Lynne Hamill
August 2007

New Communication Technologies and Travel

Draft Chapter 3: The Micro-Macro Debate and Social Simulation Models

---

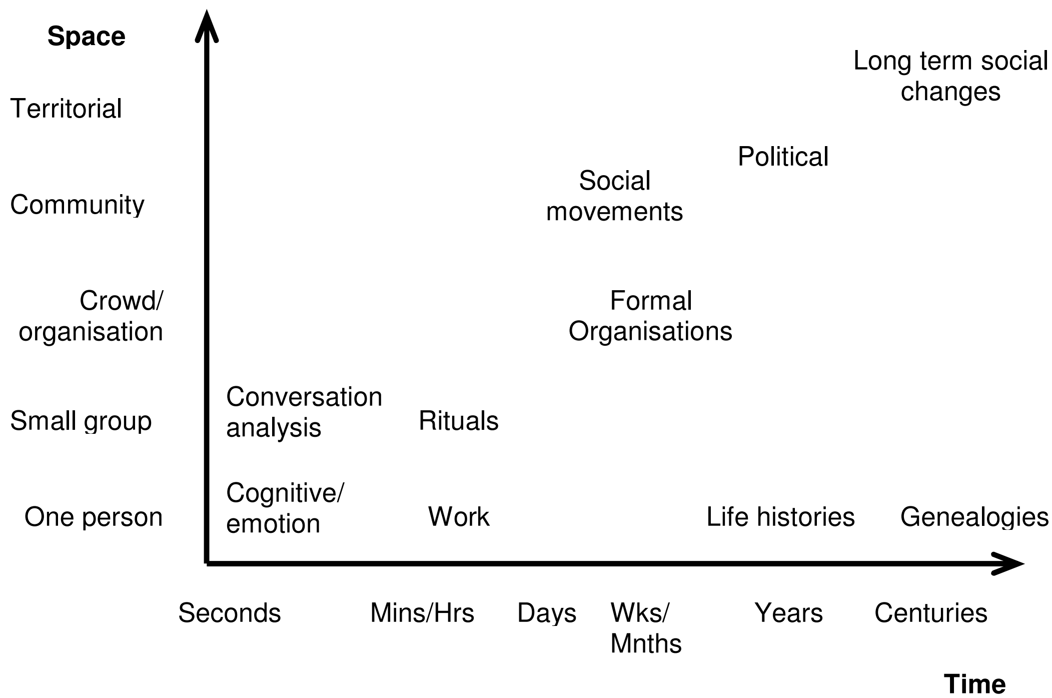
In the Introduction, it was claimed that agent-based models could bridge the gap between micro and macro. This Chapter explains that claim. Section 3.1 summarises the sociological micro-macro debate and Section 3.2 introduces the concept of emergence. Section 3.3 describes the key features of agent-based modelling and how it relates to emergence. Section 3.4 presents a simple example of an agent-based model. Finally, Section 3.5 introduces macro, top-down system dynamics modelling and compares it with agent-based modelling.

3.1 The Micro-Macro Debate

Microsociology "is the detailed analysis of what people do, say, and think in the actual flow of momentary experience" (Collins, 1981) and focuses on "social interaction and communication" (Blau, 1987: 71). Macrosociology is the analysis of large-scale and long-term historical social processes, dealing with concepts such as class, organization, state, economy, and society (Collins, 1981: Sawyer, 2005: 210). Macrosociology "analyzes the structure of different positions in a population and their constraints on social relations. The focus is on the external limitations of the social environment on people's relations" (Blau, 1987: 71).

Elias (1970/1978: 129) argued that the individual and society are "two different but inseparable levels of the human world" and Collins (1981) illustrates the micro-macro spectrum as a continuum in time and space. Fig.3.1 summarises his analysis: the time period under consideration varies from seconds to centuries; the space from that occupied by a single individual to a 'territorial society': the lower and further left, the more micro; the higher and further to the right, the more macro. The micro-macro question is how to bring together analysis of activity at the bottom left-hand corner of the diagram with analysis of activity at the top right-hand corner.

Fig. 3.1: Summary of Collins (1981) view of micro-to-macro spectrum



Ritzer & Goodman (2004: 483-536) distinguish between the micro-macro duality and the agency-structure duality on the grounds that they differ in their theoretical roots and that the micro-macro duality provides a vertical hierarchy – micro at the bottom, macro at the top and meso in between – while the agency-structure duality does not "since both agency and structure can be found at any level of social analysis". So, for example, Coleman, Collins and Elias are characterised as micro-macro writers while Giddens and Habermas are placed in the agency-structure camp. However, for the purpose of this thesis at least, the distinction is not, in my view, important.

Attempts at bringing together the micro/agency and the macro/structure duality are discussed below, starting with Coleman and then briefly covering Giddens, Collins and Habermas before going on examine the claims that the debate has been overtaken by the arrival of the networked society.

In the 1960s Coleman noted that by observing the interactions of many gas molecules, scientists derived Boyle's Law that described aggregate behaviour and reasoned that social scientists could do the same, moving from individual behaviour to group behaviour (Smith, 1997: 52). Smith goes on:

> "The issues involved in addressing this problem are still with us and remain at the heart of the so-called micro-macro problem in social theory".

Coleman argues that "the proposition system begins and ends at the macro level, but in between it dips to the level of the individual" (1994: 8) which is the only place that action occurs because:

> "the system level exists solely as emergent properties characterizing the system of action as a whole. It is only in this sense that there is behavior of the system. Nevertheless, system-level properties will result, so propositions may be generated at the level of the system" (1994: 28).

