then combined with Tract and Public Use Microdata Area (PUMA) level data from the 1990 US Census to control for relevant demographic factors and create the basic geographic unit of analysis used in the later chapters: Exchange Areas. The motivation for this methodology and its limitations conclude this chapter.

2.1 Computerised Bulletin Board Systems (BBSs)

BBSs were developed by American hobbyist communities interested in computer hardware in the 1970s and 80s. A technology that at the time was designed to make socialising within these small and localised groups easier, BBSs would eventually spread across the world, connecting people with a wide range of interests beyond just computing. Echoes of these communities and innovations are still present in the online forums, social networking sites, technical standards bodies,[3] and programming communities that are now commonplace on the internet. This section summarises their history.

2.1.1 History: Hardware and Hobbyists

On 5 March 1975, the Homebrew Computer Club had its first meeting in a garage in Menlo Park, California. They met to discuss and 'swap' early computer hardware and ideas. At that time, personal or 'micro' computers did not exist as we know them; consumers could only obtain primitive hobbyist kits, requiring a great deal of technical knowledge to build, operate and maintain. Hobbyist clubs like the Homebrew Computer Club grew up around the intense interest and difficulties

[3]Most internet protocols have been designed by publicly posting Request for Comments (RFC), just as the python programming language has developed by posting Python Enhancement Proposals (PEP). These posts then have responses, and through these public discussions, where anyone can participate, successful standards have emerged.

these machines fostered, much like amateur or ham radio operator clubs that began in the 1910s (Haring, 2008).

The ranks of the Homebrew Computer Club heavily influenced the basic design of computers that are ubiquitous today: screens, standardised keyboards and means of exchanging data (Freiberger and Swaine, 2000).[4] This last element—data exchange—has ultimately led to much of the sociologically significant aspects of computer use: sharing files, sending messages and collaboration in a wide variety of areas such as software development[5] and visual art.[6] BBS communities would provide a key step in developing the networking technologies and social organisations now taken for granted, just as the Homebrew Computer Club left its mark on hardware.

While physical means of transferring data, such as disks and cartridges, were much easier and cheaper to implement and use, very early modems were also available to consumers by the 1970s. At that time there was no internet a consumer could connect to—ARPANET and JANET, the networks from the US and UK governments that would eventually lead to the internet—were only accessible from universities and government institutions. This meant that—like ham radio operators—early consumer modem users were quite limited in who they could communicate with and mostly just called each other. And again like ham radio operators, many modem users hoping to find others to connect to would do so via their local hobbyist group.

It is worth highlighting that the lack of a centralised network like the internet meant that finding other like-minded computer and modem owners was highly dependent on the network of friends and acquaintances an individual already had, be they through family, work, education etc. The modern combination of internet service providers and search engines allows websites and the people and groups

[4]That first meeting famously included Apple's Steve Wozniak, and he was inspired to build the Apple][from a presentation that night (Wozniak, 2007).

[5]`http://github.org` is a modern and very popular site for software collaboration.

[6]`ASCII` and `ANSI` art would eventually have vibrant communities across FidoNet and websites such as `http://sixteencolors.net` are still active today.

associated with them to easily be found by anyone with an internet connection. In the late 1970s to mid 1980s however, people with modems were very much bound by their personal ties and their geographic proximity, both in terms of their likelihood of meeting others and because of the basic cost of calling and maintaining a connection with another modem.[7]

Analogies can again be made to ham radio radio operators whose transmission distance is limited, but also to finding employment opportunities via pre-existing social ties (Granovetter, 1973) and the significance of geographic proximity on the spread of social movements (Hedström, 1994; Andrews and Biggs, 2006). These similarities will feature prominently in the analysis of later chapters, but for now it is worth remembering the importance of pre-existing ties and distance in the social systems of early computer users. Arguably, BBSs were developed to reduce these barriers, and the costs of long-distance communication would eventually lead to the creation of FidoNet.

In 1977 the ideas of the Homebrew Computer Club others like it bore fruit: the '1977 trinity'[8] of personal computers were released, finally combining all the basic components expected from modern desktop computers: screens, keyboards and cassette or disk drives. Brands were incompatible in terms of hardware and software, but now computers could be used without assembly, and subsequently attracted a broader set of users.

Note that at this point, modems were one means of sharing information that bridged the gap of disk drive incompatibility. While the cassettes or disks of one brand might not work in another, text could be sent via modems connected to one of the few standard ports across the different brands.[9] While many hobbyist groups would be focused on specific types of hardware, which enabled them to share data and hardware modifications more readily (Sadofsky, 2005), this ability for modems

[7]Landline phone calls tend to charge extra for calls outside their local call area, and these premiums tend to increase with distance.

[8]The September 1995 issue of *Byte Magazine* would retrospectively call the computers released that year—the Apple][, TRS-80 and the PET 2001—the '1997 Trinity'.

[9]This was the serial RS-232 port.

to connect different computer models and brands would prove crucial in the spread of BBSs, and by extension FidoNet.

The great blizzard of 1978 brought with it the first BBS. Two members of a Chicago hobbyist club—Ward Christensen and Randy Seuss—snowed in by a blizzard for two weeks in January, decided to create a system for members of their club to send text and files to one central hub. There the received messages could be stored for other callers. Because modem connections at that time were essentially primitive phone calls between computers, only one person could be connected at a time. While connected, a person could read the messages and files posted by others and post their replies. After they disconnected, the next person could do the same. The first computer forum was born.

They called it a Computerised BBS, inspired by the bulletin boards—often made of cork with plastic and metal pins—found in public places meant for people to post information. Community members might pin up posters for events, lost dogs, music lessons, selling cars etc. These are often found at local libraries or schools—places where community information can reach many passersby. Anyone can freely post whatever they might want the community to know, and anyone can respond in kind.

Christensen and Seuss decided to publish the source code for their BBS in *Byte*: a popular computer hobbyist magazine (fig. 2.1). This allowed anyone to type up the code and compile it for their own computer, modifying as much as they liked. This process of working on publicly available computer code, then publishing those changes for others to benefit from, is now a core component to the modern software industry. The various open source licenses available today[10] have codified and protected this process. However, at the time, sharing computer code was a more informal practice, very much along the lines of swapping hardware.

From then on, BBSs spread across the US and the rest of the world, initially around computer hobbyist communities. By the mid-1980s a variety of different

[10]the GNU's not Unix (GNU) Public License (GNU Public License) and Berkley San Diego License (BSD) are the most popular of these.

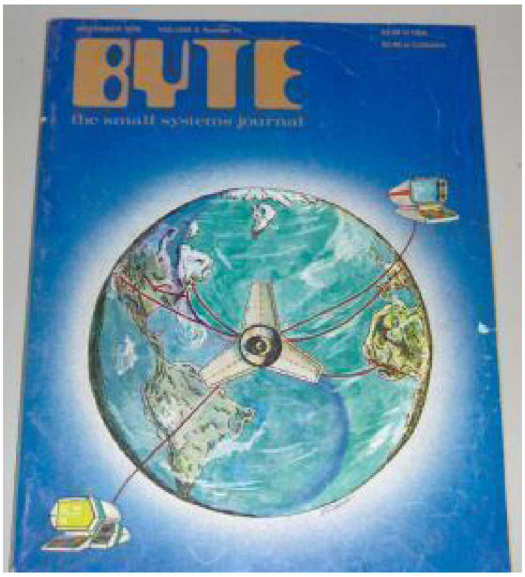


Figure 2.1: *Cover of Byte Volume 3, Issue 11 (November 1978)—the issue which published the source code to the first BBS.*

BBS programs had emerged, each with differing features and standards. FidoNet was one of these and the most popular—with the largest global reach of any single BBS network. Other popular BBSs became commercial ventures and would advertise to increase their users.

However, by 1995 the internet began to expand at a rapid pace,[11] and many of the commercial BBS companies began losing money as many of the features they provided were available via the internet, as well as many they could not provide. Some went bankrupt, others were bought out by other companies, and some went on to becoming internet service providers. The free BBSs also began to decline though perhaps at a slower rate.[12]

BBSs today are accessible via a variety of more modern protocols like Telnet, email, File Transfer Protocol (FTP) and IRC even radio, but many still maintain dialup capabilities. While the number of people using BBSs fell dramatically in the 1990s with the rise of the internet, Taiwan and Russia maintain especially popular

[11] PACE

[12] This estimate can be made with reference to FidoNet—which was largely free access though many accepted donations—but it is very difficult to estimate the relative decline of commercial and free BBSs. FidoNet's continued activity to this day is perhaps a testament to its resilience as compared with its competitors.

BBS communities (Sadofsky, 2005).

2.1.2 A BBS Community: Users and Sysops

As mentioned above, BBS community members fall into two categories: sysops and users. Both types socialise, play games and meet face-to-face in gatherings of their particular BBS community. There are, however, significant social status distinctions between sysops and users, as well as differences in costs—both financial and personal—and motivations. Alongside this dichotomy is the local community as a whole: an at times intimate and supportive group—at others quite volatile—these might meet regularly and often foster friendships and romantic relationships, as well as a sense of belonging to individuals have have felt marginalised in more traditional social contexts.[13]

A user is someone who connects to a BBS. To do so they need a modem, a computer and a BBS telephone number. Getting that number is a non-trivial task: unlike the internet there is no way to effectively search the telephone network for BBS numbers.[14] As a result, most users started BBSing via a social tie who gave them a number and explained how to connect. In the late 1970s that was again likely a hobbyist club connection, but as BBSs spread the diaspora of users became far more diverse and friends, family, work and schools became sources of BBS awareness, as did, eventually advertising (more on this below).

A sysop is an operator of a BBS. Sysops provide the computer with the requisite software, the modem and telephone number for a user to connect to. Once a user has that number, their own computer and modem, they can dial in and—if the phone line is free—connect to that sysop's BBS.

Once connected, a user can leave messages for other users, read previously posted messages, run software on that specific BBS, and chat live with the sysop.

[13]Most of the claims made in this section derive from BBS *The Documentary*Sadofsky (2005), and rather than repeated cite that documentary I shall merely make this statement here.

[14]Random Digit Dialing (RDD)—randomly calling telephone numbers searching for BBSs—is one way of trying to find BBSs, but this is especially expensive and time consuming, especially in the early days of BBS use.

Games, commonly called Door games were a major draw in the 1980s, especially those that allowed users to interact with each other or the sysop. Some games were asynchronous, meaning players would 'play' at separate times, like in a correspondence chess match: each player makes a 'move', sends that move (like a message), to the other player and waits for them to respond. Others were synchronous, meaning like live conversation, players could play together at the same time, either user and sysop (easiest and cheapest) or with other users provided the sysop had multiple modems and phone lines.

Multiple phone lines and modems are another source of potentially significant financial costs for sysops. If they have an especially popular BBS they might pay to have multiple phone lines in part because of needing normal voice-based telephone service (potentially for the rest of the household) or so that they could support more users at a time. Obtaining, maintaining, upgrading and potentially expanding the computer hardware can also be expensive, not to mention money spent on BBS software which, while some of it was freely available, some was not.

Financial costs also apply to users. Telephone call cost structures vary across countries and providers, but by and large it is free to receive calls: it is the caller who pays. Thus sysops do not face charges for the length of connections or the distance to the caller because they receive calls and do not make them.[15] Users, as callers, are charged, and how much depends on the cost structure. I will only consider US phone billing here,[16] which usually means that local calls are of unlimited length with no extra charge, but calls outside the local call region cost extra. The surcharge is based on length of call and distance to recipient, so if a BBS a user wishes to connect to is not within their local call area, BBS use can become expensive. This is demonstrated in fig. 2.2. Particularly avid users may also invest in extra phone lines (again to avoid inconveniencing other house members), and this can also add to financial costs.

[15]FidoNet alters this structure as described in section 2.2.

[16]European phone billing may have impacted some of the variations in BBS use, as exemplified by FidoNet Nodelists in Europe.

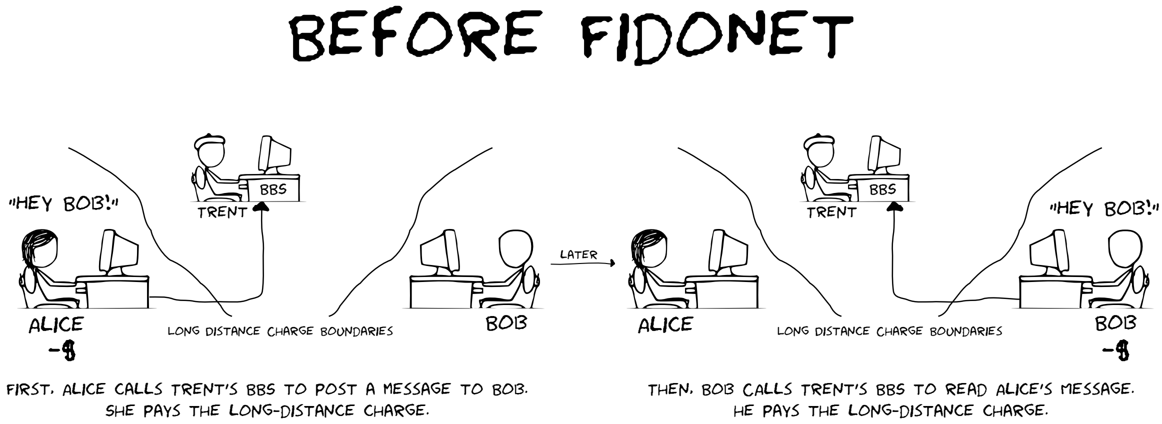


Figure 2.2: *Alice wishes to send a message to Bob, but neither of them are within the local call zone of the nearest BBS (run by Trent, their sysop). So first Alice calls Trent's computer and pays a long-distance charge to send her message. Sometime later Bob calls Trent's BBS (again paying a long-distance charge) and downloads Alice's message (perhaps even posting a reply.)*

Sysops often enjoy a higher social status than users (*ceteris paribus*) particularly within their individual BBS community. A sysop provides a service to their users, and the community may feel indebted to them for that service, especially since many BBSs were free of charge (particularly those in FidoNet). Coming again from the context of hobbyists, sysops may be perceived as having greater technical skill or knowledge (since they are able to run a BBS), and within that community such ability is highly regarded. Users are also likely to have the most direct social contact with their sysop, because the sysop is the one person who, when at their computer, can chat live with whoever is connected.

Users tended to live near their sysop, especially in the late 1970s-80s. This was probably due to two factors. First, is again the cost of a phone call which is significantly higher if it is outside a local call area.[17] At a time when modem speeds were extremely slow,[18] calls could take a great deal of time to transfer what by today's standards would be tiny amount of information, and thus quickly become expensive if they were too distant. An example of how early BBSs could be expensive for users is demonstrated in fig. 2.2.

[17]US local call areas are highly heterogeneous. For our purposes, they can be thought of as areas where calls are free and unlimited with a basic monthly phone bill.

[18]Early, consumer grade modems modems could transmit text at about human reading speed.

Second, BBSs were probably a way of maintaining pre-existing social ties, and as such were a reflection of an extant offline community (like a hobbyist club). Much of Barry Wellman's research supports this hypothesis for far more modern ICT communities (Wellman, 2001; Hampton and Wellman, 2001; Mok and Wellman, 2007), and it is reasonable to suggest these effects were even stronger back when computers were so comparatively primitive. Instead of instant photos and real-time conversations, people had monochrome, green screens only capable of text transferred at very slow speeds.

2.2 FidoNet

In the early 1980s a San Francisco programmer named Tom Jennings was sending messages to a friend from Baltimore, MD (the other side of US). This was, understandably, expensive, and Jennings imagined solving this problem by having a network of BBSs across the US, each a local call away. That way, each hop would be free and he could send messages to his friend without paying long distance charges.

Jenning's innovation was to allow BBSs to forward messages to other BBSs. Figure 2.3 demonstrates how a user (Alice) could send a message to her local BBS (run by Carlos), who could forward that message on to her friend Bob's local BBS without either of them paying a long-distance charge. Bob could send a reply the same way.

Suddenly, each spatially isolated BBS community could be connected and conversations could span longer distances without costing users any more money. If we think of each isolated BBS community as a source of social capital[19] for its users, connected BBSs consolidate that social capital into one larger, multifaceted community, much like Granovetter's (1973) theory of the importance of weak bridging ties and Burt's (2005) theories of brokerage.

In 1983 Jennings, released his own BBS program called Fido. Fido BBSs had the ability to forward messages and files to each other. Jennings called his network of

[19]For good reviews on social capital see Portes (1998); Lin et al. (2001).

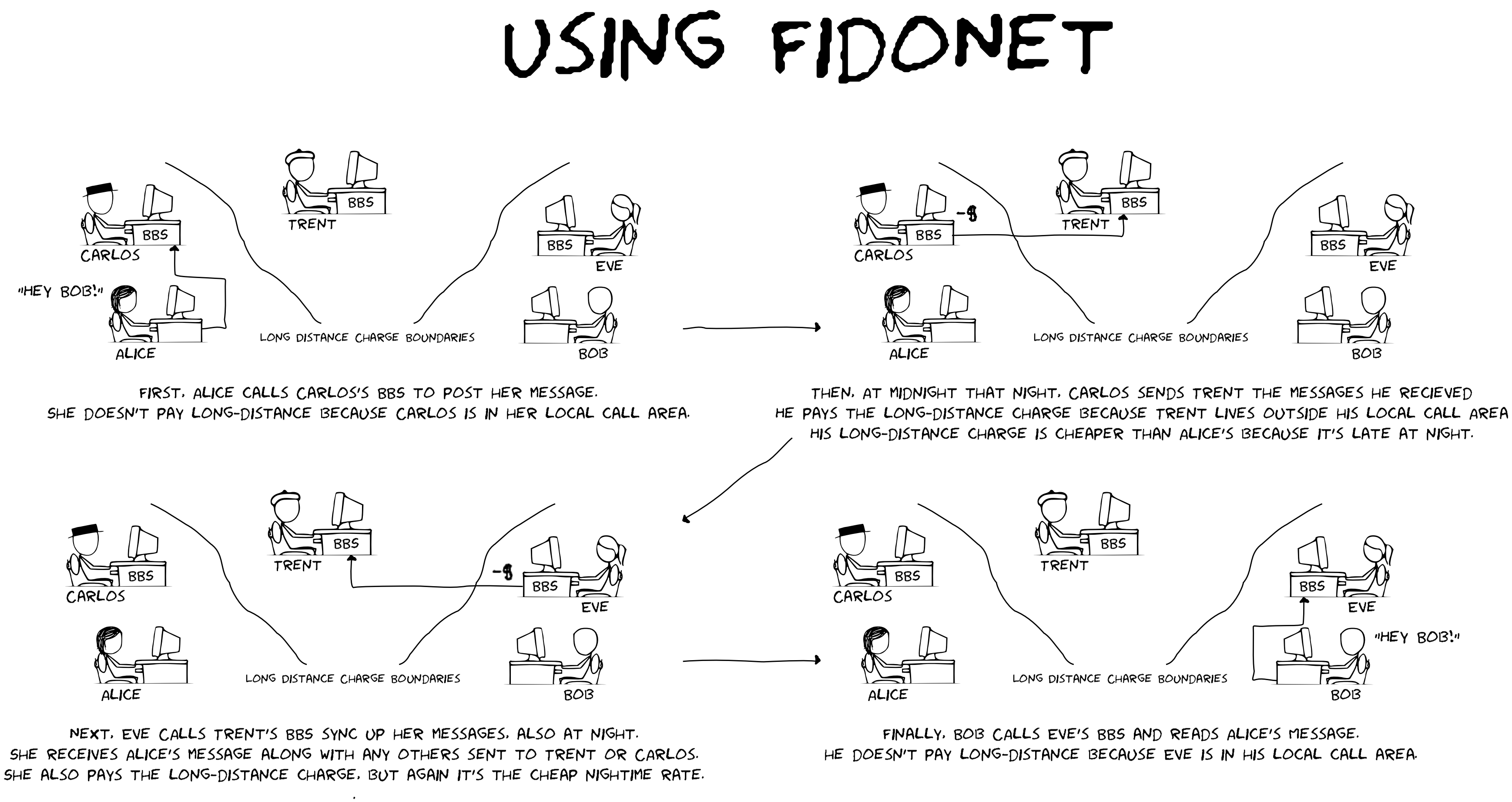


Figure 2.3: *Alice still wishes to send a message to Bob, but now they both have local BBSs connected via FidoNet. These BBSs will forward the messages they receive in batches when long-distance calls were cheapest so as to minimise the costs to sysops. Costs were shifted from users to sysops and sysops could send all the messages at once at a cheaper time of day. Thus the total costs incurred by the system were reduced.*

BBSs FidoNet and distributed the software to people he knew. His friends were the first FidoNet sysops, and as of 9 Nov 1983, FidoNet had 12 BBSs (and by extension 12 sysops to run them).

Unfortunately, with only 12 BBSs, Jenning's dream of having each BBS a local call away could not yet be realised: there weren't enough sysops to cover the US. While users suddenly had access to messages other BBSs, the sysops had to foot the bill for the long-distance calls. To minimise costs a 'FidoNet Transfer Hour' was arranged for the time of day when calls were cheapest (12:00–1:00 AM). Thus a message might spend a few days, hopping from BBS to BBS every night before reaching the intended destination.

While this helped, costs were still hefty for sysops, even when the network became large. A sysop I corresponded with over email, who has operated a BBS since 1989 (and continues to this day) had this to say about about her financial contribution:

> *Ron used to joke that we could have had two summer homes (we've never had one <g>)*[20] *for the $$ I tossed into Fidonet.. :)*

Somehow, most FidoNet sysops refused to charge their users, preferring to take donations or cover the costs themselves. Some imposed rationing systems to prevent users from abusing the service, and still others did charge in various ways, though that appears to have been unusual. Thus FidoNet could be considered a public good, and a classic free-rider problem.[21]

Since messages were no longer simply local conversations, they were organised into 'Echoes', which, like modern forums and email lists, were divided into a host of topics such as politics, philosophy, computer science, sexual deviance and art (see figure 2.4). A sysop would subscribe[22] their BBS to the Echoes their users wished to read and respond to.

[20] <g> is a shorthand for 'grin'
[21] See Olson (1971) for the seminal work on public goods and collective action.
[22] Just as people are 'followed' on twitter or a website's RSS/ATOM feed is subscribed to

```
                  --
                 /  \
                /|oo \
               (_|  /_)
                _`@/_ \    _
               |     | \   \\
               | (*) |  \   ))
  ______       |__U__| /  \//
 / FIDO \       _//|| _\   /
(________)     (_/(_|(____/
                     (jm)
```

Figure 2.4: *John Madil's* ASCII *art logo for FidoNet.*

2.2.1 FidoNet's Growth

Within a year the Nodelist (a list of all BBSs, their sysops and telephone numbers) had grown from 12 to over 200, far beyond Jennings' expectation.[23] The network had already spread beyond the US and the Nodelist became increasingly difficult to for Tom maintain as it grew in length and the phone numbers of various countries became more complex.

Some of the more technically adept sysops decided to relieve Tom of this responsibility and standardised the Nodelist format for international use. They created the FidoNet Technical Standards Committee (FTSC), which took over the organisational and technical running from Jennings, and provided the infrastructure needed to continue growing rapidly.

From then on, a new sysop could add their BBS to the Nodelist simply by informing the current committee of their name, telephone number, hours of operation and what software they used. Nodelists were released every Friday, without interruption, for over two decades. Out of the FTSC came the International FidoNet Association (IFNA), which became a 503(c) US non-profit organisation, had formal committee positions, and ran a newsletter called FidoNews.

The success of this model is undeniable: FidoNet's initial growth (in terms

[23]'I had no idea it was going to become so amazingly large. Or long lived, that's the part that gets me.' —Tom Jennings *Creator of FidoNet*

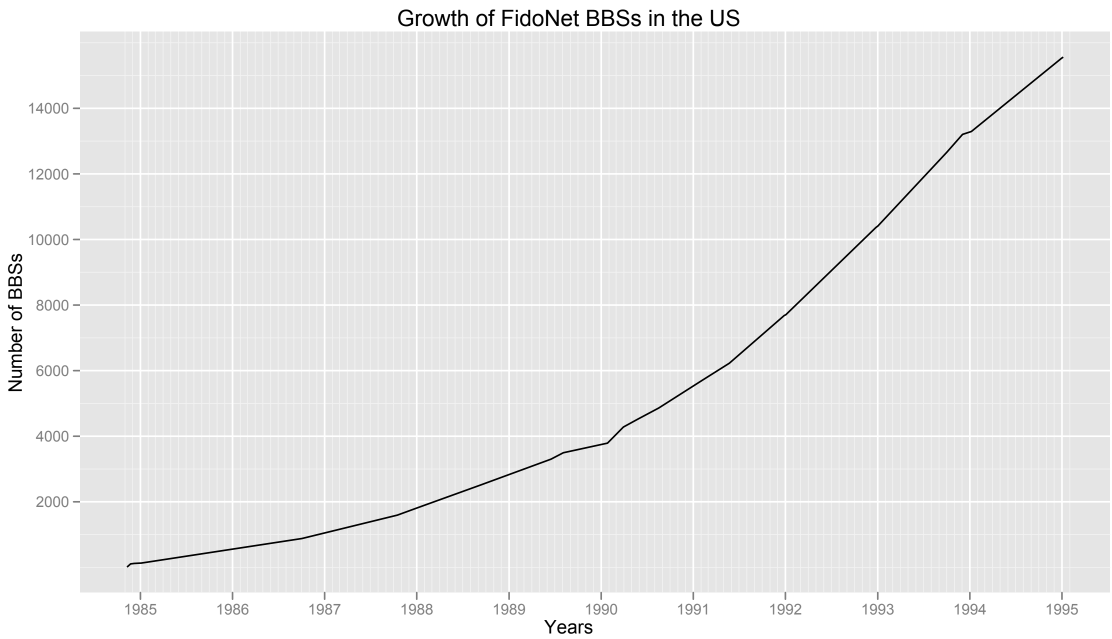


Figure 2.5: *Growth of* US *FidoNet BBSs*

of BBSs rather than users, see figure 2.5) resembles other dramatic, super-linear membership growth patterns such as participants in the Chicago strike wave of 1886 (Biggs, 2003). This is especially notable when compared with Putnam's (2001) results on the decline of US social capital during the mid-80s to 90s, suggesting that many civic organisations of the time were in decline.[24] At its peak there were 20,000 US BBSs, and 40,000 worldwide. Estimates of users per BBS are difficult, but estimates of 200 can be found by the early 1990s. An estimate in 1993 put US FidoNet users on the order of 2 million (Bush, 1993).

FidoNet's growth is even more surprising when considered alongside the passivity of FidoNet's elite: IFNA merely provided a means for registering new sysops, collecting donations and organising responsibilities within the network. As far as we are aware, no campaign was ever organised to promote the network. A ProQuest search for FidoNet yields seven results in technology magazines prior to 1991, when a blurb in the Wall Street Journal lists some BBS numbers, including a few from FidoNet (wsj, 1991). From that point on, FidoNet gets sporadic coverage in

[24]This may be evidence that some of that dissipated civic engagement was reinvested in other, newer forms of social association.

trade magazines like *Association for Computing Machinery* (Bush, 1993) with a single notable article in the Washington Post (Snow, 1993).

2.3 Data

To robustly model the geographic spread of FidoNet we combine data from a variety of sources. These include the FidoNet data itself, which provides a time series of active BBS landline numbers, coordinates of US telephone exchanges which we use to approximate the locations of those BBSs, and regional demographic data from the 1990 US Census to control for geographic heterogeneity. Each of these sources will be described in turn, particularly with respect to our chosen unit of analysis: Exchange Areas.

Exchange Areas are composed of census tracts and approximate the service areas of telephone exchanges. We use them to construct a demographic profile of the population served by each exchange. While the actual data preparation procedure is detailed in appendix A, we explain and define Exchange Areas on a basic level for ease of interpretation in §2.3.2 below.

2.3.1 FidoNet Nodelists

Prior to 1985 Nodelists were published haphazardly at the convenience of the Nodelist maintainer.[25] Once the Nodelist was taken over by a committee, the format was standardised and a new Nodelist was published every Friday throughout the year. This practice continues to this day. The Fido History Project has collected these Nodelists for posterity and graciously allowed them to be used in this research. For each week, every FidoNet BBS was listed, including the Sysop, modem speed, landline number, city and region, current software versions and whether the BBS was currently operational or not.[26]

While a small fraction of Nodelists prior to 1991 have survived, full data is available at weekly time intervals from 1991 on. However, the vast majority of

[25] Initially Tom Jennings, creator of FidoNet, maintained the Nodelist.
[26] If, for example, a Sysop went on holiday, their Node would be 'DOWN' until they returned.

```
90. MidNet UW_Madison, WI Mike Mansfield (608)-233-8449
91. The_DEAD_END Round Rock, TX Steve Sparks (512)-255-1282
92. PCUTILboard Detroit, MI Jon Tara (313)-393-0527
99. The_REAL_Fido San Francisco,CA Tom Jennings -UNPUBLISHED- RE,RT 1
104. Baby_BYTE/Net Peterborough, NH George Bond (603)-924-9820
108. Samson Rolling Meadows,IL Larry Miglore (312)-991-8304 5p - 8a WE
115. Arquimedes Washington,DC Sam Hargadine (202)-332-9512
*117.Jim_Filgo Jakarta, INDONESIA Jim Filgo 01162-21-372518
124. TimeLifeBks Alexandria, VA Eldon Ziegler (703)-833-7355 RE
125. Radar's_Node Fairfax, VA Joel O'Rourke (703)-978-0351 Irregular
```

Figure 2.6: *A segment from the Dec 28, 1984 Nodelist*

```
Region,20,Sweden,S,Mats_Knuts,46-40-549189,9600,CM,HST,XA
,10,EchoMail_Gateway,Lund_Sweden,Ulf_Nilsson,46-46-320080,9600,MO,CM,XA,HST
,260,Lightline's_Revenge,Sundbyberg,Kaj_Lehtinen,46-8-985404,9600,CM,HST,XR
,822,RFC_822_GATE_WAY,Karlstad_Sweden,Lennart_Svensson,46-54-566988,2400,CM,MO,V21,V22,V23
,999,Region_20_Gate,Sweden,Mats_Knuts,46-40-549189,9600,CM,HST,XA
;
Host,200,South_Net,Malmoe_Sweden,Mats_Knuts,46-40-549189,9600,CM,XA,HST,
,101,SSF_BBS_*HST*,Malmoe_Sweden,Mats_Knuts,46-40-549189,9600,CM,XA,HST,USDN,SDS
,104,CompuDog_BBS,Trelleborg_Sweden,Leif_Olsson,46-410-27335,9600,CM,XA,HST,USDS
Pvt,106,The_Channel,Hoellviken_Sweden,Tommy_Svensson,-Unpublished-,9600,
,108,Walkabout_Creek,Malmoe_Sweden,Tomas_Gradin,46-40-469369,2400,CM,XA,UNEC,UREC
,109,The_Newness_Fairlight_BBS,Malmoe_Sweden,Joakim_Lindstroem,46-40-231145,2400,V21
,110,Professor_Baltzar's_BBS,Malmoe_Sweden,Giovanni_Ferrari,46-40-974417,9600,HST
```

Figure 2.7: *A segment from the July 13, 1990 Nodelist*

these lists are in a complex 'difference' format that we were not able to parse. For Chapter 3 that time series was used, limited to 35 irregularly spaced snapshots of the network over 20 years. Since this chapter only considers the growth phase, this is limited to the first twelve years and 27 time steps.

However, for Chapter 4, Nodelists from Fido History Project allowed week-level changes to be analysed for the decline phase. Had time permitted, the growth chapter could have been re-analysed with this data but that proved infeasible. It is feasible to at least plot that week-level time series both at the world 2.8a and continental US levels 2.8b.

2.3.2 US Landline Geo-location

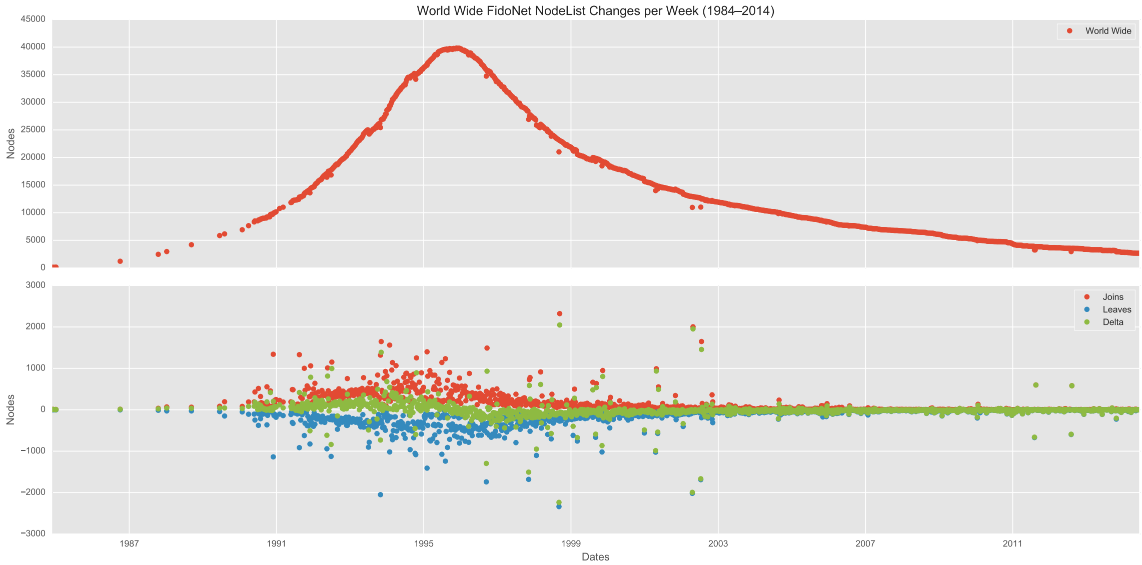
To approximate the location of each BBS we used a public domain dataset provided by Dykstra (1998),[27] which lists the geographic coordinates of US telephone exchanges in 1998. Every US landline number prior to the early 2000s conforms to the NPA/NXX standard, which specifies which telephone exchange that number is connected to via its first 6 digits (excluding the leading 1 which is the international code for North America).

Over the last decade, especially with the rise of mobile phones, the NPA/NXX system has become far more fluid—it is now possible to migrate landline numbers outside the geographic area designated for their code. However, since our analysis ends in 1995, we feel confident in assuming that the landline numbers in the FidoNet dataset are unlikely to have moved significantly enough to undermine our analysis[28].

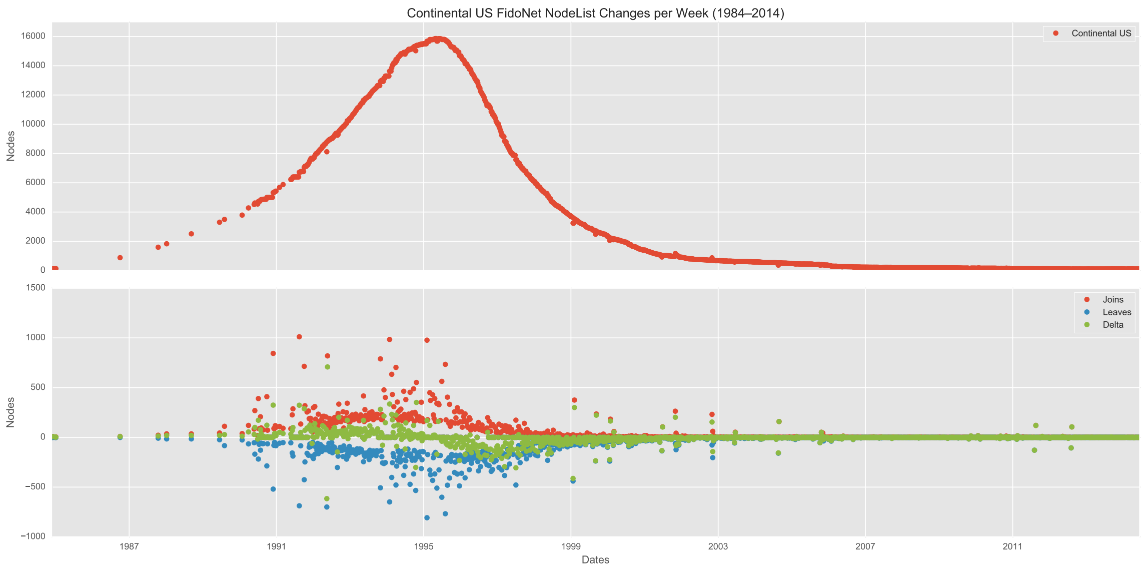
There is a larger potential problem: the number of exchanges probably increased between 1983 and 1998, which means that the set of exchanges which could

[27] Phillip Dykstra was contacted by email regarding the authenticity of his data and permission for its use. It was obtained from a company called Telcordia for a project he worked on at the Department of Defence. He believes it to be highly accurate. We are extremely grateful for Phillip's generosity.

[28] It has long been possible to move landline numbers *within* NPA/NXX areas (like moving to another house in the same city), but since we are measuring adoption at the exchange level—not the house level—this has no effect on our analysis.



(a) *World wide growth and decline of FidoNet BBSs*



(b) *Continental US growth and decline of FidoNet BBSs*

Figure 2.8: *World and Continental US (region analysed in 3 and 4) time series of active FidoNet BBSs, aligned with the weekly joining and leaving events ranging from Nov 23, 1983 to Jun 13, 2014.*

potentially adopt may be inflated for the earliest portions of the time-series.

We feel confident that this problem should not be significant for two reasons. First, the expansion of the NPA/NXX system appears to add additional codes to pre-existing exchanges rather than build new exchanges in different areas. Of the 87,622 unique US NPA/NXX codes only 21,139 have distinct coordinates, indicating that many codes are served by the same physical exchange.[29] Since our interest is in the geographic spread of FidoNet, our set of potentially adopting exchanges is only as large as the number of distinct coordinate pairs. The prevalence of multi-code exchanges suggests that new codes may often be added to existing exchanges, which would reduce the likelihood of newer codes inflating the number of geographically distinct exchanges.

Second, we further aggregate physical exchanges into Exchange Areas–our unit of analysis—in order to incorporate demographic information into our model. Our aggregation process should further ameliorate the possibility that a systematic increase in the number of exchanges over time impacts our results. This is explained after the next section. Figure 3.2 shows what the geo-located data looks like over time.

2.3.3 Demographic Data

Like Hedström (1994); Hedström et al. (2000); Andrews and Biggs (2006) we control for relevant demographic factors to account for the socio-geographic heterogeneity of the US. Spatial data from the 1990 US Census are conveniently prepared[30] by two projects at the Minnesota Population Center—the National Geographic Information System (nhg, 2011) and International Public Use Microdata Series USA (Ruggles et al., 2010). Using this data we constructed a demographic profile for each tract, based on factors we believe most salient to impacting the propensity

[29]Of those 21,139 physical exchanges, nearly half are associated with at least 2 NPA/NXX codes and the mean is 4.15 codes per exchange. The maximum number of codes assigned to one exchange is 325, and 55 exchanges have more than 100 codes.

[30]Raw census data is quite complex and difficult to analyse, and we are grateful to the Minnesota Population Center for making the data preparation much easier than it would have been.

of a sysop setting up a FidoNet BBS. These include: population density, urbanity, income, age, educational attainment and occupation.

The details of this process can be found in appendix A, but we will briefly explain the approach for interpretation of the results. The 1990 census aggregated much of its individual level data into 61,693 census tracts: geographic regions normalised to having approximately 8000 people.[31] National Historical Geographic Information System (NHGIS) also provides shapefiles—the geographic coordinates of the boundaries of each tract—which allowed us to associate[32] tracts with the exchanges they encompassed and were closest to.

However, some variables are not available at the tract level, and because we thought detailed occupation information to be especially important we turned to Integrated Public Use Microdata Series (IPUMS)' 1713 PUMAs. PUMAs are composed of census tracts and are normalised to a population of approximately 100,000. PUMAs have much more detailed information on occupation and improve on the thirteen broad groups available at the tract level into 514 separate categories. However, to avoid losing the high spatial resolution of tracts, we took the occupation variables from the PUMAs and—as a proportion—applied them to the data on each tract. Doing so maintained the spatial resolution of tracts while incorporating the more detailed occupation information from PUMAs.

Regarding our choice of the 1990 census: 1990 falls in the middle of our time-series (1983–1995). The 1980 census has markedly worse data, especially in terms of spatial information, and the 2000 census would only potentially be useful for the very last year of the study. While there are estimates of how the variables of interest changed in the intervening years, the added complexity of adding some of those variables, at far worse spatial resolution and lower data quality, seemed more trouble than it was worth. The simplicity and reliability of using the 1990 data for the whole time series was optimal given the constraints.

[31]The do vary widely though. One tract in Florida had a population of 70,000.

[32]The association process is briefly outlined in the next section and in detail in appendix A.1.3.

2.3.4 Exchange Areas

We combine PUMA-augmented census tract data, their shapefiles and telephone exchange locations into what we call Exchange Areas (EAs): spatially proximate sets of one or more exchanges and one or more census tracts. EAs approximate the region and population served by their associated telephone exchanges (and by extension the corresponding NPA/NXX codes). The algorithm for creating EAs is detailed in appendix A.1.3.

For the purposes of interpretation the finer details are less germane, but the basic concept is as follows. All exchanges within the same tract were combined into one geographic location by taking an average of their coordinates[33] and the tracts that geographically contained them were automatically associated with the one or more exchanges they encompassed. The remaining tracts (which did not have exchanges within their boundaries) were associated with the exchange closest to their centroid.[34]

This process combined the 61,693 census tracts and 21,193 telephone exchanges (and the associated 87,622 NPA/NXX codes) into 16,026 EAs: our initial risk-set.[35] Of these 4392 (27.4%) eventually adopt.

The aforementioned argument regarding the amelioration of the potential inflation problem follows from the fact that EAs combine spatially proximate exchanges into one entity. This implies that in many cases any exchanges added between 1983 and 1998 would be subsumed into the same EA. Only in cases where previously sparsely populated areas increased in population enough to warrant a new exchange far from existing exchanges should the inflation actually occur.

[33]This average was calculated as the centroid of the convex hull containing all the exchanges. This process prevents the suboptimal result of EA regions where the exchange is geographically outside the calculated region.

[34]Further correction factors were required to account for the irregularities of bodies of water, oddly shaped tracts and to maintain state boundaries. All these are explained in appendix A.1.3.

[35]Technically we begin our model with 16,017 at-risk EAs because nine begin the time-series adopted. The risk-set is the set of potentially adopting EAs at each point in time.

Chapter 3

Growth

FidoNet's spread within the US was vast and quick, at least for its time, and even more impressive when compared with modern social networking websites like Facebook. It had no organised form of self-promotion, no salaried employees or venture capitalists, and most important: no internet to make joining easy. Becoming a FidoNet user required having the right hardware, software, and a list of local BBS phone numbers. To join Facebook, one simply needs a url. Facebook grew on the internet, which provides the protocols and infrastructure; FidoNet had to provide the infrastructure itself.

To understand how this process worked we consider how people became sysops, as each new sysop corresponded to a new BBS.[1] We will primarily focus on BBSs/sysops only considering users to the extent they can help explain the presence of sysops. This is for three reasons. First: there is extremely high quality data on individual BBSs, whereas little data remains on users, much of which is only available after FidoNet's US peak.

Second, many of the methodologies we draw upon are for modelling the spread of strikes or union chapters—entities which tend to correspond to a rising local community or movement interested in providing their own outpost for a spreading social entity. BBSs closely resemble such outposts, and it is quite likely that sysops were former users who wanted to improve local access to the network. Thus areas with high number of users were likely to produce a BBS.

[1] Occasionally the same sysop would manage multiple BBSs, but this appears to have been rare.

Finally, understanding FidoNet's growth must at a minimum address the spread of BBSs: becoming a sysop was a significant investment in the social capital of FidoNet, and while a user joining was important, a new BBS was a palpable, durable expansion of the utility of the network for all its members. A user, by contrast, might have occasionally read messages and never posted anything, or participated in a huge number of discussions, the difference is difficult to discern. Thus users significance and contributions are difficult to determine, while a sysop has a clear, well-defined commitment to the community.[2]

We will first consider the individual decision-making process of becoming a sysop, and second the quantitative methodologies we can use to model the social influence, mediated by distance, which may have drove FidoNet's expansion.

