ISSN 2013-9004 (digital); ISSN 0210-2862 (paper)
http://dx.doi.org/10.5565/rev/papers.2083

Papers 2014, 99/4
553-578

Social contagion and homophily within romantic networks: A simulation analysis

Francisco Linares
Universidad de La Laguna
flinares@ull.es

Received: 24-03-2014
Accepted: 13-06-2014

Abstract

La logique du social anticipated, 35 years ago, the current conception of complex adaptive systems, where agents behave following some kind of rational behavior, interact among themselves and, as a result of those interactions, produce an unintended outcome. The aim of this paper is to illustrate Raymond Boudon's approach by means of an agent based simulation model, where agents represent teenagers who attempt to find a romantic relationship looking for a partner within their friendship ties. Partner choice is ruled by a homophilic principle which is, in the case of this artificial society, assumed to operate in a very simple way: agents look for someone who is similar to them in a given socio-cultural trait. At the same time, the value of this trait (which is assumed to be measured on a quantitative scale) for every agent is influenced by the values of other agents in the immediate environment. It is shown how these two social processes are interdependent, thus producing outcomes which are causally related.

Keywords: homophily; partner choice; social contagion; social networks; social simulation.

Resumen. *Contagio social y homofilia en una red romántica: un análisis de simulación*

La lógica de lo social avanzó, hace treinta y cinco años, la actual concepción de los sistemas adaptativos complejos, en los que los agentes actúan siguiendo alguna forma de comportamiento racional, interactúan entre ellos y, como resultado de tales interacciones, producen un efecto inintencionado. El objeto de este trabajo es ilustrar la aproximación de Raymond Boudon a través de un modelo de simulación basado en agentes, donde los agentes representan a adolescentes que intentan mantener relaciones románticas buscando a su pareja entre sus lazos de amistad. La elección de pareja se rige por un principio de homofilia que, en el caso de esta sociedad artificial, se asume que opera de una forma muy sencilla: los agentes buscan a alguien que sea similar a ellos en un cierto rasgo sociocultural. Al mismo tiempo, el valor de ese rasgo para cada agente (que, se asume, puede medirse en una escala cuantitativa) se ve influenciado por los valores de otros agentes en su entorno inmediato. Se muestra cómo estos dos procesos sociales son interdependientes, por lo que producen resultados que están causalmente relacionados.

Palabras clave: homofilia; elección de pareja; contagio social; redes sociales; simulación social.

Summary

1. Introduction

The aim of this paper is to illustrate Raymond Boudon's sociology by means of the analysis of a population of artificial agents. These agents represent teenagers who attempt to find a romantic relationship looking for a partner within their friendship ties. Partner choice is ruled by a homophilic principle which is, in the case of this artificial society, assumed to operate in a very simple way: agents look for someone who is similar to them in a given socio-cultural trait. At the same time, the value of this trait (which is assumed to be measured on a quantitative scale) for every agent is influenced by the values of other agents in its immediate environment. By building an agent-based model, I attempt to show how these two social processes (contagion of a social trait and partner choice) are interdependent, thus producing outcomes (i.e., number of relationships and variation of the trait among agents) which are causally related, even if statistical analysis would suggest the contrary.

The analysis briefly summarized here shares, not accidentally at all, a number of characteristics with Boudon's approach to social explanation. It is assumed that agents act in a certain social environment which, in this case, is fixed. It is also assumed that agents pursue certain goals (having a romantic relationship) and that, in pursuing these goals, they adjust their behavior to each other. It is also assumed that as an effect of this mutual adjustment a feedback process between agents' actions and their results happens, thus producing a number of social outcomes (properties of the system of action) which were not intended by the agents themselves.

The paper will proceed as follows: first a brief exposition of the characteristics of Boudon's sociology is provided. Then follows a small account of two simple simulation models of contagious and partner choice in which every social dynamic is considered to be independent of the other, thus making them useful as baseline models. Thirdly the social-contagion-and-partner-choice-model (SCPCM) is briefly described; details of the model can be found in the appendix. The most important results of the analysis of the model are then shown. Finally, the paper ends with a discussion of the results and a brief valuation of Boudon's sociology.

2. The sociology of Boudon

"Complex Adaptive Systems" is the term which has been coined to refer to systems – whether social, biological, or of another kind – whose aggregate behavior is the result of actions performed by different individuals who continuously adapt to an environment, which is itself constituted by other individuals also engaged in an adaptive behavior (Miller and Page, 2007). These multiple behaviors, may combine in complex ways to produce self-reinforcing dynamics that are difficult to understand. This is so because the behavior that emerges from individuals' mutual adaptations, aggregate behavior, may in fact be the product of a long chain of interactions and show patterns that are quite distant from individuals' initial motivations to act.

The conception of social systems as complex adaptive systems was advanced, 35 years ago in Raymond Boudon's pathbreaking books: *La logique du social* (1975) and *Le place du désordre* (1984). Summarizing Boudon's theorization:

1) The social explanandum, whether the development of a certain agricultural society, the spreading of the use of a new drug among doctors, or the rate of anomie in a population, is the product of the combination of individual actions.
2) These individual actions are to be explained within a Weberian paradigm, which is to say that reasons must be found to account for those actions. These "reasons" needn't be all of the same type.
3) Individuals themselves act in a given environment, which they usually cannot control and which influence their actions. Thus, as an example, doctors working in a hospital will not be influenced by each other in the same way as private practitioners. So, even if they have the same motivation to treat ill patients, the influence process will be different in each case.
4) These basic principles lead, depending on the characteristic of the system, to reproductive processes, cumulative processes or transformational processes. Cumulative and transformational processes imply feedback dynamics between the system of action and its outcomes.