He identifies "three kinds of components to any theory in which system behavior derives from actions of actors who are elements of the system" (1994: 11):

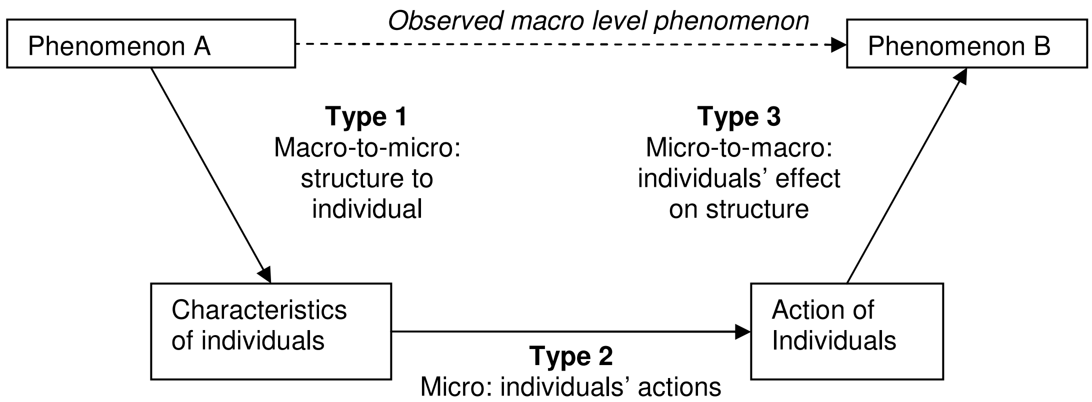
- Type 1: macro-to-micro: structure to individual:

  "the rules of the game, rules which transmit consequences of an individual's action to other individuals and rules which derive macro-level outcomes from combinations of individuals' actions" (ibid: 19).
- Type 2: micro: individuals' action:

  "gives rise to different systematic behavior – that is, different social phenomena – when located in different social contexts when different persons' actions combine in different ways" (ibid)
- Type 3: micro-to-macro: individuals' effect on structure.

Thus in Coleman's view of the world an apparent relationship between two observed macro level phenomena is underpinned by micro level characteristics and action as illustrated in Fig. 3.2 (based on his 1994: 8, Fig. 1.2). Coleman is criticised for the fact that the arrows flow in only one direction, not allowing feedback (Ritzer & Goodman, 2004: 491).

Fig. 3.2: Coleman's example of micro-macro interactions



In *The Constitution of Society* (1984), Giddens presents his theory of 'structuration'. He defines a 'social system' as having three basic elements:

> "Structure(s): rules and resources or sets of transformation relations, organized as properties of social systems" and "out of time and space"
>
> "System(s): reproduced relations between actors or collectivities, organized as regular social practices" or "comprise the situated activities of human agents, reproduced across time and space"
>
> "Structuration: conditions governing the continuity or transmutation of structures, and therefore the reproduction of social systems". (1984: 25)

Giddens continues:

> "Analysing the structuration of social systems means studying the modes in which such systems, grounded in the knowledgeable activities of situated actors who draw upon rules and resources in the diversity of action contexts, are produced and reproduced in interaction".

There is a duality, he asserts: "the structural properties of social systems are both medium and outcome of the practices they recursively organize". However, as demonstrated by these quotes, his explanation of his theory is obscure. Gilbert (1995: 147) suggests that Giddens is arguing that:

> "human action is both constrained and enabled by social structures, for this is the medium through which action is performed. Structure is at the same time both the outcome of knowledgeable human conduct and the medium that influences how conduct occurs".

Given the wide scope of his theory and its lack of clarity, not surprisingly it has been widely criticised (see for example, Craib, 1992).

Collins (1981) proposed linking macro and micro through what came to be labelled as 'situationalism'. Institutions are made up of people and thus interacting with an institution means interacting with people and thus "micro levels of interaction are not between individuals as isolated social entities, but between individuals considered as bearers of the affairs of larger social units" so that micro and macro are linked through "interaction ritual chains" (1981, 1987). Collins identifies (1987: 196) the "'mesostructure', the network of repeated encounters". The idea of a mesostructure fits with "classical sociological theory" which "has from the very beginning….incorporated the notion of a middle level of social connectedness that generates both individuals and institutions" (Rawls, 1992).

Habermas, too, acknowledges that interacting with institutions means interacting with people but builds on it differently. He sees "society as a system that has to fulfil conditions for the maintenance of sociocultural lifeworlds" where the lifeworld is "a context forming background of processes of reaching understanding" by communication (Habermas, 1987: 126 & 151-2). Essentially, in the lifeworld, people interact with people as people, for companionship and emotional support. It is the basis of personal life. This contrasts with interacting with people as representatives of the state and the economy, the 'system', when interactions are task-driven and functional (Myerson, 2001: 30; Edgar, 2006: 89).

However, this micro-macro view of the world is now being directly challenged by those who argue instead that as a result of new communications technologies, the social world should be seen as intersecting networks. Castells argues (2000: 500) that "networks constitute the new social morphology" because of the new information technology has made information "the key ingredient of our social organization" and "flows of messages and images between networks constitute the basic thread of our social structure" (ibid: 508). And Meyrowitz (2004) argues that:

> "external perspectives are no longer minimal. Today's consciousness of self and place is unusual because of the ways in which the evolutions in communication and travel have placed an interconnected global matrix over local experience. We now live in "glocalities." Each glocality is unique in many ways, and yet each is also influenced by global trends and global consciousness."

Furthermore:

> "The media-networked glocality also affords the possibility of having multiple, multi-layered, fluid, and endlessly adjustable senses of identity. Rather than needing to choose between local, place-defined identities and more distant ones, we can have them all, not just in sequence but in overlapping experiences."

He concludes that "We are witnessing both macro-level homogenization of identities and micro-level fragmentation of them". Yet, as noted in Chapter 1, Simmel (1922/1955: 140-2) argued long ago that the many different groups an individual belongs to in effect defines an individual's identity. Thus Meyrowitz appears to be saying that this effect is stronger than in times past and not so tied to physical location.