3.1 Joining FidoNet

Setting up and running a BBS was complex, expensive, time-consuming and required esoteric technical knowledge. To understand this process, we can think in terms of *opportunity* and *desire* following Elster (2007). Opportunity relates to the knowledge, resources and skills required to set up a Fido BBS, and desire relates to what motivates a sysop to volunteer to do so.

Given the 'outpost' analogy, it is useful to consider what will drive new users to join FidoNet as well as what will lead to more sysops. It is likely most sysops were originally users and so whatever leads to more users, leads to a greater supply of potential sysops.

3.1.1 Opportunity

Opportunity can be split into knowledge and resources; we will consider each in turn. In order to use FidoNet or run a BBS, one must have some knowledge of the

[2] Wikipedia has a similarly vast disparity between the contributions of its users, most of whom only read articles. A very tiny minority, however, contribute huge amounts of research (Ortega et al., 2008).

network and the relevant technical details: what software and hardware, what are the phone numbers of the nearest BBSs etc.

In short, a potential user almost certainly had to know someone who was already a user or sysop to join, especially since FidoNet did not advertise. That person could then explain what software and hardware was needed to connect. In some cases face-to-face interaction was the only means by which said software could be obtained: without a pre-existing structure like the internet, there is no way of distributing the necessary software for connecting—someone must physically provide it.

Already the importance of space is clear: knowing another person who uses FidoNet is probably the most limiting factor in the opportunity to join. If we accept Wellman's findings on the spatial embedding of relationships, those living near other FidoNet members should tend to have a higher probability of knowing or encountering one, *ceteris paribus*, and by extension should have a higher opportunity to join. Living nearby also means the FidoNet user could more easily assist a new potential user in setting up their computer. Equally, a potential sysop living near another sysop could more easily have support in setting up their own BBS.

Other useful sources of knowledge are experience with and understanding of computers, both of which might arise from high levels of education or highly technical/computer-related occupations. Computer literacy is particularly useful in maintaining a BBS, which requires a much higher level of robustness and reliability than your average personal computer.

The resources required to run a BBS are considerable as well, and traditionally far more than those required to be a user. A user merely needs a computer, a modem, and an available phoneline. A sysop needs these, often one or more dedicated phonelines,[3] potentially multiple modems, a more expensive than average computer and extra software. Since most BBSs were just in sysops' houses, a quiet room in their basement often proved key. Phone bills could become exorbitant as well,

[3] A phone line just for FidoNet so that other members of the household could still receive calls while users were connected.

especially if sysops have to connect to other Fido BBSs outside their local call area. Finally, sysops need free time to maintain their BBS and support their users. Thus high incomes and salaried job contracts (which may yield more flexible hours) should increase the opportunity of becoming a user or a sysop.

3.1.2 Desire

While the factors affecting opportunities for potential users and sysops were relatively similar, motivational factors probably differed. It is one thing to want to participate in a community, it is another to want to spend time and money providing a service to other members.

Becoming a FidoNet user was similar to adopting a new technology: at the time FidoNet was a novel use of computers, and provided a new form of communication and social intercourse. Young (2009) distinguishes three types of innovation diffusion mechanisms which relate to desire: contagion, thresholds and social learning.

Contagion is the human tendency for imitation: we have an innate instinct or desire to copy each other. Whether instinct counts as desire is a matter for debate; it will suffice as a form of motivation. Thus being around others using FidoNet should increase the likelihood of wanting to join, and the greater the contact with FidoNet the stronger the incentive. Nearby FidoNet use, therefore, is likely to lead others to join.

Threshold models take the effect of other people joining as approaching a discrete limit, beyond which a person is motivated to join. Once five of your friends—or a proportion[4]—join FidoNet you will want to as well. Here the effect of network externalities is motivating: the benefits of joining are tied to the utility of socialising with the other people who have already joined, and if enough of your friends have joined you know you can at least enjoy socialising with them.

[4] Most models seem to favour Granovetter's (1978) proportional approach though the advent of social network analysis approaches to threshold models (Macy, 1991; Valente, 1996) have considered absolute thresholds.

Again: spatial proximity to FidoNet users, assuming the spatial embeddedness of relationships, should increase the desire to join.

Finally, Young's (2009) idea of social learning, which he cobbles together from a variety economic works, is the strategic updating of information about an innovation that can ultimately motivate someone to adopt. This information gathering happens via social interactions with people who have already adopted, and once enough information is provided to overcome an individual's concerns they are motivated to adopt as well. Much like the other two mechanisms, increased contact with FidoNet users and sysops will increase one's information about the community and eventually motivate one to join (or not join). What differentiates social learning from the other two is the flexibility: greater amounts of information could cause positive or negative motivation.

Nevertheless, assuming a general inclination towards something like FidoNet, greater exposure will help convince potential users. In turn, the spatial proximity to others is likely to be motivating. Whichever mechanism is at work, higher densities of nearby FidoNet users and sysops is likely to lead to more users.

But what of sysops? Assuming most sysops are former users, the above mechanisms are less likely to apply: largely because sysops are far more rare than users. If estimates of 200 users per sysop (in 1993) are correct (Bush, 1993), it ias unlikely to be high concentrations of sysops leading to other sysops. In fact it is more likely the dearth of sysops or rather too many users per sysop, that drives users to volunteer.

We can break the likely motivators for becoming a sysop into three categories: instrumental, social status/altruism and interest. Instrumental reasons relate to BBS availability: if there are too many users for one BBS, another will be needed to balance the load. Only one user can connect to a phoneline at a time. A sysop could pay for more phonelines to increase their capacity, but it would be highly unusual (and extremely financially prohibitive) to have more than five.

A sysop has unlimited access to their own BBS, so a user who wants more access could set up their own. Thus an overwhelming concentration of FidoNet users per BBS impel more users to volunteer to be sysops. Similarly, a user in an area far away from the nearest BBS could save money by running their own and transferring data more cheaply during the FidoNet transfer hour (see §2.2). Here higher concentrations of users (a larger population of potential sysops) in areas too far from the nearest BBS to use local calls should develop their own BBS.

Notice of course, that doing so will have the added benefit to others in the area of reducing their costs as well. This relates to the social status/altruism effects. Since sysops incurred great costs, generally to the benefit of their users, their sacrifice could be seen as charitable and noble. Such behaviour could be 'motivated by a desire to avoid scorn of others or receive social acclaim' (Becker, 1974) (desire for social status) or for the purely altruistic 'warm glow' (Andreoni, 1990) of doing something for others while expecting nothing in return, or as Andreoni notes, some 'impure' mixture of the two. Such motivation is again driven by the needs of users, so again, a lack of service for users should be a driving factor.

Finally, technical interest has resonance with work on contributions to open source projects. Results from Hertel et al. (2003) indicate that 'scratching a developer's personal itch' (Raymond, 2001) is a significant motivator for contribution. That itch could be a desire to make FidoNet work better or—just as open source programmers find it 'fun to program' (Torvalds and Diamond, 2001)—sysops might just find the technical challenge of running a BBS intrinsically enjoyable. This motivator is far less likely to be driven by concentrations of users, more possibly correlated with education or occupation.

Thus we have covered a variety of factors which could have contributed to the spread of FidoNet users and sysops. We have covered both users and sysops because many of the factors driving up the number of sysops/BBSs directly relate to the presence of and proximity of users. We aim to provide a sort of condensation image: as concentrations of users increase, eventually a sysop will emerge from them, just as water vapour 'condenses' into droplets. A new sysop will reduce the costs of using FidoNet for their users and increase their access. A new sysop will take on costs (time and financial) in exchange for unlimited access to their own BBS, greater social status, altruistic satisfaction and/or a technical challenge.

3.2 Diffusion Modelling

The hypothesis that behavioural contagion is mediated by distance was first posed in Hägerstrand (1967), an intensely detailed analysis of the diffusion of technologies through Sweden. He found that innovations tended to be adopted by neighbours of current adopters, and accordingly he hypothesised that innovations tend to diffuse through pre-existing networks of communication. He also suggested that variations in the structure of communication networks would have a significant impact on the speed and trajectory of adoption.

His hypothesis has been taken up by many in the network diffusion literature (Valente, 1995; Centola and Macy, 2007), but his influence is even more palpable in the geographic diffusion literature. A corollary of Hägerstrand's network diffusion hypothesis is that communication networks are mediated by geography, and that a channel of communication is more likely to exist between individuals living near each other than individuals living further away *ceteris paribus*.

Both Hägerstrand's core thesis—of the importance of pre-existing communication channels in the spread of behaviour and his hypothesis about the importance of distance in particular—have been shown by a number of papers in a variety of contexts. Hedström (1994) was one of the first, combining Strang and Tuma's (1993) method of using a Cox Proportional Hazards model, Hägerstrand's theory, and a novel dataset on the geographic spread of Swedish trade unions. Hedström used Swedish districts called *härads* as his unit of analysis, modelling the time to adoption—appearance of a trade union—in each. He focused on the cumulative effect of already adopted härads weighted by distance and controlled for heterogeneities such as labour force, urbanity, and membership in other related organisations.

Andrews and Biggs (2006) take a similar methodological approach to studying the spread of sit-ins across the US during the civil rights movement. Taking cities with a minimum total population of 10,000 and a minimum black population of 1,000 as their unit of analysis, Andrews and Biggs assessed a variety of hypotheses for why certain areas held sit-ins when they did. Again using a proportional hazard model, their results indicate spatial proximity was significant (and a better predictor of sit-ins than athletic association between majority black universities) as was news coverage.

Three other recent papers build on this foundational work. For the first two, distance mediated contagion plays an important but less central role. Vasi and Strang (2009) study the spread of Municipal Bills of Rights in response to the USA PATRIOT act. Looking at all cities in the US with a population greater than 25,000, they assess how cities influenced each other as the movement spread, finding that initially nearby Bill of Rights legislation had a significant positive effect but as national coverage became more prevalent the local effect became insignificant.

Ingram et al. (2010) include spatial proximity in their study of a movement to limit the expansion of Wal-Mart stores across the US. They also find proximity between protests to have a significant effect, though they are considering a complex framework of strategy and do not use the same event-history models used by the other papers mentioned here. Finally, Myers (2010) re-evaluates waves of riots after the death of Martin Luther King, deriving a sophisticated model for the effect of distance on the spread of riots, which again shows a significant positive effect.

One key feature of both the Andrews and Biggs (2006) and Vasi and Strang (2009) papers was the importance of media coverage. In a marketing context (following Bass's (1969) influential model), media coverage would be termed an 'external' or exogenous effect because it is separate from the endogenous, social interaction effect. The relative importance of internal and external effects has been a subject of great contention (Coleman et al., 1966; Burt, 1987; Marsden, 1990; Strang and Tuma, 1993; Valente, 1996; Van den Bulte and Lilien, 2001). Fortunately for our purposes, the aforementioned ProQuest search (see §2.2.1) suggests that we can largely discount external effects in our study.

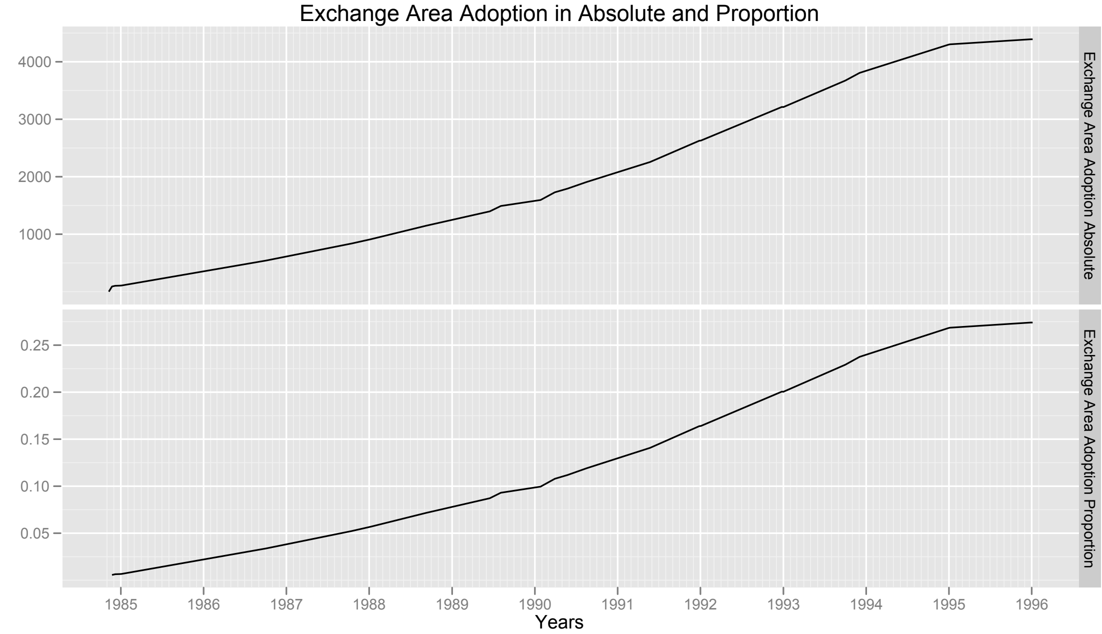


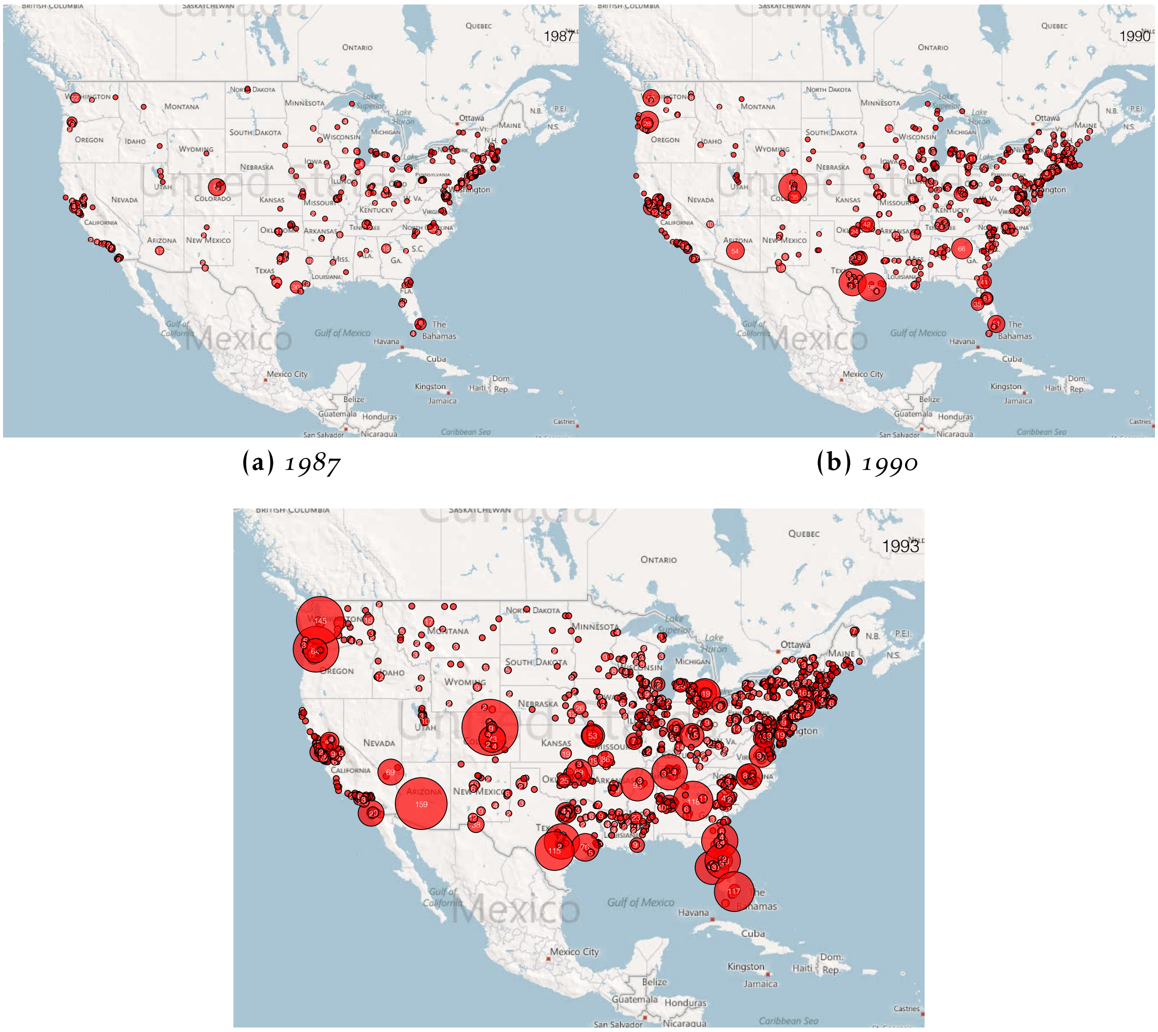
Figure 3.1: *FidoNet adoption curves for EAs in absolute and proportional scales.*

3.3 Data

While a small fraction of Nodelists prior to 1991 have survived, full data is available at weekly time intervals from 1991 on. However, the vast majority of these lists are in a complex compressed format that had not been able to parse at the time of the data analysis for this chapter. Therefore we are limited to 35 irregularly spaced snapshots of the network over 20 years. Since this chapter only considers the growth phase, this is limited to the first twelve years and 27 time steps.

3.4 Method

We use a Cox Proportional Hazards model (CPH) to assess the cumulative effect of adoption weighted by proximity in the diffusion process. However, because our data is at irregular time intervals we estimate a CPH using complementary log-log regression following a method suggested in Rabe-Hesketh and Skrondal (2008). This method controls for the length of each time interval using dummy variables. With each Exchange Area, we constructed a demographic profile as a population-



(a) *1987* **(b)** *1990*

(c) *1993*

Figure 3.2: *The spatial spread of FidoNet from 1987 to 1993. The circle diameter is proportional to the number of BBSs estimated to be at that location.*

weighted average of the data from their constituent tracts, using proportions where possible. Only median income and population density use an absolute scale.

3.4.1 Controls

We chose population density, urbanity (the proportion of households living in an urban area[5]), median income, age, educational attainment and related occupations as our control variables. These controls are chosen to account for systematic spatial variations which would encourage potential sysops to set up Fido BBSs irrespective of the cumulative distance effect. Our selections are motivated by the theoretical discussion in §2.1.2 and help isolate the spatial effect we seek to characterise by reducing the chance of spurious inference.

We combine area and population into density because we believe the emergence of a sysop is more likely to occur from a process of social interaction, both off and online. Higher population density should *ceteris paribus* increase the frequency of social interaction offline. In turn, this should increase the likelihood of information about FidoNet being spread and increase the cohesion of a local BBS community should one emerge.

The urbanity variable should have a related but different effect: information tends to diffuse through cities before rural areas,[6] and shops with computer equipment are more likely to be present in urban areas. An added benefit of preferring population density is that its correlation with urbanity is far lower (0.34) than the correlation between population and urbanity (0.618).[7] We rescaled ($\log_{10}$) population density to compensate for the skew in the distribution.

Median income is used as a single, representative measure of income. US income is notoriously skewed and thus the median is a far better approximation than the mean. We divide by $10,000 to keep the range closer to the other variables.

[5]The US Census defines an area to be urban if it is incorporated or is a Census Designated Place (a locality officially recognised by the US Census) with a minimum population of 2500 (US , 1995).

[6]This is a basic tenet of hierarchical diffusion theory. See Huang and Gould (1974); Fischer (1978); Sugiura (1986).

[7]We ran the regression with population and area as separate variables as well and the results were not substantially different though the fit was worse.

Age is stratified into groupings meant to track similar circumstances in life trajectories. 14–18 covers high school, 19–24 undergraduate college and first jobs, 25–29 second jobs and first promotions as well as starting families. Decades then continue up to 60, beyond which should account for retirement. Age should relate to disposable income, personal access to a phoneline, dwelling,[8] free time and technical know-how.[9]

Educational attainment should also have implications for living circumstances, income, and technical knowledge. Universities can provide the free time and a community of technically minded individuals to set up a BBS, as well as access to useful equipment. However, universities may also provide access to competing technologies, especially to graduate students, such as ARPANET. This issue is discussed more extensively below. Of course those with higher educational attainment and technical knowledge may not be the only ones to benefit direct or indirect involvement with FidoNet: their social ties may also be more likely or able to set up nodes as a result.

Occupations have implications for income and free-time (managers), but certain industries would require computer usage and knowledge. Engineers and technicians would have used and potentially built and fixed computer equipment as part of their job, and may have come in contact with computer magazines which covered FidoNet and BBSs. Moreover, areas with large numbers of engineers and technicians may well have better than average provision of the computer parts necessary to run a node either due to industry needs or the interests of engineers and technicians in new hobbyist technologies outside their work. One does not have to be an engineer or technician to take advantage of the commercial availability of this equipment.

Scientists would similarly use computers in research and may have used FidoNet and BBSs to share data and results with colleagues at other institutions. However, it is more likely researchers would have used one of the large research networks (like

[8] It is unlikely a sysop could run a BBS without a home, business, or flat.

[9] Different generations may be more or less likely to have come in contact with computers.

ARPANET) which eventually became the internet. ARPANET had its own message and filesharing community called USENET which was arguably a competitor to FidoNet.[10] Similar arguments about the availability of equipment can be made about places with large numbers of scientists as with technicians and engineers.

We also include teachers because as early as 1980 computers were an important component of pre-college education.[11] Teachers would have had some contact with running and maintaining computer systems at school, which might have increased their awareness of projects like FidoNet and overall technical expertise.

3.4.2 Model

Like the papers on spatial diffusion cited above, we use a CPH to estimate the hazard rate of adoption over the course of the time-series, taking geographic locations—Exchange Areas—as our risk-set. Cox models are of the form found in equation 3.1, where $\Lambda(t|\mathbf{X}_{it})$ is the hazard rate of Exchange Area i at time t.

Λ_0 is the baseline hazard and $\boldsymbol{\beta}$ is the vector of estimated coefficients for the effects from $\mathbf{X}_{it}$. We include t as an index for the vector of explanatory variables $\mathbf{X}_{it}$ because some of these variables—namely the effect of distance mediated contagion to other adopted[12] Exchange Areas and a dummy variable for previous adoption—change over time.

$$\Lambda(t|\mathbf{X}_{it}) = \Lambda_0(t)\exp(\boldsymbol{\beta}\mathbf{X}_{it}) \tag{3.1}$$

Previous Adoption

Before delving into the cumulative adoption effect we shall address our decision to include a previous adoption variable. Unlike most diffusion models (Rogers, 2003; Young, 2009) we take into account the potential for de-adoption: when an Exchange Area loses all its BBSs it returns to the risk-set and no longer contributes to the

[10]However, a gateway was eventually built between USENET and FidoNet, and this would have increased FidoNet awareness among USENET users.

[11]55% of American schools had some form of computer access by 1975 (Molnar, 1997).

[12]For those not well versed in survival modelling: an Exchange Area remains in the risk-set—the set of potentially adopting Exchange Areas—until it adopts, at which point it is excluded from the model and joins the set of contagious Exchange Areas affected the risk-set.

cumulative adoption effect. As Allison (1984) notes, with event-history models repeated events can easily incorporated.[13] We include a dummy variable to account for the effect of a residual community of FidoNet users who persist after their local BBSs are gone, and from whom a sysop may emerge. This variable is 1 if the Exchange Area has ever adopted and 0 if it has not.

Distance Mediated Cumulative Adoption Effect

We have chosen not to use the term 'contagion' to describe the cumulative effect of adoption events mediated by distance because contagion implies an unambiguously positive effect. As our results show, there may be circumstances where increases in adoption may decrease the likelihood of a sysop emerging. This will be discussed in greater detail below, but to remain agnostic on the directionality of the effect we will use the term Distance Mediated Cumulative Adoption (DMCA).

DMCA is assumed to decay with distance: the closer an at-risk Exchange Area is to an adopted Exchange Area, the greater the effect. The functional form of this decay is the subject of some discussion. Both Hedström (1994) and Andrews and Biggs (2006) use the reciprocal of the square root of distance:

$$\kappa_{it} = \beta_\kappa \sum_j \frac{\gamma_{j,t-1}}{\sqrt{d_{ij}}} \tag{3.2}$$

where κ_{it} is the DMCA effect on Exchange Area i at time t, d_{ij} is the distance between the centroids of Exchange Areas i and j, β_κ is the estimated coefficient (from the $\boldsymbol{\beta}$ vector in equation 3.1) and

$$\gamma_{j,t-1} = \begin{cases} 1 & \text{if } j \text{ has adopted by } t-1 \\ 0 & \text{if } j \text{ has not adopted by } t-1 \end{cases}. \tag{3.3}$$

However, both papers use this form because the fit was best, rather than from any theoretical preference. Ingram et al. (2010) uses $\kappa_{it} = \beta_\kappa \sum_j \frac{\gamma_{j,t-1}}{d_{ij}}$ without apparent justification, while Myers (2010) uses a far more complex function[14] derived directly

[13] See Myers (2010) for a sophisticated application of this possibility.

[14] Myers appears to derive $\kappa_{it} = 0.998 \exp(\frac{3.48}{\exp(.00252 d_{ij})} - \frac{3.2}{\exp(.00265 d_{ij})} - .000128 d_{ij}) \gamma_{j,t-1}$ from his data by fitting a Gompertz distribution to the distribution of distances between adopting areas, though the details of this process are not well-specified in the paper.

from the data and selected for its superior fit. We follow the thus far accepted methodology of selecting based on goodness-of-fit, and do so by testing a variety of κs.

Missing data and Complementary Log-Log Regression

As discussed above in §3.3 there are significant sections of the time series which are missing. Normally this would be considered interval-censoring: a common problem with panel data where observations of a continuous time process—such as cancer relapse—are taken at regular time intervals while a relapse event may occur at any point between observations. However, in our case we have *missing* data because Nodelists were published every week but prior to 1991 many of them have been lost, and from 1991 on many are in a format we cannot yet use. The missing-censoring distinction is largely semantic though: techniques for accounting for interval censoring should be applicable to our case.

Traditional CPH modelling cannot account for interval censoring or irregular time intervals. Rabe-Hesketh and Skrondal (2008) suggest using complementary log-log regression with dummy variables for each time interval. This is an alternative means of estimating a CPH which controls for the variations in duration of each time-interval. The basic form of a cloglog regression is:

$$\Lambda(t_i|\mathbf{X}_{it}) = 1 - \exp(-\exp(\Lambda_0(t) + \boldsymbol{\beta}\mathbf{X}_{it})). \tag{3.4}$$

Including the time interval dummies, DMCA and previous adoption effects (the time dummies and previous adoption are incorporated into the $\boldsymbol{\beta}\mathbf{X}_{it}$ term) we come to equation 3.5: the model we use.[15]

$$\Lambda(t_i|\mathbf{X}_{it}) = 1 - \exp(-\exp(\beta_\kappa \kappa_{it} + \boldsymbol{\beta}\mathbf{X}_{it})). \tag{3.5}$$

[15] Note the lack of Λ_0, which is now incorporated as a time interval dummy variable into the $\boldsymbol{\beta}\mathbf{X}_{it}$ term.

3.5 Results

Our results suggest that DMCA did have a significant impact on the diffusion of FidoNet during its growth phase in the US. We also find a κ of $\beta_{\kappa 1} \sum \frac{\gamma_{j,t-1}}{d_{ij}} + \beta_{\kappa 2}(\sum \frac{\gamma_{j,t-1}}{d_{ij}})^2$ superior to the others we tried.[16] However, analysing this κ suggests that the DMCA effect changes as the effect increases, potentially becoming negative as the FidoNet growth phase reaches its peak. We will first present the results which led us to select this κ and then discuss its implications.

3.5.1 Model Without DMCA Effect

	No Adoption Effects		Previous Adoption	
$\log_{10}$ Population Density	0.732***	(27.69)	0.683***	(27.79)
Urbanity	0.856***	(15.12)	0.818***	(15.38)
Median Income ($10,000)	−0.00972	(−0.41)	−0.0169	(−0.80)
Age 14 to 18	−9.293***	(−8.48)	−10.53***	(−10.49)
Age 19 to 24	−0.488	(−1.23)	−0.739*	(−2.01)
Age 25 to 29	−1.411	(−1.37)	−1.976*	(−2.11)
Age 30 to 39	−1.324	(−1.95)	−1.982**	(−3.19)
Age 40 to 49	0.892	(1.12)	1.134	(1.56)
Age 50 to 59	−0.875	(−0.98)	−1.753*	(−2.14)
Age 60 and Over	−3.287***	(−10.80)	−3.526***	(−12.61)
High School Graduate	−0.217	(−0.87)	−0.144	(−0.63)
Undergraduate Degree	1.833***	(9.13)	1.770***	(9.66)
Post-Graduate Degree	−1.639**	(−2.63)	−1.708**	(−3.01)
Managers	−0.131	(−0.10)	−0.694	(−0.56)
Engineers	10.94*	(2.22)	13.07**	(2.93)
Scientists	−13.65	(−1.11)	−13.21	(−1.17)
Teachers	−2.122	(−0.47)	−4.175	(−1.01)
Technicians	44.23***	(5.38)	42.13***	(5.65)
Previous Adoption			0.553***	(17.19)
AIC	62097.9		**61737.9**	

390,270 **observations**, t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Reference Category: non-High School Graduates

Table 3.1: *Comparison of the Null and Previous Adoption CPH Models*

Table 3.1 shows the results for our null model (no adoption effects) compared with including the previous adoption effect. We use the Akaike Information Criterion (AIC)[17] as our measure of goodness of fit. Adding the previous adoption

[16]We will use $\beta_{\kappa n}$ to denote the coefficient for the nth term of a given κ, since some of those we tested have multiple terms.

[17]AIC compares the goodness-of-fit for different models on the same dataset, penalising for extra

variable improves the fit significantly—$\Delta_{\text{AIC}} = 360$—and the coefficient is highly statistically significant and positive, suggesting that Exchange Areas which de-adopted were significantly more likely to have another BBS emerge than Exchange Areas that never had a BBS.

The coefficients on the control variables are relatively unsurprising, with the exception of the negative effects from Scientists, Teachers, Post-Graduates, and the 30–39 age bracket. With the exception of ages 30–39 and Post-Graduates, these are statistically insignificant results. We suspect that Scientists and Post-Graduates may be more inclined to use USENET (a competing technology discussed above in §3.4.1) as they are more likely to be employed by large institutions and in particular research groups—the primary early users of USENET. We expected the 30–39 age group to have the stable income and household needed to become sysops, but perhaps that happens in the only positive though statistically insignificant age bracket (40–49).

Income is similarly puzzling, but its very low statistical significance indicates that coefficient is almost indistinguishable from 0. Perhaps this is due to variations in the cost of living, which may render this variable a problematic estimate of excess disposable income.

The caveat to all these interpretations is the classic ecological fallacy: the quantitative profile of a population does not constitute the profile of individuals who may or may not choose to join FidoNet. Our results suggest that regions that had these demographic profiles were more likely to have a sysop, but we cannot make a strong claim about the individuals without data from them directly. We have tried to focus on regional attributes which may correlate with people who would tend to be keen on FidoNet during its growth phase, but making the claim that therefore these results demonstrate these were the people who did use FidoNet is much harder to demonstrate without interviewing those participants directly.

variables. Lower values are better and a change in AIC (Δ_{AIC}) of 2–7 is modestly significant, while a change of 10 or greater is highly significant.

3.5.2 Distance Mediated Cumulative Adoption Effect

Table 3.2 compares the fit for different ways of calculating the distance decay of the DMCA (κ). All of these models improve on the null model significantly, though only two of them improve on the Previous Adoption model significantly. Hedström (1994); Andrews and Biggs (2006) both find $\kappa = \beta_\kappa \sum \frac{\gamma_{j,t-1}}{\sqrt{d_{ij}}}$ to provide the best fit to their data, and relative to $\beta_\kappa \sum \frac{\gamma_{j,t-1}}{d_{ij}}$ and $\beta_\kappa \sum \frac{\gamma_{j,t-1}}{d_{ij}^2}$ this is true for our data as well, though only by a relatively insignificant amount (Δ_{AIC} =1–2). However, adding a second order term provides a marked improvement, especially with $\kappa = \beta_{\kappa 1} \sum \frac{\gamma_{j,t-1}}{d_{ij}} + \beta_{\kappa 2}(\sum \frac{\gamma_{j,t-1}}{d_{ij}})^2$, which improves on the $\beta_\kappa \sum \frac{\gamma_{j,t-1}}{\sqrt{d_{ij}}}$ model by $\Delta_{\mathrm{AIC}} = 82.3$. Our next best model, which also includes a second order term ($\kappa = \beta_{\kappa 1} \sum \frac{\gamma_{j,t-1}}{d_{ij}} + \beta_{\kappa 2} \sum \frac{\gamma_{j,t-1}}{d_{ij}^2}$) lags behind by $\Delta_{\mathrm{AIC}} = 73.3$. Additionally, the coefficients on the first and second order DMCA effects are both highly statistically significant. Of particular interest is the negative coefficient on the second order term, which indicates the DMCA effect peaks and then declines. This effect will be explored in the next section.

We also considered a density effect, whereby the effect of each adopted Exchange Area was multiplied by the number of BBSs it contained at that point in time. These results had a considerably worse fit ($\Delta_{\mathrm{AIC}} = 25.6$) so we chose to focus on the quadratic model. The results and methodology for the density effect can be found in appendix A.4.

We have strong evidence to suggest that two separate adoption effects—previous adoption and the DMCA effect—were significant factors in the spread of FidoNet across the US. Under our preferred model, the coefficients on both the linear and quadratic effects are highly statistically significant and the model improves on the null model by $\Delta_{\mathrm{AIC}} = 446.5$, an extremely significant result. Had we not considered a second order effect, we might have discounted the importance of the DMCA effect because the single order models only improve on the AIC marginally. This suggests that the second order effect may be of particular importance.

$\kappa =$	$\sum \frac{\gamma_{j,t-1}}{d_{ij}}$		$\sum \frac{\gamma_{j,t-1}}{\sqrt{d_{ij}}}$		$\sum \frac{\gamma_{j,t-1}}{d_{ij}^2}$		$\sum \frac{\gamma_{j,t-1}}{d_{ij}} + (\sum \frac{\gamma_{j,t-1}}{d_{ij}})^2$		$\sum \frac{\gamma_{j,t-1}}{d_{ij}} + \sum \frac{\gamma_{j,t-1}}{d_{ij}^2}$	
$\log_{10}$ Population Density	0.674***	(27.20)	0.682***	(27.72)	0.690***	(27.75)	0.682***	(27.49)	0.680***	(27.24)
Urbanity	0.821***	(15.44)	0.811***	(15.20)	0.813***	(15.25)	0.789***	(14.85)	0.814***	(15.27)
Median Income ($10,000)	−0.0191	(−0.90)	−0.0183	(−0.86)	−0.0194	(−0.91)	−0.0275	(−1.30)	−0.0254	(−1.20)
Age 14 to 18	−10.97***	(−10.71)	−11.52***	(−10.60)	−10.44***	(−10.38)	−13.32***	(−12.47)	−11.14***	(−10.83)
Age 19 to 24	−0.908*	(−2.42)	−1.012**	(−2.63)	−0.699	(−1.89)	−1.617***	(−4.18)	−0.969*	(−2.57)
Age 25 to 29	−2.001*	(−2.14)	−2.081*	(−2.22)	−1.933*	(−2.06)	−2.336*	(−2.48)	−1.940*	(−2.06)
Age 30 to 39	−2.229***	(−3.52)	−2.410***	(−3.70)	−1.867**	(−2.99)	−3.414***	(−5.24)	−2.211***	(−3.48)
Age 40 to 49	1.022	(1.41)	0.911	(1.25)	1.168	(1.61)	0.316	(0.43)	0.997	(1.37)
Age 50 to 59	−1.943*	(−2.35)	−2.031*	(−2.45)	−1.625*	(−1.98)	−2.499**	(−3.02)	−1.857*	(−2.26)
Age 60 and Over	−3.648***	(−12.71)	−3.791***	(−12.48)	−3.505***	(−12.51)	−4.419***	(−14.82)	−3.708***	(−12.91)
High School Graduate	−0.219	(−0.95)	−0.288	(−1.22)	−0.154	(−0.67)	−0.375	(−1.61)	−0.298	(−1.28)
Undergraduate Degree	1.834***	(9.90)	1.820***	(9.88)	1.747***	(9.54)	1.911***	(10.27)	1.842***	(9.95)
Post-Graduate Degree	−1.843**	(−3.23)	−1.881***	(−3.29)	−1.645**	(−2.90)	−1.904***	(−3.36)	−1.833**	(−3.22)
Managers	−0.852	(−0.68)	−0.846	(−0.68)	−0.459	(−0.38)	−0.651	(−0.53)	−0.537	(−0.44)
Engineers	13.55**	(3.04)	13.52**	(3.04)	12.55**	(2.81)	12.93**	(2.92)	12.93**	(2.91)
Scientists	−12.42	(−1.10)	−11.55	(−1.02)	−13.50	(−1.19)	−9.005	(−0.81)	−12.32	(−1.09)
Teachers	−4.067	(−0.98)	−4.216	(−1.02)	−4.192	(−1.01)	−4.218	(−1.02)	−4.007	(−0.97)
Technicians	41.57***	(5.59)	41.07***	(5.52)	41.74***	(5.59)	40.09***	(5.42)	40.38***	(5.43)
Previous Adoption	0.548***	(17.00)	0.550***	(17.10)	0.555***	(17.27)	0.548***	(17.13)	0.549***	(17.03)
$\sum \frac{\gamma_{j,t-1}}{d_{ij}}$	0.0282*	(2.09)					0.395***	(8.53)	0.0511***	(3.34)
$\sum \frac{\gamma_{j,t-1}}{\sqrt{d_{ij}}}$			0.00299*	(2.32)						
$\sum \frac{\gamma_{j,t-1}}{d_{ij}^2}$					−0.176	(−1.68)			−0.333**	(−2.72)
$(\sum \frac{\gamma_{j,t-1}}{d_{ij}})^2$							−0.0356***	(−8.12)		
AIC	61734.9		61733.7		61735.5		**61651.4**		61724.7	

Reference Category: non-High School Graduates

390,270 **observations**, t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.2: *Comparing Fit of Different* κ*s for the* DMCA *Effect*

3.6 Discussion

Our results robustly confirm our hypotheses but pose significant questions. In particular: why is the second order term negative? What does that result imply about the spatial component of FidoNet's diffusion? How can we interpret this variable? Have we overfit our data to an arbitrary equation?

To assess these questions we analyse our selected κ to assess the importance of its shape given our data. It could, for example, be the case that the implied inflection point (where the effect would turn negative) lies outside the range of values ever reached by our data and therefore has no actual bearing. Our conclusions pose more questions and provide few answers, but if these results are robust they may prove significant in modelling other geographic diffusion processes.

We first assess the form of the distance decay by analysing the effect of an additional adopting Exchange Area at a range of distances. We then consider the overall shape of the DMCA effect given the range and distribution of the DMCA variable. We find the inflection point and assess the extent to which that point is reached by the data. We finish with concluding remarks and suggestions for further work.

3.6.1 The Shape of the Distance Decay

Figure 3.3 plots the effect of an additional adopted Exchange Area with respect to distance using the top three κs. In this plot and those below, the hazard rate is calculated with respect to the median values of the control variables and a Previous Adoption of 0. Only the DMCA effect varies. Each of these curves is a ratio relative to a baseline DMCA effect of 0. The solid curve is the κ of best fit, which has a slower decay than the κ used by Hedström (1994); Andrews and Biggs (2006) ($\beta_\kappa \sum \frac{\gamma_{j,t-1}}{\sqrt{d_{ij}}}$) which is the smaller dashed line.[18] However, the form of $\kappa = \beta_{\kappa 1} \sum \frac{\gamma_{j,t-1}}{d_{ij}} + \beta_{\kappa 2} \sum \frac{\gamma_{j,t-1}}{d_{ij}^2}$ is opposite of the other two, implying a repulsion effect where nearby events have

[18]The curve of $\kappa = \beta_\kappa \sum \frac{\gamma_{j,t-1}}{d_{ij}}$ has a similar form to $\kappa = \beta_\kappa \sum \frac{\gamma_{j,t-1}}{\sqrt{d_{ij}}}$ and we chose not to include it to leave the plot uncluttered.

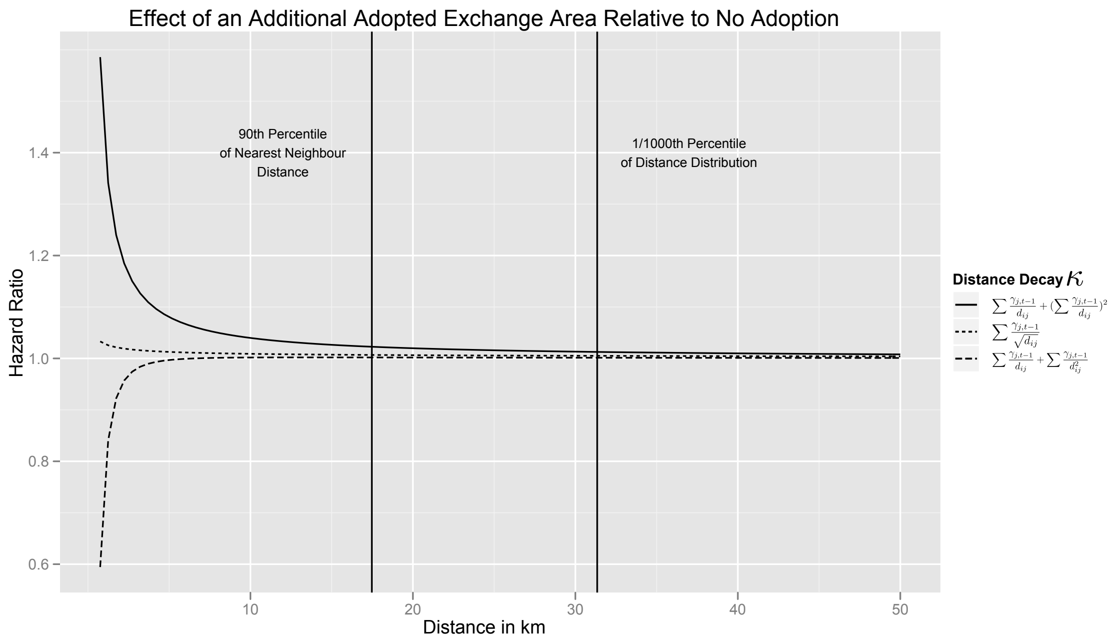
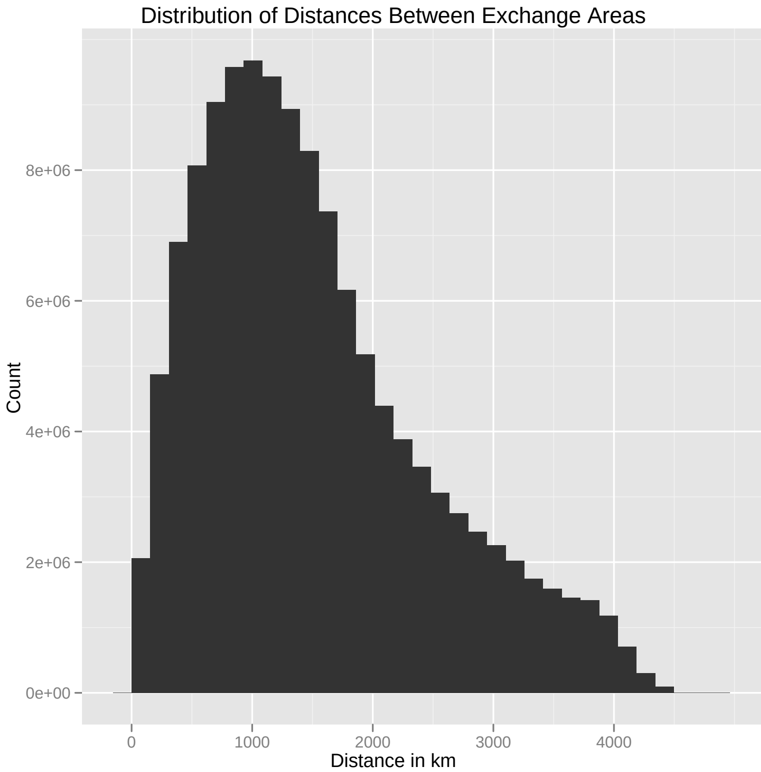


Figure 3.3: *Hazard ratio with respect to a baseline hazard of no* DMCA *effect for the three best* κ*s.*

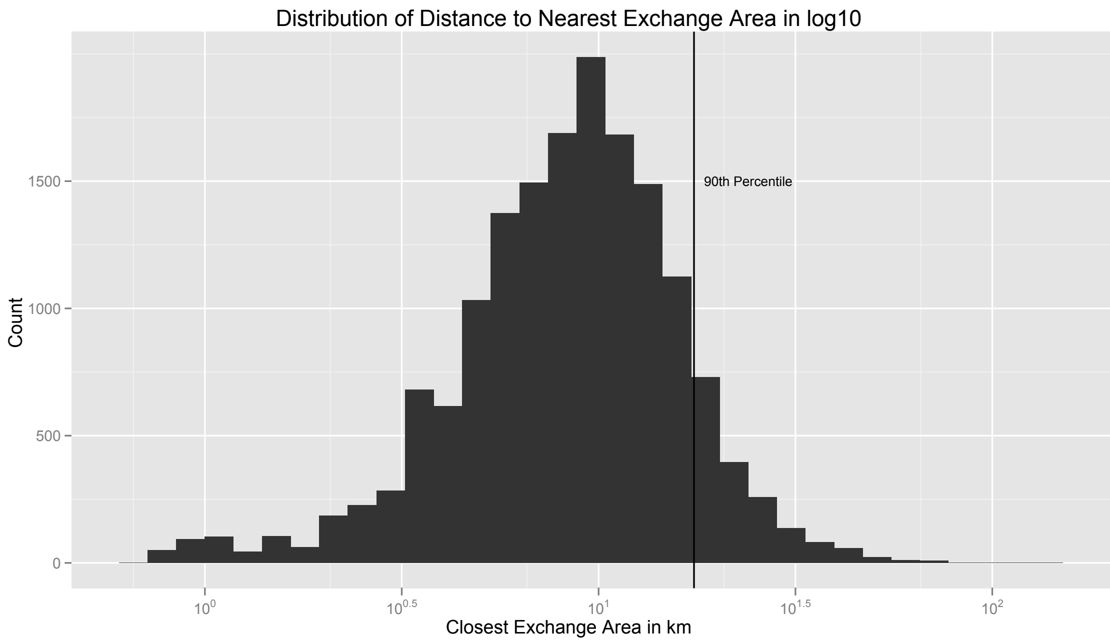
a strong negative effect on the hazard rate which declines with distance. The fact that this model is our second best candidate, and improves significantly on models which have a positive effect for nearby adoption is strange given our best model also indicates a positive proximity effect.