Wellman (2000: 7.5) makes much the same point, drawing on the gemeinshaft-gesellshaft dichotomy:

> "Since the Agricultural Revolution, community has traditionally been based in agricultural villages/towns, itinerant bands, and urban neighbourhoods. People walked to visit each other in spatially compact and densely-knit communities. These communities were bounded, so that most relationships happened within their gates rather than across them. They were not necessarily immobile, but even in big cities and trading towns, much intercourse stayed within neighbourhoods. Most people in a settlement knew each other. They were limited by their footpower in whom they could contact. When they visited someone, much of the neighbourhood knew who was going to see whom. Contact was between households as much as between individuals, with the sanction - or at least the awareness - of the settlement."

Like Meyrowitz, Wellman sees digital technologies resulting in a fundamental change in that connections are now made between people rather than places. Wellman argues that this weakens "attachment to the family or workgroup that inhabits that place" because in the place-based, gemeinshaft society people interacted with one group at a time whereas in a computer supported social networked, gesellshaft society there is much greater scope to move between social circles (2000: 8.3). However, the recent development of social networking websites such as *Facebook* (www.facebook.com) bring together different social circles, allowing them to overlap in a way not previously seen, the opposite direction to that supposed by Wellman.

Urry (2003a: 122) has taken this further, arguing that "there is no 'structure' and no 'agency', no 'macro' and no 'micro' levels, no 'societies' and no 'individuals', and no 'system world' and no 'lifeworld'". Instead, he says, there is 'relationality' brought about "through a wide array of networked or circulating relationships that are implicated within different overlapping and increasingly convergent material worlds". There is, he proposes, a "global interdependence" with no tendency towards equilibrium but containing "pockets of ordering" and "'global fractals', the irregular but strangely similar shapes that are found in very different scales across the world, from the household say to the UN". These pockets form "the strange attractor of glocalization" (ibid: 86, 102-103). Using the language of 'complexity theory' – such as his reference to strange attractors – Urry appears to be arguing both that world has changed and that complexity theory provides a new ways of looking at the world.

The extent to which the 'world has changed' due to new technology is discussed in Chapter 8. I now turn to the discussion of complexity theory and emergence.

3.2 Emergence

The key difference between the micro and the macro views is whether society has properties that are not just the sum of individuals or their actions i.e. whether there is emergence. However, the term 'emergence' has been used in many ways (Sawyer, 2005: 3; Cederman, 2005), and often loosely (see for example Urry (2003a: x) – quoted above – who admits he uses the terms qualitatively "with no attempts to apply the mathematics of chaos and complexity"). So I shall start by clarifying the terminology.

It is generally recognised that J.S. Mill first identified emergence in his *Logic*, which was first published in 1843. Given this, together with the clarity with which he expresses the ideas, I make no apology for setting out his position at some length.

Borrowing from mechanics, Mill identified what he called ‘the Composition of Causes’ in which:

> “one cause never, properly speaking, defeats or frustrates another; both have their full effect. If a body is propelled in two directions by two forces, one tending to drive it to the north and the other to the east, it is caused to move in a given time exactly as far in both directions as the two forces would separately have carried it; and is left precisely where it would have arrived if it had been acted upon first by one of the two forces, and afterwards by the other” (Mill, 1843/1973: Book III, Ch VI, 370-371)

In contrast, chemical reactions, Mill suggested, operate differently:

> “The chemical combination of two substances produces, as it is well known, a third substance with properties different from those of either or the two substances separately, or of both of them taken together” (ibid)

Thus for ‘mechanical’, but not ‘chemical’ causes, Mill says:

> “we can compute the effects of combinations of causes, whether real or hypothetical, from the laws which we know to govern those causes when acting separately; because they continue to observe the same laws when in combination which they observed when separate” (ibid).

In conclusion:

> “This difference between the case in which the joint effect of causes is the sum of their separate effects, and the case in which it is heterogeneous to them; between laws which work together without alteration, and laws which, when called upon to work together, cease and give place to others; is one of the fundamental distinctions of nature” (ibid: 373)

In modern parlance, Mill’s mechanical process can be described as reductionist, because the whole can be ‘reduced’ to its components. Other writers have invented their own terminology. For example Smith (1997:58) uses the term ‘aggregative’. Cilliers (2000) talks about complicated processes which can have many stages but the relationships between them are fixed and can be defined by rules.

In contrast, the chemical mode can be described as emergence and is now intimately connected with the idea of non-linear, dynamic, complex systems (see Goldstein, 1999). Cilliers (2000) says that complex processes are:

> "constituted through a large number of dynamic, non-linear interactions…..Living things, language, cultural, and social systems are all complex… complex things have emergent properties, complicated things do not".

Complex systems display emergent features arising out of interaction and these features are "neither predictable nor deducible from lower or micro-level components" (Goldstein, 1999).

It is worth briefly noting here the difference between complex and chaotic systems. Strogatz (1994: 3) defines chaos as occurring when

> "a deterministic system exhibits aperiodic behavior that depends sensitively on the initial conditions, thereby rendering long-term prediction impossible".

Chaotic systems are highly non-linear; equilibrium is never obtained although irregular aperiodic oscillations called 'strange attractors' can emerge. The weather system is the classic example. While complex systems are also non-linear, they are less sensitive to the initial conditions than chaotic systems. Chaotic systems are then special cases of complex systems.

How do these concepts apply to society? Mill, however, was quite clear that the chemical mode did not apply to society: in Book VI of his *Logic* (1843/1974: Book VI, Ch VII: 878) Mill wrote:

> "The laws of the phenomena of society are, and can be, nothing but the laws of the actions and passions of human beings, united together in the social state. Men, however, in a state of society, are still men; their actions and passions are obedient to the laws of human nature. Men are not, when brought together, converted into another kind of substance, with different properties; as hydrogen and oxygen are different from water…Human beings in society have no properties but those which are derived from, and may be resolved into, the laws of the nature of individual man. In social phenomena the Composition of Causes is the universal law".

This is the micro-agency view that society is the 'simple' sum of its parts. It is echoed today by symbolic interactionists such as Goffman (1983: 6) who argue that most of the rules that frame society are created and maintained "from below".

In contrast, Durkheim was the first sociologist to talk about emergence, although he did not actually use the term (see for example, Sawyer, 2005: 100-108). Durkheim's *sui generis* argument is essentially that the interactions of people create society but that society in turn shapes them:

> "Because individuals form a society, new phenomena occur whose cause is association, and which, reacting upon the consciousness of individuals, for the most part shapes them. That is why, although society is nothing without individuals, each one of them is much more a product of society than he is the author" (Durkheim, 1893/1984: 288 fn16).

Later Mead (1938, quoted in Mihata, 1997) said:

> "When things get together, there then arises something that was not there before, and that character is something that cannot be stated in terms of the elements which go to make up the combination. It remains to be seen in what sense we can now characterize that which has so emerged."

Echoing Mead, Gilbert & Troitzsch (2005: 11) say:

> "Emergence occurs when interactions among objects at one level give rise to different types of objects at another level. More precisely, a phenomenon is emergent if it requires new categories to describe it which are not required to describe the behaviour of the underlying components".

And Sawyer (2005: 98-99) adds:

> "most social properties on the face of it seem to meet the criteria of non-reducibility…most social properties are nonaggregative, many social systems are not decomposable, most are not functionally localizable".

Examples of emergence can be found throughout social systems. Indeed, Gilbert (1995: 149) argues that societies as a whole are "best considered as emergent phenomena arising from the interaction of social institutions", defining 'an institution' as:

> "an established order comprising rule-bound and standardized behaviour patterns. Examples include the family, tribes and other collectivities, organizations, legal systems and so on" (ibid, 151-152).

These social institutions are examples of emergent phenomena because their properties cannot be identified with an individual. By way of illustration, Sawyer suggests (2005: 13) a 'church' (in the sense of an organisation, not a building) is an example as "'being a church' cannot be localized to any of the individuals belonging to the church". Taking an example from economics, corporations are similar in that they are regarded as entities in their own right in law, independent of their directors, shareholders and employees and with characteristics that are independent of those people, such as profit. At the most micro level, a conversation could be said to be emergent as it arises out of the interaction of at least two people and has properties that are separate from those taking part, such as duration.

Furthermore, people can observe and react to emergent properties. This reaction to emergent properties in social systems has been labelled 'second-order emergence', 'intrinsic emergence' or 'immergence' (Gilbert & Troitzsch, 2005; 10: Cederman, 2005; Sawyer, 2005: 172). This is an important feature of social systems. Norms are generated and maintained, 'reproducing society'. Prices emerge from the operation of markets and people react to those prices resulting in changes in supply and demand, changes in the market. In terms of the subject matter of this study, communication and transport, Sawyer (2005:3) cites the traffic jam as an emergent feature. More fundamentally, Root (2000: 439) talks about transport and communication modes creating complex, non-linear patterns that can create new social connections. More will be said about this in Chapter 4.

Strogatz (1994: 8-9) explains:

> "…linear systems can be broken down into parts. Then each part can be solved separately and finally recombined to get the answer. But many things in nature don't act in this way. Whenever parts of a system interfere, or cooperate, or compete, there are nonlinear interactions going on. Most of everyday life is nonlinear."

It is therefore essential to be able to look at the whole system as well as the individual components, the macro as well as the micro. As Bak (1996: 60) observes that if faced with a "complex, critical system":

> "Studying individual grains under the microscope doesn't give you a clue as to what is going on in the whole sandpile."

By a 'complex, critical' system, Bak means a large system in which minor changes may have major consequences (ibid: 1-2). The idea is that if you pour grains of sand into a pyramidal pile, as the pile grows, landslides will occur. These landslides are unpredictable and can be very large. This sort of effect is thought to underlie stock market crashes (Gilbert & Troitzsch, 2005: 10)

Simmel saw society not as a "'substance,' nothing concrete, but an event" in which "languages, social structures, norms, and conventions are created through 'societal production, according to which all these phenomena emerge in interactions among men'" (Cederman, 2005 quoting from Wolff). Elias, too, saw a process and coined the term "figuration" to describe it. He said "If four people sit around a table and play cards together, they form a figuration" (Elias, 1970/1978:130). The game is the outcome of the interaction, figuration is "the changing pattern created by the players as a whole" (ibid).

Thus society should be seen as a dynamic, emergent social phenomenon – i.e. with characteristics independent of individuals – yet created by the interaction of individuals. As Elias (1970/1978:132) said "if sociology has to investigate figurational processes resembling complex games, then statistical aids must be developed which will be suited to this task". Agent-based models appear to be just such an aid and to these I now turn.

3.3 Introduction to Agent-Based Models

Agent-based models, however, can simulate emergent societies created by the interaction of individuals. This Section explains in general how these models work (leaving detailed, technical discussions to Chapter 5).

An agent-based (AB) model – strictly speaking, a multi-agent-based model – is an artificial world of heterogeneous agents in which each agent can interact both with other agents and with its environment (Gilbert & Troitzsch, 2005: 173). The interaction between these agents generates complex, emergent behaviour that can be compared with macro phenomena (See e.g. Gilbert & Troitzsch, 2005: 11-12). AB models "explore the simplest set of behavioral assumptions required to generate a macro pattern of explanatory interest" and implement Coleman's view (discussed in the previous Section) combined with emergence (Macy & Willer: 2002).