Before considering this puzzle, we draw the reader to the two vertical lines designed to indicate how important the shape of this function is given the data. The vertical lines relate to quantiles in the distribution of distances between Exchange Areas. The left vertical line is the 90% percentile for the distribution of *minimum* distances between Exchange Areas (figure 3.4b is a histogram of these distances). This means that the nearest neighbour of 90% of Exchange Areas is less than 17.47 km away, indicating that for the vast majority of Exchange Areas, if their nearest neighbour adopted the effect would significantly impact their hazard rate.

However, the other vertical line is the one-one-thousandth quantile of the entire distribution of distances between Exchange Areas (figure 3.4a), meaning that 99.9% of the distances between Exchange Areas are *greater* than 31.34 km. These lines



(a) *Distribution of Distances*



(b) *Distribution of Minimum Distances*

Figure 3.4: *Histograms of Distances Between Exchange Areas*

indicate that while nearly every Exchange Area has at least one neighbour closer than 20 km, the vast majority are much further away. Given how steep the decay for all three κs, it would seem that the effect of an additional adoption is only highly significant in circumstances of very close proximity. It is worth considering however, that this is a cumulative effect: while long-distance adoption is individually negligible, the sum of many long-distance adoptions could be significant.

3.6.2 The Shape of the Distance Decay as DMCA Increases

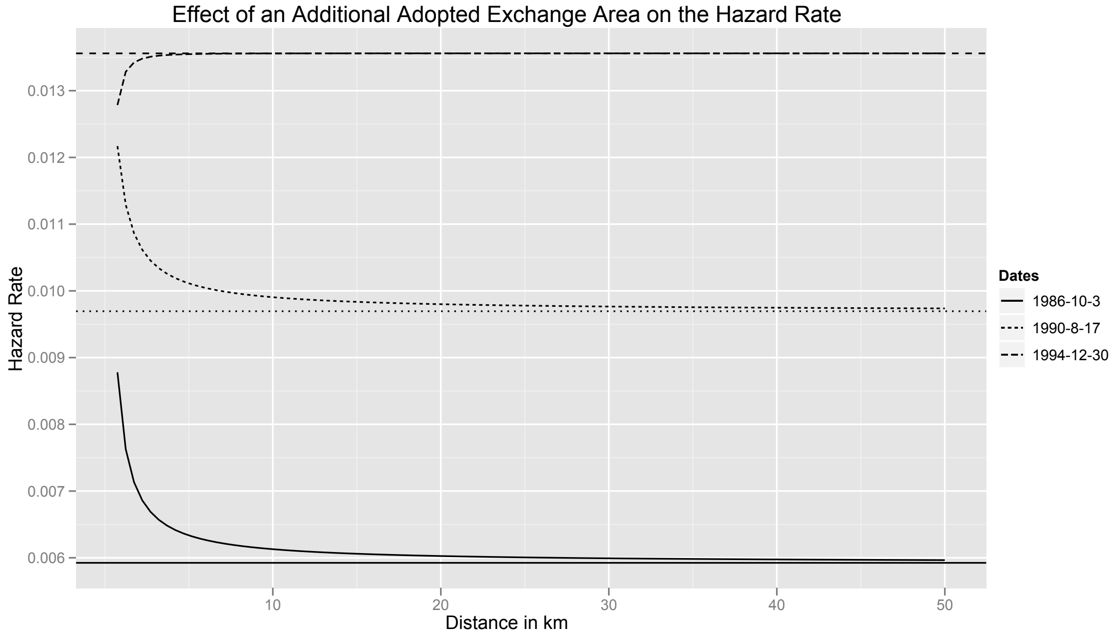
Figure 3.3 demonstrates the effect of an additional adoption (a distances as indicated along the x-axis)relative to a DMCA level of 0, but increasing the DMCA level changes the shape of the decay function for our κ-of-best-fit.[19]

Figure 3.5a shows the distance decay with respect to the hazard rate for an additional adoption event at the average DMCA level for three points in the time series. The horizontal lines show the baseline hazard rates for those points in time. This plot demonstrates that as the DMCA increases the decay curve shifts, eventually flipping over to resemble the repulsion effect implied by the $\kappa = \beta_{\kappa 1} \sum \frac{\gamma_{j,t-1}}{d_{ij}} + \beta_{\kappa 2} \sum \frac{\gamma_{j,t-1}}{d_{ij}^2}$ model. This may explain why that model had a superior fit relative to the other models with a decay similar to that at the beginning of the time series, and that this flip may be a significant effect. Figure 3.5b demonstrates the relative shift more clearly by showing the same time series as a hazard ratio (relative to the baseline). As DMCA level increases with time, the baseline curve shifts up. However, the distance decay curve (relative to the shifting DMCA baseline) flattens, eventually becoming a repulsion curve.

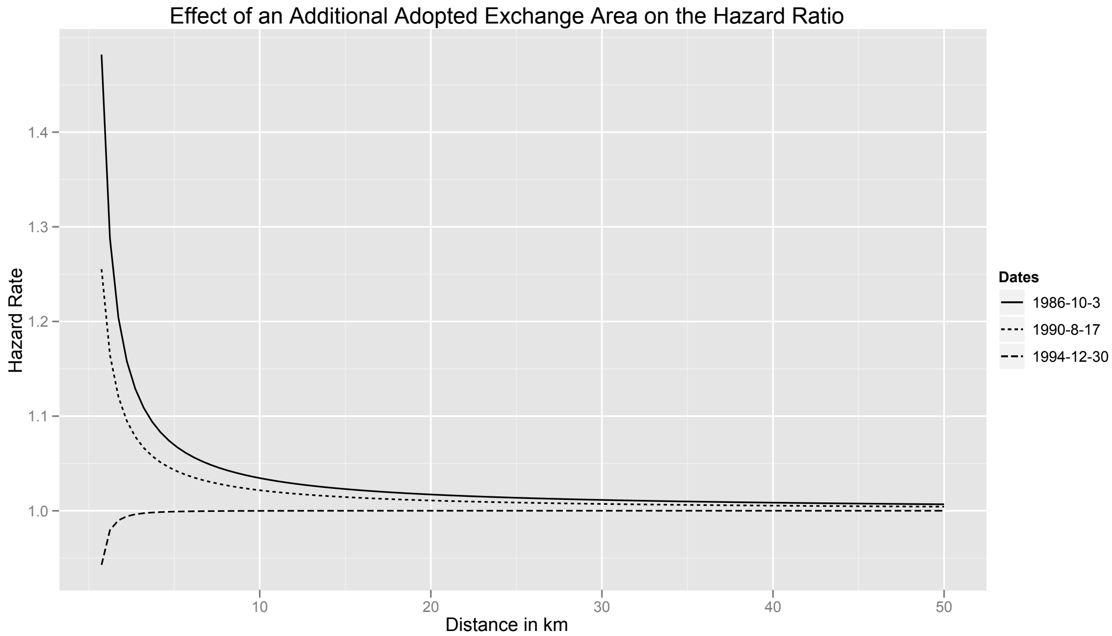
Figure 3.6 shows the shape of this effect relative to the range of values the DMCA variable takes in the data.[20] These plots show the same information in different ways. Figure 3.6a shows the change as a 3-dimensional surface where the Z-axis is the hazard ratio and the corner of the surface closest to the Z-axis shows the

[19]The rest of our analysis focuses on the κ which fit the data best: $\beta_{\kappa 1} \sum \frac{\gamma_{j,t-1}}{d_{ij}} + \beta_{\kappa 2} (\sum \frac{\gamma_{j,t-1}}{d_{ij}})^2$.

[20]The range of distances is 0.73 km (the minimum distance between Exchange Areas in our database) and 10 km. We use 10 to show the change in colours more clearly than if we had chosen 50.



(a) *Hazard Rate*

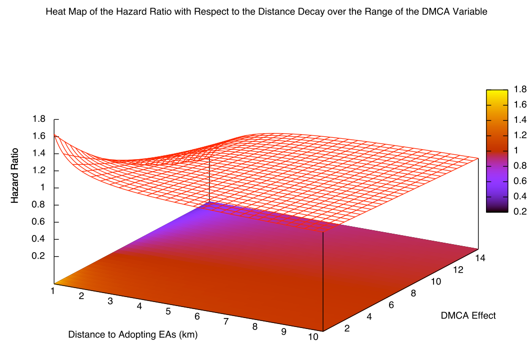


(b) *Hazard Ratio*

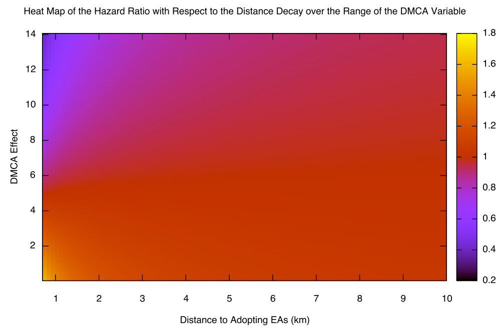
Figure 3.5: *Change in the distance decay as* DMCA*s increase, using the average* DMCA *level at three points in the time series.*

effect with low DMCA (early in the time series) while the far corner—pointing down—shows the effect for the highest DMCA levels (at the very end of the time series).

These levels are projected onto a heatmap which is reproduced as figure 3.6b for convenience. Here the hazard ratio is indicated by colour, with yellow as the highest level (initial peak about 1.63), red at about 1, and purple as the lowest level (minimum of 0.39). The flip seems to occur for a DMCA level of between 4 and 6.



(**a**) *3d Surface with Heatmap Projection*



(**b**) *Heatmap Projection*

Figure 3.6: *The distance decay with respect to the range of the* DMCA *variable, demonstrating the flattening to repulsion effect.*

What is potentially misleading about this plot is the extent to which the maximum DMCA level—14.066—is an outlier. If the distribution of the DMCA variable remains lower than 6, then the repulsion effect would only apply to a few isolated cases, and indicate that it is in fact comparatively negligible.

Accounting for Shifts in the DMCA Level

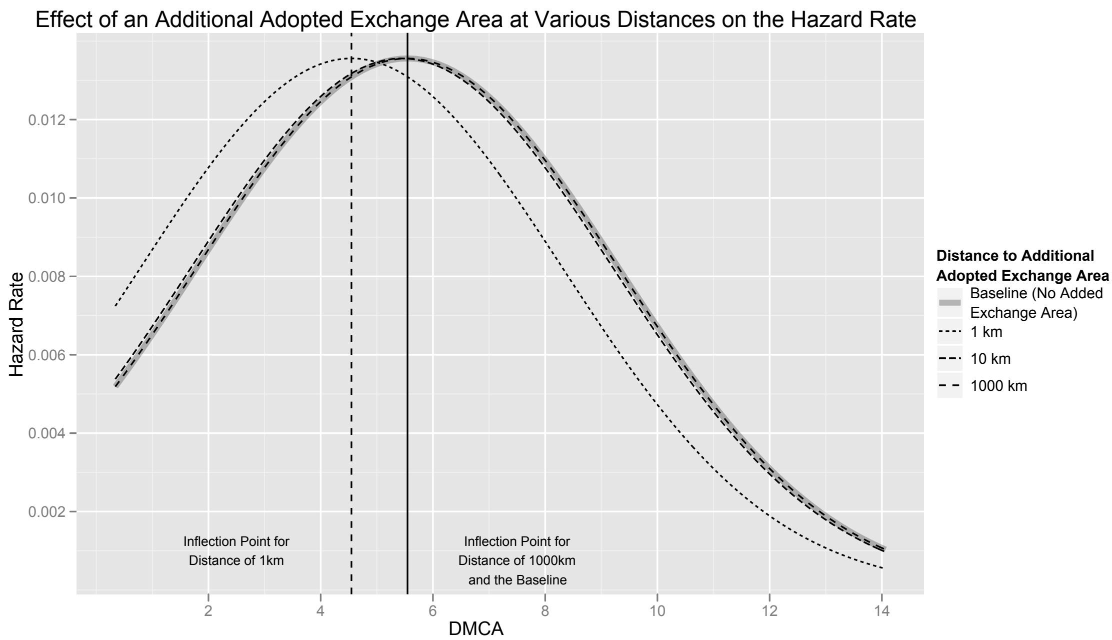
Before assessing this possibility, we would like to consider another effect implied by figure 3.5a: the shift of those curves upward as the DMCA level increases. Not only is the distance decay shifting as the DMCA effect increases, but the entire baseline hazard rate is shifting up as well, indicating that while the effect of distance on additional adoption events is flattening, the overall cumulative effect of additional adopting Exchange Areas is positive and significant, shifting the entire curve upwards.

Figure 3.7a shows the relationship between the hazard rate and the DMCA over its range in the data [.036,14.066]. The baseline—the relationship without additional Exchange Area adoptions—is the light grey line furthest to the right. The dashed lines show the effect of an additional Exchange Area adoption at 1, 10 and 1000 km. This is a family of curves, translated along the x axis by an Exchange Area adoption event at that distance. An additional Exchange Area 1 km away is a translation of 1 ($\frac{1}{1}$) toward the origin, 10 km leads to a 0.1 ($\frac{1}{10}$) translation and so on. Since $\lim_{x\to\infty}\frac{1}{x} = 0$, greater distances tend toward the baseline, which coheres with the *decay* of distance: the further the adoption, the less effect it has on a potentially adopting Exchange Area.

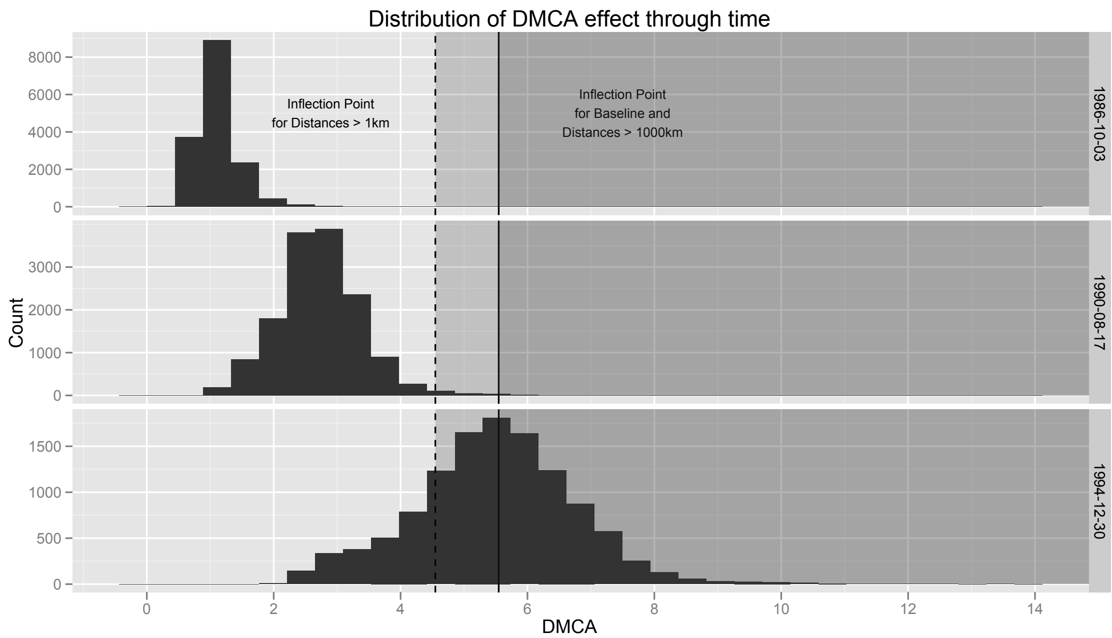
The two vertical lines indicate the inflection points for the 1 and 1000 km curves (DMCA = 4.548 and 5.548, respectively). To the left of those points, additional adoption events have a positive effect on the hazard rate. Once the DMCA level passes the inflection point for the curve associated with a particular distance, any additional adoption events up to that distance will have a negative effect on the hazard rate.

Further, past the baseline inflection point (which to three decimal places is no different from the 1000 km inflection point) any additional Exchange Area adoptions will be negative, irrespective of their distance away. Their impact may be negligible if they are especially far, or great if they are very close, but it will be negative. This seems to suggest an overall peak and decline, but without comparing it to the actual distribution of the DMCA variable, the significance of the effect is difficult to assess.

[21] The curves on this plot are of the function $\kappa = 1 - \exp(-\exp(.395(\mathrm{DMCA} + 1/d) - .0356(\mathrm{DMCA} + 1/d)^2 - 5.3894831))$ where DMCA is the current state of the region (the x axis), d is the distance to a new adoption (one for each curve as demonstrated in the caption) and k is the hazard rate (the y axis). 0.395 and −0.0356 are the linear and squared coefficients estimated in the model, and −5.3894831 is the sum of the averages of all the demographic effects multiplied by their respective coefficients. These curves would vary for each region dependent on its proximity to other adoptions and its demographic characteristics.



(a) *Inflection points for additional adoption events*[21]



(b) *Time series of* DMCA *distribution*

Figure 3.7: *The distance decay with respect to the range of the* DMCA *variable, demonstrating the flattening to repulsion effect.*

DMCA Distribution Over Time

Figure 3.7b shows the distribution of the DMCA variable at three time points. As time passes, the distribution shifts higher (to the right), eventually crossing the two inflection points from 3.7a. For Exchange Areas in the sections of the distribution beyond the 1 km points, additional Exchange Area adoptions may have a negative effect on the hazard rate depending on their distance. Exchange Areas past the baseline (1000km) inflection point have an unambiguously negative effect, though the strength of that effect remains a function of the distance (closer is more negative than further away).

3.7 Implications and Conclusions

These results are unusual and surprising: while for the majority of the time series the DMCA effect was positive[22], which coheres with other results on spatial diffusion processes (Hedström, 1994; Andrews and Biggs, 2006; Ingram et al., 2010; Myers, 2010) at least in terms of directionality, thus far no results we are aware of show a negative effect.

Vasi and Strang (2009), in their study of the spread of locally passed Bills of Rights in response to the US PATRIOT act, takes a different approach: instead of a distance decay function, they use a 60 mile cut-off, beyond which they assume there is no effect.[23]

They also find that for the early part of their time series nearby adoption had a significant positive effect, but towards the end the effect becomes non-significant. They interpret their findings as indicating that nationwide awareness of the Bill of Rights movement superseded the nearby adoption effect. Early on, local adoption propagated information about the movement but as national media coverage paid greater attention the local effect lost significance.

[22] Positive in the sense that additional adoptions increased the probability of further adoptions.

[23] Our model predicts that, even for extreme DMCA values, a distance of 60 miles (96.54 km) should have almost no marginal effect on the hazard rate, but cumulative effects make make a difference.

With respect to FidoNet, the possibility of a broadcast effect was discounted due to a general lack of media coverage. However, the aforementioned Wall Street Journal blurb was released in 1991, near when the DMCA distribution begins to pass the inflection point (when the effect becomes negligible) and into the negative region. Perhaps if their research covered a longer time series they would have seen a similar pattern. It is a possibility worth exploring, though without some way of distinguishing which Exchange Areas experienced more or less coverage it may be difficult to pin-point whether this can explain our results.

Another broadcast-like effect we hoped to include was the number of world-wide BBSs, which may increase incentives for users to join (seeing just how many people were involved) and may relate to overall congestion: the more people in the network as a whole, the more messages sent and potentially a greater need for sysops. An unfortunate limitation of our CPH estimation process, which includes dummy variables for each time period, currently precludes the inclusion of a global adoption variable because doing so creates a singular matrix, rendering the model inestimable. We are exploring other options at present, in particular the iterative convex minorant algorithm (Pan, 1999).

The cut-off approach is also worth considering as an explanation for our negative effect. Long-range adoption events may be acting as outliers, forcing the curve into a spurious shape. A distance cut-off point prevents long-range effects from having excess weight, which may explain the difference in our results.

Further on the subject of more sophisticated data analysis: incorporating the higher time resolution data which is available but at present in accessible could make a difference as well. This only applies post-1990, but would increase the number of time points by at least an order of magnitude. Coupled with the spline approach, this should yield extremely robust results.

We also chose not to allow for multiple adoption: once an Exchange Area has adopted, it leaves the risk-set until it no longer has a server, at which point it re-enters. Multiple adoption would allow an Exchange Area to remain in the risk-set and adopt a second BBS (or more), as it were modeling the density effect (see

§3.5.2 and appendix A.4). Whether this would have made a difference is unclear: since no Exchange Area would ever leave the risk set, presumably some areas with the greatest underlying propensity for FidoNet would continue adopting but it's unclear what impact that might have on the estimated κ coefficients.

Competition from the Internet might be a plausible explanation. At the time of FidoNet's peak US internet adoption was soaring, and if early internet adopters also tended to be FidoNet users, early internet adopters would in effect be broadcasting a competing technology, thus causing a negative effect near areas of high FidoNet concentration. Unfortunately Internet adoption data is quite poor until 1996, the year after FidoNet peaks in the US, and even then it is very difficult to get good regional data.

Another possible effect was a kind of over-popularity: following the sort of threshold arguments of Rogers (2003) and Granovetter (1978), we can imagine early adopters who, in Granovetter's words, were 'radical' and found FidoNet's eventual popularity alienating. Thus areas that may have adopted early become increasingly less inclined, causing a sort of collapse from the core which might again cause a reverse effect for areas near BBSs. This seems unlikely given the highly positive effect of previous adoption, which would suggest residual BBS communities are more a positive effect than a negative one.[24]

The last effect considered here is over-crowding: a related but different explanation to the over-population explanation. Adding new BBSs near existing ones fails to significantly increase the geographic dispersion of the network. Perhaps this was somewhat frowned-on by other sysops, which would suggest a kind of territorial possessiveness. As the network became larger, this sort of pressure would intensify as the US became more covered, such that the number of suitable regions far enough away from existing Exchange Areas would reduce significantly. However, this certainly runs contrary to Jennings' goal of BBSs close enough to avoid long-distance charges. Furthermore, a number of Exchange Areas had more than

[24] It is worth noting that according to the *BBS Documentary*, most of the original BBS creators (including Jennings) left the community by the early 1990s.

one BBS, and the coefficient on the quadratic term of the density model (see A.4) is also negative, suggesting the sign of that model is similar.

Similarly, we also tested the robustness of our results through splines following Snijders and Bosker (2011) in Appendix A.5. By splitting at points near the point of inflection estimated in the model we have focused on, we control for the possibility of over-fitting. The goodness-of-fit seems to peak at a point of inflection of 5, very close to the peak in the model tested Appendix A.5. This suggests our results are robust in this regard, and visually the shape of the best curve is quite close to the quadratic model we have focused on Figure A.3. We then compare the best spline model to the quadratic model, including the linear term, and the sign and magnitude for the coefficients of interest are very similar. The Δ_{AIC} for the splines model is slightly better than the quadratic model, but the Δ_{BIC}[25] is slightly improved on the quadratic. These results suggest both models they are very similar in goodness-of-fit and shape.

Even with this check, further study into the negative effect is needed. Controlling for global FidoNet adoption could be helpful, but this will require an improvement in our estimation procedure. Accessing the rest of the FidoNet data (the week level changes that we are currently unable to access) could also make quite a difference. The second order κ is also worth investigating with respect to other data: it may be that others have not considered this possibility in their estimation, and second-order effects occur in phenomena outside FidoNet as well. Vasi and Strang's (2009) results add some support to this hypothesis.

Finally and as motivation for the coming chapter: obtaining and processing data at the week level for Chapter 4 led to evidence that as the quantity of FidoNet's US sysops reached its peak in 1995, many US sysops were already leaving as shown in Figure 2.8. This implies that leaving events at the Exchange Area level had begun, and that contagion of the rise of leaving events at least in part account for the peak.

[25]Bayesian Information Criterion (BIC) is similar to AIC, except it has a higher penalty for a large N.

The rise of leaving events in the run up to the peak suggests that modelling the growth of a social system may require considering micro-level leaving events as well, even though the macro-level is only growing. Comparing this with the most standard social system contagion models—which only considers joining or adopting events—has significant implications for what is required for understanding the timing of a peak. The models most commonly employed may be appropriate in certain examples such as a well defined crowd that fully adopts a behaviour (Granovetter, 1978) or the classic hybrid corn case of full adoption of a community (Young, 2009), but the data shows that in FidoNet both leaving and joining events occur throughout the time series.

This has implications for the analysis presented in this chapter. In cases were only one sysop was active in an Exchange Area, these would count as leaving events in the unit of analysis in both this and the following chapter. It is plausible that what we observe as a lower likelihood of joining is actually affected by sysops leaving FidoNet. This might be true if the places with the highest density of sysops were also the places where FidoNet had been running their node(s) the longest times—and therefore sysops were more likely to leave due to burnout or simply a decline in interest.

3.7.1 Conclusions

While some of our results are puzzling, the core result that DMCA had a significant impact on the diffusion of FidoNet during its growth phase is well supported. As compared with a null model, the goodness-of-fit of our preferred model is very highly improved, with high statistical significance levels on both the DMCA variables. Additionally, the effect of previous adoption was highly significant and positive.

[26]These plots assume that at least three out of four aspects of each Nodelist entry—sysop name, city and region, BBS name and US telephone number—are the same. The noisy aspects of the leaving and joining events in this plot may in part be due to this metric, and a more sophisticated comparison per week may smooth these plots.

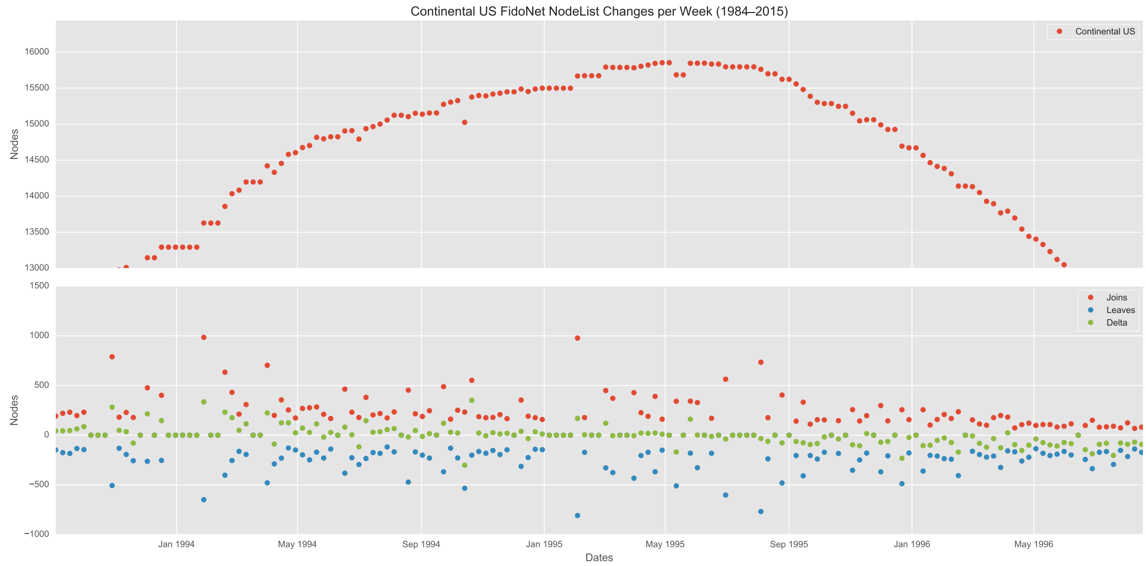


Figure 3.8: *Peak of FidoNet's growth when the leaving events began to outweigh the joining events (as demonstrated by the green Delta points which are the joins subtracted from the leaves of a week)*[26].

What sets our results apart are the second order effects, which imply that our DMCA is initially positive, but gradually reduces, eventually becoming negative. However, that transition occurs just as the growth phase is ending, and may be the result of other effects we cannot currently control for. We hope that further research will disentangle these effects.

Chapter 4

Decline

The number of FidoNet sysops peaked in 1995 with 15,854 in the US alone (April 28) and 39,827 worldwide (November 11). Estimates of then-active users in the US are in the millions.[1] However, within ten years FidoNet's US sysop population reduced to the low hundreds[2], a 98% reduction. Today it is a small yet resilient community, a tiny fraction of its previous success.

FidoNet is not unique in this regard; social systems of many kinds decline. Studied examples include trade union membership (Biggs, 2003), networks of firms from the New York garment industry (Saavedra et al., 2008), church attendance (Bruce, 2013), US civic organisations (Putnam, 2001) a variety of ancient civilisations (Turchin, 2003; Diamond, 2005; Axtell et al., 2002) and online social networks (Garcia et al., 2013). In each of these cases an aggregate measure of social activity—membership, number of firms, population etc.—reduces, largely monotonically, over time.

What drives decline processes and can a micro-mechanismal modelling approach apply to decline in the way it is applied to growth? Social system decline can—like in the previous chapter—be approached from a micro-mechanismal perspective, treating each decline event—events like a sysop leaving FidoNet or a garment firm filing for bankruptcy—as the result of micro-level interactions between agents. But with the exceptions of Axtell et al. (2002); Saavedra et al. (2008)

[1] Bush (1993) estimates that there were 'on the order of 2 million FidoNet users [who] read or write eNews', a popular newsletter about FidoNet.

[2] 309 as of June 3, 2005.

and to a lesser extent Turchin (2003) none of the above listed papers approach decline at the micro-level, focusing solely on the number of decline events in the aggregate, rather that *which* agent leaves when, and why that agent then.

The arguments in favour of a micro-approach apply equally well to decline as growth, and the FidoNet data provide an excellent opportunity to test a micro-level theory, especially as the data on its decline is—unlike growth—uninterrupted to the time of writing.

In this regard FidoNet is far from unique—particularly when compared with early online social networks such as Friendster and MySpace (Garcia et al., 2013)—and in a broad sense social systems of all sorts decline, often rapidly[3] and without recovery. Notable examples include Chicago's 1870's trade union movement (Biggs, 2003), the network of firms in the New York garment industry (Saavedra et al., 2008), British car manufacturing (Church, 1995), church attendance (Bruce, 2013), civic organisations in the US (Putnam, 2001), and numerous ancient civilisations (Turchin, 2003; Diamond, 2005). Taking a broader notion of decline, further examples include the spread of smoking cessation (Mercken et al., 2010) and divorce (Åberg, 2009).

These works are a testament to the prevalence of social system decline and its sociological significance, but with the exception of Saavedra et al. (2008), Mercken et al. (2010), Åberg (2009) and Garcia et al. (2013), few[4] consider or attempt to model the underlying micro-mechanisms that may explain the given instance of decline. Most only approach decline as macro effect driven by macro forces.

Drawing inspiration from the methodology employed in the previous chapter and the analytic arguments (Hedström and Bearman, 2009) for theories of the middle range,[5] I below aim to address this gap in the literature on decline. My hope is that over time a model framework will emerge with a well-tested set of micro-mechanisms that researches can employ to understand various decline processes,

[3]Rapidly with respect to growth and/or the duration of the system's existence.

[4]Putnam (2001) often refers to micro mechanisms but makes little attempt to directly test those claims, relying on broad demographic variables as proxies.

[5]Merton (1968) is famous for coining this phrase and advocating for this approach to sociology.

just as is the case with the growth literature. This chapter is meant to be an initial step towards that goal, focusing on FidoNet's specific circumstances while leaving the model simple enough to be applicable beyond FidoNet's idiosyncrasies.

To wit: I seek to address the following questions. How and why did such a large and quickly growing community contract as spectacularly as it expanded?[6] How can we quantitatively model FidoNet's decline, and can the growth models used in the previous chapter—drawn from the relatively well-developed literature on geographic social system growth—guide our approach? Turning from the macro (decline in general) to the micro: what makes an individual FidoNet sysop decide to shut down their node? Can deadoption events be explained by a spatially mediated contagion mechanism similar to the mechanism used in the previous chapter to model adoption?

Put simply: if someone I have a social tie to leaves FidoNet, does that increase my likelihood of leaving? Such a possibility would imply that social contagion mechanisms can apply to both growth and decline of the same system, and perhaps understanding the transition between those two phases of a system's dynamics involves some competing ratio between them.[7] That possibility will not be addressed here, but the possibility of modelling the entire life-cycle of a community with consistent mechanisms throughout the time-series is more theoretically sound than modelling growth and decline separately.

Assuming social contagion can explain the decline of FidoNet, does it behave symmetrically to growth or are there differences between contagion driving growth and contagion driving decline? For example: does an individual leaving a community only affect others in the short-term, or does deadoption contagion—like adoption contagion—continue to affect others long after the deadoption occurred?

Taking advantage of the uninterrupted time-series of data points covering the entirety of FidoNet's decline, coupled with demographic data used in the previous

[6] This question will be touched on but not directly addressed in this chapter. It will be explored further in the next.

[7] Mercken et al. (2010) demonstrate this in the spread of similar smoking behaviour.

chapter and survey data on FidoNet's most obvious competitor—the Internet—I test the hypothesis that FidoNet's decline was driven by a contagious, distance-mediated deadoption mechanism. I consider two possible time components to this mechanism: short-term: where after a given length of time a deadoption event no longer affects others in the system; and long-term where the event continues to impact other potential deadopters (current sysops) for the remainder of the time-series.

In doing so I develop a simple model of geographic community decline which draws inspiration from models of growth and tests how temporal duration may be a fundamental difference between adoption and deadoption contagion processes. My results suggest that two effects may be in operation at different time scales. In the short-term, deadoption appears to be contagious as hypothesised. However, in the long-term nearby deadoption appears to decrease the likelihood of further deadoption, and I hypothesise that that may be due to one or more of a few different effects. The significance of both versions and differing signs suggests that competing mechanisms may be operating at different time scales. Alongside, my results suggest that internet adoption strongly discouraged deadoption, implying that perhaps the contagion mechanism tested was in fact more of a driving factor in FidoNet's decline than the increasing popularity of its most obvious rival.

4.1 Literature on Decline

Social decline—be it in the broad sense of the decline of values and cohesion like in Tönnie's fear of the rise of *Gesellschaft*, or in a narrow quantitative sense of fewer and fewer firms in the New York garment industry —has been a common theme in sociology and anthropology. What is rare is quantitative modelling at a micro level, attempting to understand the individual events that contribute to overall decline, be they sysops leaving FidoNet or companies filing for bankruptcy.

Here I provide a short overview of work on decline at the macro level, initially summarising the history of work on broad notions of decline, and subsequently emphasising the quantitative, empirical work that seeks to model decline in a specific system. I demonstrate some of the problems with only considering macro effects, and conclude by covering a paper which looks at competition in the aggregate, something I will return to in my model later.

Next I cover the micro literature which more directly inspires the model I employ. There are very few papers that approach decline at the level of individual action, and the majority of these use ABM, an approach I will explore in chapter 5.

For the sake of simplicity, and because decline has different meanings in different contexts, I will use the term decline event to refer to an event where an agent in a system is removed from that system. That may be that agent's decision (like a sysop leaving FidoNet) or a consequence of a particular condition (a firm closing due to an unfavourable market). The term is meant to allow comparisons of micro decline events across different systems.

4.1.1 Macro Studies of Decline

The study of social decline in a broad sense motivated much of the pioneering work that led to the development of sociology as an academic discipline. From the early empirical work of French criminologists (Guerry, 2002; Quetelet, 1842) and their British (Rawson, 1839; Booth, 1889) and American (Park and Burgess, 1969) descendants, to more broad social theoretic notions like Durkheim's anomie (Durkheim, 1997) and social disorganisation theory (Shaw and McKay, 1972; Kornhauser, 1978; Bursik, 1988), sociologists and anthropologists have always been interested in the extent to which civilisations (Tainter, 1990; Diamond, 2005), countries (Putnam, 2001, 2007), cities (Sampson, 2012) or communities (Whyte, 2012) are thriving or decaying. They may disagree on theoretical or ideological grounds[8] but the sheer volume of such studies indicates the enduring importance of decline to sociologists.

This is especially clear given the motivation for the work just mentioned is broadly prescriptive, as in: how can decline be averted or mitigated? Durkheim's

[8] Whyte's (2012) work famously critiqued traditional studies that suggested crime necessarily constituted social disorganisation, demonstrating that the Italian community in North End Boston thrived through illegal organisations.

(1997) *Suicide* is devoted to understanding how decline in social integration leads to anomie, which in turn leads to higher levels of suicide. Putnam (2001) devotes two chapters to 'What Is to Be Done?' and Snodgrass (1976) writes of Shaw and McKay:

> *The aim was to 'stimulate' community organisation without engineering and controlling it and to 'spark' the latent potential for community control.*

Many of these works do not, however, directly approach the narrow question of membership decline—where decline events outnumber growth events over a sustained period of time, because they are concerned with the growth of an activity which is perceived to be negative (crime or suicide). The most relevant works to FidoNet are anthropological works on the decline of civilisations and those on decline in voluntary organisations.

Of the latter arguably the most significant work on membership decline of communities and voluntary organisations is Putnam's (2001), which tracks the decline of membership in US voluntary organisations from 1980 to 2000. Putnam seeks to explain a macro pattern of decline across a whole host of social organisations[9] which he believes are a primary source of aggregate American social capital.

Putnam's approach only considers exogenous causes, concluding that macro-level factors such as intergenerational change and the rise of television adoption explain the majority of the social capital decline he observes. He does consider some micro-effects—such as social network structures—and liberally refers to anecdotes of specific communities or organisations losing members, but these factors are presented as reacting to macro-level changes, rather than being directly causal themselves. Thus declines in 'schmoozing'—when individuals provide social brokerage across disparate parts of a social network[10]—are driven by changes in urban sprawl, and these in turn become indirect causes of declines in voluntary organisation membership.

[9] Examples include Bridge Clubs, Religious Organisations, the 4H Club.
[10] See Burt (2005) for an analysis of brokerage.

These complex causal chains are difficult to demonstrate empirically, especially with purely macro data. Putnam's quantitative results solely test macro effects to explain macro changes, and the remaining links in the causal chain are less concretely demonstrated. Arguably, Putnam risks a classic ecological fallacy by making substantive claims about micro-properties and behaviour almost exclusively on the basis of macro results.

What seems a missed opportunity in Putnam's work is the failure to quantitatively analyse the many individual cases he goes to great lengths to describe. These small systems could be used to model the actual process by which social ties decline and/or fail to be replenished over time. With enough data, such an approach could be interwoven with his macro models through a multi-level model.[11] Treating each bridge club or rotary chapter as a system to be studied would, if feasible, provide an excellent comparative study that could with far greater certainty demonstrate some of the more micro claims he makes.

Work on secularization (Stark and Iannaccone, 1994; Bruce, 2013) also considers macro decline, and again employ an exclusively macro approach. Bruce (2013) goes so far as to call his secularization theory a 'paradigm' rather than a set of mechanisms or hypotheses, drawing a complex web of causal relations across broad changes in western society, including inequality, education and technology. These are difficult theories to test beyond general correlations. They also provide little assistance in modelling individual decline events.

Biggs (2003) considers an interdependence mechanism in modelling how memberships in a trade union in Chicago spiked, then rapidly declined. He posits that a combination of interdependence—striking alongside others reduces individual fear of retaliation, greater expectation of success, and fulfills a moral obligation—coupled with inspiration—knowing others have striked in a similar condition—can generate a feedback mechanism that induces a quick and extensive wave of strikes. His explanation is specific and tractable and he supports his claim with a combination of qualitative and quantitative data. He is hindered by data limitations but is

[11] See Snijders and Bosker (2011) for approaches to multi-level analysis.

able make a plausible case and draw a link with later strikes in the American Civil Rights movement.

Again however, the only individual level data he employs is anecdotal, and he is unable to track individual-level decision-making in a systematic way. His model suggests a contagion-like mechanism in that interaction with other striking union members will tend to increase the probability of striking, and by extension interacting with other members who are not striking should discourage striking, potentially generating decline events after the peak. However, his primary focus is on growth, and does not make strong claims about the strikes subsiding.

Norton and Bass (1987) take the classic Bass (1969) model of product diffusion and generalise it to model price competition for a series of substitutable innovations. Their work demonstrates the importance of cost learning and how the market share of firms of varying sizes, market entry times, and cost advantages will converge as time $t \rightarrow$ inf. They demonstrate their theories with data from the US semiconductor market in the 1970s.

This last paper, and variations on this model (see (Rogers, 2003)) is the most relevant of the macro papers to modelling FidoNet's decline because it posits a quantitative model competition which will yield an overall decline of market share for the less successful company. The comparison is somewhat tenuous, since it is presented from the point of view of the producing companies as opposed to consumers, and is primarily analysing strategic market entry behaviour. However, it is a quantitative, aggregate level model of use of a product in exchange for a substitute.

Considering the possibility of cost advantages from the consumer's perspective, the relationship becomes even more relevant as the internet eventually provided a greater geographic reach and diversity of content than FidoNet. In the model I propose below, the internet is included as a potential technological competitor, and in fact the structure of the model, which largely mirrors that used in the previous chapter, is in effect Strang and Tuma's (1993) individual-level, geographic reorganisation of the basic Bass (1969) diffusion model.

4.1.2 Micro Studies of Decline

A smaller collection of papers model decline at the micro level. I take a broad view of what counts as decline, including smoking cessation and divorce in as much as they can be thought of as a ending of a behaviour (a decline event), because they employ models similar to my approach and because there is very little work on micro decline mechanisms.

Saavedra et al. (2008) model the decline of the New York garment industry over 19 years, tracking how firms die out, rewire their business transactions, and how new firms enter the system. They find that firms with low degrees (quantities of other firms they transact with) tend to die out more than firms with high degrees. They create a simulation of the system's decline which includes adding and removing firms, as well as link (transaction) addition and removal.

Unfortunately network information could not be taken into account in this thesis so their mechanisms are not easily applicable. Were we to include a measure of the position of a particular Exchange Area in the FidoNet network, then some of their mechanisms could be tested in our case, in particular the extent to which low degree sysops, or sysops with low centrality (meaning few shortest paths through the network include them) are more likely to leave than those with high degree and/or centrality.

Axtell et al. (2002) also uses a simulation approach to modelling the decline of the Anasazi inhabitants of Long House Valley in Colorado between 800 and 1300 AD. Their model is quite complex with a great variety of parameters that describe their agents (households) and the relative mechanismal dynamics. Some of these include age of household members, household location, crop yield, nutritional requirement and household fission age (when a daughter is married to another house).

Here decline occurs via two mechanisms. First, household members age and eventually die, and if all household members die that household is removed from the system. Second, households choose to leave the valley if they cannot find a location with a high enough crop yield to support themselves. Neither of these have especially useful analogues in the case of FidoNet's decline, other than perhaps my control variable for income, and the survival model's inherent inclusion of duration of involvement.

I will return to these simulation models in chapter 5 as their methods become more relevant there. These two papers are the only I am aware of that deal with a direct form of social system decline. Calling these systems social is somewhat tenuous as their units of analysis are not socialising as such, but are instead business transactions or marriages. But in both cases decline is of the sort I am modelling—where the number of agents is reducing over time—and the approach is micro: they model events at the agent level which depend on interactions with other agents.

If we consider decline in a broader sense two recent papers hold useful insights. First is Åberg (2009), which models the spread of divorce as a contagious process amongst coworkers in a random selection of firms in Sweden. Severing a marital tie is similar to a sysop leaving FidoNet (or at least shutting down their BBS) in as much as a particular social connection between individuals is ended. In both cases nothing prevents other social connections to remain between the actors—sysops may simply become users just as divorcees may stay friends.

The motivations are likely to be fundamentally different, as in the case of divorce potential alternative partners is quite significant, suggesting people are inclined to consider replacing a partner. I am unaware of a direct competitor to FidoNet in terms of replacing the social connections, though if we treat the internet as a competing technology then that could be thought of as replacement. This possibility is discussed in section 4.2.3.

More theoretically salient is the contagion mechanism, which Åberg presents as potentially operating in a variety of ways, including: a coworker divorcing triggers someone to consider divorce, or demonstrates the advantages, or becomes a friend who removes the danger of potential social isolation. I provide a similar list regarding sysops in section 4.2. Her results suggest that the hypothesised contagion effect is significant in the direction expected.

Mercken et al. (2010) consider influence and selection effects in smoking behaviour for adolescents in Finland: both adoption and cessation. They find both influence and selection to be significant, though the latter more so. Taking Åberg's (2009) approach to contagion following a variety of social pathways, I think it fair to consider their results as further evidence of a contagious process.

It is questionable whether smoking cessation is a form of social system decline, whereas with marriage at least some social tie is severed, even if the system is just a dyad or a nuclear family. While smoking is often a social activity, it is not exclusively social. By contrast it is difficult to imagine a use of FidoNet[12] or membership in a volunteer organisation which is not intrinsically social, and therefore leaving implies a severing of a tie (or at least a severing of a particular context of socialising). Tenuous though the comparison may be, their results demonstrate that the cessation of a behaviour driven can be driven by interaction with others who have also ceased (or the converse) and that even starting or ceasing decisions may be driven by desires to form new social ties.

The lack of detailed micro analysis of decline is surprising given the historical interest in and theoretical prominence of social decline in general, as well as the number of papers on growth and diffusion. Possible reasons for this include a preference or selection bias towards novelty, which means growing systems are often more prominent or attractive than declining ones. Perhaps data on declining systems tends to be more difficult to obtain, as joining or purchasing is more likely to be well defined and recorded, whereas decreases and cessation of use or leaving may be less so. Finally, of the few papers mentioned above, three employ highly novel and technically sophisticated quantitative methodologies modelling both growth and decline simultaneously. If such methods are especially difficult to use,[13] then that may be a further discouraging factor.

[12] Even users who only read messages and never post themselves are engaging in an asymmetric social activity.

[13] While this author strongly considered including growth during decline, he decided against it for the sake of tractability.

In sum: there is a small collection of work on quantitative modelling of decline, and an even smaller collection considering micro decline events. Of that handful, the most pertinent effects to a system like FidoNet are contagion (in various guises) and potentially competition.

4.2 Mechanisms

The two mechanisms I highlighted in the literature were contagion and competition, and these form the core of my diffusion model. The former is inspired by—but fundamentally different from—the mechanism studied in the previous chapter because of two ways in which time since the decline event (leaving in the case of sysops) can be considered. These two mechanisms—long-term and short-term contagion—will be discussed first, followed by internet competition.

The long-term contagion mechanism is a direct mirror of the growth contagion mechanism, again cumulative and mediated by distance. I call this the long-term contagion because past leaving events continue to be as influential in the future as they were in the past, irrespective of how it has been since the decline event.

The second effect I call short-term contagion, and this reflects the possibility that the effect of a sysops leaving FidoNet decreases in influence over time. This is because a sysop who has left is likely to have severed a main source of social interaction with the remaining sysops, and after some time has passed, the impact of their leaving will subside.

The last is competition from the internet: a technology that began to flourish as FidoNet declined. The internet—through the websites and social networks that operate on it—now provides many of the features that initially made FidoNet attractive, as well as a host of others FidoNet cannot at present provide and is unlikely to in the near future.

These mechanisms constitute my hypotheses for this chapter, which I will explain in further detail as I cover each one.

4.2.1 Long-Term Contagion

The various models cited in the previous chapter treat geographic contagion as a cumulative mechanism. As one individual or region adopts, they join the pool of actors thought to be broadcasting awareness and (potentially) the advantages of using the innovation or joining the diffusing social system. In some models it is assumed they continue to broadcast through to the saturation point and often that adoption is permanent (Ryan and Gross, 1943; Bass, 1969; Young, 2009). My previous chapter relaxes that second assumption to include leaving events during the growth phase, such that areas that leave FidoNet no longer are included in the cumulative contagion effect and can subsequently re-adopt.

However, even with this correction the model implies that any adopter or member who is current (has not deadopted) is continually broadcasting to potential adopters for as long as they are members. It is not the case that their effect on others changes over time, 'wearing off' after they have been involved for some time, spiking when they adopt followed by a leveling off, or an increase in intensity over time as one becomes more involved with the system or more experienced with the innovation.

It is this possibility of temporal fluctuation that will motivate the crux of my two contagion hypotheses. Since my growth model assumed a long-term time scale in effect[14] I call a model that is similarly long-term in scale 'symmetric' because it treats the growth and decline contagion mechanism as operating in the same way.

The mechanism, therefore, is that EAs leaving FidoNet will influence EAs that are still FidoNet members to leave as well. That influence effect decays with distance and—like in the growth process—is assumed to persist for the remainder of the decline phase. In that sense the process in cumulative, because leaving sysops join

[14] I do argue that there are temporally correlated changes in the directionality of the contagion effect, but those are mechanistically driven by adoption past a certain threshold and as a function of time. It is because the model is cumulative that that the effect on some EAs passes this threshold. Thus the temporal effect here is a function of accumulation that is correlated with time. The base underlying modelling assumption is that the individual influence of an adopted EA on a susceptible EA is constant over time.

the nearly-monotonically increasing pool of left EAs: an EA that left in 1998 still affects a remaining EA in 2000 or 2008 to the same extent as when the deaoption first occurred.

Equation (4.1) specifies how this effect is calculated.[15] Like the cumulative adoption mechanism from the growth model, we weight the effect of each deadopted EA on all EAs that still have sysops by the distance between them.

$$\kappa_{it}^{(p)} = \sum_{j} \frac{\gamma_{j,t-1}}{d_{ij}} \tag{4.1}$$

where κ_{it} is the cumulative adoption effect on EA i at time t, d_{ij} is the distance between the centroids of EAs i and j, vector in equation 4.5 and

$$\gamma_{j,t-1} = \begin{cases} 1 & \text{if } j \text{ has deadopted by } t-1 \\ 0 & \text{if } j \text{ has not deadopted by } t-1 \end{cases}. \tag{4.2}$$

As the set of deadopted EAs grows, so will the effect on all remaining FidoNet-using EAs. The effect from each deadopted EA will remain constant for the rest of the decline phase, weighted by the distance to the EAs still susceptible to deadoption. This is all in keeping with the mechanism tested in the previous chapter, and those tested in nearly every sociology paper considering a contagion mechanism to model diffusion.[16]

Those papers include, as mentioned above, work on deadoption such as Åberg (2009); Mercken et al. (2010). Both papers can be interpreted as modelling leaving or deadoption events, the former being divorce and the latter smoking. In both cases the contagion effect was deemed temporally invariant, though it should be noted that their data covers comparatively short time-scales and has data on personal ties rather than geographic proximity. To my knowledge only David Strang and his co-authors have approached the problem of temporal variation in contagion effects and this will be considered in the next section.

[15]This same equation can be found in the previous eq. (3.2).

[16]The two exceptions—(Strang and Tuma, 1993; Vasi and Strang, 2009)—are explored in more detail in the next section.

This background—coupled with wanting to compare to the growth process studied in the previous chapter and with other geographic diffusion models—is why I chose to explore this mechanism. There are, however, reasons to doubt its plausibility. With growth it is arguable that the time invariance assumption is reasonable: EAs which have adopted are presumed to contain active users of FidoNet, and that activity is assumed to continue to broadcast awareness and the benefits of FidoNet to other potential users, both locally and in other EAs. Intuitively, a lack of a FidoNet community would imply an absence of activity, and that absence may not broadcast the way a presence of activity tends to. In simpler terms: can a *lack* of activity produce a similar contagious effect indefinitely over time as a *presence* of activity?

Two interpretations—not mutually exclusive—are possible. First, that any individual who will have left will leave a 'social vacancy', which the remaining members of the system are constantly aware of.[17] An analogy: a close knit set of friends from school have a reunion every year, even after they are grown and have families. However, gradually over time people choose not to come to the reunion, and because of their close knit past, their absence continues to be felt long after they first left, and that absence induces other members of the group to consider not attending the next reunion.

This is amplified if the leavers continue to maintain social bonds with those who still attend *outside the context of the reunion*, and thus their reasons for no longer attending, and the fact that they have left, is frequently presented to remaining attendees, and in particular those attendees who are socially closest to those who have left. Taking the analogy to FidoNet: if the communities that were associated with an EA that deadopts maintain social bonds with communities that remain in FidoNet, they may provide a constant pressure to deadopt. Somewhat similar (though less plausibly) if the area which has deadopted no longer has significant

[17] Perhaps the closest similar idea is Burt's (2001) notion of structural holes, where dense cliques of social relationships are isolated, and structural holes are the absence of ties between these dense cliques.

social ties to remaining FidoNet communities, but when that area was still part of FidoNet it had very strong ties to some, its absence may, like a sort of ripped out part of a social fabric, continue to be felt long after their departure.

Another interpretation (which could work in tandem with the previous explanation) is that FidoNet leavers now actively discourage others from joining and encourage current members to leave. Again this would depend in part on the maintenence of social bonds after sysops (and users) have left the system with still active members, and their repeated interactions continually discouraging current members from continuing. In order for this interpretation to be consistent with this mechanism, it must be assumed to happen continue to happen at approximately the same level for each deadopter for the duration of the system's decline.

Perhaps related in both cases is the possibility that these effects would also undermine the possibility of new adoptions during the decline phase, and the sense in which a lack of replacement new users and sysops leads to decline of the system as a whole. These ideas will be explored in further detail in the next chapter, but again this could help undermine the user population, and it could continue to impact potential new users long into the future in as much as they could interact with old users who suggest FidoNet may not now be worth their time.

It is important to consider that EAs are the unit of analysis, and they correspond to one or more sysops who are likely to reflect an underlying community (or communities) of users. If that sysop is to feel her efforts are worthwhile, her local community may be a source of that motivation. Whatever prevents new potential users from replacing users who leave may eventually stagnate a local community, and induce the sysop to leave as well.

In sum: this long-term contagion mechanism is meant to test the possiblity that contagion effects in decline are consistent with the assumptions traditionally applied in the context of growth: namely that their effects persist from the event (in our case leaving FidoNet) for the duration of the model.

4.2.2 Short-Term Geographic Contagion

The other possibility is that deadoption events could decrease in contagiousness over time, and therefore their effect is (comparatively) short-term. The intuition here is that once a deadoption event is old enough, its effect on remaining active FidoNet members has declined to the point of negligibility, and thus it no longer has an effect. This is designed to account for arguably the most problematic component of the long-term mechanism under a decline process, and would apply to the contagiousness of leaving or deadopting but potentially not to (or to a lesser degree) joining or adopting.

Just as in the long-term mechanism, a sysop de-adopting is expected to have a greater impact on those they are socially closest to than those that are socially further away. Again taking the assumption of geographic proximity correlating with social closeness,[18] the effect is assumed to affect those geographically closer more than those geographically further. However, as time proceeds the effect of that specific leaving event eventually ceases to discernibly impact current FidoNet members, and thus that particular deadoption event is seen as no longer being contagious.

Under this regime, deadoption events eventually fade from memory, and it is the immediacy of the event in both time and space that is the primary concern. Taking a social group analogy again—someone leaving a tight-knit clique is like a shock to that system: those closest to the leaver may feel abandoned or rejected, and may choose to join them in leaving. This is especially likely if they wish to maintain ties with the leaver, and if leaving will make it easier for them to maintain those ties. If the ties they have with others in the system are less important to them, that may increase their likelihood of leaving as well.

However, as time continues, those left in the system either eventually forget the impact the leaver had, or come to terms with it such that the effect becomes negligible. *Ceteris paribus* the system reaches an equilibrium of sorts until the next

[18] Arguably our results from the previous chapter lend credence to this assumption.

leaving event occurs. If those events overlap in effect-time, then the effects of those events are cumulative in the way that long-term effects are assumed to be.

Applying the analogy to FidoNet: the time dimension is measures in weeks, which is how often Nodelists were published. I consider a variable l, which corresponds to effect-time as described above: how long a deadoption event continues to impact other remaining FidoNet members. p may be related to both the sort of close-knit ties explanation mentioned above, but it may take into account the time needed for leaving events to be communicated through FidoNet, as the system does not transfer information instantaneously.

Whether it is the absence of a sysop from the Nodelist which informs someone of a departure, or a farewell message targeted at specific friends, or indeed just their absence from discussions they would normally partake in via Echolists, it will take time for that information to propagate. Once the information is conveyed, that will affect those socially connected to the leaver for a limited amount of time, and either they will choose to depart as well or the effect will eventually become negligible, either because it has been forgotten or has been taken into account and ceases to demonstrably affect the dynamics of the system.

This mechanism is more plausible than the long-term version in the context of de-adoption because of assumptions about activity. In adoption, the contagiousness mechanism is assumed to be long-term because an adopter is an active member of that system.[19] Being an active (adopted) member of FidoNet implies participation. That activity is likely to be present in one's mind and come up in the course of conversation. That activity may be more visible in terms of someone's affinity (say having a large, dedicated computer server in one's home), and if someone becomes increasingly involved then their contagiousness for encouraging others to adopt may even increase over time.

This would imply a different form of long-term effect—increasing rather than

[19]This may be less true of certain technologies, where having purchased one does not imply constant use, and use may decay over time. That would suggest the contagiousness of certain types of products may decay in the form I suggest here.

decaying—but even if this effect is significant, approximating it with a long-term effect is at least more conceptually defensible than in modelling decline. This is because despite adopting some time ago—potentially years—that *current* activity should continue to be contagious. In general: we expect that contagiousness to still be significant and positive *irrespective of the time of adoption*, and thus the time-invariance assumption is plausible.

De-adoption—by-contrast—may be a worse approximation if time-variance is a significant factor. If the effect of de-adoption contagion decreases over time, potentially significantly, assuming a long time scale will become increasingly inappropriate, especially if the effect is *relatively systematic*. If most cases tend to be decaying, especially in the case of a long time-series, then the possibility should be taken into account. If they are not (especially likely in the case of a short time-series), then the approximation may be valid.

The crux of the argument therefore hinges on the systematic nature of the time-dependence. If the majority of de-adoption events decline in contagiousness over time, then we cannot rely on the long-term mechanism. This is because the effect is *systematic in direction*—reducing in strength—and the importance of this effect will increase with the length of the time-series. If the contagion effect is more randomly distributed in its time dependence—which I find more plausible in the case of adoption contagiousness because I expect some adopters will increase in contagiousness over time while others decrease or stay relatively constant—then the long-term mechanism should be less problematic. However, if the effect becomes especially strong with long time-series, this may also warrant consideration. Arguably, the non-linear results from the previous chapter support this hypothesis.

To my knowledge, the prior work on long-term mechanisms within contagion processes is largely connected with David Strang in two papers: Strang and Tuma (1993); Vasi and Strang (2009). In the former he presents his then novel, now highly influential cox-model approach to diffusion processes, which allowed individual level events to be modelled. He then presented the possibility that long-term time scales could be relaxed suggesting that events could reduce in contagiousness over time, potentially following an exponential decay.

He tested this hypothesis empirically using Coleman et al.'s (1966) much revisited[20] work on the diffusion of a medical innovation amongst prescribing doctors. He however found no evidence of such a decay. To date this appears to be the only paper to specifically attempt to empirically test the possibility of time-decay on contagion mechanisms at the level of individual events. While various papers—particularly within the network contagion literature—are heavily time-dependent and individual-level, and may imply time-decay in their methods,[21] none of these explicitly consider this possibility. By extension: none demonstrate it empirically.

The one empirical result I am aware of is Vasi and Strang (2009), which tracks the spread of municipal bills of rights in response to the USA PATRIOT act. There, again using the cox-model approach, a cumulative contagion mechanism, inversely weighted by distance, is modelled. This is again the diffusion of adoption rather than deadoption, but its geographic elements—both in the way the mechanism is modelled and the dataset being the United States—are highly relevant to this work.

However, rather than testing the effect of time on individual adoption events, the authors split the studied time-series into two periods. In the first their contagion effect is significant. However, in the second it loses significance, suggesting that the overall contagion effect decreased in influence over time. Their explanation is that as media coverage of the movement increased, the awareness provided by distance-mediated contagion was drowned out, leaving whatever persistent contagion mechanism less significant.

While this is indeed an empirical demonstration of a short-term effect, it is arguably far different than the individual-level effect postulated in Strang and Tuma (1993) because it is only considered at a macro-level. Further adoption events are seen as having a negligible effect on the system in the second-half of the time

[20]Papers that have reanalysed his work include Burt (1987); Marsden (1990); Van den Bulte and Lilien (2001).

[21]Examples include (Valente, 1996; Centola and Macy, 2007; Åberg, 2009; Mercken et al., 2010).

series, not because of *how long it has been since their adoption* but because of the overwhelming impact of an external force (media attention). For the duration of the first half of the time-series, a long-term effect is again assumed, and thus I argue that individual-level time-variance has yet to be demonstrated empirically within the sociology literature.

My intention is to test this hypothesis in the vein proposed in Strang and Tuma (1993): at the level of individual events. Equations 4.3 and 4.4 present this idea mathematically.

$$\kappa_{it} = \sum_j \frac{\gamma_{j,t-1}}{d_{ij}} \tag{4.3}$$

where again: κ_{it} is the cumulative adoption effect on EA i at time t, d_{ij} is the distance between the centroids of EAs i and j. However, it is in expression 4.4 that the short-term effect is specified.

$$\gamma_{j,t-1} = \begin{cases} 1 & \text{if } j \text{ deadopted between } t-l \text{ and } t-1 \\ 0 & \text{if } j \text{ did not deadopt between } t-l \text{ and } t-1 \end{cases}. \tag{4.4}$$

Like the long-term contagion mechanism there is still a cumulative effect, but that is limited by time period l which corresponds to the time until the deadoption no longer is included in the summation. As before the deadoption event affects the rest of the system from the week after it occurs up until p weeks have elapsed. Thus the effective contagiousness period is from $t-l$ to $t-1$.

I could have considered a decline effect as well, taking for example the exponential decay suggested in Strang and Tuma (1993). However, as this is the first paper to test this hypothesis, I feared a decay parameter would require an excessive estimation processes which would in turn prove difficult to interpret. In many ways this resembles the debate surrounding the distance decay function studied in the previous chapter, and were there more literature present on either system decline, especially geographically, or long-term contagiousness that might have been a worthwhile pursuit.

At present I thought it best to take the simplest approximation, which has the advantage of providing the clearest interpretability and comparison with the long-term hypothesis. My hope is the results below justify further, more sophisticated work on this element of contagiousness.

4.2.3 Internet Competition

An obvious potential reason for FidoNet's decline is competition from the internet: a technology that provided a more user-friendly, advanced alternative to FidoNet's services. US internet adoption probably eclipsed FidoNet's user population by 1995,[22] and today it is estimated 40% of the world is connected to the internet.

By 1997 the internet had a host of advantages (the year we begin our analysis of the decline phase). First and foremost was its ease of use. Internet Service Providers (ISPs) such as America Online and Comcast went to great lengths to ensure non-technical computer owners could connect with minimal knowledge and effort.[23]

By this point, the World Wide Web and web browsers—particularly Internet Explorer and Netscape (eventually renamed Firefox)—provided a graphical platform for users to engage with the quickly growing network of websites. Initially email and email clients provided the main source of online socialising (at least for non-technical users), but quickly instant messaging, chat rooms, blogs and eventually social media sites like MySpace, Facebook and Twitter rose in popularity, as have more specialised sites like LinkedIn and OkCupid, which focus on job opportunities and dating respectively. All of these have now far eclipsed estimates of the peak userbase of FidoNet, and (with the exception of MySpace) continue to grow (Garcia et al., 2013).

It is, therefore, obvious to see the internet as a competitor to FidoNet, particularly for less technically literate users. Both services provide a low-cost, electronic means of socially interacting with others, potentially over great distances. For those

[22]Based on Bush (1993) we might expect FidoNet's user population to peak around 10 million in 1995.

[23]An unofficial estimate suggests America Online spent upwards of $300 million sending out install disks in the mail to potential customers http://www.quora.com/How-much-did-it-cost-AOL-to-distribute-all-those-CDs-back-in-the-1990s (accessed 9th October 2014).

more interested in reading public discussions like those provided by the various FidoNet echomail topics—which ranged from philosophy and politics to computer programming and art[24]—the internet offers a much greater variety through websites—both personal and from mainstream news and intellectual institutions—not to mention full-color images, video and audio. FidoNet's technology dates back to an era when computers were only suitable for displaying monochrome text, and has not evolved to take advantage of new graphics, audio or video capabilities in computers or mobile phones.

However, many of these effects would apply to potential new users or sysops—less so to pre-existing FidoNet members. And while a lack of new recruits may have undermined FidoNet's growth, that would perhaps have lead more to a levelling off of growth rather than a sharp decline. To understand the internet as a direct competitor driving deadoption, we must consider what would actively motivate existing sysops and users to quit.

All the reasons mentioned above would certainly have weight, especially if FidoNet users and sysops tended to be more tech-savvy than average, wanting to have access to the latest and greatest capabilities computers can provide. For the early adopting crowd, FidoNet would perhaps appear old fashioned and crude—certainly by today's standards—and that may have driven this category of FidoNet members away.

Rogers (2003) would call this group 'innovators' or 'early adopters' in his schema of innovation adoption categories. He describes such people as 'risk taking' and 'venturesome,' willing to spend more to be the first to have something new and potentially unsuccessful. Rogers divides the remaining population into early and late majority and 'laggards'. These three are the remaining portions of the population, whose personalities are (in order I listed) increasingly more risk averse and/or suspicious of innovations, and generally are the last to adopt.

Considering FidoNet users and sysops who fall into these categories, they may have waited until the internet became more successful before switching. Leaving a

[24]In its heyday, FidoNet had vibrant ASCII and ANSI art communities.

community one is comfortable with and knows well, especially since it is highly self-selected,[25] for a new and far more mainstream alternative may have been less attractive, especially to FidoNet's equivalent of Roger's 'laggard' category.

Regardless of these possibilities, it is crucial to note that the internet and FidoNet were not mutually exclusive: any user or sysop could use both, and for some the internet was another means of connecting to FidoNet. As early as 1991, the internet and FidoNet were connected to save sysops money[26] and that connection has developed considerably since then. Today the internet provides an archive of most echomail[27] and any internet user can participate in most FidoNet conversations without the need of a BBS and a modem.

This fact perhaps would account for the greatest source of competition: replacement. Since the internet can provide all the services a sysop once did as well as its own features, why not just use the internet for both? Doing so completely removes the need for sysops: anyone with an internet connection can be a FidoNet user, and long-distance charges have been completely removed as an impediment (provided one has an internet service provider). No more congestion problems, and even far-reduced communication lag: normally a message would have to be routed through the various layers of the FidoNet routing hierarchy before reaching a local BBS—the internet bypasses that problem entirely.

A final point on this issue is technology incompatibility. As the internet grew and provided more sophisticated media, connection speeds increased as well. Telephony based 'dialup' modems were unable to keep up with speed demand and have largely been replaced with much faster 'broadband' modems such as cable and Digital Subscriber Line (DSL). These are however, incompatible with FidoNet BBSs, and so users without a dialup modem can only use FidoNet via the internet, and sysops must have a dialup modem to run a BBS. Dialup modems are increasingly rare:

[25]FidoNet's comparative lack of user-friendliness implies that the community was likely to be more computer literate than the average internet user, certainly by 1997.

[26]The first experiments were to reduce the cost of intercontinental communications between FidoNet Zone nodes (Bush, 1993).

[27]See `http://fidonet.ozzmosis.com/` (last accessed 9th October 2014)

only 2% of Americans had one as of 2013, down from 34% in 2000.[28]

Incompatibility extends to operating systems as well. FidoNet was originally developed to be used with text or terminal-based operating systems like DOS and OS/2, and as operating systems have advanced much of FidoNet's software have been rendered obsolete.[29] While FidoNet software does still work on Linux and the community continues to fix bugs, it seems no new features have been added for some time.[30] While these technological developments are not a simple resultant of the internet, they point to the increasing disparity between the internet and FidoNet as a functional and useful means of socially engaging via computers.

All these effects should undermine the set of users in need of local sysops to connect to and reduce their likelihood of becoming new sysops. In turn, sysops would likely see a downturn in demand for their telephony services, as well as an increasing difficulty of maintaining a node on modern hardware. If a sysop merely keeps an old computer running, with an older operating system and dialup modem, perhaps the decline in demand will allow them to continue running a BBS with lower costs, and the internet further reduced their costs by providing a free way of sending them the nodelist.

On balance it seems highly likely that the internet would have a major impact on the FidoNet sysop community. Specifically it is likely it in part drove FidoNet's decline as more and more members would have switched to using the internet for their FidoNet communication and/or to using the internet's myriad other socialising opportunities. Thus my first hypothesis is that the competition from the internet partially drove FidoNet's decline.

It is worth noting however, that if we were studying FidoNet's usage outside the US then other factors would be important to consider. In particular: at the time of this writing, Russia appears to have the most active FidoNet community, and

[28] `http://www.pewinternet.org/fact-sheets/broadband-technology-fact-sheet/` (last accessed 9th October 2014)

[29] For a list of FidoNet software, see `http://www.z1.fidonet.org/fidosoft.html`.

[30] See `http://sourceforge.net/p/makenl/wiki/Home/` (last accessed 9th October 2014) for recent changes to the main Nodelist compiler.

speculation suggests that this is partly due to the lack of censorship and surveillance of FidoNet relative to the internet (Sadofsky, 2005) in that part of the world. To my knowledge this does not apply to the US.

A final point regarding the question of endogeneity: the internet's development may well be intertwined with FidoNet's, and if I could obtain local internet adoption at the EA or sysop level it would be an excellent way of robustly testing that hypothesis. Sadly data limitations in terms of quality and comparability restrict us to accounting only for regional—Northeast, South, Mid-West and West—variations, at semi-biannual time resolution. This forces me to treat the internet as an exogenous effect, and the remained of this section will explore the advantages and disadvantages of doing so.

It is reasonable to expect areas of heavy FidoNet use may have been among the first to adopt the internet since they were likely to have tech-literate populations and already own the necessary hardware needed to connect to early ISPs. Further: if users realised they could use FidoNet just from the internet, then that may have driven them to switch to the internet as well, especially if their nearest sysop was a long-distance call away. Finally, there is some evidence to suggest that FidoNet BBSs may have actually been some of the first local ISPs in the US. If we assume the internet was a significant driver of FidoNet decline, but FidoNet was a significant amplifier in early internet adoption, there is a problem of causal circularity.

To my knowledge there is no publicly available dataset that with enough spatial or time resolution to test this possibility, and in particular no good data exists on US internet adoption prior to 1997. It is for that reason that our decline analysis begins in that year. By that point US nationwide internet adoption was at 22.6% and FidoNet was already two years into its decline by that point.

While ideally I would test this possibility, I think it reasonable to treat internet adoption as exogenous under the circumstances. FidoNet's impact on internet adoption was probably strongest in the first half of the 1990s when internet awareness was still nascent. In 1993, having a modem and computer would probably have increased the likelihood of trying and using the internet substantially. By 1997 I expect that effect will have lessened somewhat, while general public awareness of the internet—which in 1993 would likely have been more prevalent among FidoNet members than the rest of the population—will have expanded considerably.

In essence my argument is similar to the results from Vasi and Strang (2009), which suggest that for the first half of the diffusion of municipal bills of rights, geographically mediated contagion was significant but as their process gained more widespread attention, the local effect lost significance. While I cannot test this claim, it is the best approach I am aware of under the data constraints available. This issue will be returned to again in the results section.

This leads us to the other mechanisms I test: geographically mediated contagion processes. Crucially the question—as explored theoretically in §4.1.2—is whether contagion can apply in the de-adoption of an innovation or social system, and if so whether it applies in a way symmetric to the diffusion of the system or whether some fundamental modification is required—in our case temporal dependence. I will begin with the symmetric approach—which I will describe as cumulative contagion—then proceed to the asymmetric form, which I describe as short-term contagion.

To summarise: I propose three hypotheses. First that competition from the Internet was a significant source of FidoNet's decline. Second and third that over and above that competition effect, decline was contagious in either a long-term or short-term manner.

4.3 Sysops Leaving FidoNet

For analysis purposes, a leaving event occurs when a sysop is present in the Nodelist at week $(t-1)$, then absent the following week (t). A sysop leaves by informing their local Nodelist maintainer—a sysop who keeps track of changes to the Nodelist for their local region—that they are closing down their BBS and it is removed from the list.[31]

[31] Nodelist maintainers may also periodically check on sysops under their jurisdiction and if after a few weeks they appear to no longer be operational they may also remove them from the Nodelist.

This is akin to someone deleting their Facebook account or stepping down from running a book club: they no longer interact socially with the members of that group *in the capacity they once did*. They may however continue to interact with others in the system through others ways: a sysop may become just another FidoNet user or maintain ties with FidoNet members via the internet or face-to-face interaction. What is important is that they will no longer do so *via* the position of being a sysop.

That implies three things: first that they no longer provide a service to their local users. No longer can someone call their computer to connect to FidoNet and receive mail or read echolists[32]. If other BBSs are still operational in their local area, users will still be able to connect to FidoNet for without paying long-distance charges. However, the demand on the remaining BBSs in that area will now be spread across a smaller number of nodes and thus reducing the available service *ceteris paribus*. If there are few users in this area, this impact will be negligible, but for an area with a dense user population this may undermine available connectivity.[33]

Second: from a purely logistical perspective they will have less direct communication with the FidoNet users who used their BBS as their connection point. When a user connects to a FidoNet BBS, the sysop who operates it may chat to them directly via text in real time (much like modern chatrooms). As a result, a sysop leaving FidoNet is likely to reduce their interaction with other users *ceteris paribus* because users could not easily have that same interaction with each other[34] and the casual frequency with which that was likely to occur is significantly different from calling someone directly.

Third and final is the social status of being a sysop. This derives from the responsibility of running a BBS, the publicity that may come from having one's name on a Nodelist and potentially signing messages with one's BBS name, and

`http://www.fidonet.org/policy4.txt` (last accessed 9th October 2014)

[32]Echolists were the equivalent of modern online forms

[33]Most BBSs only allow one user to connect at a time. At peak usage hours, areas with high user density and few sysops may not be able to provide a connection for all users wishing to use FidoNet.

[34]It was possible for one user to call another directly via computer or just as a normal phone conversation.

the respect and appreciation—or pressure and criticism—that may accompany that position. Sysops are the volunteers that make FidoNet operational, and the higher their level in the routing hierarchy, the more responsibility they have and ostensibly the greater power and respect. Their technical ability may be assumed greater than the average user, and they may enjoy greater admiration and clout on public echolists.

However, the responsibility of being a sysop may attract criticism from unhappy users who are frustrated with their connectivity, be it from overcrowding, software bugs or user error. A local sysop essentially becomes tech support for their user community, and depending on their position in the routing hierarchy, may be responsible for other sysops and maintaining their part of the nodelist. These responsibilities were at times coupled with political conflict—some quite brutal[35]—and certainly these may have motivated some sysops to leave.

These motivating factors stem from the specific circumstances of being a sysop, but of course other factors wholly separate from FidoNet could apply. Family, employment and illness are just three of the wide variety of sources that could encourage a sysop to step down. Similarly, moving to a new location even if a sysop then set up a new BBS at their new location, would appear as a leaving event to their former community and within the analysis below.

I include demographic controls to account for systematic geographic effects relating to age distribution, relevant industries and education (as per the previous chapter). The various other factors listed here, if largely random with respect to space and time, should have little to no correlation with the behaviour of others in the system *at the time of their departure*. Given the size and detail of our dataset, it is hoped that those random events should be drowned out by systemic factors.

In particular, a sysop's leaving may influence others to do the same in the weeks after the departure.[36] That, at least, is two of three hypotheses I test below. However,

[35] FidoNet earned the nickname 'Fight-o-Net' for its infighting (Sadofsky, 2005).
[36] Nodelist time resolution is in weeks.

first I will explore the implications of analysing leaving events at the Exchange Area level, rather than the individual sysop level.

4.3.1 Exchange Areas

As per the previous chapter, my unit of analysis is not sysops but Exchange Areas (EAs): combinations of telephone exchanges and US census tracts. These allow me to spatially aggregate de-adoption events into geographically coherent regions that I can map to us landline numbers, thereby mapping the geographic dynamics of FidoNet's decline. Additionally, I can control for salient demographic features provided from the 1990 US census.[37]

This aggregation process has non-trivial implications for the dynamics of the system and the modelling process in general. I will not repeat the reasons given in the previous chapter but highlight those that relate specifically to decline.

Perhaps most crucial is the possibility of EAs having more than one sysop at a point in time, and the extent to which that keeps an EA within FidoNet under the scheme I have chosen longer than otherwise. As mentioned in the previous chapter, my motivation for this decision—beyond its data convenience—is the extent to which EAs resemble regional communities of FidoNet users and sysops. If sysop A leaves, but sysops B and C are still within that local call area, the community survives albeit with less service capacity. Sysop A may well switch to being a FidoNet user, still engaging socially but without the responsibility and costs being a sysop entails.

The advantage of telephone exchanges in this context are they correspond to areas of unlimited call service: beyond a normal monthly phone bill, calls are unlimited. Outside that area, long-distance charges apply. The ideal then is for there to be at least one sysop per local call area, as that means any user can connect without paying long-distance charges[38].

[37] Census tracts were redrawn for the 2000 census, and reconstructing EAs for the newer census would ruin our ability to compare our geographic results with the previous chapter

[38] This was in fact the major motivation for FidoNet's creation as detailed in the previous chapter

Having more than one sysop for a local call area then tends to provide alternative connection points for the same community; in other words greater connection capacity. If more users wish to use the service simultaneously, then more sysops are needed. As FidoNet grew, it is likely that greater sysop densities arose from capacity constraints and users wishing to have more access, some of whom may have decided to become sysops partly to increase their own access.

However, as FidoNet declined, demand for capacity would also have declined. Internet availability—which was an alternative means of connecting to FidoNet—would also have driven down demand as users could still connect via their internet provider, irrespective of their local sysop population. Internet effects will be discussed in greater detail in the next section, but what is important is the possibility that as the system declined, 'excess' capacity may have been the first to leave, pruning high density EA communities down to one or two sysops who would keep the local community alive and the service still available.

In keeping with the method of the previous chapter, we ignore this additional variation, treating a de-adoption event as when the last sysop in an EA leaves the nodelist. Doing so masks churn due to demand variation, which is akin to waiting for a local rotary club finally closing its doors: it is unlikely that anyone can now use FidoNet in that region in the traditional phone-to-phone way without paying long-distance charges.[39] This may in fact induce a kind of 'stickiness' for the last sysop: if there is a small but resilient local community, that last sysop may keep the service running with minimal effort, and only when there really are no more local users would the sysop finally consider pulling the plug.

Finally it is worth acknowledging that factors that affect users—and in particular change how many there are—are salient to the behaviour and potential leaving of sysops as well. Reductions in users will both reduce the possible social interactions for the remaining users and sysops and reduce the demand on sysops for FidoNet access. Less demand may ease the responsibilities of sysops to a point, but beyond

[39]The one possible exception is extremely large metropolitan areas where population is so dense that they have multiple EAs. These are few (New York has 4)

a certain level a sysop may feel it not worth continuing, especially if maintaining their node is costly. Further: less users mean less potential new sysops. Since our analysis is at the EA level, that means fewer replacement opportunities for a flagging community. It is for this reason that I freely discuss the impact of various factors on sysops and users in the sections below.

4.4 Data

Like the previous chapter I consider adoption from the perspective of Exchange Areas (EAs): geographic aggregates of 1990 US Census tracts and US telephone exchanges. The method by which these were constructed is in the previous chapter and has not changed here. Similarly, I use a time-series of FidoNet Nodelists which are week-level in resolution. What differs is our incorporation of internet adoption data at a regional level to test internet competition, the completeness of our Nodelist time-series, the contagion mechanisms tested and the fact that the susceptible population has changed to EAs with currently active FidoNet sysops, rather than EAs currently without sysops. I will summarise these in turn.

4.4.1 Exchange Areas

These have not changed from the previous chapter. My reasons for doing so are largely a question of comparability and time-constraints. The 2000 census redraws census tract lines significantly, so incorporating that data would require recalculating the EA borders significantly. The 2000 census also changes the questions asked, particularly in terms of occupational categories, and thus the controls would differ not only spatially but substantively.

Were it the case that internet data were available in a time series dependent on the geographic boundaries of the 2000 census then redrawing these may have been worth the time. However, as I detail below, internet data is unfortunately only available in variable spatial resolution, and because the questions changed over time, their results are not easily comparable.

To most directly compare our long-term and short-term effects with the related mechanism of adoption, keeping these demographic controls and level of analysis constant is more advantageous than investing the extra time required in recreating the unit of analysis.

That said, there are significant downsides: if US demographics, particularly with respect to age distribution and employment categories radically changed between 1990 and 2010, then our geographic controls may be misleading and undermining the quality of the model, potentially causing spurious geographic effects. Those possibilities will be discussed with respect to the results below.

4.4.2 Internet Data

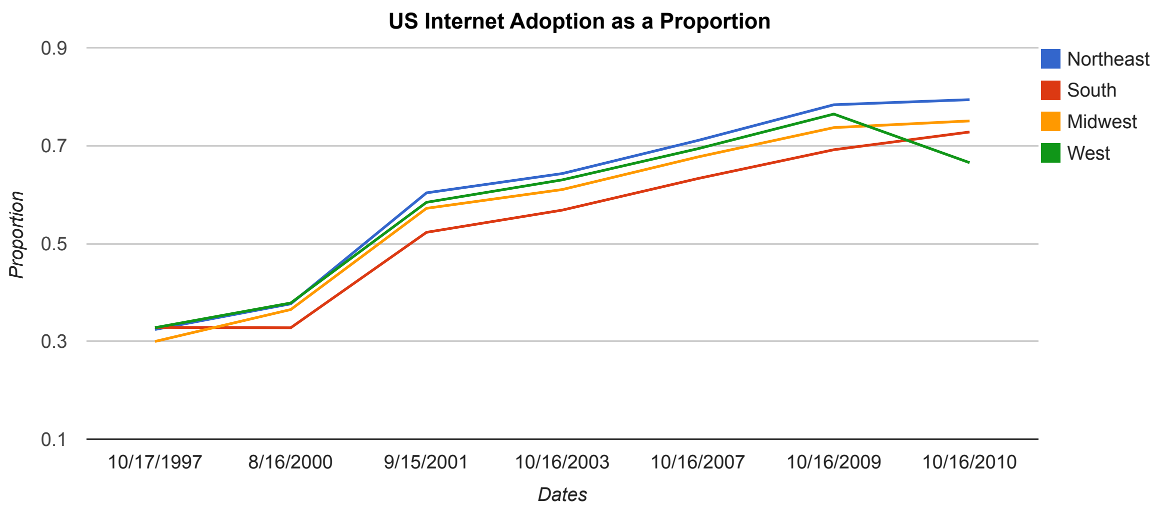


Figure 4.1: *Regional Internet Adoption*

Internet data was compiled from data also provided through the US Census via the Current Population Survey (CPS), which at varying time intervals, tracked US home internet usage from October 1 1997. While the dataset includes state-level data, unfortunately the nature of the questions asked varied from year to year, and the only consistently asked questions from August 2000 on were at the regional level, meaning Northeast, South, Midwest and West.

The question was: 'Was the internet used at home'? I also included the 1997 year which used the question: 'Who in the household uses the Internet at home?', and included any households which had an affirmative in any category as internet users. I then converted these values into proportions in keeping with our other demographic variables and interpolated both linear and logistic time-series in between each observed time point to generate the data used. These were then applied to the relevant EAs in the cox-model.

It is unfortunate that higher quality time and spatial internet data appear unavailable. Were it possible to consider adoption at a level comparable to our other demographic variables it might have been possible to consider the competition effect with greater accuracy.

Additionally, the way the internet is used has changed significantly over time, and since the question we felt best addressed our interest was household based, we were unable to consider new devices such as mobile phones. This is thought to explain the decline in internet use in the west exhibited in figure 4.1, where the technology savvy west-coast is perhaps using mobile devices more than home devices (File, 2013). As such this effect might have been misleading for this last year, so we conclude our analysis in 2009.

4.4.3 Nodelists

As before we use FidoNet Nodelists as our dataset, and how they are processed remains the same.[40] However, what differentiates this analysis from the previous is that all Nodelists, without interruption, are available from 1992 to the present day. Subject to human error in Nodelist creation, that entails a complete, week-level time-series which, in the case of decline, begins in 1995 and covers nearly 19 years (if we include part of the present year).

However, due to the limitations of internet adoption data, which are necessary to test one of our key substantive hypotheses, we have limited our study to begin on October 1 1997: the day of our first internet adoption data point. I conclude October 16, 2009, the Friday closest to the last internet survey data-point before the

[40]See previous chapter for details.

Western region downturn. That corresponds to 12 years worth of uninterrupted week-level data.

4.5 Method

Our method largely resembles that of the previous chapter, with the exception of no longer needing to account for variable time-intervals. Thus a normal cox-model can be used without modification. As per the previous chapter, the hazard-rate is the probability of deadoption, which is assumed to be a function of the included demographic controls and mechanisms we test for.

4.5.1 Controls

The demographic controls are those used from the previous chapter with a few slight moderations, again to retain comparability. Those modifications are intended to make the model easier to interpret and more robust. Their interpretations are different, however, as their impact on leaving is likely to differ from their impact on adoption.

Population density and urbanity should again relate to the contagiousness of the area, both likely to increase it as the probability of interaction is likely to be higher. Population was log transformed as in the previous chapter to account for skew, and again the relatively low correlation between the two variables should be helpful in distinguishing density effects and urbanity effects. What shifts here is that we have used made the population from 16 over as that better aligns with the ages included in the occupation question (which are 16 and up).