The agents themselves are models – that is, simplifications – of actors in the sociological sense. Usually, an agent represents a person, but they can represent a household, a firm or even a nation (Gilbert, forthcoming: 15). According to Wooldridge and Jennings (1995, quoted in Gilbert & Troitzsch, 2005: 173) these agents "typically" have "autonomy", "social ability", "reactivity" and "proactivity". Gilbert (forthcoming: 37) and Macy & Willer (2002) explain these terms as follows:

- autonomy, in that there are no authorities telling the agent what to do, rather the agent follows the rules that are programmed: they are self-organised
- social ability: agents interact with one another, for example, influencing and imitating
- reactivity: agents react to their environment, sometimes indirectly influencing one another by changing their shared environment
- proactivity: agents can take the initiative to pursue their own goals.

These ideas are, however, rather difficult to implement. Gilbert (forthcoming: 37) suggests instead agents have:

- perception: agents "can perceive their environment", including other agents in the vicinity
- performance: agents have "a set of behaviours that they are capable of performing" such as moving and communicating
- memory of their "previous states and actions"
- policy: "rules, heuristics or strategies" that determine what they do next. These rules are simple. Macy & Willer (2002) point out that:

  > "we follow rules, in the form of norms, conventions, protocols, moral and social habits, and heuristics. Although the rules may be quite simple, they can produce global patterns that may not be at all obvious and are very difficult to understand."

The key feature that differentiates AB models is that interaction between heterogeneous agents is allowed, unlike microsimulation which permits heterogeneity but not interaction and systems dynamics which permits neither. How AB models work will hopefully become clearer with the example presented in the next Section.

3.4 A Simple Example of an Agent-Based Model

This Section describes a very simple AB model designed to explore the take-up of fixed-line phones. Following the modelling procedure set out in the Introduction, this Section first sets out the observed phenomenon and the data, then the theory and finally the model and the results.

The Observed Phenomenon and the Data

Britain's first public telephone exchange opened in London in August 1879 to serve eight subscribers. The first phone directory was issued in 1880 and "contained details of over 250 subscribers" plus details of 16 provincial exchanges (BT Archives, 2007). By 1882, there was one phone for every 3,000 people in London: by 1890, the ratio was up to one in about 800 but it did not reach one per 100 until 1905 (Perry, 1977). However, these numbers included business as well as private subscribers. In fact, examination of the 1880 Directory shows that virtually all the subscribers were businesses (Telephone Co., 1880). However, the data series for households only starts in 1964, by which time 21.6% of British households had fixed line phones. This percentage peaked at 95% in 1999, after which it started to drop as mobiles were substituted (CSO, 1994; ONS, 1998; ONS; 2007). Thus phones took some 80 years to spread from virtually no households to around 20% and almost 120 years to reach 95%.

Initially, there was limited geographical coverage: London's first trunk line was not opened until 1884 and it was not possible for Londoners to call "the Midland and Northern Counties" until 1890 even though the first line to Paris opened in 1891 (BT Archives: 2007). However, 'a unified national phone system' did not become available until 1912 (ibid). Rural areas were not well covered and phones were concentrated in London: writing in 1910, Casson (1910/2007) says that a quarter of phones "are now" in London and Perry (1977) reports that it was one third in 1913.

Perry (1977) suggests that this slow take-off was due to high price and the way the nascent phone industry was regulated. Phones were expensive: in 1901 "When you could employ a maid for £20 per year, having unlimited phone use for £17 per year did not seem to be a bargain" (Perry, 1977). In today's money, this is equivalent to about £1,600, but it is better to look at what else £20 could buy in the late nineteenth century as you could not employ a maid for £1,600 today. For example, in 1901 £20 would pay for the postage of 4,800 letters (as it still cost only one penny to send a letter and there were 240 pennies to the pound). In 1885, the United Phone Company (1885: xiii) answered the complaint that the rental charge of £20 a year was too high by comparing it to £24 that was charged in Paris and Boston, £25 in Chicago and between £32 and £42 in New York. But the pricing structure was complex too and often depended on distance from the exchange (United Telephone Co., 1880: 5-6: National Telephone Co., 1891: 6, 11). It was, however, not necessary to have a phone at home to make a call: public phones were available from 1886 (BT Archives: 2007). This facilitated making phone calls to businesses but not receiving them or making person-to-person social calls.

The Theory

This simple AB model focuses on just two important factors that underlie the adoption of phones: the network effect and affordability.

- **The network effect.** In his seminal paper Bass (1969) argued that except for first adopters, take-up of new "generic classes of products" (as opposed to new models of older products) is related to the number of previous buyers. If the new product happens to be a link to a communication network, this effect is particularly important. Metcalfe's Law "is said to be true of any type of communications network" and states that "the value of a communications network is proportional to the square of the number of its users" (Briscoe et al: 2006). Essentially "the idea is that a network is more valuable the more people you can call" (Briscoe et al: 2006). Fischer (1992: 262) noted that when phones were introduced in the US they were used "to widen and deepen existing social patterns rather than to alter them". Valente (1996) argued that there are two processes at work: one reliant on the "entire social system" and the other on "an individual's personal network". The first is a matter of following changes in society in general, such as opinion leaders who are not personally known, while the other implies that an important determinant of phone adoption is whether your family and friends, that is, those in your 'personal network', already have phones. This is discussed in more detail in Section 4.3.)