Urban areas may also increase the probability of de-adoption because—following the hierarchical innovation diffusion hypothesis (Huang and Gould, 1974; Fischer, 1978; Sugiura, 1986)—innovations like the internet and similar may diffuse there first, and by extension they may be the first to deadopt. I may, following Rogers (2003), expect a higher density of innovators here and potentially a lower density outside urban areas.

Median income—again divided by $10,000 to keep the range similar to other variables—should again be potentially correlated with Roger's notion of innovators and early adopters, thus leading to potentially quicker FidoNet leaving. If however the costs of running a BBS are prohibitive, these sysops may actually last longer as the costs may impact them less.

Age is once again stratified into groupings, however I changed these to decades rather than half-decades to make them less fine-grained. I hope this reduces sensitivity to potential variations in age distribution across the 19 years since the US census. The motivations for these categories are grounded in life changes: 16–19 covers high school, 20–29 early adulthood, career establishment, marriages and family creation; 30–39 mid-career, higher income and life-stability, family expansion; 40–49 later-career, still higher-income or income plateau, children nearing young-adulthood; 50–59 children leaving the home and approaching retirement; 60 and above retirement.

It is important to consider that since the population here is of areas that have used FidoNet affecting other areas using FidoNet, the filtering effect should produce non-trivial results. We must also consider the extent to which each category used FidoNet, and potentially *for those in these categories that did*, however unlikely, what is the implication?

The first of these categories—16–19—is far less likely to ever use FidoNet from 1997 as the internet was quickly adopted among young adults, and assuming sysops were initially users they would have to be using FidoNet from a relatively young age to be a sysop after FidoNet's peak. Computer hardware will have come down in price and availability, and if internet connectivity made keeping a BBS running cheaper, perhaps that would help. However on balance it seems very unlikely that this would be a significant predictor and if so a positive one: maintaining a BBS may have been difficult for a student in the late 90s and 2000s just as it would have been for the growth phase, and increasingly young students are less likely to have seen FidoNet during it's dialup glory days. It is much more likely that if they used FidoNet, they used it via the internet.

The 20–29 group would also have had to be relatively young FidoNet users to have used it in a dialup capacity, but less so. This group would again have a problem of relative instability in early career and family formation. If they were young sysops, they might have left FidoNet as they went on to higher education or moved jobs. Those late in the time-series would again have missed FidoNet's dialup days and tend to be internet FidoNet users. I expect this group to potentially be more significant than the previous but more likely positive (increasing the likelihood of deadoption) than negative.

The 30–39 group is the reference category, and their likelihood of involvement would perhaps be more resilient, having likely seen the end of FidoNet's growth and reached a point of relative prosperity and security during the decline phase. They are considerably more likely to have used FidoNet via dialup and thus keep the old system alive. The 40–49 and 50–59 groups should be much the same, only more prosperous and stable, and even more significant. In all three cases I would expect a negative effect (discouraging deadoption) increasing in strength with age.

The 60 and over group however, is possibly too old for FidoNet, having been in their late-40s or 50's during FidoNet's growth and possibly missing the first bandwagon.[41] Those who did use FidoNet are perhaps likely to keep things going, especially if it provides a low-intensity form of social stimulation with low financial and logistical cost, especially when coupled with increased disposable time. I expect this age group to have a less significant effect than the previous group, possibly weakly negative.

Educational attainment may again prove significant, and rather than taking the non-highschool graduate reference category I chose highschool graduates as the reference category, which I expect will be easier to interpret. As non-highschool graduates seemed a comparatively positive predictor for growth, I would expect them to decrease deadoption likelihood here, strange as the result was last time. Perhaps FidoNet was a community that welcomed those outside the traditional

[41] See the previous chapter's results.

employment ladder, and that effect would potentially continue to be strong as the rest of the community left to join the mainstream.

Highschool grads, our new reference category, in keeping with the previous chapter's results, may be less FidoNet keen, and as that population ages may again be drawn to the internet. Still, I expect this category to be negligible in strength in keeping with the previous chapter.

Those who have finished an undergraduate course may be one of the strongest predictors of FidoNet resilience again based on previous chapter's results. This would again imply a negative coefficient. My expectation here is not that these are current undergrads but those who would have graduated some time ago and are now in the 30–50 year age range but had the right knowhow back when FidoNet was in its heyday. Their increased income expectation would translate into the security necessary to keep a BBS running.

Postgraduates were expected to be Usenet users in the previous chapter, and in that sense they may prove negligible in strength here. I expect any keen postgrads who used FidoNet in its heyday may again have the right life circumstances and knowhow to continue running a BBS for some time, but I expect the effect to be negligible.

Occupations should again have a different implications. Managers I expect to still be insignificant, though potentially mildly negative (discouraging deadoption) due to their income and security. Teachers I expect would similarly be negligible as the schools hypothesis proved largely unsupported.

Areas where a greater proportion of the population were scientists I would expect to accelerate decline just as these areas discouraged growth. Though: scientist sysops may again have the security and knowledge required to keep a BBS running into the 2000s. As many scientists may have used large computers which require ease with a terminal interface, perhaps they would have been more comfortable with FidoNet's increasingly antiquated user interface. Still given the strength during decline is probably unlikely that the effect is too strong as there were probably few FidoNet scientists.

Engineers and technicians, arguably the strongest positive categories in growth, I would expect to be again strong and now negative (discouraging deadoption). Having been such a strong category during growth, I expect a great deal of social similarity applies to those in FidoNet that fit these categories and given the likelihood of their large population, I expect they would have a relatively strong community even during decline, which would discourage leaving.

These demographic controls are motivated by individual-level hypotheses. There are also potential spillovers of individual effects onto social ties, or correlated industrial characteristics that may also affect FidoNet leaving. These different processes are lumped together in the coefficients for the proportion of the population in different groups, so care will be needed in the interpretation of these results.

4.5.2 Statistics

Freed from the difficulty of varying time-intervals, I was able to use a traditional cox proportional-hazards model without modification. As stated previously the risk set at time t is the set of currently adopted EAs, and the set of explanatory variables are the demographic factors just detailed, the internet adoption time-series, and the two contagion effects.

A cox model follows equation 4.5, where $\Lambda(t|\mathbf{X}_{it})$ is the hazard rate of EA i at time t, Λ_0 is the baseline hazard and β is the vector of estimated coefficients for the effects from $\mathbf{X}_{it}$. As before t is an index for the vector of explanatory variables $\mathbf{X}_{it}$ because some of these variables change with time (specifically those corresponding to our hypotheses).

$$\Lambda(t|\mathbf{X}_{it}) = \Lambda_0(t)\exp(\beta\mathbf{X}_{it}) \tag{4.5}$$

The two contagion effects are calculated via the equations specified in their respective sections above. The long-term calculation is relatively trivial and is in keeping with that from the previous chapter. However, the short-term effect requires an additional level of complexity.

To adequately account for the various lengths of l, the early parts of the time series must be taken into account. For example, at the start of the time-series modelled in the cox-model, an l of 4 would imply including adoption events that occurred within the previous month to potential adopters in that week. As l increases, earlier and earlier deadoption events must be considered in order to accurately cumulate the time-effective span of each.

The implications of this are discussed after the results section below.

4.6 Results

Section 4.6 details the results for the base demographic model and internet adoption. I will begin with the controls model as most of these effects are statistically weak. The only significant controls are population density—which is highly significant and negative, meaning discouraging deadoption—and the age groups 50 and over, which are mildly significant and strongly positive, indicating the prevalence of those age groups encourage deadoption.

However, including internet adoption changes the model radically. Population and the 50–59 category lose their significance, while all education categories, income, urban, and all job categories save technicians become significant, teachers and income being the most. I will return to the substantive implications of these results and move on to the most surprising variable of all: internet adoption.

It is first worth pointing out the relatively extreme improvement of goodness-of-fit as measured by the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). It is generally thought that an improvement of 10 or greater is significant. Here we have an improvement of 6075.49 (AIC) and 6064.565 (BIC). The internet adoption variable—here using the linear interpolation between time points[42]—is highly statistically significant and has a large coefficient.

However, what continues to surprise and puzzle is the sign: negative. That implies that internet adoption strongly discouraged FidoNet deadoption, soundly

[42]The logistic interpolation had a worse fit. This can be found in the appendix.

	Demographic Controls	Internet Adoption	Long-Term Contagion
pop16in1000	−0.009***	−0.001	−0.001
	(0.002)	(0.002)	(0.002)
UrbanCombined	0.009	−0.268*	−0.077
	(0.165)	(0.130)	(0.139)
inc10000	0.005	0.192***	0.340***
	(0.076)	(0.061)	(0.067)
Age16to19	0.757	1.074	1.718
	(2.261)	(1.688)	(1.834)
Age20to29	2.400	1.416	3.015*
	(2.192)	(1.670)	(1.738)
Age40to49	2.537	0.321	−1.224
	(3.201)	(2.426)	(2.520)
Age50to59	5.945*	1.636	3.069
	(2.633)	(2.032)	(2.068)
AgeOver59	2.864*	3.391**	3.101**
	(1.536)	(1.190)	(1.270)
BelowHighSchool	1.234	−2.108**	−3.453***
	(0.985)	(0.747)	(0.765)
Undergrad	1.528	−1.791**	−4.815***
	(0.949)	(0.732)	(0.732)
GraduateDegree	−1.118	−2.886*	−0.598
	(1.706)	(1.386)	(1.417)
Managers	−0.176	−6.865**	−3.583
	(3.696)	(2.811)	(2.912)
Engineers	−14.635	29.472**	20.912*
	(14.589)	(10.685)	(10.367)
Scientists	−44.916	−63.676*	−51.949*
	(34.993)	(28.652)	(26.352)
Teachers	3.529	37.755***	32.577***
	(15.328)	(9.866)	(10.036)
Technicians	31.940	−4.199	23.814
	(26.283)	(20.966)	(19.541)
Internet		−36.395***	−36.060***
		(0.589)	(0.708)
Long-Term Contagion			−0.337***
			(0.022)
Log-likelihood	−35712.953	−32674.208	−32387.673
AIC	71457.906	65382.416	64811.346
BIC	71632.702	65568.137	65007.991
N	410333	410333	410333

Table 4.1: *Deadoption baseline model (1997-2009) comparison with Internet Competition and Long-Term Contagion*

rejecting my hypothesis. I have consulted others with respect to my methodology[43] and I will return to this result in the discussion section.

Moving on to the long-term contagion mechanism, we have yet another highly significant improvement in AIC (571.07) and BIC (560.146), and a highly statistically significant effect on the hypothesised mechanism, but again the sign is opposite to that which was expected. This leaves this hypothesis unclear in its interpretation, and this will be returned to in the discussion section. Note as well that while some of the control variables shift in significance though none in sign, internet adoption is unaffected by the inclusion of long-term contagion, both in magnitude, sign and significance.

We turn finally to the short-term results presented in table 4.2. These have two interesting properties: first, as l increases the goodness of fit mildly improves; second, past an l of two weeks the sign on the contagion effect switches from negative to positive and becomes strongly significant. The model thus coheres to expectation in terms of sign and significance, leading cautious support to the hypothesis. However, the goodness of fit is considerably worse as compared with the long-term mechanism. These results are discussed further below.

4.7 Discussion

The results are somewhat puzzling, particularly the internet effect. I will respond first to the demographic factors, then the internet effect, and finally the contagion results.

4.7.1 Demographic Controls

The demographic effects with statistical significance have coefficients which are relatively large as compared with the growth model. This perhaps has to do with the pre-selection of EAs which already contain FidoNet members, which may amplify

[43] Dr. David Barron graciously met with me multiple times to discuss these results and could find nothing at fault.

	1 Week	2 Weeks	3 Weeks	4 Weeks
pop16in1000	−0.001	−0.001	−0.001	−0.001
	(0.002)	(0.002)	(0.002)	(0.002)
UrbanCombined	−0.267*	−0.267*	−0.278*	−0.285*
	(0.130)	(0.130)	(0.129)	(0.129)
inc10000	0.193***	0.193***	0.182**	0.177**
	(0.061)	(0.061)	(0.061)	(0.061)
Age16to19	1.088	1.085	1.004	0.974
	(1.689)	(1.690)	(1.681)	(1.677)
Age20to29	1.424	1.423	1.360	1.330
	(1.671)	(1.671)	(1.665)	(1.662)
Age40to49	0.312	0.314	0.394	0.441
	(2.427)	(2.427)	(2.417)	(2.413)
Age50to59	1.638	1.637	1.618	1.604
	(2.033)	(2.033)	(2.025)	(2.021)
AgeOver59	3.396**	3.395**	3.370**	3.365**
	(1.191)	(1.191)	(1.184)	(1.182)
BelowHighSchool	−2.124**	−2.121**	−2.019**	−1.980**
	(0.747)	(0.748)	(0.745)	(0.745)
Undergrad	−1.816**	−1.811**	−1.628*	−1.543*
	(0.733)	(0.734)	(0.731)	(0.731)
GraduateDegree	−2.882*	−2.882*	−2.923*	−2.948*
	(1.387)	(1.387)	(1.378)	(1.375)
Managers	−6.875**	−6.872**	−6.805**	−6.781**
	(2.813)	(2.812)	(2.798)	(2.792)
Engineers	29.541**	29.525**	29.051**	28.861**
	(10.689)	(10.688)	(10.665)	(10.655)
Scientists	−63.770*	−63.751*	−63.153*	−62.882*
	(28.664)	(28.662)	(28.570)	(28.533)
Teachers	37.829***	37.806***	37.388***	37.226***
	(9.871)	(9.870)	(9.842)	(9.831)
Technicians	−4.198	−4.201	−4.051	−3.931
	(20.976)	(20.973)	(20.923)	(20.906)
Internet	−36.467***	−36.451***	−35.945***	−35.721***
	(0.599)	(0.602)	(0.597)	(0.597)
Short-Term 1 Week	−1.056			
	(0.977)			
Short-Term 2 Week		−0.408		
		(0.587)		
Short-Term 3 Week			2.157***	
			(0.283)	
Short-Term 4 Week				2.438***
				(0.260)
Log-likelihood	−32673.406	−32673.976	−32659.498	−32645.968
AIC	65382.811	65383.952	65354.997	65327.935
BIC	65579.456	65580.597	65551.642	65524.580
N	410333	410333	410333	410333

Table 4.2: *Short-Term Contagion Models*

demographic effects that would be weaker were the majority of EAs, especially non-FidoNet adopting ones, included.

The results on education were reasonably close to expectation: areas with more undergraduate degree holders and those below highschool were by and large significantly less likely to deadopt relative to areas with more highschool diploma holders. Those with graduate degrees had a mildly significant effect as well.

Most age categories proved insignificant, though the upper age categories, particularly over 60, proved significant and positive suggesting that retirees may have been more likely to leave FidoNet. This coheres with expectation, though I expect more significance in the 30–50 age range.

Population was only significant before internet adoption was included, and urban surprisingly was mildly negatively significant, though the results are fairly weak. Hierarchical diffusion, like internet competition, was not as expected. Similarly with income—security does not, *ceteris paribus*—translate to increased sysop resilience, the opposite seems to be the case.

However most surprising among the demographic factors are employment categories. Like growth these coefficients were much higher, which in part relates to their relatively low range[44] but nevertheless the effects are significant and large.

Most surprising is the coefficient for the proportion of teachers in the area, which I expected to be insignificant as during growth but become a very strong predictor of deadoption, both statistically and in terms of magnitude. Perhaps the set of teachers who did once use FidoNet, was relatively small (given the result during growth) but prevalent enough to strongly drive deadoption. Again, changes in school technology use may have helped drive this process, though that seems a weak explanation.

The number of engineers in an area is the second most surprising effect—in contrast with before this is now a strong predictor of FidoNet deadoption. Perhaps they fit Roger's innovator's criteria: quick to adopt the newest thing but also

[44] The proportion of people in different age categories far outweighs the proportion of people in occupation categories.

quick to leave for its replacement. This combined with the relative insignificance of technicians is also surprising given the original significance of both and the likelihood that there may be similarity between these groups.

Finally, the proportion of managers and scientists are mildly surprising in their minor significance but also negative sign, suggesting that perhaps the security effect I considered above applies in an employment capacity rather than an income capacity.

It should of course again be noted that these controls are from 1990, and despite aggregating to the EA level and broader age categories, significant shifts, particularly in the technology sector, may have a strong role to play that is inappropriately controlled for here. Even assuming the demographics of these regions remain relatively similar over two decades, the ecological fallacy mentioned in the previous chapter (§3.5.1) is far more of a risk. While again we are theoretically aggregating behaviour to the point of all sysops in an Exchange Area, hoping that somewhat reflects the activity of the underlying community, it is ultimately a claim about the attributes of individuals that may be heavily unrepresentative.

4.7.2 Internet

This result is most surprising. It suggests that the internet in fact kept FidoNet alive, helping sysops keep costs down by providing communication shortcuts and potentially kept the community active through online users, many of whom may never have used their local BBS. Perhaps the cost of running a BBS decreased so much such that even without a local dialup userbase, there was no downside to continuing to run a node, and perhaps within the community the status associated with being an active sysop was incentive enough to keep going.

It is again worth considering data quality. In contrast to the Nodelists I have regional, semi-biannual data on internet adoption, and if I had much more detailed time or spatial information I would be far more confident in making this claim. However, the result is strong, robust to different effects, and has a not implausible interpretation. Thus I feel some confidence in rejecting my hypothesis, and suggesting that the internet helped sustain FidoNet.

4.7.3 Long-Term Contagion

Another puzzling result, again relating to sign, is the long-term contagion effect. The goodness-of-fit is a huge improvement over the demographic controls and the internet model, four orders of magnitude in fact, but the fact remains that the sign is inverse of what the theory would require. This suggests that as time progressed, if former FidoNet members continued to interact with persistent members indefinitely, or the remaining areas remained a beacon in people's minds of their past social connections, that reduced the likelihood of leaving.

I two potential interpretations. Previously sysop-occupied EAs may have had persistent user populations that may have stopped using dialup BBSs once their sysops shut their BBSs down, but remained involved via the internet. This coupled with the internet effect above, would suggest that for the remaining sysops the internet in fact kept their social community alive online, despite various other factors driving the decline of sysops.

Testing this hypothesis would require data on users or potentially messages. User data is notoriously difficult to obtain, while a decent amount of message data is available on internet archives. It is unfortunately difficult to estimate the completeness of this data—unlike Nodelists—but perhaps controlling for for traffic might allow us to understand what may explain these results.

I did attempt to include a variable for the number of sysops in the US as a global effect, but unfortunately that led to severe convergence issues in the model estimation process, and so I was unable to include that effect (much as I was in the previous chapter). In both cases, some measure of the overall health of the FidoNet community could lead to useful insights into its dynamics.

This interpretation certainly runs counter, however, to the plausibility of long-term contagion as the explanatory mechanism and the sign would suggest that this effect cannot explain why FidoNet declined. Especially with the internet result, this is particularly puzzling as even if the effect were short-term, one would perhaps expect a weak effect still in the correct direction. The fact that the goodness-of-fit so stronger, however, leads us to consider this an important effect to investigate with future research.

The other, potentially complementary interpretation is that during the decline phase remaining sysops want to maintain a certain level of geographic coverage. This might arise from a sense of duty to provide an infrastructure that efficiently cover the country—especially because FidoNet has been from its creation a means of sending messages over long distances and if those outputs are finally gone, the costs (for those still dialing up) could be prohibitive. Sysops may feel a responsibility to maintain the system, even with a very tiny set of local users—perhaps a form of nostalgia for the inception of a community they have had strong ties to.

Coupling this with positive short-term contagion effect (covered in the next section), the initial spike in leaving events just after FidoNet's peak followed by slower decline makes sense: when Exchange Areas and their respective local call areas had large quantities of sysops, one sysop leaving may have had less of an impact on FidoNet's geographic coverage. There may have been fewer local BBSs leading to less available local call numbers, but the local community could still persist. As the number sysops of that Exchange Area reach zero, the impact on the users increases as when the very last local sysop quits, they will at least have to pay long distance charges to connect, and may also feel like the local off-line component of FidoNet has also come to an end. The last sysop for a local community may then feel their responsibility grow in the long term, and thus the community still persists. Importantly, sysops who left while the system was still saturated do not bear the same responsibility for maintaining coverage that the last sysops standing.

4.7.4 Short-Term Contagion

Our final results pertain to the short-term contagion effect, which—unlike long-term contagion—actually has the expected sign for ls greater than 2. This provides empirical support for the hypothesis that short-term contagion is both a factor worth considering in contagion processes, and in particular that it applies to deadoption in the way expected.

However, the goodness-of-fit is considerably worse than that of the short-term effect, which suggests that more is going on than either model can account for. Under the short-term approach no mechanism—other than those captured by the demographic controls—can explain FidoNet's decline but the goodness-of-fit is comparatively high. With the short-term mechanism, FidoNet's decline is accounted for provided l is greater than 2, but the goodness-of-fit is much worse.

Ideally we could combine the models in some way, but doing so renders the results difficult to interpret as we are in theory measuring the same thing in different ways. Another effect we have considered which will be considered in the next chapter is a lifetime effect, where the deadoption time is largely dependent on the average length a sysop stays in the system and when they adopt. Such a mechanism would imply that the aggregate decline is simply a time-lagged reflection of the adoption curve.

Given the trend of AIC improvement in the short-term effect, I tried running the model for increasing ls as long as the AIC and BIC continued to improve. The results are shown in figure 4.2, which plot the relationship between AIC and weeks of l all the way to two years.

Increasing l to approximately 1.5 years (65 weeks) improves AIC by 877.31[45] a significant improvement, though still considerably less than the goodness-of-fit improvement to the long-term contagion option. Importantly, the coefficient remains significant and positive for the entirety of the ls above 2 weeks.

This lends support to the idea that short-term contagion may prove a useful concept, particularly in modelling decline processes, but that for short time-series long-term contagion may be a good-enough approximation. An l of 1.5 years suggests that leaving events affect the system for a long time, and were our time-series shorter perhaps the two contagion mechanisms would yield more similar results.

[45] 64672.594 from 65549.903

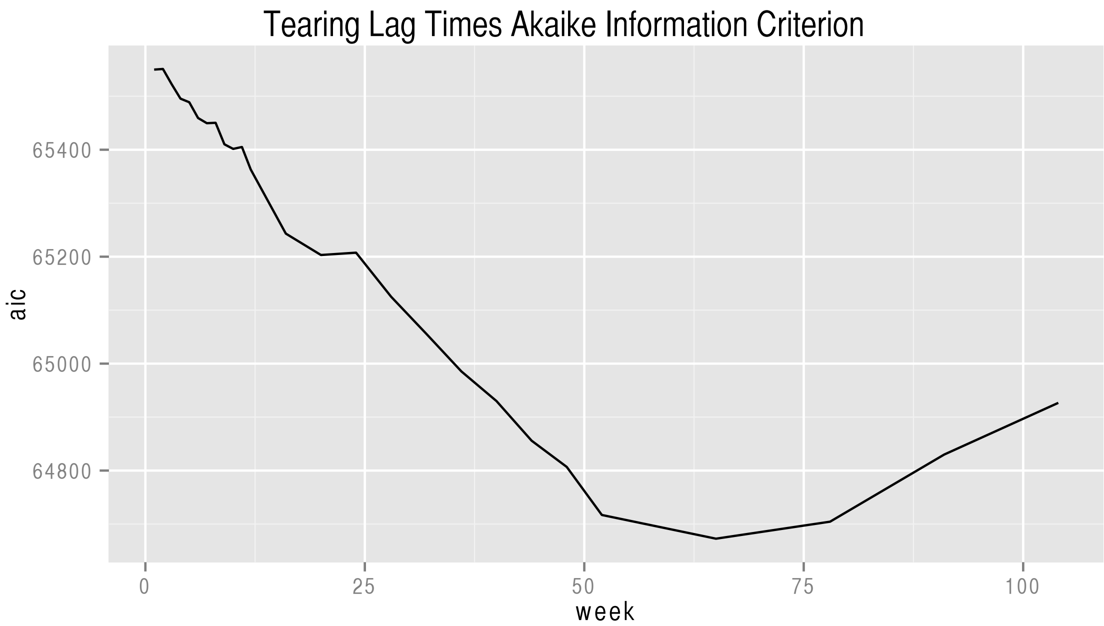


Figure 4.2: *AIC improvement with increasing l*

At present I believe these questions must be answered with a simulation and that a cox-model cannot tease out the difference between these mechanisms, especially if the lifetime hypothesis best explains the decline, and the two contagion mechanisms are merely operating around that more dominant effect. This question is addressed in the next chapter.

4.8 Conclusions

Although it has not been modeled explicitly as such here the data imply a multilevel structure. Sysops are in fact a middle level—further up the hierarchy are Exchange Areas which may contain more than one sysop, and lower down are individual users of which there would almost always be more than one for each sysop. The correspondence is not exact however—individual users may well access different sysops across time, not necessarily within the same Exchange Area. Rather than looking directly at sysops or individual users, this chapter studies Exchange Areas. As described in section 4.3.1, the rationale behind this is that Exchange Areas are the best approximation of a geographical community that is available in the data.

With our informal understanding that FidoNet relies on local communities for the expansion of the network, it is appropriate to focus on the closet approximation of the geographical units that would match to these local communities. This also allows us to identify the demographic characteristics of the surrounding area.

However, a more sophisticated analysis would try to take account of the multilevel issues in the study. One difficulty in this of course would be that each individual user might access multiple sysops, possibly across different Exchange Areas, so the levels do not neatly nest in one another.

Such an approach would allow us to look more correctly at the effect of the information from the Census on sysop leaving—these effects should be interpreted as the effect of living in an area with (e.g.) an high level of engineers, rather than as describing individual-level effects of being an engineer on a person's likelihood of becoming a sysop. The effect observed at the Exchange Area level could operate in different ways at the individual level—it could be that engineers are more likely to move onto another technological fascination and therefore stop being sysops, but it could also be that people with social ties who are engineers are more likely to have been sysops and drop out when their engineer family member or friend loses interest. These are hypothetical examples, the main point being that the calculated effects of different local characteristics need to be interpreted appropriately, with awareness that there are multiple different individual-level mechanisms that might lead to the same observed correlation.

A related limitation of the current analysis is that it fails to take account of the geographical size of Exchange Areas. Smaller Exchange Areas are in urban areas—often there are multiple Exchange Areas within a single city. If the last sysop in a small urban Exchange Area deadopts, users are less likely to be negatively affected as the nearest Exchange Area is not far away and will not incur a long-distance call charge. The current models include population density which will capture some of this effect, but a more sophisticated multilevel analysis that grouped Exchange Areas within cities (and sysops within Exchange Areas) would more clearly delineate this effect.

Chapter 5

Decline Simulation

The previous chapter's results are unexpected and difficult to interpret, and speak to the potential complexity of the underlying mechanisms governing the dynamics of a system such as FidoNet. They also speak to the limitations of the modelling approach employed, partly in terms of data availability but also the extent to which modelling multiple mechanisms interacting with each other is problematic. Arguably the results from chapter 3 suggest similar complexities even for what appears to just be one mechanism.

In light of these concerns, I use an agent based model (ABM) to explore the implications of the two mechanisms proposed in the previous chapter and what dynamics can arise from their interaction. The model studied is very simple, and at present is only considered on a square lattice, following many other works that have chosen that method for its simplicity and tractability (Schelling, 1969; Nowak and May, 1992). The model is also constructed so that future empirical work can employ a MLE to scale up the simulations for use with the entire US and the demographic variables included in chapters 3 and 4. My results suggests that the model tends to have interesting behaviour when the mechanisms are of opposing sign.

The purpose of the agent-based model is threefold. Firstly, an agent-based model allows us to explore the dynamics of a system based on the regularities observed in the previous chapter. Therefore the first purpose is to explore the dynamics of a system with certain similar properties to FidoNet if the coefficients for the long-term and short-term deadoption events calculated in the previous chapter were to reflect behaviours at the individual-level (rather than being limited to describing what happens at the Exchange Area level). The extent to which the dynamics of the ABM match up with those observed in FidoNet is in this way a test of the conclusions of the previous chapter.

The second purpose is that an ABM permits us to vary the direction, strength, and relative strength of the long-term and short-term contagion effects in a way that is impossible outside a simulated system. This is useful because it guides our interpretation of the contagion effects in the previous chapter and how they might apply to the decline of other social systems—what is the likely relative importance of the different effects, and in what circumstances will one outweigh the other? The third purpose of the ABM is that we can remove the geographical and demographic context of FidoNet by placing the simulated system on a simple lattice grid. It should be evident in this simplified system that the effects we observe are not artefacts of or dependent on US geography or demography.

5.1 Literature Review

The January 2005 issue of the *American Journal of Sociology* (AJS) was a watershed in the application of ABM to sociology: an entire special issue devoted to the topic. Nigel Gilbert, editor of *Journal of Artificial Societies and Social Simulation* (JASS), co-authored with Andrew Abbott, editor of AJS, the foreword (Gilbert and Abbott, 2005):

> *But the most important changes in [the last twenty years of] social science computation have come in the use of computers to 'think through' the implications of human actions within social structures—action in networks.*

Since that issue, ABM has become more prevalent within the social sciences, but not perhaps as prevalent as the editors would have hoped, at least within high-impact social science journals like AJS. ABM seems to be gaining ground in more hard science journals like *Science*, *Nature* and *Proceedings of the National Academy of Sciences* (PNAS) in a variety of fields as well as social science published in those journals

(Saavedra et al., 2008; **?**). Meanwhile within sociology, ABM has been strongly supported by the Analytical Sociology movement (Hedström and Swedberg, 1998; Hedström, 2005; Hedström and Bearman, 2009), and a number of textbooks have been published to introduce ABM to various disciplines.

These recent developments can trace their inspiration back to theoretical work that dates back to the late 1960s (Schelling, 1969; Gardner, 1970), and a variety of important works that came between (Nowak and May, 1992; Huberman and Glance, 1993; Hedström, 1994; Epstein and Axtell, 1996; Axtell, 2001; Axtell et al., 2002; Macy and Willer, 2002). I summarise these developments, focusing on works that are most pertinent to the model I develop later in the chapter.

5.1.1 Precursors to ABM

Before ABM, ideas of simple interacting elements generating complex aggregate behaviour were explored theoretically, and many of these works have heavily influenced the development of ABM.

Schelling's (1969) segregation model casts a long shadow over the field of ABM. It is not an ABM itself but rather a theoretical demonstration of how a simple model with interacting agents can exhibit complex behaviour, and has been re-analysed in a variety of contexts since (see (Clark, 1991; Clark and Fossett, 2008)).

Schelling's segregation model is perhaps best described as a precursor or prototype ABM model, e.g. no heterogeneity between agents. It is a demonstration of how a very simple model can generate complex behaviour which on the surface seems quite difficult to explain. The basic model uses a square lattice (as I do below), to explain how segregation can emerge in a system where agents only have a slight preference for homophily. That slight preference is self-reinforcing over-time, amplifying the separation between agents, leading to a highly segregated system despite having agents who are not as extreme in the preferences individually as the system as a whole would suggest.

Similarly, Conway's game of life, first presented as a game in *Scientific American* (Gardner, 1970), also uses a lattice structure and follows simple rules designed to replicate basic features of population dynamics, where too many agents create overpopulation, too few underpopulation, and some balance produces a new agent provided there is available space. The range of time-series this very simple model can generate is astonishing despite the fact that it is entirely deterministic from its initial starting point. It is often used more for creativity today, but it often cited as inspiration to ABM research (Macy and Willer, 2002)

5.1.2 Computational

Reynolds's (1987) work on flocking behaviour is perhaps the first fully-formed ABM. He models the way birds move together with 'boids', simple agents that try to match the direction and speed of their neighbours. His work is one of the first to generate time series of simulations computationally for comparison with real-world data.

Moving on to social science, Nowak and May (1992) study the classic prisoner's dilemma from game theory, again on a lattice. Their results generate fascinating patters, especially when visualised, but their results rely entirely on synchronous updates, where at each point in time agents choose whether to defect entirely independent of each other.

Huberman and Glance (1993) noted this problem showing that the majority of their results were artifacts of this synchronous action structure. It meant that in effect they were not interacting as the theory required and Axtell (2001) explores a variety of solutions to this problem.

Hedström (1994) uses an ABM to assess the implications of a spatial contagion model like those studied in this thesis, using a random spatial arrangement. However, his methods are problematic in two ways. First he is subject to the same synchronicity criticism of Nowak and May (1992). Second, despite appearing to generate a probabilistic model, where a probability is calculated for each agent at each point in time and then a random number generator randomly draws between 0 and 1 that value is compared with the calculated probability, Hedstrom merely treats probabilities under .5 as a no adoption and probabilities above .5 as adoption.

Finally, Epstein and Axtell (1996) generate one of the first highly complex ABMs: an artificial ecosystem they call Sugarscape. Sugarscape also operates on a lattice, where locations have resources and agents attempt to use those resources to support themselves. The variations of how Sugarscape has and can be used are extensive, allowing multiple organisms to compete and the effects of various exogenous forces to be modeled. Sugarscape is like a a laboratory, allowing a host of different features to be experimented with.

This approach breaks with the approach of the early very simple models allowing a host of possibilities to be explored. It does however, reduce the parsimony of the system and the effects are added, the more difficult it is to make strong claims about the predictability of the model's behaviour.

5.1.3 Developments in Sociology

Since these early models ABM has become much more prevalent in the social sciences, and within sociology three major strands have strongly evangelised the approach. First is Macy and Willer (2002), a review article which advocates greater use of ABM in sociology, noting the advantages it has in the exploration of social theories and the dynamics of micro-mechanisms.

Promoters of the Analytical Sociology approach Hedström and Swedberg (1998); Hedström (2005); Hedström and Bearman (2009) also encourage the use of ABM, which goes hand-in-hand with their advocacy of theorising about social systems dynamics as the interaction of social mechanisms. Social mechanisms are often well-suited to ABM because they are individually simple and easy to represent in computer code, but the implications of their interactions are non-trivial. ABM allows their interactions to be explored in an isolated way, much like an artificial social experiment.

Finally, the aforementioned *AJS* issue (Gilbert and Abbott, 2005) was a major point in the advancement of ABM in sociology, including work on modelling norms (Centola et al., 2005), voter turnout (Fowler and Smirnov, 2005), and knowledge diffusion (Chang and Harrington, 2005).

5.1.4 ABM and Declining Systems

The lack of studies of declining systems in general extends to the ABM literature, and as mentioned in chapter 4, most of the papers that have modeled decline in a quantitative way have done so via ABM. I mention them again here in the extent to which they relate to the modeling I do in the remained of this chapter.

Axtell et al.'s (2002) is a follow-on from the Sugarscape approach, in that it is a highly complex model, again of population dynamics, with a wide range of mechanisms and components. What distinguishes it from Sugarscape is its connection with real-world data on the decline of the Anasazi, and the extent to which the data can be replicated by their model. Some of the structural decisions made below are designed to potentially allow similar exploration in future, but I have avoided the difficulties implied by a model of that complexity, which increase the model parameter space considerably and undermine the clarity of interpretation.

Saavedra et al. (2008) model the collapse of the New York garment industry, again with respect to real-world data. Their decline has a number of components but far less than Axtell, arguably a middle-range between Sugarscape and Schelling's contagion model. Their work is on a network rather than a grid, and the mechanisms derive from economic competition rather than contagion.

Both papers model decline with respect to real-world data, incorporating mechanisms that allow for removal of agents but also introduction of agents. My model explores the basic components of the former but at present does not consider the latter. My hope is future work will consider growth mechanisms and properly incorporate the FidoNet data.

5.2 Model Specification

The model itself is relatively simple, largely motivated by parsimony and future tractability for use with my data. It attempts to strip away the various other effects which may be interfering with the results of chapter 4, providing a laboratory for testing the implications of how these two mechanisms may interact.

I first describe the basic components of the model: agents embedded in a square lattice. I next explain how I address the issue of synchronous updates mentioned above. This determines which agents are active at each time point, and I discuss the plausibility of this assumption in the context of FidoNet. Next I cover the baseline hazard of leaving for agents, which for the results below is set constant but could in future be relaxed to incorporate various demographic factors. I then present the distance effects modelled in the previous chapter, and finish the section by specifying how the above effects are combined in a logistic function following Hedström (1994) among others.

5.2.1 Agents

Agents are meant to be representative of active FidoNet sysops. I focus on sysops rather than Exchange Areas as our unit of analysis because ideally all of the analysis would be conducted at the individual level, and it is individual sysops or users who decide to join or leave FidoNet—Exchange Areas cannot make decisions. Were we to scale the model up in a multi-level fashion, agents could be aggregated together into Exchange Areas to maximise comparability with the empirical data. For now, we shall leave Exchange Areas to one side but will return to this possibility in future work.

Agents have four attributes:

- Position
- Demographic profile
- Status of being a sysop
- Status of being active

A position is a pair of coordinates on a square lattice following many other theoretical models (Schelling, 1969) and ABMs (Nowak and May, 1992; Epstein and Axtell, 1996). Coordinates are integers beginning at $(0, 0)$ and ranging up to the square root of the number of agents (n) in the simulation assuming full occupancy.

A demographic profile is a floating point number which represents demographic variation in individual propensity for agents to randomly stop being sysops. In the results below this value is equal for all agents for a given simulation run, and serves as a baseline probability of leaving.

This effect means that any currently active sysop —even without the short or long-term contagion effects—has a non-zero probability of leaving in that week (t). This effect heavily influences the overall time-scale of the decline process for a given run, and the relative strength of the parameters of the contagion effects to the baseline will determine whether the set of sysops all leave very quickly, almost never leave[1], or leave in a decline curve more similar to the actual FidoNet data.

More realistic values could be input in future work to this variable, corresponding to the control variables included in chapters 3 and 4. A time-series of each agent's profile, including changes in age, education and occupation, could be included either by simulating these attributes or inputting actual data were it available at either the Exchange Area or sysop level.

However, for each of these effects an additional parameter would have to be included in the model to set the relative strengths of these attributes just as those strengths are estimated in the Cox model coefficients. Exploring the set of possible strengths purely by simulation would imply an even larger parameter space because two reasons. First, if the attributes themselves were selected from a distribution—say, the mean and variance of a Gaussian distribution—then each of the parameters to that distribution (2 in the case of a Gaussian) would add to the parameter space to be explored, not to mention the parameter required to input the relative strength of each effect.

Second and more fundamentally: the implications of random draws on attributes are likely to be highly non-trivial, implying a great number of runs and tests to explore the implications of the possibilities adequately. As is demonstrated in section 5.3.1, even with just a uniform baseline constant for all agents, a variety of dynamic regimes is implied, which means that because these parameters so

[1] Regardless of the parameters to the model, as $t \to \infty$, the all agents will leave.

fundamentally impact the underlying dynamic properties of the system, much more time than is feasible under the constraints of this thesis would be required.

There are ways of solving this problem, first using the hybrid or bootstrapping approach in Hedström and Swedberg (1998), where coefficients are estimated statistically from real data and then a simulation is run using those coefficients. Alternatively, a MLE approach can, provided the model is suitable, take empirical data and estimate the parameters by maximising the log-likelihood. I have designed the model in such as way as to make this feasible in future, and this issue will be covered in section 5.2.5.

Finally, the model becomes less straightforward to interpret the more components are added. If we can replicate the desired model behaviour with the absolute minimum of components, then that model forms an excellent starting point for future investigation. Each non-essential effect can then be added systematically to assess how they will modify the underlying behaviour, while keeping that core behaviour in the model.

The status of being a sysop is simply a boolean True or False indicating whether, at the current point in time, the agent is a sysop or not. Once they have left, under the present model the sysop cannot rejoin, though were a joining mechanism added in future this constraint could be relaxed.

For the purposes of the model below we will represent the baseline to by the parameter C.

5.2.2 Contagion Effects

Just as in chapter 4, we define the contagion effect for agent i in a system with n agents at time t with effective time period p to be

$$\kappa_{it}^{(p)} = \sum_{j \neq i}^{n} \frac{\gamma_{j,t-1}^{(p)}}{d_{ij}} \tag{5.1}$$

which is the sum of 1 divided by the distance between agent i and agent j (d_{ij}) for all other agents j. This is multiplied by $\gamma_{j,t-1}^{(p)}$ where

$$\gamma_{j,t-1}^{(p)} = \begin{cases} 1 & \text{if } j \text{ left between } t-p \text{ and } t-1 \\ 0 & \text{if } j \text{ did not leave between } t-p \text{ and } t-1 \end{cases}. \tag{5.2}$$

$\gamma_{j,t-1}^{(p)}$ is 1 if j left within the contagion effectiveness period p, and 0 otherwise.

As in the previous chapter, $p \in \mathbb{Z}_{\geq 0}$. If $p \to \infty$, then eq. (5.1) is the long contagion mechanism and for $0 < p << \infty$ it is the short term contagion mechanism.

On the square lattice, d_{ij} is measured as Euclidean distance[2]. Given the lattice specification in section 5.2.1

$$\{d_{ij} \in \mathbb{R}_{\geq 0} | d_{ij} \leq \sqrt{2n}\} \tag{5.3}$$

where n is the number of nodes. Summing $\frac{1}{d_{ij}}$ within this range will quickly lead to values greater than 1, and thus following many other ABMs, a logistic function is used to combine these effects (see section 5.2.4).

Each contagion mechanism has a β term: β_S for short-term contagion and β_L for long-term. These β parameters function like the coefficients in the statistical models, determining the strength of the given effect relative to the others.

5.2.3 Time and Asynchronous Events

Throughout this thesis weeks have been the basic unit of time. This is because the dataset is a raw archive of FidoNet Nodelists, and these are published every Friday. Thus far this discretization of time has been appropriate as Cox modelling software requires covariates to be input in regular intervals, and given the time-resolution of the data I have interpolated variables as necessary (internet adoption in chapter 4 for example).

But FidoNet—like any social system—is in constant flux, and messages can be sent day or night, provided a node is on and has an available connection. Indeed, the social connections that are made and/or maintained via FidoNet can also be

[2] Vincenty's distance used in the previous chapters accounts for the Earth's curvature.

expressed through face-to-face or other ICTs like email or text messaging. This all relates to the extent to which the Nodelist is the primary means by which sysops are notified of sysops leaving, and therefore if it is most sensible to model the contagion of leaving events in this discretized manner, especially in the framework of a simulation where data constraints can be relaxed.

Theoretically this is a difficult possibility to address, though it is worth exploring. Åberg (2009) suggests that when an someone divorces, that triggers thoughts about divorce for their social ties, increasing their likelihood of divorce. Anyone discussing leaving FidoNet, even if they never do, could have a similar trigger effect on their social ties, and as such waiting until the actual leaving event to model the impact may miss the impact of the social process that leads up to the decision.

Considering this aspect of FidoNet's dynamics has implications for how to solve the synchronous problem discussed above (see section 5.2.3). If every week, all currently active agents are given the opportunity to decide to leave, making that decision solely on the state of the system at the previous t, then they are making that decision synchronously and the problems of Nowak and May's (1992) Prisoner's Dilemma model will apply.

One possibility I considered was having a random ordering draw of active agents each week, and so for each time step there would be on average n_t (number of still active agents at time t) mini time-steps for each agent's chance to decide to leave. With each agent's time-step, only they are given the opportunity to leave, and they make that decision purely based on the state of the system at the mini time-step previous, when another agent was given the opportunity to leave.

This method effectively removes the weekly structure from the agent's perspective and allows agents to influence each other in an arguably more realistic way, communicating outside the confines of the weekly Nodelist schedule. The Nodelist then becomes a signpost for potential comparison with real-world data.

However, this method introduces difficulties in how the influence effects are calculated. The mechanism tested in chapter 4 has a time-period p, which determines how long an agent's leaving event actively affects the remaining agents.

The effect of p with a random ordering introduces different artifacts depending on implementation.

First, if the weekly signposting is used as the basic calibration for duration of effect, then agents who are randomly selected earlier in the week will artificially affect others longer than those selected towards the end of the week. Second, if instead the effect lasts for a constant number of agent time-steps, then as the system shrinks the p duration effectively changes implied length. More formally: the average number of times an agent should be selected to be active over a given p should not change as n decreases. But since as $t \to \infty$, $n_t \to 0$ monotonically but p is constant, with a smaller n_t, the average number of active selections for the remaining sysops will increase. This last problem could be controlled for by dividing by the remaining n_t at each leaving event, but that seems increasingly convoluted, and may introduce other artifacts as yet unconsidered by this author.

Thus a hybrid approach, following Axtell (2001); Axtell et al. (2002), is employed below. At each week time-step t, the probability of agent i being active is set to:

$$P(A_{it}) = s \tag{5.4}$$

where A_{it} is the event of agent i being active at time t and s is a constant parameter for the model. For each time-step, agents are selected to be active and then separately have a chance to leave based solely on the state of the system at the previous t. This is exactly the same as the synchronous problem system described above, except s determines on average what proportion of remaining agents n_t are active for each t. Provided $s < 1$, the synchronous problem is reduced.

This allows the short-term effect to be calculated in the same way it was in chapter 4, and for the results to retain direct comparability with the data for future analysis. This arrangement also facilitates the use of MLE in future work, a topic addressed in section 5.2.5.

5.2.4 Combining Effects

In order to generate a probability of an agent leaving, we need to combine all the aforementioned effects. These effects may be on very different scales and will not be bounded between zero and 1 ($[0, 1]$). To accomplish this, following (Hedström, 1994) I use a logistic function (eq. (5.5)):

$$f(x) = \frac{1}{1 + \exp(-x)} \tag{5.5}$$

It is well-suited to use as a probability function because—like a probability—it is bounded between $[0, 1]$ and because it approaches those bounds asymptotically, regardless of the value of x, the function still returns a valid probability. It also means that as our combined effects becomes larger, the probability increases, just as the probability decreases as the combined effects decrease.

I define the event of i leaving at t to be L_{it} and by extension the probability of that event is $\mathrm{P}(\mathrm{L}_{it})$. Since an agent must be active at t in order to leave, I will focus on the probability of leaving given the agents is active: $\mathrm{P}(\mathrm{L}_{it}|\mathrm{A}_{it})$. This probability will take the logistic form, combining effects in an additive manner similar to the way effects are combined in a Cox model.

$$\mathrm{P}(\mathrm{L}_{it}|\mathrm{A}_{it}) = \frac{1}{1 + \exp(-(\mathrm{C} + \beta_\mathrm{L}\kappa_{it}^{(\infty)} + \beta_\mathrm{S}\kappa_{it}^{(p)}))} \tag{5.6}$$

where as defined above: C is the baseline effect (extensible to a demographic profile in future), $\kappa_{it}^{(\infty)}$ is the long-term effect on i at time t and $\kappa_{it}^{(p)}$ is the short-term effect on i at time t with influence period p. β_L and β_S determine the strengths of the long and short-term contagion effects (respectively), much like coefficients for effects in a Cox model.

We combine the probability eq. (5.6) with the activation procedure A_{it} defined above in eq. (5.4) to obtain:

$$\mathrm{P}(\mathrm{L}_{it}) = \mathrm{P}(\mathrm{A}_{it})\mathrm{P}(\mathrm{L}_{it}|\mathrm{A}_{it}) = \frac{s}{1 + \exp(-(\mathrm{C} + \beta_\mathrm{L}\kappa_{it}^{(\infty)} + \beta_\mathrm{S}\kappa_{it}^{(p)}))} \tag{5.7}$$

where s is the probability agent i is active at t ($\mathrm{P}(\mathrm{A}_{it})$).

5.2.5 MLE

While I do not use MLE in the analysis below, I have specifically designed the model such that it will be possible to combine the model with the FidoNet data in future. This would allow the attributes in the actual data to be leveraged, and the simulation to be scaled up and compared to what actually occurred.

MLE is a method of estimating the parameters of a statistical model. Technically, it maximises the likelihood of a dataset—given a model—by an algorithm that searches for likelihood peaks under different parameter regimes. Whatever the highest peak found, those parameters are then returned, and can be used to simulate the events taking into account the available data. Assuming high quality data, this is a means of avoiding some of the difficulties of excessive parameter spaces discussed above.

In order to compute the MLE, the likelihood of the data given our model needs to be well-defined and computationally tractable (at least to a reasonable degree of precision). In our case deriving an appropriate likelihood function requires conditioning on previous states, and in fact a more precise version of eq. (5.7) is

$$P(L_{it}|S_{t-1}, X_{t-p} \dots X_{t-1}) = \frac{s}{1 + \exp(-(C + \beta_L \kappa_{it}^{(\infty)} + \beta_S \kappa_{it}^{(p)}))} \tag{5.8}$$

where S_{t-1} is the state of the system at time t and X_{t-1} is the number of agents who left at $t-1$. We must condition on these states because the current long and short-term effects are dependant on these previous states. This implies this is a higher-order Markovian process up to p weeks prior (as implied by the short-term effect).

Now that this has been established, I can demonstrate how this makes an MLE more tractable. An MLE is a method of estimating the parameters of a model which will maximise its likelihood given a dataset. Therefore, to compute the MLE, we must compute the probability of all the observed events given the model. Were all the events entirely independent, this would be a simple product of all the probabilities. But, because of the dependencies highlighted in eq. (5.8), some extra work is required.

Within a given week, each agent only takes into account what has happened by and including the previous week, and does not include the decisions of other active agents at the current week. In terms of FidoNet, this is akin to each sysop only being aware of the Nodelist of the previous week, and expressly not communicating with each other before making their leave or stay decision for that week. Only the decisions of sysops from previous weeks (within p weeks to be precise) and the sysops who have left the system since the start of the simulation, impact their decision.

Mathematically: for agent a who is active at t, and agent b who is active at t, $L_{at}|S_{t-1}, X_{t-p} \ldots X_{t-1}$ is independent of $L_{bt}|S_{t-1}, X_{t-p} \ldots X_{t-1}$. By contrast: had the finite fraction time-scale suggested and rejected in section 5.2.3 been used, these events would not be independent.

Having demonstrated this independence, we can now consider modelling the likelihood of observations given the model. The probability of observing all the events in a week from the dataset—given the model—is

$$P(\bigcap_{i \in \alpha_t} L_{it}|S_{t-1}, X_{t-p} \ldots X_{t-1}) = \prod_{i \in \alpha_t} P(L_{it}|S_{t-1}, X_{t-p} \ldots X_{t-1}) \tag{5.9}$$

where α_t are the set of agents who are still sysops at t. Since we have established that within a given t all events are independent, we can write the probability of the events in a given week as the product of all the observed events using eq. (5.8).

For the sake of convenience, I will define W_t as the set of observed events at time t:

$$W_t = \bigcap_{i \in \alpha_t} L_{it} \tag{5.10}$$

We can now express the probability of observing all the events in the dataset

$$P(\bigcap_t W_t) = \prod_t P(W_t|W_0 \ldots W_{t-1}) \tag{5.11}$$

where $W_0 \ldots W_{t-1}$ are all the previous weeks observed.

It would appear to be necessary to account for dependencies all the way back to the start of the time series because of the long-term contagion effect. However, the I argued above that the model is Markovian up to p weeks prior to t, provided we know the set of all agents who have left since the start of the time-series.

With this fact, we can rewrite eq. (5.11) as

$$P(\bigcap_t W_t) = \prod_t P(W_t | S_{t-1}, W_{t-p} \dots W_{t-1}) \tag{5.12}$$

which is the likelihood of the observed events given the model. Taking the log of eq. (5.13) leads to a nicely computable summation: the log-likelihood of eq. (5.11) as

$$\log(P(\bigcap_t W_t)) = \sum_t \log(P(W_t | S_{t-1}, W_{t-p} \dots W_{t-1})) \tag{5.13}$$

This expression can then be written as a function in a language such as R, and then passed into a standard MLE package[3] with the appropriate data. The extra work under this regime is considerably smaller and computationally tractable than that required under the alternative time structures discussed in section 5.2.3, because those regimes implied a greater dependence between events.[4]

Thus the model parameters can—with some work—be estimated in a tractable way via MLE, allowing future work to incorporate the observed FidoNet data for larger, more empirically grounded simulations.

5.2.6 Implementation

The model is implemented in the Python programming language, heavily leveraging the Numpy vector library for computing effects and the multiprocessing library within the Python standard library. This combination, on a large clustered server, kept RAM usage from exceeding capacity and run times short. As mentioned before, the model was designed with application to real data in mind, and the data

[3](Bolker and Team, 2016) is one such contributed package to R.

[4]If the model were entirely synchronous (such that there was no activation parameter), then techniques from a standard logistic regression could be used. But then the issues outlined in Huberman and Glance (1993) would apply.

structures are such that the model can be scaled up beyond the parameters ranges explored below, potentially adding in more demographic attributes.

However, the computing time scales with the approximately the square of the number of agents[5] (n), so increasing the number of agents quickly slows things down. Adding additional demographic attributes would not significantly slow the system down, even if they were a time series, as that vector could be computed in advance and reused across multiple runs. This would of course increase the parameter space (as mentioned above), so below only one attribute is considered.

Plots use the Matplotlib library (Hunter, 2007), with some modification using the Inkscape vector image editing project.

5.3 Results

In the plots below (unless otherwise indicated) decline curves and heat map results are from 20 runs under the same parameter regime. For all runs the time-series was capped at 1000 steps, such that the model finishes either when all agents have left or 1000 weeks is reached to keep the problem computationally trackable.

Decline curve coordinates (except where indicated) correspond to the mean at the t on the x-axis across all runs, with error bars at 1 standard deviation from that mean.

The coloured lattices demonstrate a single run as a time series, colouring agents with respect to when they leave FidoNet under that run. These are selected to be demonstrations of particular behaviours rather than representative of the sample as a whole.

I begin with a basic validation of the underlying baseline, and then a heat map demonstrating the parameter combinations that produce interesting behaviour with respect to a given baseline. These heat maps suggests that as t grows, regimes with opposing coefficients on the two contagion parameters tend to produce more interesting results. I then present results where the contagion mechanisms are in

[5] The big O is actually non-trivial to compute even in the simple case of the stripped down model here, and will depend on p as well as n.

isolation as a basic understanding of their dynamics, and then finish by covering areas where they interacting. The implications of the results are discussed in greater detail in sections 5.3.4 and 5.3.5.

5.3.1 Baseline Validation

The first plots fig. 5.1 are a simple demonstration of the model regime without either contagion effect and the extent to which it behaves as would be expected.

Figure 5.1a plots the decline curves for 100 separate runs with a baseline $C = -5$ as the blue curves. Their stochasticity is indicated by the noise on each of those curves as contrasted with the smooth red curve, which is an estimation of what the curves should look like (and as the number of trials approaches infinity, should average out to being).

The red curve plots the expected decline given that the baseline model equates to a geometric distribution. This is because each actor's leaving probability is effectively a Bernoulli trial with a fixed probability of

$$\frac{1}{1+\exp(-C)}. \tag{5.14}$$

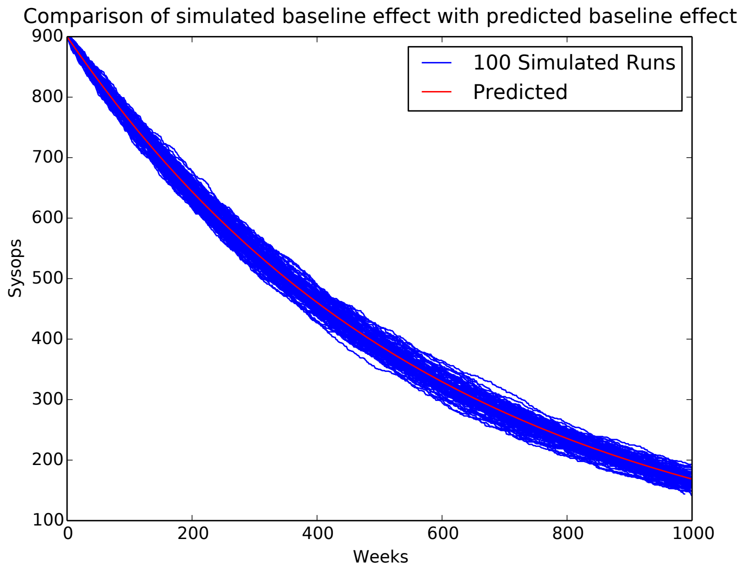
Therefore, the expected survival time of each agent should be geometrically distributed, such that

$$f(t) = n\left(1 - \frac{1}{1+\exp(-C)}\right)^{t}. \tag{5.15}$$

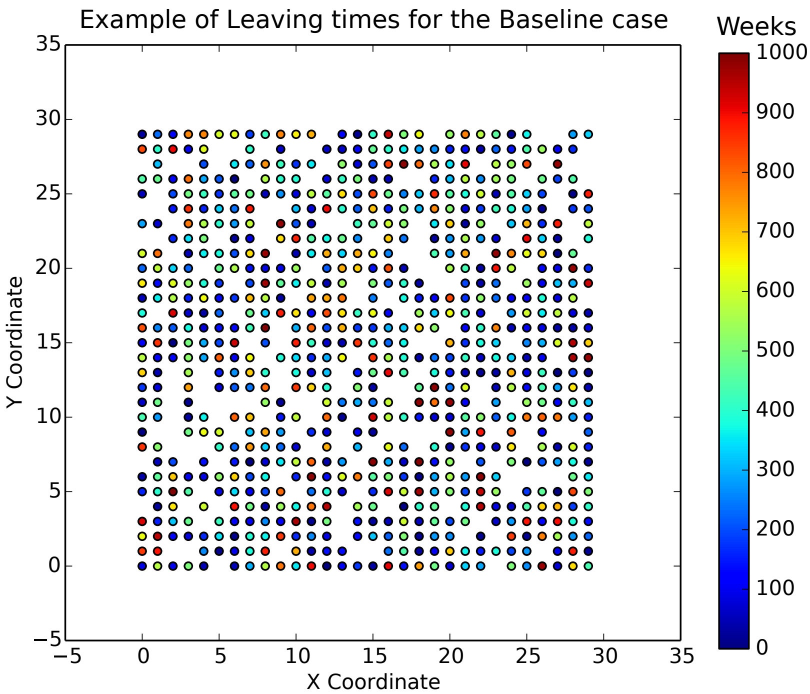
The red curve is simple a plot of the function in eq. (5.15), and it visually approximates the mean of the blue curve trials around it.

Figure 5.1b is as single run from the set of blue trials in fig. 5.1a. Each circle corresponds to an agent on the lattice, and their colour indicates when in the trial they left the system: blue for early, green for partway through and red for late. Gaps in the lattice indicate agents that never leave.

Figure 5.1b demonstrates the random spatial arrangement of the time series, as should be expected without either contagion mechanism. Each agent has an entirely independent probability of leaving the system, and that probability has



(a) *Baseline Decline Curve*



(b) *Baseline Grid*

Figure 5.1: *Baseline Probability Plots with* $C = -5$

no dependence on space. These results are not surprising, but are meant as a demonstration that the basic model framework behaves as should be expected.

5.3.2 Fixing Lattice Size

One parameter that is worth fixing across the remainder of the analysis is lattice size, as the length of the simulation run is very dependant on the number of agents.

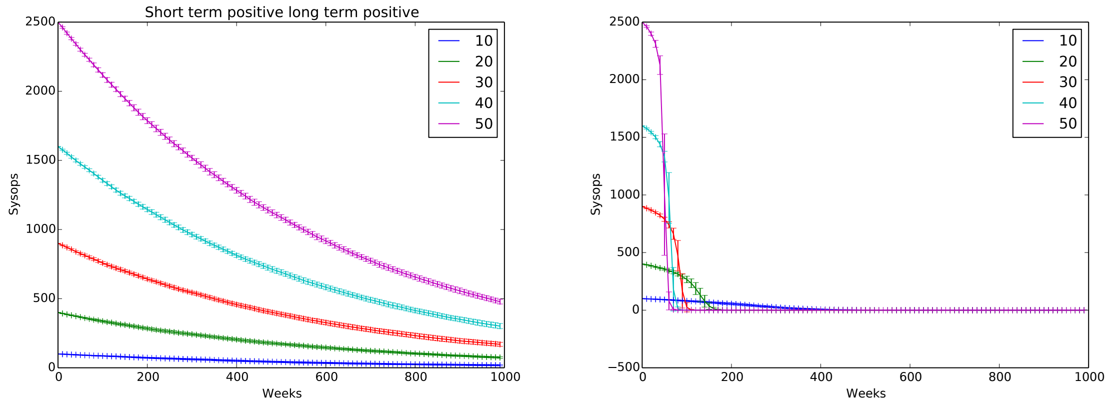
Figure 5.2a demonstrates the basic behaviour variation without contagion effects (baseline at -5). The curves are similar, starting at different initial adoption states (because they vary in the n) but all exhibiting the declining curve without inflection points. The distance between the curves compress non-linearly as the grid size reduces, but in terms of the basic dynamics of the system, it is a question of scaling. Thus taking a smaller n is justifiable as the simulation time will be shorter while keeping similar dynamics

Figure 5.2b however demonstrates the complexities induced with the contagion mechanisms. While we again have a family of similar curves (a reverse S-curve), they flatten at different times, declining slower as the n of the system decreases. This is a surprising result: one would perhaps expect smaller systems to decline faster as fewer agents need to leave for the system to be depleted. Instead, the contagion effects are amplified by larger ns, meaning the system actually accelerates beyond the increase in agents to stabilise near zero faster than for smaller systems.

However, the underlying shape of the curve is still consistent across the different lattice sizes, and therefore, arguably, the system has not fundamentally changed. We will therefore choose a mid-range between these extremes—30—which is both computationally tractable and large enough for the contagion effects to be strong.

5.3.3 Parameter Search

Having demonstrated the simple features of the model framework, and fixed the lattice size at 30, I then explore the various parameter regimes to find which combinations tend to produce interesting output. By interesting I have two criterion. First, regimes where the decline curve does not stabilise quickly with either all



(a) *Baseline -5, Short-Term 0, Long-Term 0* (b) *Baseline -5, Short-Term 0, Long-Term 0.1*

Figure 5.2: *Effects of Lattice Size*

agents having left or very few agents ever leaving (within the 1000 week window). Second, regimes where the spatial arrangement appears to indicate some non-random structure to the time-series.

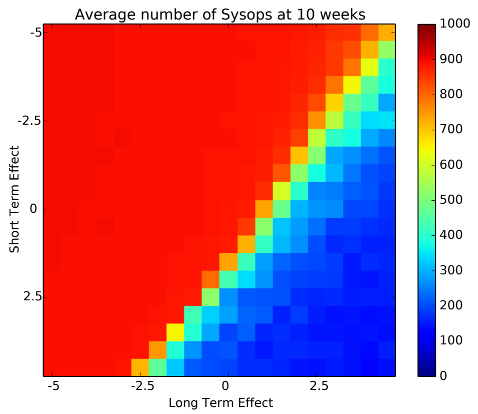
I conducted this search via heat maps of parameter spaces under 3 different baselines: -5 fig. 5.3, -7 fig. 5.4, and -10 fig. 5.5. Each of these groups of plots corresponds to 8,000 simulations: 20 per coordinate on a 20 by 20 surface. I chose these 3 baselines because for positive Cs all agents tend to quickly leave, and from -5 the system begins to last long enough to generate interesting behaviour.[6]

Heat map colors correspond to the mean number of remaining agents under that parameter regime after a given time has elapsed. Blue indicates few of the original n agents are remaining, green is around $\frac{n}{2}$ and blue to little or no agents remaining.

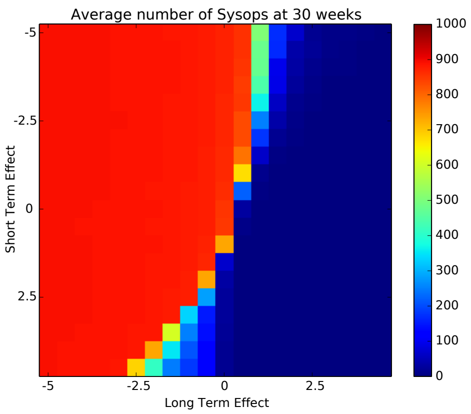
The results demonstrate the variety of behaviours that can emerge even with a very simple model with few parameters. While all three regimes quickly stabilise to no leavers in the top left quadrant—where both short-term and long-term effects are negative—the other quadrants have more variability.

Focusing on the results for $C = -5$ and $C = -7$, the majority of the bottom

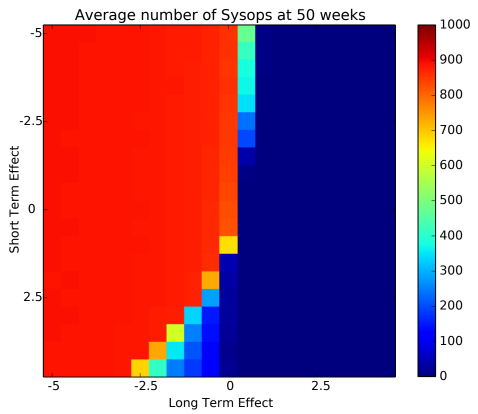
[6] It is always possible to offset a C with contagion parameters that are strong enough, such that the system does not immediately collapse. However, for interesting behaviour at the time scale of 1000 weeks *without other effects*, -5 is a useful starting point.



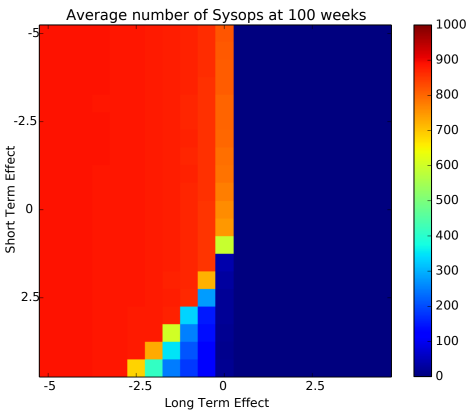
(a) *Average Number after 10 Weeks*



(b) *Average Number after 30 Weeks*

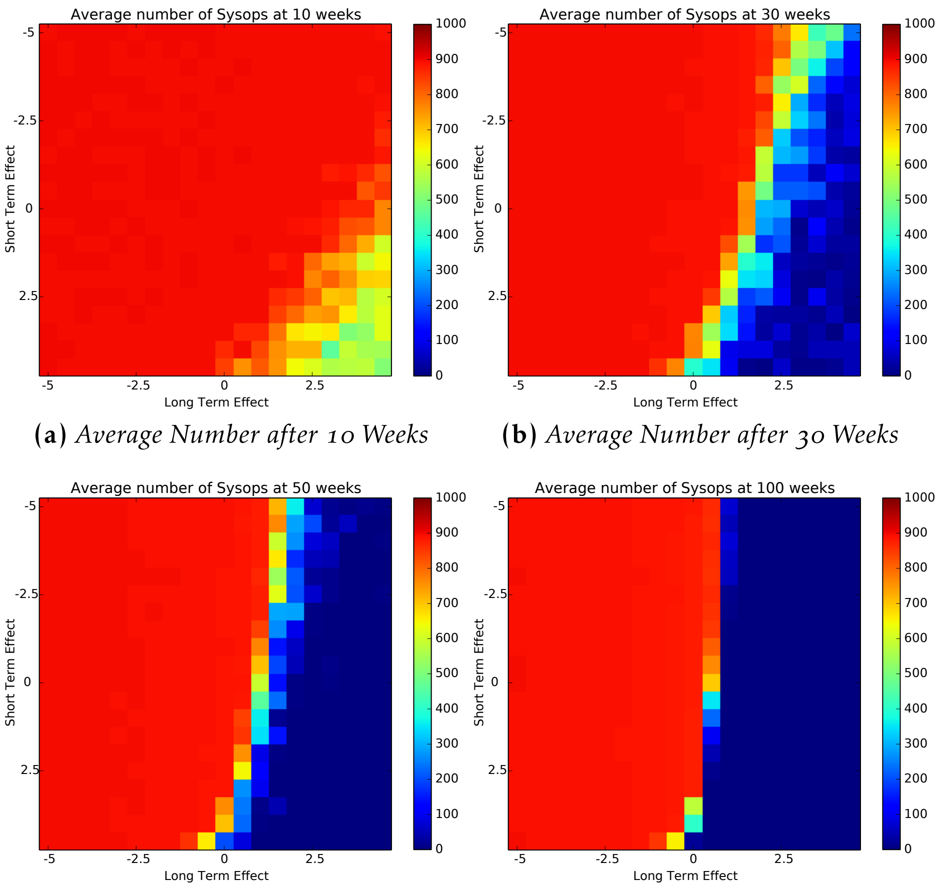


(c) *Average Number after 50 Weeks*



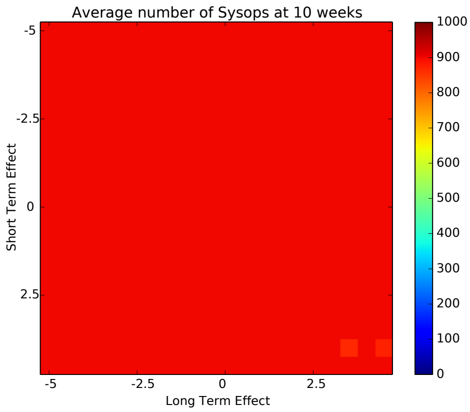
(d) *Average Number after 100 Weeks*

Figure 5.3: *Time Series Heat Maps for* $C = -5$

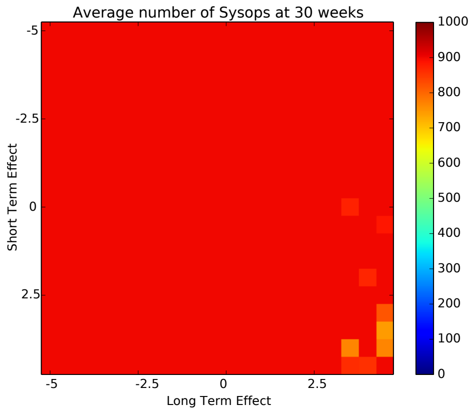


(a) *Average Number after 10 Weeks*
(b) *Average Number after 30 Weeks*
(c) *Average Number after 50 Weeks*
(d) *Average Number after 100 Weeks*

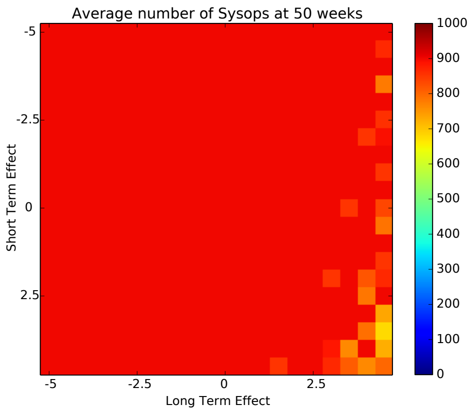
Figure 5.4: *Time Series Heat Maps for* $C = -7$



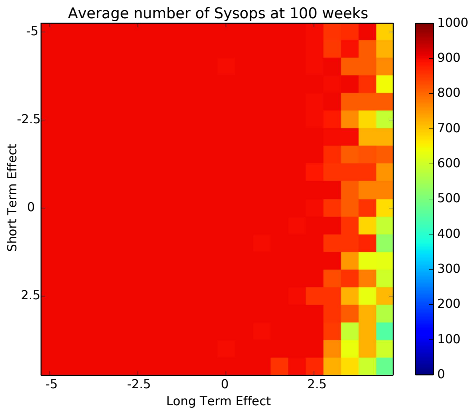
(a) *Average Number after 10 Weeks*



(b) *Average Number after 30 Weeks*



(c) *Average Number after 50 Weeks*



(d) *Average Number after 100 Weeks*

Figure 5.5: *Time Series Heat Maps for* $C = -10$

right quadrant—where both effects are positive—generally stabilises within 50 weeks to most agents leaving, with an edge near $(0, 0)$ surviving longer. That edge corresponds to relatively low values of both parameters $(C < 1)$.

These results are not surprising: in both quadrants the directionality of the effects is the same, meaning the effects are reinforcing each other. As the magnitude of the parameters become larger, the combined effect will non-linearly send the system to extremes of joining or leaving. Very low values will lead to more variation because it will take longer for the interacting effects to amplify each other.

The fact that the upper left quadrant shows little to no variation is also unsurprising because the underlying baseline is negative and both effects are negative, meaning any leaving event is extremely rare. Were the underlying baseline positive, we would expect the sliver of variation around $(0, 0)$ to be in the upper left quadrant, rather than the bottom right.

The fact that the majority of the variation occurs in the other two quadrants speaks to the variability that arises from mechanisms of opposing direction. Similarly, the bottom right quadrant has more variability than the upper left because here the contagion effects are of opposing direction to the underlying baseline, rather than all three reinforcing each other as in the upper left quadrant. For the $C = -10$ plot, the one area with variability within the 1000 week time scale is the bottom right because the baseline is so strongly negative that only when both contagion effects are reinforcing each other, are they able to generate leaving events. Thus the remained of the section will focus on these areas of opposing effects.

Other features of note within the heat maps are the curve like edge between the left and right halves of the plots towards the 50 week time point. For $C = -5$ the curve extends significantly into the region where the long-term effect is negative and the short-term effect is positive while in the $C = -7$ plot the curve just barely reaches into that quadrant. Conversely, the top end of the curve more prominently extends into the top right quadrant on the $C = -7$ plot than the $C = -5$ plot.

I focus on the remainder of the analysis on the $C = -5$ baseline due to the constraints of time and interest in the bottom left quadrant behaviour.

5.3.4 Long Term

The Long-Term contagion effect decline curves cover a relatively small parameter range ($[-.1,.1]$) because of the orientation of the border range on fig. 5.3 along the vertical axis. Small changes generate a great degree of variation since the effect is cumulative over time, growing slowly then becoming much stronger. Note, this section and the Short-Term section after are just these effects in isolation section 5.3.6 covers the interaction between the effects.

For positive values—encouraging leaving—this generates a reverse S-curve shape, slowly decreasing, then accelerating the decline, then asymptotically approaching 0. The cumulative effect grows rapidly, and for all but the lowest values of β_L do we see runs that last longer than 300 weeks.

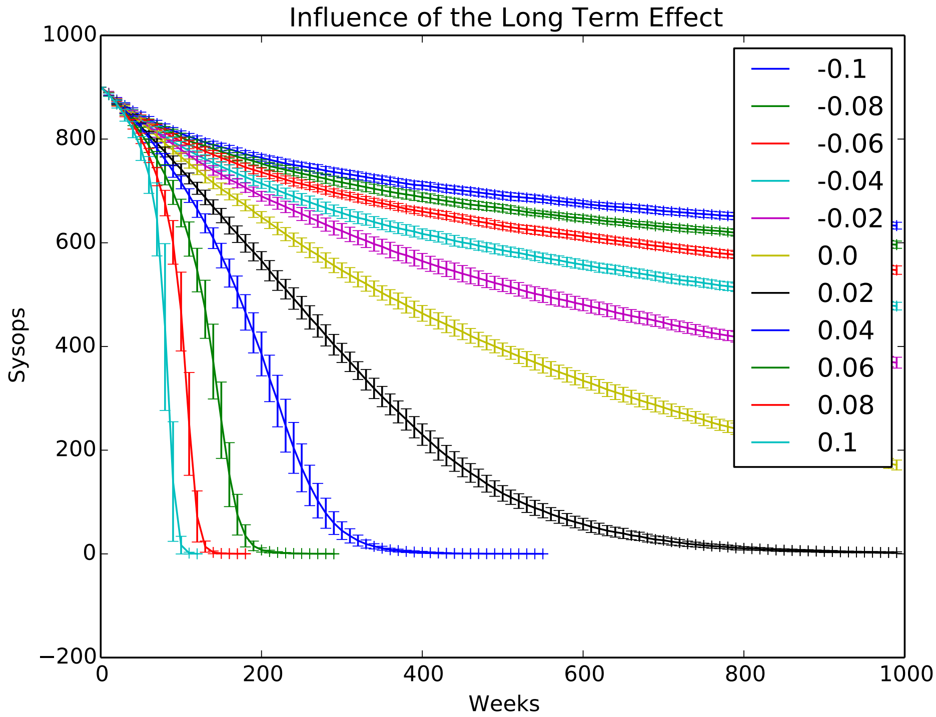
Negative values—discouraging leaving—generates a slow, monotonic decline curve that will outlast the 1000 week barrier. The quick initial decline reflects the cumulative nature of the effect: once enough agents have left, the effect is strong enough to slow down the baseline effect (included for comparison as the yellow middle curve) significantly.

We see also, for both positive and negative values, the non-linear effect of changes, such that as the absolute value of β_L increases, the curves bunch together and conversely as $B_L \to 0$, the difference in the curves widens considerably. Finally, the error bars on these points are quite tight, suggesting the results are highly stable.

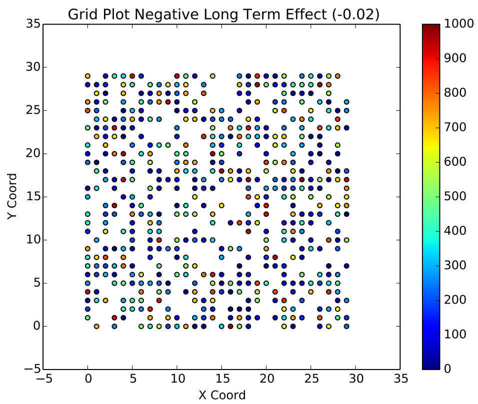
The grid plots are less obviously significant, with the arrangement not qualitatively distinct from random, other than perhaps a small bias for $\beta_L < 0$ for leaving events to be far from the centre as the non-uniformity of a grid (rather than a torus) would imply that the cumulative effect will be on average stronger in the centre.

5.3.5 Short Term

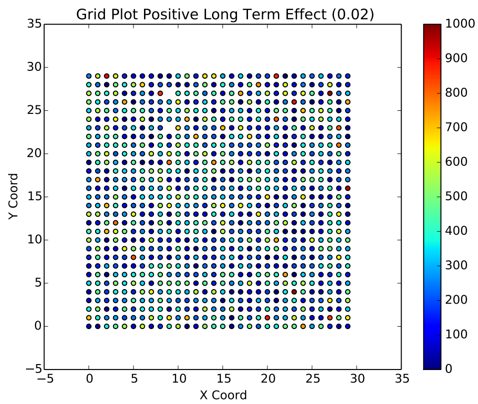
The short term effect has two parameters: p for the length of time a leaving event is still the contagious and β_S for the relative strength of the effect.



(a) *Long-Term Decline Curve*

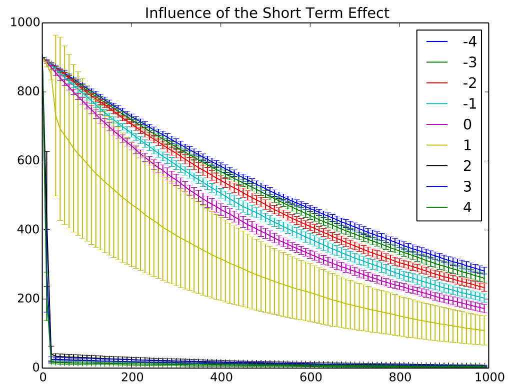


(b) *Long-Term -.02 Grid*

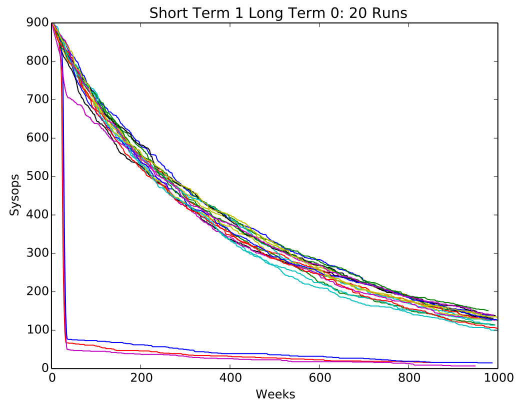


(c) *Long-Term .02 Grid*

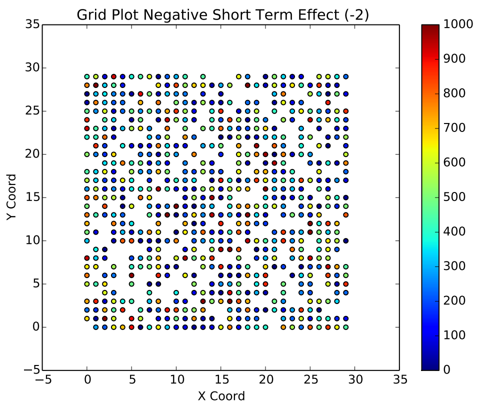
Figure 5.6: *Long-Term Effects with* $C = -5$



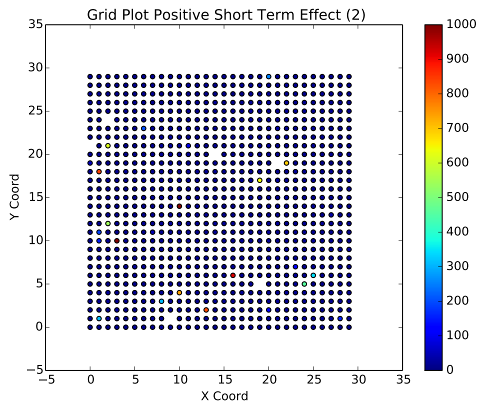
(a) *Short-Term Decline Curve*



(b) *Short-Term Decline Curve for* $\beta_S = 1$



(c) *Short-Term -2 Grid*



(d) *Short-Term 2 Grid*

Figure 5.7: *Short-Term Effects with* $p = 2$, $C = -5$

If we hold p constant at 2 weeks, we can assess the effect of varying β_S. For $\beta_S < 0$, the curves are similar to the long-term curves, again non-linear in the difference between each, quite tight on standard errors and generally slowing the baseline effect depending on their magnitude.

For $\beta_S > 0$, behaviour becomes more unstable. For values greater than 1, the system collapses quickly, much earlier than the long-term curves. This may be due to magnitude. However, when β_S is 1, the curve teeters between the sort of decline curves that we see for negative values and the collapse that occurs for higher positive values. The error bars reflect this variation, and so I include a second plot demonstrating that those correspond primarily to two outcomes rather than a family of options in between.

Once again the grid has no apparent qualitative spatial significance under a negative value. The positive values, however, seem to leave islands of survivors who do not get caught in the initial cascades, leaving some survivors. This relates to the behaviour in the curves where even for high positive values, the system does not completely collapse, and seems to stabilise with isolates that persist.

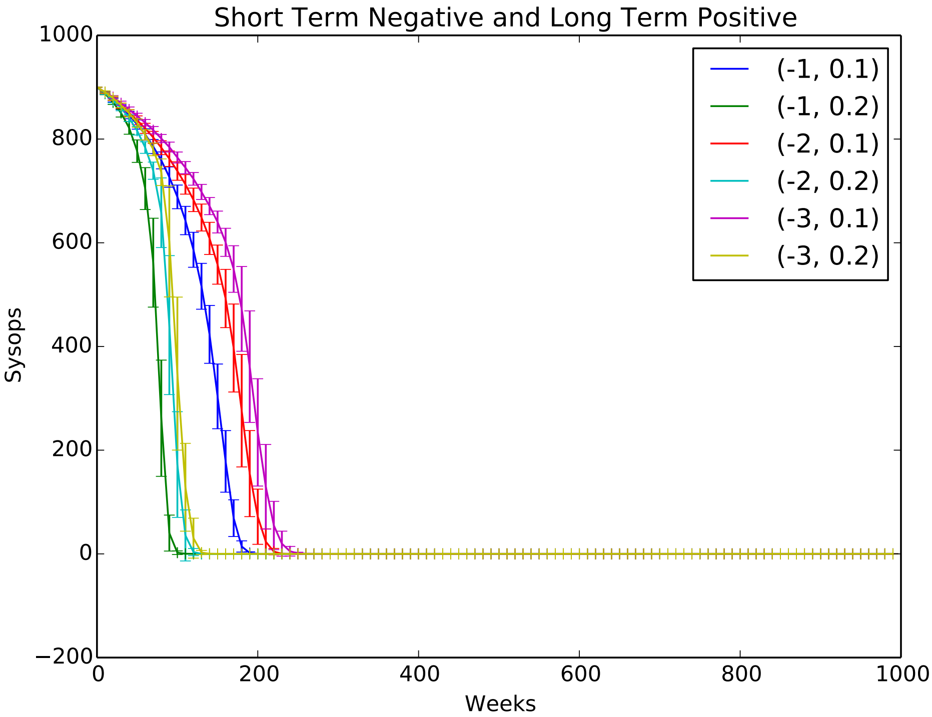
5.3.6 Interacting Effects

[

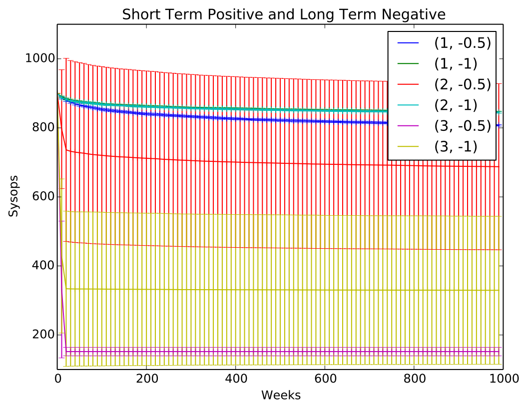
Having covered the effects in isolation, we can now consider the contagion effects interacting with each other.

With the Short-Term effect negative and the Long-Term effect positive we have a stable family of similar curves, each declining in a fairly swift S-Curve fashion and then for especially strong short-term effects a small set of agents survive through to the end, not dissimilar to the behaviour of positive the Short-Term effect. The error bars here are fairly tight again, suggesting these results should be stable, and like in the previous sections the distance between each curve appears to vary non-linearly

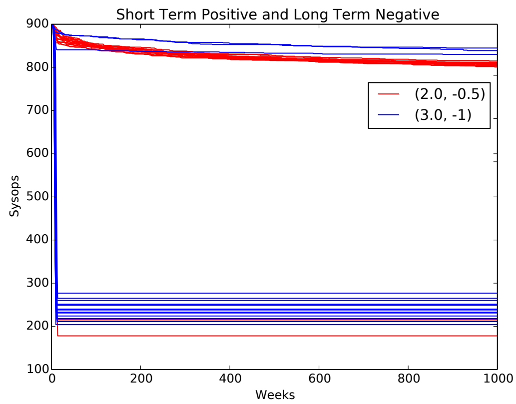
With the short-term effect positive and the long-term effect negative, things are less stable. The cascade behaviour that allows short-term effects to overwhelm the system at high β_S is interrupted by the cumulative long-term effect. The long-term



(a) *Short-Term Negative, Long-Term Positive*



(b) *Short-Term Positive, Long-Term Negative*



(c) *Individual Runs*

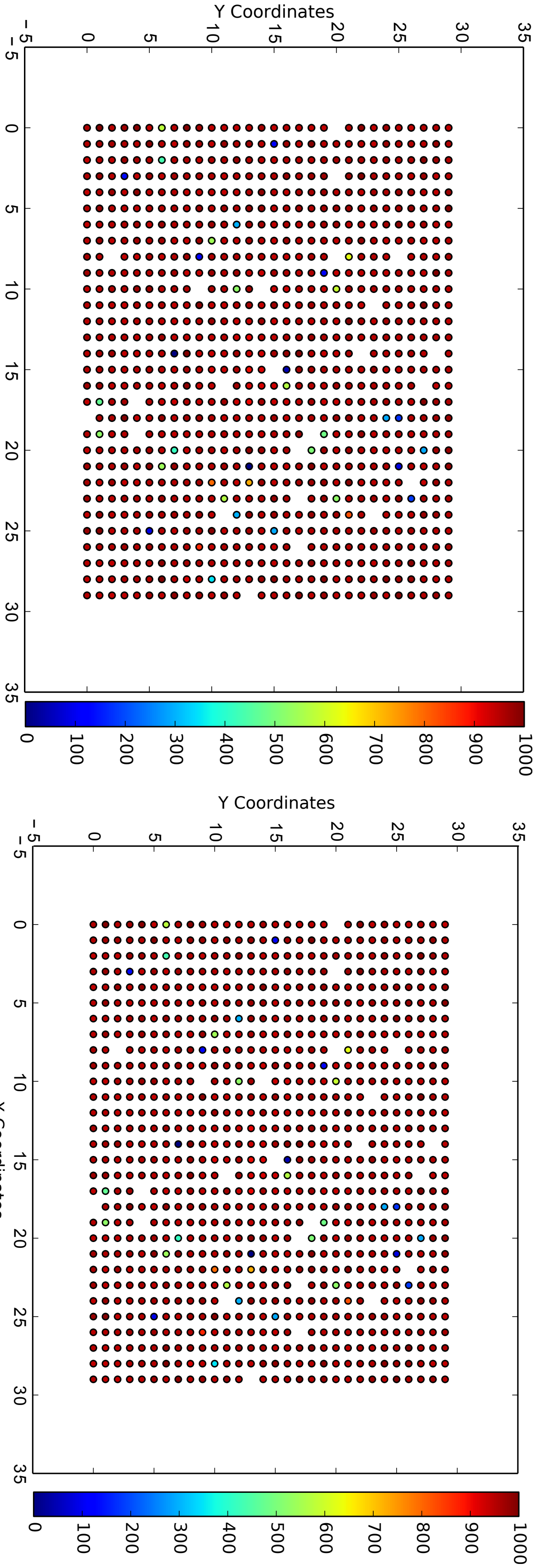
Figure 5.8: *Interaction between Short and Long-Term Mechanisms*

effect begins relatively weak, hence the initial drops, but leavers accumulate it stabilises into a slow asymptotic behaviour that persists through the 1000 week window.