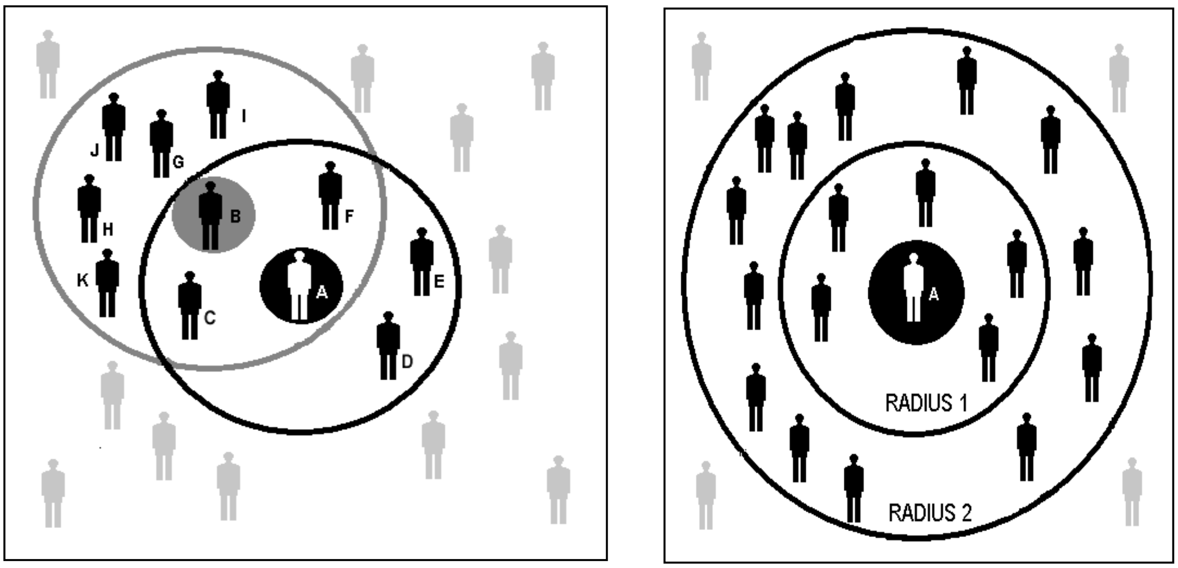
- **Affordability.** Initially, only the better off could afford phones. In the US "the more affluent households were the earliest subscribers" (Fischer, 1992: 261). In Britain, adoption was led by the aristocracy: Queen Victoria was an early adopter (Casson, 1910/2007: 86). They were followed by bankers and others in the City of London (Casson, 1910/2007: 86; Telephone Co. Ltd, 1880). Brief examination of nineteenth century directories suggests that phones spread to homes from businesses, at least partly as a result of discounts being given to owners of businesses who had phones at their work places (Telephone Co. Ltd, 1880; United Telephone Co. 1880 & 1885).

The Model and Results

In this case, let us assume agents represent households and try to create an analogous model. There are assumed to be 10,000 agents. (The exact number is not important as long as it is large.) Furthermore, these agents are divided into two types: 'Blues', who represent the affluent early subscribers and 'Greens', who represent the rest of the population. The agents are randomly distributed across their 'world' but the Blues are confined to one quadrant to reflect their geographic and social proximity; while the Greens are spread randomly throughout the whole world. On each run, the agents are distributed differently.

The size of each agent's personal network of family and friends is defined by the 'personal network radius'. This radius is the same for all agents on each run although it can be changed between runs. Nevertheless, each agent is unique in having a different personal network, depending on how the agents are distributed. This is illustrated in the left hand diagram in Fig. 3.3. Agent A has a personal network of five: agents B, C, D, E and F; while agent B has a personal network of eight: agents A, C, F, G, H, I, J and K. Increasing the personal network radius will usually increase the size of each agent's network, as illustrated in the right hand picture in Fig 3.3: with Radius 1, agent A has a personal network of five; with Radius 2, the personal network is 18.

Fig. 3.3: The 'personal network radius" used in the AB model



One of the Blues is designated the first adopter at random. In each time period the adopting Blues 'persuade' another Blue within their personal network radius to adopt. Once the adoption rate among the Blues stops rising, adoption spreads to the Greens on the same basis.