However, for some cases the short-term cascades are able to break through the cumulative effect fig. 5.8c and drop dramatically, stabilising around a new very low point, again probably coinciding with the island behaviour described above in the short-term effect. Once the islands are established, the lack of possible influence from the short-term effect returns these agents to the competition between the baseline probability, and the resilience of the negative long-term effect.

5.4 Conclusion

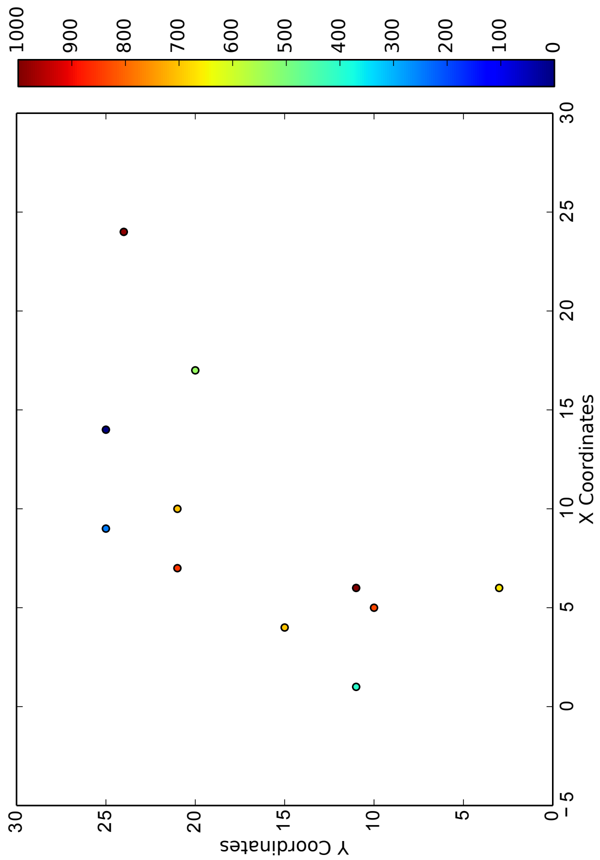
Despite the simplicity of the model proposed, a variety of different behaviours can emerge. In particular the island arrangement that can emerge under the short-term effect seems especially, when positive and combined with the resilience of a negative long-term effect, interesting and worthy of future study. What remains is to further analyse the dynamics and incorporate actual data, potentially following the MLE method proposed in section 5.2.5.



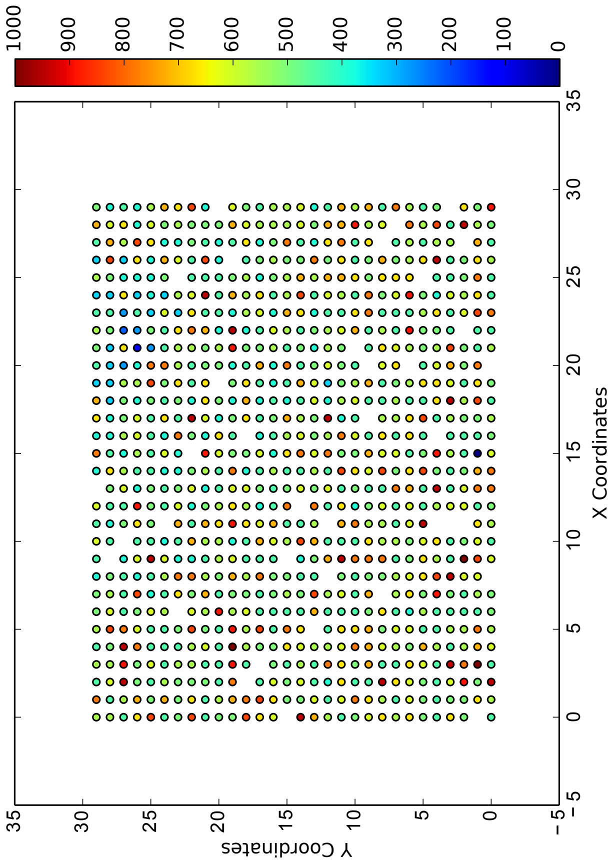
(a) C = −3, *Short-Term 3, Long-Term -0.1*

(b) C = −3, *Short-Term 3, Long-Term -0.4*

Figure 5.9: *Short-Term stable at 3, two examples varying Baseline (C) and Short-Term*

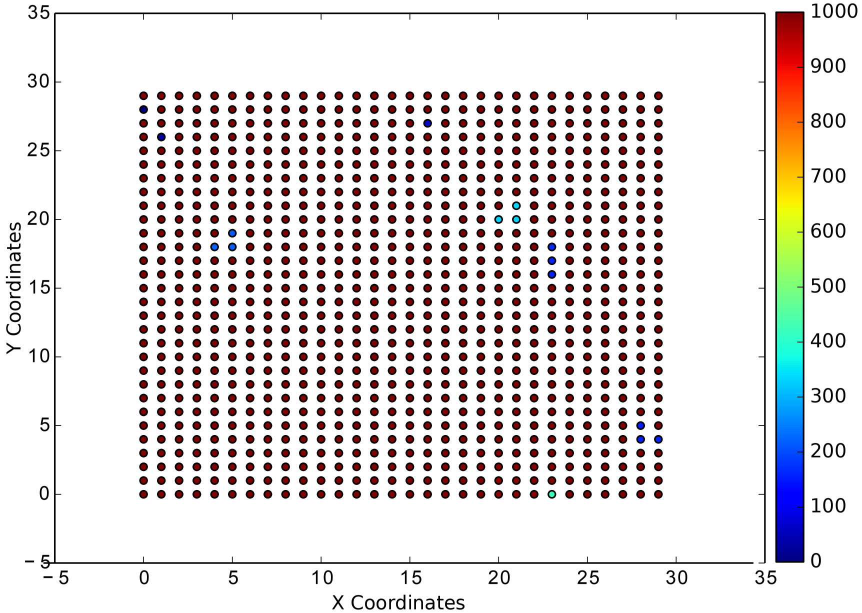


(**a**) C = −5, *Short-Term 5, Long-Term -0.6*

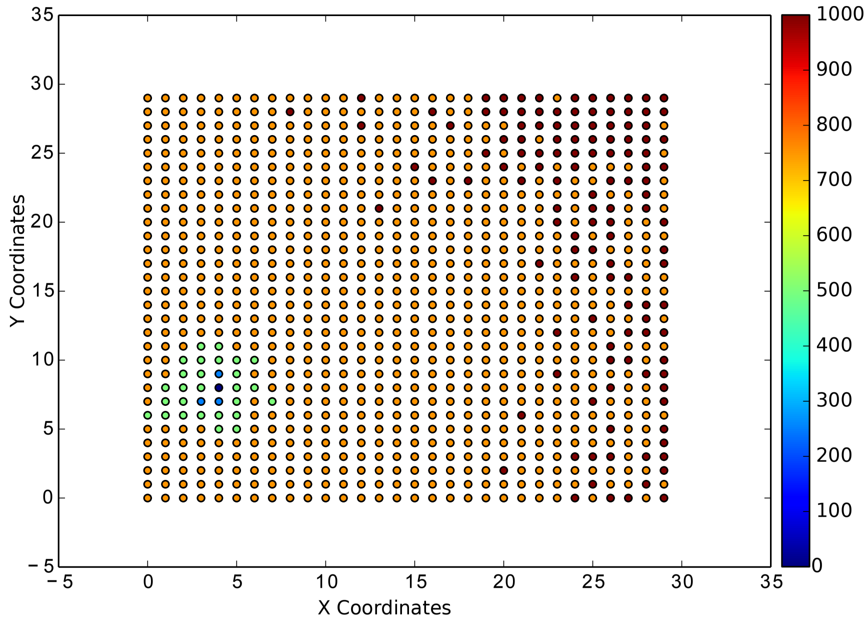


(**b**) C = −5, *Short-Term 5, Long-Term -0.4*

Figure 5.10: *Short-Term stable at 5, two examples varying Baseline (C) and Short-Term*



(a) C = −10, *Short-Term 10, Long-Term -1, Lag 1*



(b) C = −10, *Short-Term 10, Long-Term -0.5, Lag 1*

Figure 5.11: *Short-Term stable at 10 and lag set to 1 (rather than 4), two examples varying Baseline* C *and Short-Term*

Chapter 6

Conclusion

Throughout this thesis I have approached social systems that are inherently dynamic via a micro approach. I think that sociology as a discipline has much to gain from the explosion of micro-level data on human behaviour from various technologies like ICTs and from government institutions that increasingly strive to accurately understand the social systems they govern (Ruggles et al., 2010). With this data, and the computational power that can process and model it, we finally have a chance to model human behaviour as individuals making decisions, rather than as broad categories of people correlated with outcomes. In other words: I firmly believe that micro methods—where possible—are the best way to quantitatively model human behaviour. Below I summarise the problems of macro approaches, give an overview my results, and discuss what future work can build on the contribution of this thesis.

6.1 Micro Over Macro

Schelling (2006) uses an anecdote to neatly demonstrate a key problem with approaching a dynamic social system exclusively at the macro-level. An audience sits in an auditorium, waiting for him to give a lecture. As far as he can tell, the seats are entirely filled from the back, leaving the thirteen closest rows empty. After the lecture, he asks the organiser why seating was restricted in this manner. The organiser replies that there was no restriction; the audience freely chose where to sit.

Why leave so many seats between the lectern and the first occupied row? Schelling posits six alternative hypotheses—all with plausible accompanying narratives—that would require film of the audience sitting down, perhaps a survey, or even a series of double-blind experiments to confirm. In particular: the sequence of individual action—what order seats were chosen and by whom—is extremely helpful in distinguishing between different explanatory mechanisms.

Data on sitting sequence allows one to consider action in a purposive capacity. If I assume actors behave in rational, preference-guided ways that are often dependent on the behaviour of others then I can consider the environment each individual was faced with when they chose to sit and test hypotheses regarding the preferences that governed their choice.

For example: if most people prefer to sit to as far as possible from the lectern, the sequence would show the rows filling up in reverse order from the furthest to the closest occupied. If, however, everyone prefers to sit further than everyone already seated, then the system will highly depend on the choice of those who arrive first.

Of course one mechanism may be insufficient to explain every actor's behaviour: there may be a distribution of preferences that lead to differing behaviour patterns, and attempting to explain the entire system with respect to one mechanism may be misleading. It may be that as time progresses differing mechanisms take over in systematic ways.

Reid and Novak's (1975) results suggest that men tend to select urinals by spatial convenience if alone, but if at least one urinal is already occupied men try to ensure some spatial separation to their nearest neighbour. This suggests convenience and personal space are important factors that potentially conflict, personal space being more important on average. Once again, having data on the circumstances of individual action allow those preferences to be uncovered. Even greater detail—say the foot patterns prior to selection—would provide data on potential indecision, which could indicate that cleanliness also plays a role.

Thus if the objective is to understand what governs human behaviour, focusing exclusively on the macro results robs the analysis of the underlying dynamics of the system which, if understood, could allow for a more accurate extrapolation to other circumstances. If the attendants at Schelling's lecture were keen to sit behind the first person to sit down, getting the audience to sit closer would simply require enticing the first arrival to sit as close as possible. If however, everyone sat as far as possible from the lectern, then the organizers would have to cordon off the back rows of the auditorium to ensure the audience was moved forward *en masse*.

These issues are not purely pragmatic either: they imply strong claims about social norms and interaction in that particular environment, claims worth testing rigorously. If everyone sits as far as possible, that may imply a general disaffection of the student body—perhaps only in attendance because it is compulsory—with the teacher, curriculum, university administration or all three. If however the students are largely keen to sit behind whoever has already sat down, perceived social status within the student body would be more salient.

My claim is not that macro studies are unhelpful, but that—where possible—it is worth understanding a system via its underlying micro-constituents and their dynamics. This requires high quality empirical data which can be expensive to obtain and process, but the potential benefits are considerable.

6.2 My Results

This thesis is an ambitious combination of dataset creation, high time resolution modelling of a very large and long-lived social system, and an exploration into ABM approaches to understanding such systems. My results pose more questions than answers, but suggest that at a minimum the micro-mechanisms that may govern complex systems may require considerably more rigorous and detailed analysis than has been conducted to date.

I will summarise the results of each chapter and then discuss how they relate.

6.2.1 Growth

Chapter 3 is perhaps the most traditional substantive chapter in this thesis, both in terms of topic and methodology. It carries on from the ample work on spatial contagion in growth processes, and following Strang's (1991) utilises the now standard Cox modelling framework to do so.

But what started as a chapter attempting to both explain FidoNet's growth and answer a niggling question found across the discipline—how contagion affects decay with distance—led to a quite complex analysis of the implications of the results. The question became: what effects are implied by the functional form that best fits the data, especially because the effect prominently switches sign before the end of the time-series. It is one thing to have variations in distance decay. It is quite another to have a mechanism that changes direction over time.

Ideally that very question will be pursued in greater detail work to come. The possibility of my proposed repulsion effect, or indeed an entirely distinct effect that also happens to be distance and time dependent, is of particular significance because they hold implications for why social systems stop growing. I expect it is quite likely that leaving mechanisms of the sort implied in chapter 4 will have already begun to slow FidoNet's expansion long before it peaked, and it is the interaction of those mechanisms that most interests me about these results.

The primary research goal of this thesis was to identify micro-level processes and mechanisms that correspond to the macro-level phases of growth and decline in the system. This chapter contributes to this research goal by testing the empirical predictions of hypotheses about social contagion, and how other contextual factors might influence the growth of an ICT.

6.2.2 Decline

Chapter 4 was designed to be as simple as possible. While I did initially consider a host of different distance decay functions as per my approach to chapter 3, the complexities of interpreting those variations quickly became intractable, and it became clear that simply modelling decline as a spatially contagious mechanism—given the dearth of work on the subject—was itself a significant project.

The simple result of my work is that there is evidence to support the claim that spatial contagion can encourage others to leave a systems just as ample work shows it encourages others to join. That in itself is a result worth pursuing. It may be the case, therefore, that rising and falling technologies are not always simply a case of the replacement of something better once the population is made aware of its advantages.

More sociologically salient: that the social ties that bind people into social systems—be they voluntary organisations, communities or perhaps social movements—are means by which people influence each other in a host of ways, positive and negative, encouraging and discouraging, and likely in many different ways all at once.

Of course the details are where the complexity lies. What was meant to be a study of two variants on the same mechanism—contagion in the long term or contagion in the short term—returned an especially surprising result: both significant and of opposing signs. When constructed, those effects were expected to be a question of which was more significant and a better predictor of the model. These were assumed to be variants of the same idea, and one would simply better approximate what actually happened.

As with any complex process and a statistical approach to understanding it, all sorts of other components may have interfered, and without further study it is unclear if there is something wrong with the way the model was constructed, or the database itself. It could be that US demographics changed radically enough in a decade that my control variables are doing more harm than good. It could be that some other mechanism is at work and highly correlated with the effects tested.

Amongst this uncertainty, I found the prospect of two mechanisms of opposing directions both the most plausible interpretation of the results, and the most intriguing. This led to my final substantive chapter.

Similarly to the chapter on the growth of the system, the chapter modelling the decline of FidoNet contributes to the research aim of identifying the micro-level processes that underpin the dynamics of the system. The unexpected empirical result meant that it was necessary to challenge the expectations and hypotheses that were held going into the chapter.

6.2.3 Simulation

For chapter 5 I wanted to strip away the complexity of the previous two chapters, and approach the most essential dynamics of what the implications of chapter 4. I was also motivated by the results of chapter 3, which again implied that what may have actually occurred was a complex interaction of mechanisms of opposing direction.

Therefore I used an ABM to explore the possible interaction of these two mechanisms. I observed interesting non-linearities in the parameter sets. In particular there are certain combinations of parameters under which the system is susceptible to large cascades with feedback loops. This is reminiscent of social systems which experience rapid large-scale decline, and where the leaving of other agents leads to increasing numbers of leaving agents. On the face of it, this does not appear to be the entire story of FidoNet's decline, given that the rate of leaving is not as precipitous as observed in these cascade events. The caveat is that FidoNet may be experiencing both leaving and adoption contagion during the decline phase, such that the rate of decline appears to be artificially small.

The other salient result to emerge from the decline simulation is that there are large regions in the parameter space where the system declines precipitously and also large regions where the system stabilizes such that decline is very small (effectively the baseline probability of a leaving event occurring independent of contagion). Neither of these correspond to the pattern observed in FidoNet. However, the small regions with equilibrium states where the long-term and short-term effects cancel out mostly occur when these effects are of opposite signs. This is consistent with the results observed in chapter 4, but we must be extremely cautious in inferring that these mechanisms are truly the ones at work. Nevertheless, the simulation effectively explores the interaction between the two hypothesized mechanisms.

The ABM achieved the second major research goal, which was to model non-linear interactions of the possible mechanisms that were tested against the real-world data in the previous chapters.

6.2.4 Data contribution

This thesis speaks to a wide range of literature on social decline, ICTs, communities, social contagion and ABM. The final contribution however is the dataset itself. The FidoNet data is a high quality time-series over almost three decades with highly granular geographical information and weekly observations for the majority of the time series. It is a large scale social community over an ICT, and perhaps most importantly it offers an insight into an early instance of technological cooperation. There are plans to make the data publicly available (indeed it was constructed in such a way as to make this as easy as possible), and it is my hope that it will be fully exploited by the wider scientific community.

Appendix A

Data Preparation

Preparing the data for estimating the Cox proportional hazard model was a complex and time consuming process. The data is currently in a Holl and Plum (2009) database, which I connect to using a GeoDjango (Bronn), a Geographic Information System (GIS) object relational mapper (ORM) written in Python. The ORM allows me to manipulate spatial information in Python, a robust, intuitive, and powerful language with a wealth of libraries for data analysis (SciPy (Jones et al., 2001–) and Matplotlib (Hunter, 2007)), and export to other statistical packages such as R (Team, 2011) and Stata (StataCorp, 2009).

I will here detail the method used to construct Exchange Areas and calculate their demographic profile based on data from the 1990 US Census. I will also provide an overview of the various data sources used, and a statistical summary of variables in the dataset.

A.1 Exchange Areas

Exchange Areas are regions which approximate the population served by each US telephone exchange. They are conceptually similar to cells in a Voronoi tessellation such as Figure A.1. A Voronoi tessellation (or decomposition as it is sometimes called) partitions a space (in our case a plane) into regions (cells) closest to an arbitrary set of points or 'sites'. An Exchange Area the area around a telephone exchange which is closer to that exchange than any other. Assuming that exchanges are wired to minimize the distance between homes and buildings and their exchange, this is a reasonable approximation of each exchange's 'service area'. We can then approximate a profile for each exchange using the demographic properties (age, income, occupations and the like) measured in the 1990 US Census.

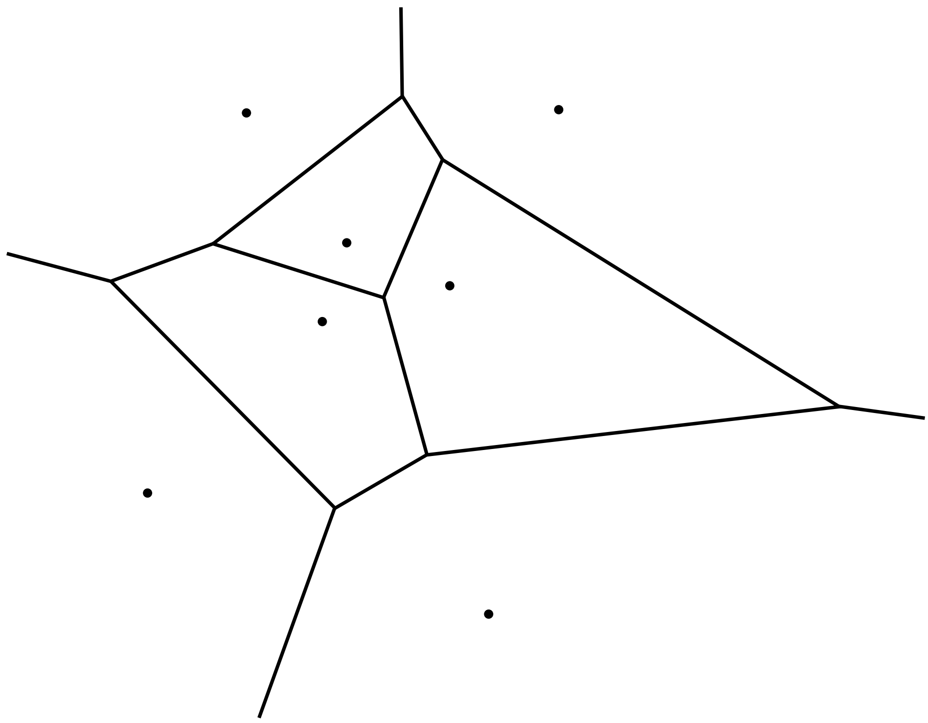


Figure A.1: *An example Voronoi diagram.*

However, spatial US Census data from 1990 is only available at two different spatial resolutions: census tracts and Public Use Microdata Areas (PUMAs). Both resolutions are spatial regions constructed along geopolitical boundaries (city/state lines, rivers and streets), ideally conceived to have approximately homogeneous populations while being large enough to obfuscate the identities of their constituents.[1] Census tracts are meant to have 8000 members and pumas 100000 members. PUMAs are spatial aggregates of tracts and have data not available at the tract level. We used PUMA data to approximate the occupational profile of their constituent census tracts.

Combining census tracts and the concept of a Voronoi cell yields an Exchange

[1] For an explanation of sample weighting for confidentiality see Fellegi (1972).

Area: the set of census tracts whose centroid[2] is closest to a given telephone exchange. Aggregating the demographic profiles of an Exchange Area's constituent tracts allows us to control for varying factors which might impact the likelihood of a BBS emerging from that region. The following sections explain how Exchange Areas were constructed in much greater detail, beginning with telephone exchanges and BBS landline numbers, then census tracts and PUMAs, and finally the actual algorithm for constructing Exchange Areas and the correction factors used to compensate for irregularities.

A.1.1 Telephone Exchanges and the NPA/NXX System

North American telephone numbers since the 1950s have conformed to a system of area codes called the North American Numbering Plan (NANP), which associate the first three digits of a landline number with a set of area codes which correspond to geographic regions called Numbering Plan Areas (NPAs). 301, 410, and 240 are all current NPA codes for Maryland. The next three digits (NXX) are the Central Office Exchange Code, which specify which physical telephone exchange that number is wired to. The final four digits determine the building or household served by that exchange.

Historically, the first six digits of a US landline number uniquely specify which telephone exchange that phone is wired to. However, over the last ten years, the advent of mobile phones and the option of migrating landlines has made telephone exchanges more difficult to determine from landline numbers. Since our time series ends in 1995, we can safely ignore this possibility. FidoNet nodelists include full landline numbers for most active BBSs, and using Dykstra (1998), we can approximate their location.

While our list of 1998 exchanges includes 95,540 unique six digit combinations, only 24,269 unique geographic coordinate pairs exist in the dataset. We therefore combined exchanges with the same coordinates into one exchange, associating

[2] A centroid can be thought of as the arithmetic mean of all the points of an object, and corresponds to the centre of mass of a three dimensional object.

the range of NPA/NXX with the same coordinates together into one entity in our database. Thus the set of potentially at-risk telephone exchanges was 24,269, and we can determine which BBS numbers correspond to these and thereby approximate the location of each BBS.

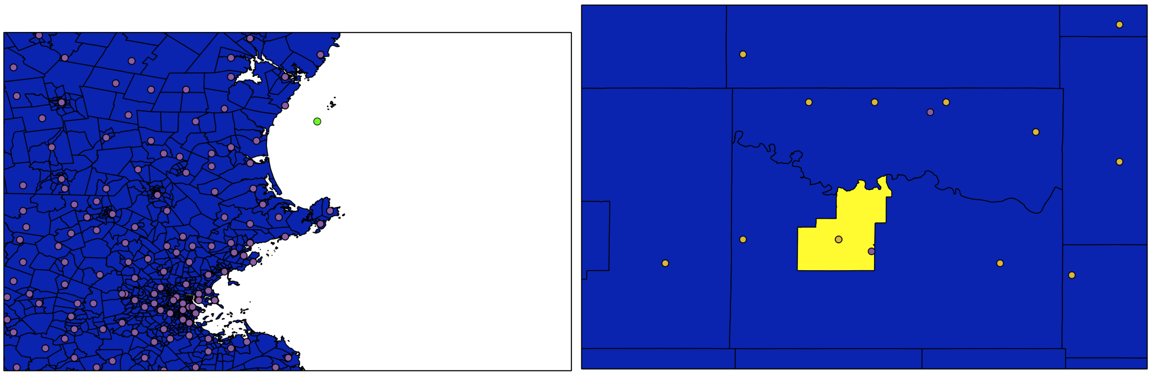
A.1.2 Census Tracts and PUMAs

However, in order to control for the various demographic factors we considered salient we used data from nhg (2011) and Ruggles et al. (2010). Both are standardized, conveniently prepared forms of the geographic data collected from the Bureau (1993). NHGIS is tract level data, meaning aggregated data for populations of approximately 8000 in a geographic region bounded by geo-political borders such as streets, rivers or state/city boundaries. The IPUMS data uses PUMAs, which are composed of tracts and have populations of approximately 100,000.

While we used tracts as our main demographic unit, tract data does not include detailed information on occupations. PUMAs provide this information and since each PUMA is composed of census tracts we approximated the proportion of individuals in each occupation category by the proportion in the entire PUMA that tract is a member of. We similarly used proportions for all other variables where possible, leaving only population density and median as absolute. Thus for each tract each variable value is the proportion of households or individuals in that tract who fulfil the relevant criteria (for example being between ages 14 and 18 or having only graduated from high school), with the exceptions of occupational categories which are approximated from the overall PUMA proportion, the median income and population density.

A.1.3 Combining Tracts, Exchanges, and Exchange Areas

Once the variables for each tract were appropriately calculated, they were associated with telephone exchanges to create EAs. The association process aimed to find the tracts whose centroids were closest to each exchange, but the data required a more complex algorithm. Because some tracts are quite large in area, it was possible for multiple exchanges to be contained within one tract. Further, some of the exchange coordinates fell outside the census shapefiles, essentially in bodies of water (figure A.2a). We could not remove this data because those exchanges have valid NPA/NXX codes and could have adoption events. Finally, some tracts were phsyically surround by a single, larger tract, and this meant that the centroid for the larger EA was actually inside another exchange area (see figure A.2b).



(a) *Exchange in Water* **(b)** *Tract Surrounded by Another Tract*

Figure A.2: *Histograms of Distances Between EAs*

To account for these irregularities we used the following algorithm:

1. Find what US State each NPA/NXX code is officially associated. Restrict all future tract/exchange associations to respect this requirement.

2. For each exchange in a body of water, associate it with the nearest in-state exchange centroid.

3. For all tracts with more than one exchange within its boundaries, combine those exchanges into one geographic point by taking the centroid of the convex hull of their locations. Associates those exchanges with that tract.

4. For all remaining tracts that have exactly one exchange within boundaries, associate that exchange with that tract.

5. For all remaining tracts (without associated exchanges), associate them with the exchange closest to their centroid.

A.2 Data Sources

Description	Source	Date	Description
FidoNet Nodelists	`http://www.textfiles.com`	1983-2004	Dated lists of FidoNet BBSs
Telephone Exchanges	Dykstra (1998)	1998	95,540 6-digit NPA/NXX codes with 24,269 unique geographic coordinates
Census Tracts	nhg (2011)	1990	61,693 geo-politically bounded regions with approximate population of 8000
PUMAs	Ruggles et al. (2010)	1990	1713 geo-politically bounded regions with approximate population of 100,000. Composed of Census Tracts

A.3 Variables

Variable	Min	q_1	$\widetilde{x}$	$\bar{x}$	q_3	Max	sd	IQR
(1) $\log_{10}$ Population Density	−8.478	−5.014	4.519	−4.402	−3.767	−1.330	0.921	1.246
(2) Urbanity	0.000	0.000	0.134	0.381	0.849	1.000	0.418	0.849
(3) Median Income ($10,000)	0.536	2.088	2.547	2.783	3.225	11.700	1.050	1.136
(4) Age 14 to 18	0.000	0.063	0.072	0.072	0.081	0.237	0.015	0.017
(5) Age 19 to 24	0.000	0.125	0.145	0.151	0.166	0.722	0.050	0.041
(6) Age 25 to 29	0.000	0.063	0.073	0.075	0.083	0.319	0.020	0.021
(7) Age 30 to 39	0.000	0.145	0.160	0.161	0.176	0.535	0.028	0.031
(8) Age 40 to 49	0.000	0.112	0.126	0.127	0.141	0.352	0.024	0.029
(9) Age 50 to 59	0.000	0.082	0.093	0.094	0.105	0.264	0.019	0.022
(10) Age 60 and Over	0.000	0.145	0.181	0.186	0.220	1.000	0.066	0.076
(11) High School Graduate	0.000	0.295	0.347	0.345	0.399	1.000	0.081	0.104
(12) Undergraduate Degree	0.000	0.247	0.318	0.334	0.407	1.000	0.116	0.160
(13) Post-Graduate Degree	0.000	0.023	0.034	0.045	0.054	0.391	0.038	0.031
(14) Managers	0.008	0.029	0.034	0.037	0.042	0.140	0.014	0.013
(15) Engineers	0.000	0.003	0.004	0.005	0.006	0.037	0.004	0.004
(16) Scientists	0.000	0.001	0.001	0.002	0.002	0.017	0.001	0.001
(17) Teachers	0.000	0.001	0.002	0.003	0.003	0.041	0.004	0.002
(18) Technicians	0.001	0.003	0.005	0.005	0.006	0.018	0.002	0.003
(19) $\sum \frac{1}{d_{ij}\gamma_{j,t-1}}$	0.034	1.312	2.597	2.860	4.267	14.066	1.950	2.954
(20) $(\sum \frac{1}{d_{ij}\gamma_{j,t-1}})^2$	0.001	1.723	6.742	11.985	18.207	197.854	13.877	16.484

Table A.1: *Summary statistics of variables.*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
(2)	0.79***																		
(3)	0.42***	0.33***																	
(4)	-0.28***	-0.26***	-0.20***																
(5)	0.45***	0.41***	0.06***	-0.07***															
(6)	0.43***	0.38***	0.15***	-0.27***	0.69***														
(7)	0.24***	0.17***	0.46***	-0.19***	0.13***	0.41***													
(8)	0.05***	-0.06***	0.58***	0.00***	-0.24***	-0.10***	0.30***												
(9)	-0.25***	-0.27***	0.04***	-0.10***	-0.48***	-0.39***	-0.29***	0.20***											
(10)	-0.19***	-0.13***	-0.35***	-0.34***	-0.48***	-0.49***	-0.59***	-0.30***	0.33***										
(11)	-0.36***	-0.45***	-0.29***	0.14***	-0.31***	-0.21***	-0.14***	-0.11***	0.16***	0.17***									
(12)	0.37***	0.45***	0.65***	-0.30***	0.25***	0.20***	0.38***	0.30***	-0.16***	-0.24***	-0.56***								
(13)	0.43***	0.40***	0.69***	-0.32***	0.18***	0.12***	0.26***	0.39***	0.01***	-0.12***	-0.58***	0.68***							
(14)	0.29***	0.24***	0.60***	-0.29***	0.11***	0.19***	0.32***	0.36***	0.01***	-0.16***	-0.29***	0.51***	0.57***						
(15)	0.35***	0.27***	0.58***	-0.17***	0.16***	0.22***	0.34***	0.33***	-0.05***	-0.28***	-0.24***	0.45***	0.44***	0.68***					
(16)	0.02***	0.09***	0.22***	-0.12***	0.07***	0.08***	0.18***	0.16***	-0.03***	-0.12***	-0.23***	0.31***	0.34***	0.46***	0.40***				
(17)	0.07***	0.05***	0.07***	-0.08***	0.18***	0.07***	0.07***	0.02***	-0.08***	-0.06***	-0.11***	0.20***	0.26***	0.26***	0.19***	0.44***			
(18)	0.18***	0.13***	0.32***	-0.12***	0.17***	0.17***	0.24***	0.18***	-0.07***	-0.22***	-0.07***	0.28***	0.24***	0.51***	0.67***	0.50***	0.40***		
(19)	0.13***	0.06***	0.07***	-0.03***	0.06***	0.05***	0.03***	0.03***	0.00	-0.03***	0.02***	-0.02***	0.05***	0.05***	0.05***	-0.03***	0.01***	0.03***	
(20)	0.14***	0.07***	0.08***	-0.03***	0.07***	0.06***	0.03***	0.04***	0.00**	-0.03***	0.02***	-0.02***	0.06***	0.06***	0.06***	-0.03***	0.01***	0.03***	0.94***

Table A.2: *Correlation Matrix of Variables in the Model*

A.4 Density Model

We also considered a density effect as part of the DMCA, whereby the effect of an adopted EA at time t was multiplied by number of BBSs in that EA at t. In this model $\gamma_{j,t-1}$ merely becomes the number of BBSs in EA j at time $t-1$. We rejected this model because the fit was substantially worse than our preferred model. Table A.3 presents these results.

	Non-Density DMCA				*Density* DMCA			
	Linear		Quadratic		Linear		Quardratic	
$\log_{10}$ Population Density	0.674***	(27.20)	0.682***	(27.49)	0.675***	(27.30)	0.679***	(27.39)
Urbanity	0.821***	(15.44)	0.789***	(14.85)	0.821***	(15.44)	0.796***	(14.97)
Median Income ($10,000)	−0.0191	(−0.90)	−0.0275	(−1.30)	−0.0189	(−0.89)	−0.0224	(−1.06)
Age 14 to 18	−10.97***	(−10.71)	−13.32***	(−12.47)	−10.89***	(−10.71)	−12.81***	(−12.12)
Age 19 to 24	−0.908*	(−2.42)	−1.617***	(−4.18)	−0.882*	(−2.36)	−1.452***	(−3.79)
Age 25 to 29	−2.001*	(−2.14)	−2.336*	(−2.48)	−1.992*	(−2.13)	−2.280*	(−2.43)
Age 30 to 39	−2.229***	(−3.52)	−3.414***	(−5.24)	−2.185***	(−3.48)	−3.182***	(−4.92)
Age 40 to 49	1.022	(1.41)	0.316	(0.43)	1.035	(1.43)	0.441	(0.60)
Age 50 to 59	−1.943*	(−2.35)	−2.499**	(−3.02)	−1.914*	(−2.32)	−2.453**	(−2.97)
Age 60 and Over	−3.648***	(−12.71)	−4.419***	(−14.82)	−3.626***	(−12.75)	−4.236***	(−14.35)
High School Graduate	−0.219	(−0.95)	−0.375	(−1.61)	−0.203	(−0.88)	−0.340	(−1.46)
Undergraduate Degree	1.834***	(9.90)	1.911***	(10.27)	1.826***	(9.88)	1.895***	(10.21)
Post-Graduate Degree	−1.843**	(−3.23)	−1.904***	(−3.36)	−1.816**	(−3.19)	−1.926***	(−3.39)
Managers	−0.852	(−0.68)	−0.651	(−0.53)	−0.810	(−0.65)	−0.762	(−0.62)
Engineers	13.55**	(3.04)	12.93**	(2.92)	13.34**	(3.00)	13.17**	(2.96)
Scientists	−12.42	(−1.10)	−9.005	(−0.81)	−12.74	(−1.13)	−10.40	(−0.94)
Teachers	−4.067	(−0.98)	−4.218	(−1.02)	−4.031	(−0.98)	−3.972	(−0.96)
Technicians	41.57***	(5.59)	40.09***	(5.42)	41.81***	(5.62)	40.50***	(5.47)
Previous Adoption	0.548***	(17.00)	0.548***	(17.13)	0.548***	(16.98)	0.544***	(16.94)
Linear	0.0282*	(2.09)	0.395***	(8.53)				
Squared			−0.0356***	(−8.12)				
Density Linear					0.00868*	(2.11)	0.0992***	(7.31)
Density Squared							−0.00255***	(−6.85)
AIC	61734.9		**61651.4**		61734.8		61677.0	
BIC	62246.0		**62173.4**		62246.0		62199.0	

Reference Category: non-High School Graduates — 390,270 **observations**, *t* statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.3: *Comparison of Density* DMCA *effect and non-density* DMCA *effect. To avoid confusion, the models are specified by their highest order term (Linear means* $k = \beta_\kappa \sum \frac{\gamma_{j,t-1}}{d_{ij}}$*, and Quadratic adds the squared term), while the variables specify whether they are density or not (Squared and Density Squared, for example).*

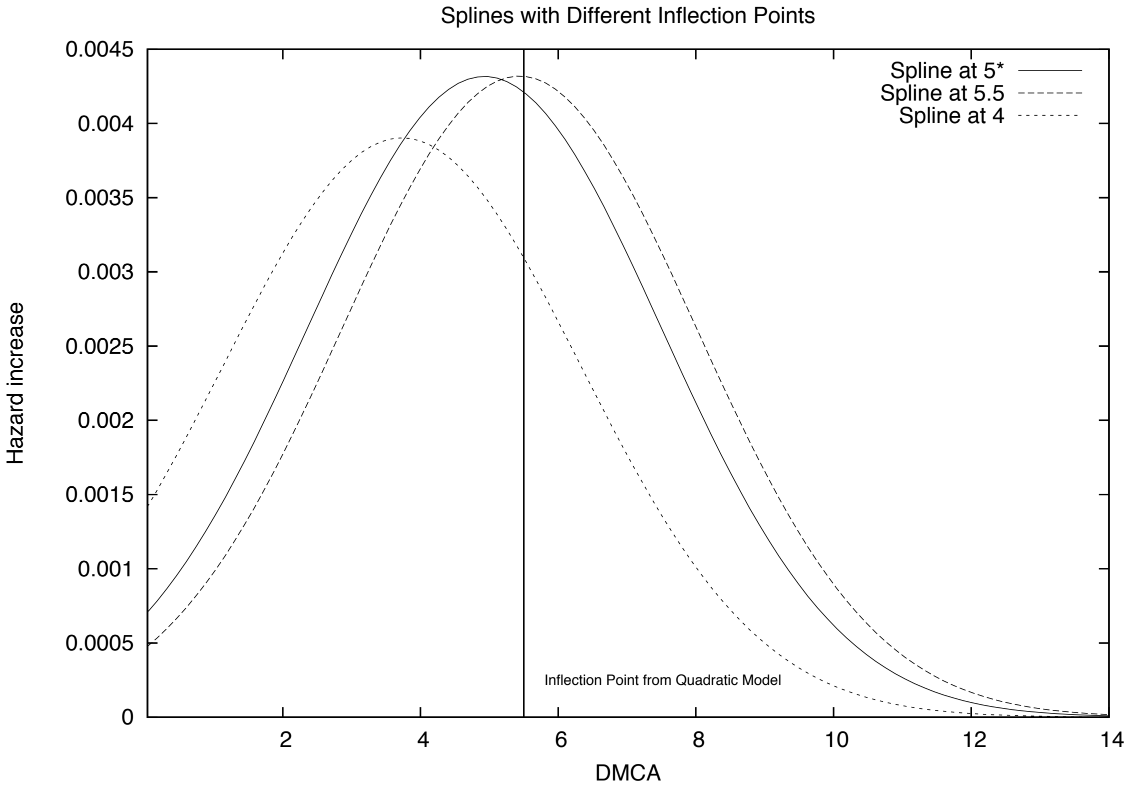


Figure A.3: *Note these curves are very similar in shape to the quadratic model, and these inflection points were chosen because they were close to the quadratic inflection point.*

A.5 Splines

We are indebted to Tom Snijders for suggesting we try quadratic splines instead of a polynomial, because polynomial functions can, especially at high orders, over-fit in spurious ways

	Quadratic		Splined at 4		Splined at 5		Splined at 5.5	
$\log_{10}$ Population Density	0.766***	(32.89)	0.719***	(30.55)	0.715***	(30.28)	0.715***	(30.30)
Urbanity	0.675***	(13.00)	0.736***	(14.15)	0.741***	(14.23)	0.741***	(14.22)
Median Income ($10,000)	−0.0537**	(−2.60)	−0.0415*	(−2.02)	−0.0405*	(−1.98)	−0.0406*	(−1.98)
Age Under 16	−6.312***	(−7.36)	−5.242***	(−6.25)	−5.147***	(−6.13)	−5.155***	(−6.14)
Age 16 to 19	−0.929*	(−2.37)	−0.452	(−1.17)	−0.411	(−1.06)	−0.413	(−1.07)
Age 20 to 29	−3.750***	(−5.90)	−2.568***	(−4.15)	−2.463***	(−3.97)	−2.470***	(−3.98)
Age 30 to 39	−0.261	(−0.35)	0.423	(0.57)	0.485	(0.65)	0.480	(0.65)
Age 40 to 49	−2.741**	(−3.22)	−2.151*	(−2.55)	−2.101*	(−2.49)	−2.105*	(−2.50)
Age 50 to 59	−3.497***	(−12.39)	−2.867***	(−10.36)	−2.812***	(−10.12)	−2.816***	(−10.14)
Age 60 and Over	−0.872***	(−4.20)	−0.511*	(−2.46)	−0.483*	(−2.33)	−0.484*	(−2.33)
Below High School	1.737***	(7.39)	1.983***	(8.40)	2.002***	(8.48)	2.001***	(8.47)
Undergraduate Degree	−1.009*	(−2.06)	−1.315**	(−2.69)	−1.344**	(−2.74)	−1.342**	(−2.74)
Post-Graduate Degree	−0.390	(−0.31)	−0.659	(−0.53)	−0.696	(−0.56)	−0.693	(−0.56)
Managers	13.34**	(2.97)	13.11**	(2.94)	13.16**	(2.95)	13.16**	(2.95)
Engineers	−5.902	(−0.52)	−6.952	(−0.61)	−7.153	(−0.63)	−7.150	(−0.63)
Scientists	−5.464	(−1.32)	−4.740	(−1.15)	−4.656	(−1.13)	−4.663	(−1.14)
Teachers	36.11***	(4.83)	38.04***	(5.12)	38.22***	(5.15)	38.21***	(5.14)
Technicians	0.542***	(16.76)	0.547***	(17.00)	0.548***	(17.02)	0.548***	(17.02)
Previous Adoption	0.313***	(6.99)	−0.0400	(−1.81)	−0.0566*	(−2.44)	−0.0553*	(−2.39)
Distance_2	−0.0297***	(−6.99)						
DistanceBelow4			−0.0627***	(−8.17)				
DistanceAbove4			−0.0115	(−1.31)				
DistanceBelow5					−0.0690***	(−8.43)		
DistanceAbove5					−0.00584	(−0.67)		
DistanceBelow5_5							−0.0687***	(−8.41)
DistanceAbove5_5							−0.00640	(−0.73)
AIC	61724.8		61676.7		**61671.7**		61672.1	
BIC	62235.9		62198.7		**62193.7**		62194.0	

390,270 **observations**, t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.4: *Comparison of Spline Inflection Points for DMCA Effect*

	Linear		Quadratic		Splined	
$\log_{10}$ Population Density	0.655***	(24.24)	0.645***	(23.77)	0.643***	(23.67)
Urbanity	0.847***	(15.03)	0.848***	(15.07)	0.849***	(15.08)
Median Income ($10,000)	−0.00997	(−0.47)	−0.0156	(−0.75)	−0.0159	(−0.76)
Age Under 16	−2.325***	(−5.98)	−3.328***	(−8.17)	−2.074***	(−5.16)
Age 16 to 19	−5.151***	(−6.30)	−6.464***	(−7.73)	−5.203***	(−6.27)
Age 20 to 29	−1.597***	(−3.73)	−2.797***	(−6.20)	−1.548***	(−3.52)
Age 30 to 39	−2.351***	(−3.87)	−3.529***	(−5.66)	−2.294***	(−3.72)
Age 40 to 49	−0.349	(−0.47)	−1.556*	(−2.06)	−0.291	(−0.39)
Age 50 to 59	−3.145***	(−3.72)	−4.177***	(−4.90)	−2.964***	(−3.49)
Age 60 and Over	−3.849***	(−12.72)	−4.956***	(−15.25)	−3.704***	(−11.65)
Below High School	0.0510	(0.22)	0.167	(0.71)	0.155	(0.66)
Undergraduate Degree	2.138***	(8.81)	2.354***	(9.64)	2.355***	(9.64)
Post-Graduate Degree	−1.723***	(−3.41)	−1.748***	(−3.49)	−1.783***	(−3.56)
Managers	−1.070	(−0.85)	−0.980	(−0.79)	−1.044	(−0.85)
Engineers	12.37**	(2.78)	11.51**	(2.60)	11.94**	(2.70)
Scientists	−12.11	(−1.05)	−8.003	(−0.70)	−8.640	(−0.76)
Teachers	−3.848	(−0.94)	−4.011	(−0.98)	−3.832	(−0.94)
Technicians	43.59***	(5.83)	42.38***	(5.70)	42.30***	(5.69)
Previous Adoption	0.552***	(16.98)	0.556***	(17.27)	0.556***	(17.28)
Distance	0.0300*	(2.22)	0.436***	(9.09)	−0.0108	(−0.44)
Distance2			−0.0391***	(−8.64)		
below					−0.0587***	(−6.99)
above					−0.0169	(−1.80)
AIC	61745.6		61649.2		**61643.6**	
BIC	62256.7		**62171.2**		62176.4	

390,270 **observations**, t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: *Comparing Polynomial and Spline DMCA Estimation*

References

Technology (A Special Report) — A Sampler, 1991. URL http://search.proquest.com/docview/398331212?accountid=13042. Copyright - Copyright Dow Jones & Company Inc Oct 21, 1991; Last updated - 2010-06-26; DOI - 27777584; 254738; 7510; WSJ; J9415527216.