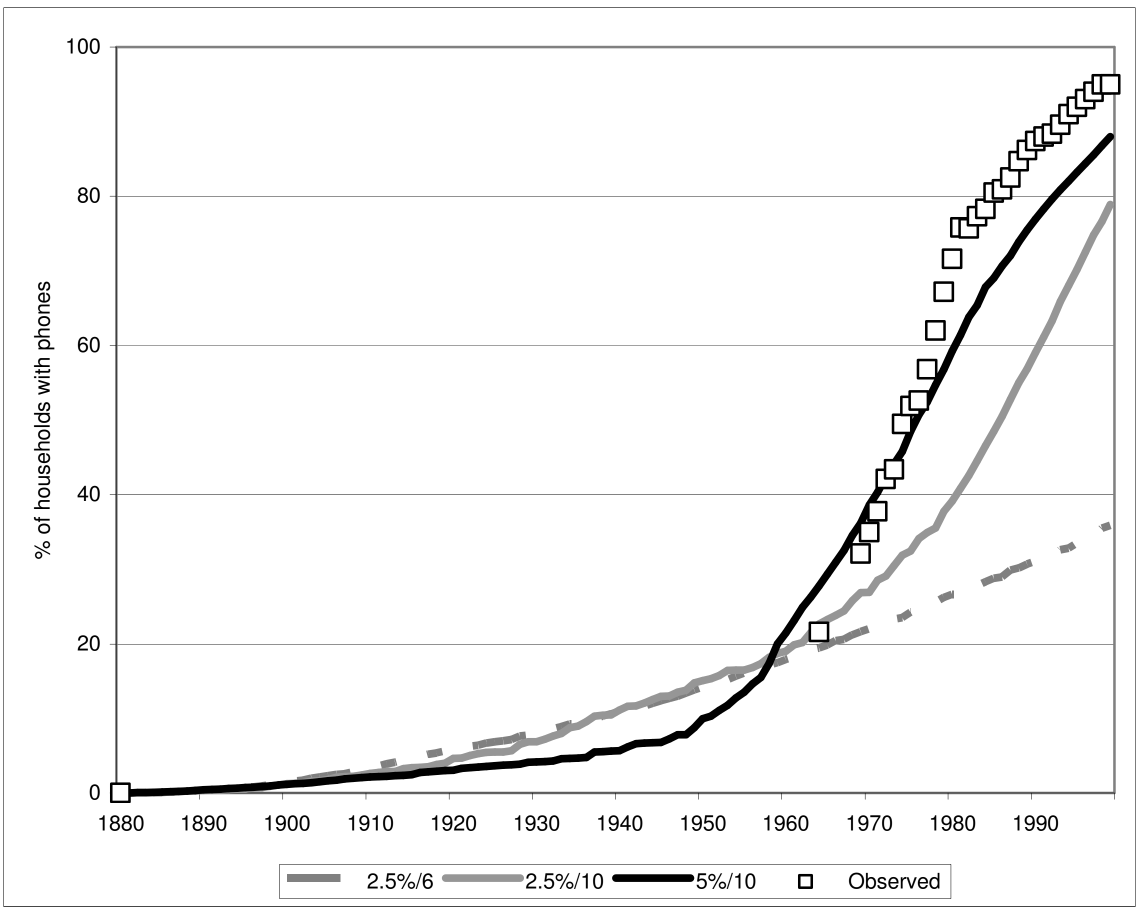
So there are just two variables: the percentage of Blues in the population and the personal network radius (PNR). The percentage of Blues in the population will, by definition, be small. The model was therefore run with the percentage of Blues set at each of five levels: 2½%, 5%, 7½%, 10% and 15%. For each percent of Blues, the personal network radius was set at 5, then 6 and so on up to 10. (Remember, this is not the size of the personal network but the radius defining the network as illustrated in Fig. 3.3.). This produced a total of 30 combinations. Only six combinations produced a curve that approximated the observed data by predicting an adoption rate of around 21 percent by 1960: these are shown in the black cells in Table 3.4. (The grey cells with black figures show where the predicted adoption rate was much too low or much too high.)

Table 3.4. Predicted adoption rate by 1960 using the AB model (%.)

Personal network radius	% Blues				
	2.5	**5**	**7.5**	**10**	**15**
5	14	9	7	6	6
6	18	10	9	6	10
7	19	10	7	8	46
8	20	11	8	45	77
9	21	8	60	82	97
10	19	21	95	100	100

Closer examination of the results from these six combinations revealed that only when it was assumed that 5% were Blues with a personal network radius of 10 did the curve approximately match the observed pattern from 1880 to 1999. Fig. 3.5 compares the observed phone adoption rate (squares) with the results from the model: with 5% Blues and a personal network radius of 10 (black line) and 2½% Blues with the radius set to 6 or 10 (grey lines). Typically, the radius of 10 implied a personal network of about 30, a realistic figure according to the literature summarised in Chapter 1.

Fig. 3.5: Observed phone adoption rate (squares) compared to output[1] from the AB model of adoption with
(a) 5% Blues and PNR of 10 (black line) and
(b) 2½% Blues and PNRs of 6 and 10 (grey lines)[2]



1. Because the agents are distributed differently on each run the result of each run varies even though the percentage of Blues and personal network radius are the same. The outputs shown here are the averages of 10 runs where the average is calculated by averaging the percent at each period. This average may never actually be seen. (See Axtell et al, 2002)
2. The intermediate radii for 2 ½% Blues produced lines between the two shown.

To sum up: this model created a simple world inhabited by agents described by just two characteristics – their social-geographical status and their personal network – who interact according to one simple rule; and from this generated approximately the observed pattern of slow adoption of phones for the first 80 years followed by a fast rise to saturation. The model has quantified the theory by showing how the adoption process might have worked, spreading incrementally from household to household, initially within a socio-economic group comprising just 1 in 20 households, and later spreading in the same way to the wider population. Thus it indicates the likely size of the 'early adopter' group, namely 1 in 20 rather than 1 in 5 or 1 in 100. And it suggests that the process did not have to change for adoption to 'take-off': the transmission process was the same for the late adopters as for the early adopters. The fact that both the theory and the numbers together tell a believable 'story' suggests that the model might be capturing something of the underlying processes despite its obvious simplicity.

3.5 Systems Dynamics Models

As noted in the Introduction, system dynamics (SD) modelling is a macro, top-down approach which uses equations to model how a system of interacting variables moves from one state to another over time. This Section uses a SD model to describe the same phenomenon described by the AB model in the previous Section, namely the adoption of fixed-line phones.

The logistic equation, devised by Verhulst in 1838 to describe the growth of populations (Strogatz, 1994: 22-23) can be used to produce a simple adoption model. Working through the mathematics and translating the equation into words, it says that the take-up at any given time depends on the take-up in the previous period and the growth rate. Specifically the increase between one year and the next is given by:

% adopting x % of non-adopters x growth rate

Thus, for example:

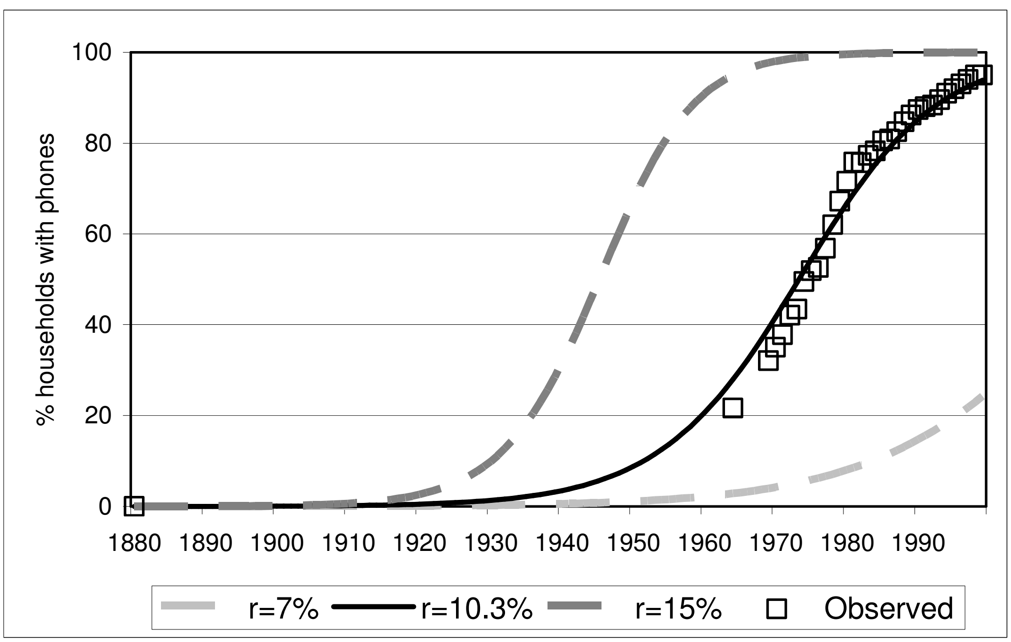
- if the adoption rate is 20 percent and the growth rate 10 percent, the adoption rate the next year will be about (20% x 80% x 10% =)1.6 percentage points higher i.e. 21.6 percent

- if the adoption rate were 50 percent, the adoption rate would be (50% x 50% x 10% =) 2.5 percentage points higher in the next year i.e. 52.5 percent
- if the adoption rate were 70 percent, the rise would be around (70% x 30% x 10%=) 2 percentage points higher i.e. about 72 percent.

Because of the way in which the change in the adoption rate depends on the level of adoption rate, the system is said to be non-linear (Strogatz, 1994: 9-10).

Now if the growth rate is set at 10.3%, this SD model can approximately reproduce the adoption pattern of fixed line phones from 1880 to 1999. This is illustrated in Fig. 3.6 which shows the actual adoption rate as squares and the result of the SD model with a growth rate (r) of 10.3% as a solid line. (The figure also shows that a growth rate of 7% would be too slow and 15%, too fast).

Fig. 3. 6: SD model of the adoption of fixed line phones



Thus the take-up of phones can be modelled in two ways: using the agent-based (AB) approach and using SD approach. The choice between SD and AB models seems often to be presented as an either/or choice, top-down versus bottom-up, macro versus micro. Yet as Fishwick (1995) noted, models of different types can be combined to answer different types of question about a given process. Van Dyke Parunak et al (1998) argued that AB modelling "is most appropriate for domains characterized by a high degree of localization and distribution and dominated by discrete decisions" and that the choice between the two approaches should be made on a case-by-case basis.

But on what basis should that choice be made? Doran & Palmer (1995) suggest:

> "A standard modelling principle is that the level and complexity of a model should be chosen so that it answers the questions and embodies the theoretical elements we are interested in, but is otherwise as simple as possible"

How can that principle be applied? I suggest three basic criteria can be used: goodness-of-fit, fitness-for-purpose and simplicity.

- Goodness-of-fit. The SD model provides a unique result for each set of parameters. Whereas the results for the AB model vary between runs because the distribution of the agents across the 'world' varies and for this reason, the average is used to measure goodness-of-fit. While there are various ways of dealing with this stochastic variation, the simplest is to compare the average for the AB model with the SD model. The SD model provides the best fit in this example. As noted in the Introduction, a good fit is a necessary but not sufficient condition for accepting a model.
- Fitness-for-purpose. As Forrester pointed out long ago (quoted by Qudrat-Ullah, 2005) the validity of a model should be judged by its suitability for a particular purpose. In this case, it could be argued that the SD model simply reflects the network effect. But it provides no insight to understanding the underlying processes whereas the AB model suggests an explanation, telling "a story" that is consistent with the literature.
- Simplicity or Occam's razor: *entia non sunt multiplicanda praeter necessitatem* or "entities are not to be multiplied beyond necessity" (*Oxford Dictionary of Philosophy*, 1994). The SD model is undoubtedly the simpler to describe and simpler to program. It is also runs faster.

To sum up: in this example, SD scores well on goodness-of-fit and simplicity but low on explanatory power. (It is therefore not surprising to see that this is the approach chosen by UK phone supplier BT to model phone uptake: see Lyons et al, 1997). The AB model scores lower on fit and simplicity but much higher on explanatory power in that it tells 'a story', suggesting how adoption might have spread from household to household (as explained in Section3.4 above).

Thus the two methods complement one another and can be used together:

- Explanation provided by the AB model can be used to refine the SD model. For example, the AB model suggested that the 5 percent early adopters were key to the process: applying this finding to the SD model by splitting to population into two, early adopters and the rest, with different rates of growth can improve the fit of the SD model.
- Conversely, the SD model can be used to refine the AB model. The simplicity of the SD model facilitates rough-and-ready testing of the impact of changing the parameters suggested by the AB model. For example:
    - using the SD model it is easy to see the impact of varying the percentage of affluent early adopters from 1% to 15% and shows that this assumption did not matter much for the first 70 out of the 120 years being studied.
    - the SD model can be used to demonstrate that if 5 percent of the population were affluent early adopters then unless their 'growth rate' is at least 7 percent it is not possible for overall take-up to exceed 90 percent in 120 years. In other words, a slow initial take-up can have very long-term consequences!

(For a technical discussion of the difference between AB and SD models, see Hamill, 2007).

Thus the two kinds of models can be used side-by-side to enrich each other and this is the approach followed in this thesis.