National Historical Geographic Information System: Version 2.0, 2011. URL http://www.nhgis.org.

Yvonne Åberg. *The Contagiousness of Divorce*, chapter 15, pages 342–364. Oxford University Press, hardback edition, December 2009. ISBN 978-0-19-921536-2. URL http://books.google.com/books?hl=en&lr=&id=Tqf718wB2CUC&oi=fnd&pg=PT415&dq=%22The+contagiousness+of+divorce%22+aberg&ots=hcPyMGdzXf&sig=8kfnZeKu89FIOnO62YA96HsuuFM.

Eric Abrahamson and Lori Rosenkopf. Institutional and Competitive Bandwagons: Using Mathematical Modeling as a Tool to Explore Innovation Diffusion. *The Academy of Management Review*, 18(3):487–517, 1993. ISSN 03637425. doi: 10.2307/258906. URL http://dx.doi.org/10.2307/258906.

Paul D. Allison. *Event history analysis : regression for longitudinal event data*. Sage Publications, 1984. ISBN 9780803920552. URL http://www.worldcat.org/isbn/9780803920552.

Benedict R. Anderson. *Imagined communities : reflections on the origin and spread of nationalism*. Verso, 1983. ISBN 9780860913290. URL http://www.worldcat.org/isbn/9780860913290.

James Andreoni. Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal*, 100(401):464–477, 1990. ISSN 00130133. doi: 10.2307/2234133. URL http://dx.doi.org/10.2307/2234133.

Kenneth T. Andrews and Michael Biggs. The Dynamics of Protest Diffusion: Movement Organizations, Social Networks, and News Media in the 1960 Sit-Ins. *American Sociological Review*, 71(5):752–777, October 2006. ISSN 0003-1224. doi: 10.1177/000312240607100503. URL http://dx.doi.org/10.1177/000312240607100503.

Robert Axtell. Effects of Interaction Topology and Activation Regime in Several Multi-Agent Systems. In Scott Moss and Paul Davidsson, editors, *Multi-Agent-Based Simulation*, volume 1979 of *Lecture Notes in Computer Science*, chapter 3, pages 33–48. Springer Berlin / Heidelberg, Berlin, Heidelberg, June 2001. ISBN 978-3-540-41522-0. doi: 10.1007/3-540-44561-7_3. URL http://dx.doi.org/10.1007/3-540-44561-7_3.

Robert L. Axtell, Joshua M. Epstein, Jeffrey S. Dean, George J. Gumerman, Alan C. Swedlund, Jason Harburger, Shubha Chakravarty, Ross Hammond, Jon Parker, and Miles Parker. Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7275–7279, May 2002. ISSN 1091-6490. doi: 10.1073/pnas.092080799. URL http://dx.doi.org/10.1073/pnas.092080799.

Venkatesh Bala and Sanjeev Goyal. Learning from Neighbours. *The Review of Economic Studies*, 65(3):595–621, 1998. ISSN 00346527. doi: 10.2307/2566940. URL http://dx.doi.org/10.2307/2566940.

A. Banerjee and D. Fudenberg. Word-of-mouth learning. *Games and Economic Behavior*, 46(1):1–22, January 2004. ISSN 0899-8256. doi: 10.1016/s0899-8256(03)00048-4. URL http://dx.doi.org/10.1016/s0899-8256(03)00048-4.

Abhijit V. Banerjee. A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, 107(3):797–817, August 1992. ISSN 1531-4650. doi: 10.2307/2118364. URL http://dx.doi.org/10.2307/2118364.

Frank M. Bass. New product growth for model consumer durables. *Management Science*, 15(5):217–227, January 1969.

Gary S. Becker. A Theory of Social Interactions. Technical Report 42, National Bureau of Economic Research, June 1974. URL http://www.nber.org/papers/w0042.

Colin Bell and Howard Newby, editors. *Sociology of Community: A Collection of Readings (New Sociology Library,)*. Routledge, 1 edition, September 1974. ISBN 0714629707. URL http://www.worldcat.org/isbn/0714629707.

Michael Biggs. Positive feedback in collective mobilization: The American strike wave of 1886. *Theory and Society*, 32(2):217–254, April 2003. doi: 10.1023/a:1023905019461. URL http://dx.doi.org/10.1023/a:1023905019461.

Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100(5):992+, January 1992. doi: 10.1086/261849. URL http://dx.doi.org/10.1086/261849.

Tony Blackshaw. *Key Concepts in Community Studies (SAGE Key Concepts series)*. Sage Publications Ltd, first edition edition, November 2009. ISBN 1412928443. URL http://www.worldcat.org/isbn/1412928443.

Ben Bolker and R Development Core Team. *bbmle: Tools for General Maximum Likelihood Estimation*, 2016. URL https://CRAN.R-project.org/package=bbmle. R package version 1.0.18.

Charles Booth. *Life and Labour of the People in London: Summary - Primary Source Edition*. London: Macmillan, February 1889. ISBN 1293612642. URL http://www.worldcat.org/isbn/1293612642. see http://booth.lse.ac.uk/static/a/3.html.

Justin Bronn. *GeoDjango*. Django Software Foundation.

Steve Bruce. *Secularization: In Defence of an Unfashionable Theory*. Oxford University Press, reprint edition, March 2013. ISBN 0199654123. URL http://www.worldcat.org/isbn/0199654123.

U. S. Census Bureau. Census of Population and Housing, 1990 [United States]: Public Use Microdata Sample: 5-Percent Sample, 1993. URL http://dx.doi.org/10.3886/ICPSR09952.

Robert J. Bursik. Social disorganization and theories of crime and delinquency: Problems and prospects*. *Criminology*, 26(4):519–552, November 1988. doi: 10.1111/j.1745-9125.1988.tb00854.x. URL http://dx.doi.org/10.1111/j.1745-9125.1988.tb00854.x.

Ronald S. Burt. Social Contagion and Innovation: Cohesion Versus Structural Equivalence. *American Journal of Sociology*, 92(6):1287–1335, 1987. ISSN 00029602. doi: 10.2307/2779839. URL http://dx.doi.org/10.2307/2779839.

Ronald S. Burt. Structural Holes versus Network Closure as Social Capital. In Nan Lin, Ronald S. Burt, and Karen Cook, editors, *Social Capital: Theory and Research*, chapter 2, pages 31–56. Aldine de Gruyter,US, 2001. ISBN 0202306445. URL http://faculty.chicagobooth.edu/ronald.burt/research/SCSH.pdf.

Ronald S. Burt. *Brokerage and Closure: An Introduction to Social Capital*. Oxford University Press, USA, October 2005. ISBN 0199249148. URL http://www.worldcat.org/isbn/0199249148.

Randy Bush. FidoNet: technology, tools, and history. *Commun. ACM*, 36(8):31–35, 1993. ISSN 0001-0782. doi: 10.1145/163381.163383. URL http://dx.doi.org/10.1145/163381.163383.

M. Castells. *The rise of the network society*. Blackwell, 1996. ISBN 9781557866172. URL http://www.worldcat.org/isbn/9781557866172.

Damon Centola and Michael Macy. Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, 113(3):702–734, November 2007. doi: 10.1086/521848. URL http://dx.doi.org/10.1086/521848.

Damon Centola, Robb Willer, and Michael Macy. The Emperor's Dilemma: A Computational Model of SelfâĂŘEnforcing Norms. *American Journal of Sociology*, 110(4), 2005. URL http://www.jstor.org/stable/10.1086/427321.

Myong-Hun Chang and Joseph E. Harrington. Discovery and Diffusion of Knowledge in an Endogenous Social Network. *American Journal of Sociology*, 110(4), 2005. URL http://www.jstor.org/stable/10.1086/426555.

Kalyan Chatterjee and Susan H. Xu. Technology Diffusion by Learning from Neighbours. *Advances in Applied Probability*, 36(2):355–376, 2004. ISSN 00018678. doi: 10.2307/1428457. URL http://dx.doi.org/10.2307/1428457.

Roy A. Church. *The rise and decline of the British motor industry*. Cambridge University Press, 1995. ISBN 9780521557702. URL http://www.worldcat.org/isbn/9780521557702.

W. A. V. Clark. Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model. *Demography*, 28(1):1–19, 1991. ISSN 00703370. doi: 10.2307/2061333. URL http://dx.doi.org/10.2307/2061333.

William A. Clark and Mark Fossett. Understanding the social context of the Schelling segregation model. *Proceedings of the National Academy of Sciences of the United States of America*, 105(11):4109–4114, March 2008. ISSN 1091-6490. doi: 10.1073/pnas.0708155105. URL http://dx.doi.org/10.1073/pnas.0708155105.

Anthony P. Cohen. *The symbolic construction of community*. E. Horwood ; Tavistock Publications, 1985. ISBN 9780853129332. URL http://www.worldcat.org/isbn/9780853129332.

James S. Coleman, Elihu Katz, and Herbert Menzel. *Medical Innovation: A Diffusion Study*. Bobbs-Merrill Co, 1966.

D. J. Daley and J. Gani. *Epidemic Modelling: An Introduction (Cambridge Studies in Mathematical Biology)*. Cambridge University Press, 1 edition, May 2001. ISBN 0521014670. URL http://www.worldcat.org/isbn/0521014670.

Jared Diamond. *Collapse : How Societies Choose to Fail or Succeed*. Penguin (Non-Classics), December 2005. ISBN 0143036556. URL http://www.worldcat.org/isbn/0143036556.

Emile Durkheim. *Suicide*. Free Press, February 1997. ISBN 0684836327. URL http://www.worldcat.org/isbn/0684836327.

Émile Durkheim and W. D. Halls. *The division of labor in society*. Free Press, 1893. ISBN 9780029079508. URL http://www.worldcat.org/isbn/9780029079508.

Phillip C. Dykstra. North American Numbering Plan, NPA/NXX's. Hosted on Website, 1998. URL http://sd.wareonearth.com/~{}phil/npanxx/.

Glenn Ellison and Drew Fudenberg. Rules of Thumb for Social Learning. *The Journal of Political Economy*, 101(4):612–643, 1993. ISSN 00223808. doi: 10.2307/2138741. URL http://dx.doi.org/10.2307/2138741.

Glenn Ellison and Drew Fudenberg. Word-of-Mouth Communication and Social Learning. *The Quarterly Journal of Economics*, 110(1):93–125, February 1995. ISSN 1531-4650. doi: 10.2307/2118512. URL http://dx.doi.org/10.2307/2118512.

Jon Elster. *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. Cambridge University Press, 1 edition, April 2007. ISBN 0521777445. URL http://www.worldcat.org/isbn/0521777445.

Joshua M. Epstein and Robert L. Axtell. *Growing Artificial Societies: Social Science From the Bottom Up (Complex Adaptive Systems)*. Brookings Institution Press MIT Press, 1st edition edition, October 1996. ISBN 0262550253. URL http://www.worldcat.org/isbn/0262550253.

I. P. Fellegi. On the Question of Statistical Confidentiality. *Journal of the American Statistical Association*, 67(337):7–18, 1972. ISSN 01621459. URL http://www.jstor.org/stable/2284695.

Thom File. Computer and Internet Use in the United States. Technical report, US Census Bureau, May 2013. URL http://www.census.gov/prod/2013pubs/p20-569.pdf.

Claude S. Fischer. Urban-to-Rural Diffusion of Opinions in Contemporary America. *The American Journal of Sociology*, 84(1):151–159, 1978. ISSN 00029602. doi: 10.2307/2777982. URL http://dx.doi.org/10.2307/2777982.

James H. Fowler and Oleg Smirnov. Dynamic Parties and Social Turnout: An AgentâĂŘBased Model. *American Journal of Sociology*, 110(4), 2005. URL http://www.jstor.org/stable/10.1086/426554.

Paul Freiberger and Michael Swaine. *Fire in the Valley: The Making of The Personal Computer (Second Edition)*. McGraw-Hill Companies, 2nd edition, November 2000. ISBN 0071358927. URL http://www.worldcat.org/isbn/0071358927.

David Garcia, Pavlin Mavrodiev, and Frank Schweitzer. Social Resilience in Online Communities: The Autopsy of Friendster. In *Proceedings of the First ACM Conference on Online Social Networks*, COSN '13, pages 39–50, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-2084-9. doi: 10.1145/2512938.2512946. URL http://dx.doi.org/10.1145/2512938.2512946.

M. Gardner. The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, 223:120–123, October 1970.

Nigel Gilbert and Andrew Abbott. Introduction. *American Journal of Sociology*, 110(4), 2005. URL http://www.jstor.org/stable/10.1086/430413.

Benjamin Golub and Matthew O. Jackson. How Homophily Affects Diffusion and Learning in Networks. November 2008. URL http://arxiv.org/abs/0811.4013.

Roger V. Gould. Collective Action and Network Structure. *American Sociological Review*, 58(2):182–196, 1993. ISSN 00031224. doi: 10.2307/2095965. URL http://dx.doi.org/10.2307/2095965.

Mark Granovetter. Threshold Models of Collective Behavior. *American Journal of Sociology*, 83(6):1420–1443, 1978. ISSN 00029602. doi: 10.2307/2778111. URL http://dx.doi.org/10.2307/2778111.

Mark S. Granovetter. The Strength of Weak Ties. *American Journal of Sociology*, 78(6):1360–1380, 1973. ISSN 00029602. doi: 10.2307/2776392. URL http://dx.doi.org/10.2307/2776392.

Guerry. *A translation of Andre-Michel Guerry's Essay on the moral statistics of France (1883): a sociological report to the French Academy of Science*. Edwin Mellen Press, 2002. ISBN 077347045. URL http://www.worldcat.org/isbn/077347045.

Torsten Hägerstrand. *Innovation diffusion as a spatial process*. University of Chicago Press, October 1967. doi: 10.1177/030913259201600403. URL http://dx.doi.org/10.1177/030913259201600403.

Keith Hampton and Barry Wellman. Long Distance Community in the Network Society. *American Behavioral Scientist*, 45(3):476–495, November 2001. doi: 10.1177/00027640121957303. URL http://dx.doi.org/10.1177/00027640121957303.

Kristen Haring. *Ham Radio's Technical Culture (Inside Technology)*. The MIT Press, February 2008. ISBN 0262582767. URL http://www.worldcat.org/isbn/0262582767.

Peter Hedström. Contagious Collectivities: On the Spatial Diffusion of Swedish Trade Unions, 1890-1940. *The American Journal of Sociology*, 99(5):1157–1179, 1994. ISSN 00029602. doi: 10.2307/2781146. URL http://dx.doi.org/10.2307/2781146.

Peter Hedström. *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge University Press, December 2005. ISBN 0521796679. URL http://www.worldcat.org/isbn/0521796679.

Peter Hedström and Peter Bearman, editors. *The Oxford Handbook of Analytical Sociology*. Oxford University Press, USA, September 2009. ISBN 0199215367. URL http://www.worldcat.org/isbn/0199215367.

Peter Hedström and Richard Swedberg, editors. *Social Mechanisms: An Analytical Approach to Social Theory (Studies in Rationality and Social Change)*. Cambridge University Press, January 1998. ISBN 0521596874. URL http://www.worldcat.org/isbn/0521596874.

Peter Hedström, Rickard Sandell, and Charlotta Stern. Mesolevel Networks and the Diffusion of Social Movements: The Case of the Swedish Social Democratic Party. *The American Journal of Sociology*, 106(1):145–172, 2000. ISSN 00029602. doi: 10.2307/3081282. URL http://dx.doi.org/10.2307/3081282.

Guido Hertel, Sven Niedner, and Stefanie Herrmann. Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel. *Research Policy*, 32(7):1159–1177, July 2003. ISSN 00487333. doi: 10.1016/s0048-7333(03)00047-7. URL http://dx.doi.org/10.1016/s0048-7333(03)00047-7.

George A. Hillery. Definitions of community: Areas of agreement. *Rural Sociology*, 20(2):111–123, 1955. URL http://search.ebscohost.com/login.aspx?direct=true&db=ehh&AN=13212774&site=ehost-live.

Stephan Holl and Hans Plum. PostGIS. *GeoInformatics*, 03/2009:34–36, April 2009. URL http://fluidbook.microdesign.nl/geoinformatics/03-2009/?page=34.

Jui C. Huang and Peter Gould. Diffusion in an Urban Hierarchy: The Case of Rotary Clubs. *Economic Geography*, 50(4):333–340, 1974. ISSN 00130095. doi: 10.2307/143060. URL http://dx.doi.org/10.2307/143060.

B. A. Huberman and N. S. Glance. Evolutionary games and computer simulations. *Proceedings of the National Academy of Sciences*, 90(16):7716–7718, August 1993. ISSN 1091-6490. doi: 10.1073/pnas.90.16.7716. URL http://dx.doi.org/10.1073/pnas.90.16.7716.

John D. Hunter. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3):90–95, May 2007. ISSN 1521-9615. doi: 10.1109/mcse.2007.55. URL http://dx.doi.org/10.1109/mcse.2007.55.

Paul Ingram, Lori Q. Yue, and Hayagreeva Rao. Trouble in Store: Probes, Protests, and Store Openings by WalâĂŘMart, 1998âĂŞ2007. *American Journal of Sociology*, 116(1):53–92, July 2010. ISSN 0002-9602. doi: 10.1086/653596. URL http://dx.doi.org/10.1086/653596.

Eric Jones, Travis Oliphant, Pearu Peterson, and Others. *SciPy: Open source scientific tools for Python*, 2001–. URL http://www.scipy.org/.

Sandeep Kapur. Technological Diffusion with Social Learning. *The Journal of Industrial Economics*, 43(2):173–195, 1995. ISSN 00221821. doi: 10.2307/2950480. URL http://dx.doi.org/10.2307/2950480.

Suzanne Keller. *Community: Pursuing the Dream, Living the Reality (Princeton Studies in Cultural Sociology)*. Princeton University Press, January 2003. ISBN 0691095647. URL http://www.worldcat.org/isbn/0691095647.

Alan Kirman. Ants, Rationality, and Recruitment. *The Quarterly Journal of Economics*, 108(1):137–156, February 1993. ISSN 00335533. doi: 10.2307/2118498. URL http://dx.doi.org/10.2307/2118498.

René König. *The community*. Routledge & Kegan Paul, 1968. URL http://www.worldcat.org/oclc/245810634.

Ruth R. Kornhauser. *Social Sources of Delinquency: An Appraisal of Analytic Models*. Univ of Chicago Pr (Tx), 1978. ISBN 0226451135. URL http://www.worldcat.org/isbn/0226451135.

Fredrik Liljeros. *The complexity of social organizing*. PhD thesis, Stockholm University, 2001. URL http://www.worldcat.org/isbn/9789172652613.

Nan Lin. *Social Capital: A Theory of Social Structure and Action (Structural Analysis in the Social Sciences)*. Cambridge University Press, 0 edition, May 2002. ISBN 052152167X. URL http://www.worldcat.org/isbn/052152167X.

Nan Lin, Ronald S. Burt, and Karen Cook, editors. *Social Capital: Theory and Research (Sociology and Economics (Paper))*. Aldine Transaction, June 2001. ISBN 0202306445. URL http://www.worldcat.org/isbn/0202306445.

Michael W. Macy. Chains of Cooperation: Threshold Effects in Collective Action. *American Sociological Review*, 56(6):730–747, December 1991. ISSN 00031224. doi: 10.2307/2096252. URL http://dx.doi.org/10.2307/2096252.

Michael W. Macy and Robert Willer. FROM FACTORS TO ACTORS: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology*, 28(1):143–166, 2002. doi: 10.1146/annurev.soc.28.110601.141117. URL http://dx.doi.org/10.1146/annurev.soc.28.110601.141117.

Vijay Mahajan and Eitan Muller. Innovation Diffusion and New Product Growth Models in Marketing. *The Journal of Marketing*, 43(4):55–68, 1979. ISSN 00222429. doi: 10.2307/1250271. URL http://dx.doi.org/10.2307/1250271.

Vijay Mahajan, Eitan Muller, and Frank M. Bass. New Product Diffusion Models in Marketing: A Review and Directions for Research. *Journal of Marketing*, 54(1):1–26, January 1990. ISSN 00222429. doi: 10.2307/1252170. URL http://dx.doi.org/10.2307/1252170.

Charles F. Manski. Social Learning from Private Experiences: The Dynamics of the Selection Problem. *The Review of Economic Studies*, 71(2):443–458, 2004. ISSN 00346527. doi: 10.2307/3700633. URL http://dx.doi.org/10.2307/3700633.

Peter V. Marsden. Network Data and Measurement. *Annual Review of Sociology*, 16(1):435–463, 1990. doi: 10.1146/annurev.so.16.080190.002251. URL http://dx.doi.org/10.1146/annurev.so.16.080190.002251.

L. Mercken, T. A. B. Snijders, C. Steglich, E. Vartiainen, and H. de Vries. Dynamics of adolescent friendship networks and smoking behavior. *Social Networks*, 32(1): 72–81, January 2010. ISSN 03788733. doi: 10.1016/j.socnet.2009.02.005. URL http://dx.doi.org/10.1016/j.socnet.2009.02.005.

Robert K. Merton. *Social theory and social structure*. Free Press, 1968. ISBN 9780029211304. URL http://www.worldcat.org/isbn/9780029211304.

D. Mok and B. Wellman. Did distance matter before the Internet? Interpersonal contact and support in the 1970s. *Social Networks*, 29(3):430–461, July 2007. ISSN 03788733. doi: 10.1016/j.socnet.2007.01.009. URL http://dx.doi.org/10.1016/j.socnet.2007.01.009.

Andrew Molnar. Computers in Education: A Brief History. *T. H. E. Journal*, June 1997. URL http://thejournal.com/Articles/1997/06/01/Computers-in-Education-A-Brief-History.aspx.

Daniel Myers. Violent Protest and Heterogeneous Diffusion Processes: The Spread of U.S. Racial Rioting From 1964 to 1971. *Mobilization: An International Quarterly*, 15(3):289–321, September 2010. URL http://www.metapress.com/content/F16204108631474V.

John A. Norton and Frank M. Bass. A Diffusion Theory Model of Adoption and Substitution for Successive Generations of High-Technology Products. *MANAGEMENT SCIENCE*, 33(9):1069–1086, September 1987. doi: 10.1287/mnsc.33.9.1069. URL http://dx.doi.org/10.1287/mnsc.33.9.1069.

Martin A. Nowak and Robert M. May. Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829, October 1992. doi: 10.1038/359826a0. URL http://dx.doi.org/10.1038/359826a0.

Mancur Olson. *The Logic of Collective Action: Public Goods and the Theory of Groups, Second printing with new preface and appendix (Harvard Economic Studies)*. Harvard economic studies, v. 124. Harvard University Press, revised edition, January 1971. ISBN 0674537513. URL http://www.worldcat.org/isbn/0674537513.

F. Ortega, J. M. Gonzalez-Barahona, and G. Robles. On the Inequality of Contributions to Wikipedia. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, page 304, Washington, DC, USA, January 2008. IEEE. doi: 10.1109/hicss.2008.333. URL http://dx.doi.org/10.1109/hicss.2008.333.

Wei Pan. Extending the Iterative Convex Minorant Algorithm to the Cox Model for Interval-Censored Data. *Journal of Computational and Graphical Statistics*, 8 (1), 1999. ISSN 10618600. doi: 10.2307/1390923. URL http://dx.doi.org/10.2307/1390923.

Robert E. Park and Ernest Burgess. *Introduction to the Science of Sociology*. Univ of Chicago Pr (Tx), 3 revised edition, 1969. ISBN 0226646041. URL http://www.worldcat.org/isbn/0226646041.

Raymond Plant. *Community and Ideology : An Essay in Applied Social Philosophy*. Routledge Revivals. Routledge, 1 edition, July 1974. ISBN 0415557682. URL http://www.worldcat.org/isbn/0415557682.

Alejandro Portes. Social Capital: Its Origins and Applications in Modern Sociology. *Annual Review of Sociology*, 24, 1998. ISSN 03600572. doi: 10.2307/223472. URL http://dx.doi.org/10.2307/223472.

Robert D. Putnam. *Bowling Alone: The Collapse and Revival of American Community*. Touchstone Books by Simon & Schuster, 1st edition, August 2001. ISBN 0743203046. URL http://www.worldcat.org/isbn/0743203046.

Robert D. Putnam. E Pluribus Unum: Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture. *Scandinavian Political Studies*, 30(2): 137–174, June 2007. ISSN 0080-6757. doi: 10.1111/j.1467-9477.2007.00176.x. URL http://dx.doi.org/10.1111/j.1467-9477.2007.00176.x.

Adolphe Quetelet. *A treatise on man and the development of his faculties*. Edinburgh: W. and R. Chambers, 1842.

Sophia Rabe-Hesketh and Anders Skrondal. *Multilevel and Longitudinal Modeling Using Stata, Second Edition*. Stata Press, 2 edition, February 2008. ISBN 1597180408. URL http://www.worldcat.org/isbn/1597180408.

Rawson W. Rawson. An Inquiry into the Statistics of Crime in England and Wales. *Journal of the Statistical Society of London*, 2(5):316+, October 1839. ISSN 09595341. doi: 10.2307/2337821. URL http://dx.doi.org/10.2307/2337821.

Eric S. Raymond. *The cathedral and the bazaar : musings on Linux and Open Source by an accidental revolutionary*. O'Reilly, 2001. ISBN 9780596001315. URL http://www.worldcat.org/isbn/9780596001315.

Edward Reid and Patricia Novak. Personal space: An unobtrusive measures study. 5(3):265–266, 1975. doi: 10.3758/bf03337629. URL http://dx.doi.org/10.3758/bf03337629.

Craig W. Reynolds. Flocks, Herds and Schools: A Distributed Behavioral Model. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, volume 21 of *SIGGRAPH '87*, pages 25–34, New York, NY, USA, July 1987. ACM. ISBN 0-89791-227-6. doi: 10.1145/37401.37406. URL http://dx.doi.org/10.1145/37401.37406.

Howard Rheingold. *The virtual community : homesteading on the electronic frontier*. MIT Press, 2000. ISBN 9780262681216. URL http://www.worldcat.org/isbn/9780262681216.

Roland Robertson. *Glocalization: Time-Space and Homogeneity-Heterogeneity*, chapter 2, pages 25–44. Sage Publications, 1995. ISBN 9780803979475. URL http://www.worldcat.org/isbn/9780803979475.

Everett M. Rogers. *Diffusion of Innovations, 5th Edition*. Free Press, 5th edition, August 2003. ISBN 0743222091. URL http://www.worldcat.org/isbn/0743222091.

Jean-Jacques Rousseau and G. D. H. Cole. *The social contract : and discourses*. E.P. Dutton and Company, Inc., 1762. URL http://www.worldcat.org/oclc/179547.

Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]., 2010.

Bryce Ryan and Neal C. Gross. The Diffusion of Hybrid Corn in Two Iowa Communities. *Rural Sociology*, 8(1):15–24, 1943.

Serguei Saavedra, Felix Reed-Tsochas, and Brian Uzzi. Asymmetric disassembly and robustness in declining networks. *Proceedings of the National Academy of Sciences*, 105(43):16466–16471, October 2008. doi: 10.1073/pnas.0804740105. URL http://dx.doi.org/10.1073/pnas.0804740105.

Jason S. Sadofsky. BBS: The Documentary. Documentary Film, May 2005. URL http://www.bbsdocumentary.com/.

Robert J. Sampson. *Great American City: Chicago and the Enduring Neighborhood Effect*. University of Chicago Press, reprint edition, March 2012. ISBN 0226734560. URL http://www.worldcat.org/isbn/0226734560.

Thomas C. Schelling. Models of Segregation. *The American Economic Review*, 59 (2):488–493, 1969. ISSN 00028282. doi: 10.2307/1823701. URL http://dx.doi.org/10.2307/1823701.

Thomas C. Schelling. *Micromotives and Macrobehavior*. W. W. Norton & Company, revised edition, October 2006. ISBN 0393329461. URL http://www.worldcat.org/isbn/0393329461.

Clifford R. Shaw and Henry D. McKay. *Juvenile delinquency and urban areas : a study of rates of delinquency in relation to differential characteristics of local communities in American cities*. University of Chicago Press, 1972. ISBN 0226751252. URL http://www.worldcat.org/isbn/0226751252.

Lones Smith and Peter Sorensen. Pathological Outcomes of Observational Learning. *Econometrica*, 68(2):371–398, 2000. ISSN 00129682. doi: 10.2307/2999431. URL http://dx.doi.org/10.2307/2999431.

Tom A. B. Snijders and Roel Bosker. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. SAGE Publications Ltd, second edition edition, December 2011. ISBN 184920201X. URL http://www.worldcat.org/isbn/184920201X.

Jon Snodgrass. CLIFFORD R. SHAW AND HENRY D. McKAY: CHICAGO CRIMINOLOGISTS. *THE BRITISH JOURNAL OF CRIMINOLOGY*, 16(1):1–19, January 1976. URL http://bjc.oxfordjournals.org/citmgr?gca=crimin%3B16%2F1%2F1.

Steve Snow. FidoNet Enthusiasts Share a Low-Cost Way to Communicate, July 1993. URL http://search.proquest.com/docview/140810914?accountid=13042. Copyright - Copyright The Washington Post Company Jul 19, 1993; Last updated - 2010-05-28; DOI - 1648390892; 23900971; 48611.

Margaret Stacey. The Myth of Community Studies. In Colin Bell and Howard Newby, editors, *Sociology of Community: A Collection of Readings*, New Sociology Library, chapter 3, pages 13–26. Routledge, 1 edition, September 1974. ISBN 978-0714629704. URL http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0714629707.

Rodney Stark and Laurence R. Iannaccone. A Supply-Side Reinterpretation of the "Secularization" of Europe. *Journal for the Scientific Study of Religion*, 33(3):230–252, September 1994. ISSN 00218294. doi: 10.2307/1386688. URL http://dx.doi.org/10.2307/1386688.

StataCorp. *Stata Statistical Software: Release 11*. StataCorp LP, College Station, TX, 2009. URL http://www.stata.com.

David Strang. Adding Social Structure to Diffusion Models: An Event History Framework. *Sociological Methods Research*, 19(3):324–353, February 1991. doi: 10.1177/0049124191019003003. URL http://dx.doi.org/10.1177/0049124191019003003.

David Strang and Sarah A. Soule. Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills. *Annual Review of Sociology*, 24(1):265–290, 1998. doi: 10.1146/annurev.soc.24.1.265. URL http://dx.doi.org/10.1146/annurev.soc.24.1.265.

David Strang and Nancy B. Tuma. Spatial and Temporal Heterogeneity in Diffusion. *The American Journal of Sociology*, 99(3):614–639, 1993. ISSN 00029602. doi: 10.2307/2781285. URL http://dx.doi.org/10.2307/2781285.

Yoshio Sugiura. Diffusion of Rotary Clubs in Japan, 1920-1940: A Case of Non-Profit-Motivated Innovation Diffusion under a Decentralized Decision Making Structure. *Economic Geography*, 62(2):125–143, 1986. ISSN 00130095. doi: 10.2307/144087. URL http://dx.doi.org/10.2307/144087.

Joseph A. Tainter. *The Collapse of Complex Societies (New Studies in Archaeology)*. Cambridge University Press, reprint edition, March 1990. ISBN 052138673X. URL http://www.worldcat.org/isbn/052138673X.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. URL `http://www.R-project.org/`.

Ferdinand Tonnies. *Community and Society*. Dover Publications, November 1887. ISBN 0486424979. URL `http://www.worldcat.org/isbn/0486424979`.

Linus Torvalds and David Diamond. *Just for Fun: The Story of an Accidental Revolutionary*. HarperCollins, 1st edition, May 2001. ISBN 0066620724. URL `http://www.worldcat.org/isbn/0066620724`.

Peter Turchin. *Historical dynamics : why states rise and fall*. Princeton University Press, 2003. ISBN 9780691116693. URL `http://www.worldcat.org/isbn/9780691116693`.

Urban Rural Definitions. US Census Bureau, October 1995. URL `http://www.census.gov/geo/reference/docs/ua/urdef.txt`.

Thomas W. Valente. *Network Models of the Diffusion of Innovations (Quantitative Methods in Communication Subseries)*. Hampton Press (NJ), January 1995. ISBN 1881303225. URL `http://www.worldcat.org/isbn/1881303225`.

Thomas W. Valente. Network models of the diffusion of innovations. *Computational & Mathematical Organization Theory*, 2(2):163–164, June 1996. ISSN 1381-298X. doi: 10.1007/bf00240425. URL `http://dx.doi.org/10.1007/bf00240425`.

Christophe Van den Bulte and Gary L. Lilien. Medical Innovation Revisited: Social Contagion versus Marketing Effort. *American Journal of Sociology*, 106(5):1409–1435, March 2001. doi: 10.1086/320819. URL `http://dx.doi.org/10.1086/320819`.

Ion B. Vasi and David Strang. Civil Liberty in America: The Diffusion of Municipal Bill of Rights Resolutions after the Passage of the USA PATRIOT Act1. *American Journal of Sociology*, 114(6):1716–1764, May 2009. ISSN 0002-9602. doi: 10.1086/597177. URL `http://dx.doi.org/10.1086/597177`.

Phillip E. Wegner. *Imaginary communities : utopia, the nation, and the spatial histories of modernity*. University of California Press, 2002. ISBN 9780520228283. URL `http://www.worldcat.org/isbn/9780520228283`.

Barry Wellman. *Networks in the global village : life in contemporary communities*. Westview Press, 1999. ISBN 9780813311500. URL `http://www.worldcat.org/isbn/9780813311500`.

Barry Wellman. Computer Networks As Social Networks. *Science*, 293(5537):2031–2034, September 2001. ISSN 1095-9203. doi: 10.1126/science.1065547. URL `http://dx.doi.org/10.1126/science.1065547`.

Barry Wellman and Caroline A. Haythornthwaite. *The Internet in everyday life*. Blackwell Pub., 2002. ISBN 9780631235071. URL http://www.worldcat.org/isbn/9780631235071.

William F. Whyte. *Street Corner Society: The Social Structure of an Italian Slum*. University of Chicago Press, 4 edition, April 2012. URL http://www.worldcat.org/isbn/9780226922669.

Steve Wozniak. *I, Woz*. Headline Review, August 2007. ISBN 0755314085. URL http://www.worldcat.org/isbn/0755314085.

H. Peyton Young. Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning. *The American Economic Review*, 99(5): 1899–1924, December 2009. ISSN 0002-8282. doi: 10.1257/aer.99.5.1899. URL http://www.atypon-link.com/AEAP/doi/pdf/10.1257/aer.99.5.1899.

Glossary

Apple][

One of the first successful personal computers, first introduced at the West Coast Computer Faire April 16, 1977. One of *Byte Magazine's* 1977 Holy Trinity.

Decline event

An event in a system whereby a unit of analysis is removed from that system. That removal may be a sysop quitting or a company closing down or a simulated Anasazi household ageing the point of death. What is constant is that it is an event, at the agent level, that reduces the size of the system.

Door game

Door games were popular in the heyday of BBSs, and were games that allowed multiple users to play. Many were asynchronous, reflecting the costs and difficulties of having more than one user connected at a time, but some were asynchronous, allowing the sysop to play simultaneously, or for BBSs with multiple phone lines and modems, these allowed users to play with each other directly, just like modern, internet multi-player computer games.

Echolist

A series of community discussion groups, very similar to the separate sections of Usenet.

ENews

A regular FidoNet newsletter.

Exchange Area

A combination of geographically proximate 1990 US Census Tracts and telephone exchanges devised as the basic geographic unit of analysis for this thesis see.

Facebook

A popular social networking website (ICT)which lets users designate each other as friends and socialise via messaging and posts on each others' pages.

Fido History Project

A carefully maintained and corrected list of Nodelists and messages hosted at `http://ambrosia60.dd-dns.de/fidonet/nodelist.php`.

FidoNet

A network of BBSs created by Tom Jennings in 1983.

Gemeinschaft

Tonnies's (1887) theory of community that combines locality, strongly reciprocal social ties, a sense of belonging and a pastoralist narrative.

Gesellschaft

Tonnies's (1887) theory of a rising urban social disconnection, the opposite of *Gemeinschaft*.

Ham radio

Amateur radio, operated by hobbyists that has been going since the turn of the 20th century. Its operators formed informal communication networks and clubs (Haring, 2008), much like FidoNet. The term ham is a pejorative that was originally meant to describe someone as being unskilled (like ham-fisted), but the term was eventually claimed by the community and remains to this day.

Homebrew Computer Club

A influential computer hobbyist club started in Silicon Valley in March of 1975. Its members (such as Apple's Steve Wozniak) heavily influenced the development of personal computers. The club ran until 1986 (Freiberger and Swaine, 2000).

Modem

A MOdulator DEModulator which traditionally converts (modulates) a computer file into sound, sends that sound over a phoneline to another modem, which is then decoded (demodulated) back into information for the recipient.

Node

Another name for a specific FidoNet BBS. The term derives from its use in graph theory as another name for vertex.

Nodelist

A weekly-published list of sysops, their telephone numbers and position in the FidoNet structure.

PET 2001

Also known as the Super PET and announced in January 1977 at the winter Consumer Electronics Show, was Commodore International's first fully-featured personal computer. One of *Byte Magazine's* 1977 Holy Trinity.

Sysop

A BBS SYStems OPerator maintained their BBS, making it available to users at specified times, usually also socialising with whoever connected.

Telephone exchange

A local provider of landline telephone connection routing to the wider phone network. In North America these conform to the North American Numbering Plan (NANPA) standard.

Telnet

A protocol first proposed in RFC 15, an early means of allowing a bi-directional text-based interface between one computer and another over a network.

Tract

Geographic region defined for the 1990 US Census intended to coherently group approximately 4000–8000 demographically related individuals enclosed by natural borders (rivers, roads etc.). Tract data is provided by the NHGIS.

TRS-80

The Tandy/Radio Shack, Z-80 was released in November 1977, marketed as a microcomputer rather than a personal computer. One of *Byte Magazine's* 1977 Holy Trinity.

Usenet

An early ICT that initially connected universities connected to ARPANET, and was connected to FidoNet in 1985.

User

To be distinguished from a sysop, a BBS user connects to their local BBS to send and receive messages to and from other users and sysops.

WhatsApp

A modern ICT that provides an means of messaging via smartphones.

Acronyms

ABM

A modelling approach where agent interactions are simulated, usually over time. ABM is often used in cases where statistical models are unsuitable, often due to computational intractability, or a lack of a suitable statistical model for understanding the behaviour of interest.

AIC

A measure of the relative quality of statistical models for a given set of data.

AOL

An early ISP that significantly spread internet adoption in the US in the 1990s.

ARPANET

The US government and university research project, first operational in 1969, that eventually led to the internet.

BBS

A computer and modem run by a sysop so that users can connect via their modems and phone lines, then post and read messages.

BIC

A criterion for model selection among a finite set of models.

BSD

BSD license is like the GNU Public License, except it allows new modifications to be relicensed as proprietary, thus enabling it to be sold as a commercial product. Many recently released open source projects use the BSD license instead of the GNU Public License to allow companies to work on the code and profit financially.

FTP

A standard interface for transferring data over a network first proposed in RFC 114 in 1970.

GNU

The GNU project was started by Richard Stallman in YEAR to replace all the components of the proprietary Unix operating system. He licensed it under the GNU Public License to ensure the work would forever be freely available for use and modification. GNU is a recursive acronym, in that it includes itself in its name.

GNU PUBLIC LICENSE

The GPL is one of the most popular open source licenses and most famously is how the Linux kernel and many of the core components of the GNU project which Linux requires to operate, is licensed to forever enforce the free availability of the source code for future modification and free distribution.

ICT

A technology that can be used as a means of communication and potentially social interaction.

IPUMS

A public database maintained by the University of Minnesota of international survey data, largely generated from national censuses and combined with high resolution geospatial information (Ruggles et al., 2010).

IRC

A text-based chatroom protocol originally based on a BBS initially developed in the late 1980s. IRC remains quite active at the time of writing, especially within the open source community.

ISP

A company that provides access to the internet. Early ISPs include CompuServ and AOL.

JANET

The UK university research project that had major influence on the development of the internet.

MLE

A method of fitting the parameters of a model to a dataset.

NHGIS

A public database derived from the US Census and maintained by the University of Minnesota that provides high quality geospatial data and integrated time series of survey questions (nhg, 2011).

PEP

Like RFCs, PEPs are the public means by which proposals for changes to the python programming language are made.

PUMA

PUMAs are composed of tracts and have approximately 100,000 people. The data available at the PUMA level is more detailed than at the tract level, and allows the census to protect the anonymity of US citizens while providing considerable detail in the demographics of a wide variety of variables. PUMAs are primarily helpful here in controlling for occupations, and are provided by IPUMS.

RDD

RDD or Wardialing as it is know in malicious contexts, is the practice of randomly calling telephone numbers with a modem in the hopes of getting a response from another modem. This is one way BBSs could be found but is highly time consuming, impractical and potentially costly. Wardialing is RDD for the purpose of breaking into the computer attached to the modem.

RFC

Originally developed as a way of recording unofficial notes related to the development of ARPANET, RFCs have become the proving ground for the standards the internet is now based on.