Besides these basic guidelines for theory construction, it is interesting to remark on two further features of Boudon's work: the use of well-established sociological knowledge, often from the classic period, to develop his arguments, and the use of simple formal models in order to better present his argument and make his conclusions easier to understand (and difficult to refute). A clear example is his examination of the relative-deprivation effect found in Stouffer's *American Soldier* by means of a simple lottery model where rewards depend on the number of people buying a ticket. This allows him to show in a quite straightforward way how higher opportunities of upward mobility (environmental conditions) make it reasonable for a larger amount of people to "buy a lottery ticket" (Weberian paradigm), thus producing widespread frustration (social fact to be accounted for). In this way, Boudon shows how a very simple mechanism operating at the motivational level of action, together with certain structural conditions, is enough to account for a perverse unintended social effect which has previously appeared as a sort of contradiction: the higher the opportunities the higher the frustration.

Although the potential of this paradigm is successfully demonstrated in his writings, one may wonder whether this method would work with more complex systems, where less stylized facts are to be explained. What if different individuals have different attitudes towards risky options, hold different assets, etc. And, even more, what if the decisions in realm A, let's say working opportunities, is somehow interconnected by decisions in realm B, let's say friendship relationships, which is itself ruled by different mechanisms.

Of course these questions are not easy to answer in a rigorous, deductive way. Fortunately, there is nowadays a tool that may help: agent-based models (ABMs). This tool was not available when Boudon wrote his masterpieces, although he was able to use a different type of simulation (micro-simulations) in addressing the issue of educational choices. An ABM is a formal and simple representation of the reality which, unlike other formal and simple representations (such as differential equations), can easily deal with heterogeneity in a population of individuals (that is, individuals may differ in many traits), as well as with decision rules other than rationality. ABMs are thus nicely suited for analyzing complex adaptive systems.

These models have several applications. By means of empirically calibrating their parameters, they have been used to explain real data, such as fertility trends in France (González-Bailón and Murphy, 2013), local youth unemployment rates in Stockholm (Hedström, 2005), or educational achievement in France (Manzo, 2013). Whatever the empirical applications, in the realm of sociology ABMs are also a formal tool for developing and exploring the implications of middle-range theories (Gilbert, 2008). According to this aim, which is also the aim of this paper, ABMs are mainly used to explore the logical consequences following from a set of assumptions about the characteristics of agents, their rules of interaction, and the characteristics of the environment. The main theoretical and methodological implications of these kinds of models, related to the sociology of Boudon, are: bottom-up explanations, the analysis of cumulative systems and the production of artificial experiments.

2.1. Explanations, generative social science and mechanisms

As Epstein and Axtell (1996) nicely put it at the end of their pathbreaking *Growing Artificial Societies*, where the now well-known Sugarscape model is analyzed:

> From an epistemological stand point, what "sort of science" are we doing when we build artificial societies like Sugarscape? Clearly, agent-based social science does not seem to be either deductive or inductive in the usual senses. But then what is it? We think generative is an appropriate term. The aim is

> to provide initial micro-specifications (initial agents, environments and rules) that are sufficient to generate the macrostructure of interest. We consider a given macrostructure to be "explained" by a given micro-specification when the latter's generative sufficiency has been established. (p. 177)

This "generative" approach, which implies that a given social pattern is explained from the bottom-up, provides a sensible answer to the micro-macro problem masterfully traced by Coleman (1986) ten years earlier. It is also very close to Elster (1989) and others' (e.g., Hedström and Swedberg, 1998) defense of "mechanisms" as the building blocks of sociological explanations. According to Hedström and Bearman (2009: 5), a mechanism "refers to a constellation of entities and activities that are organized such that they regularly bring about a particular type of outcome." Although a common unit of analysis in sociology is the individual, nothing in the concept of "mechanism" precludes the unit either being a "supra-individual" entity, such as a collective, or a "sub-individual" entity, such as the components of an individual decision-making process (e.g., attitudes, values, emotions, etc.). The concept of "mechanism" does not exclude a rational conception of action either. What the concept of "mechanism" does imply is that whatever the entities and their rules of behavior, it has to be shown that they must regularly produce the outcome that is to be explained. This conception of social explanation is essentially the same as that proposed in points 1 to 4 summarized above.

2.2. Emergence and cumulative systems

One of the most intriguing characteristics of society is the strong stability of many social patterns. Despite the fact that we all have the experience of living in an era of change, certain characteristics of society seem either to change very slowly or not to change at all. The distribution of wealth among different social classes, rules of domestic labor assignment and school achievement rates of students of different backgrounds are just a few examples. Assuming that all that happens in society is a result of individual actions, the question to answer is, how is it that individuals act in ways that produce such aggregate patterns, which are often unintended, undesired, and even detrimental to many of them?

The answer lies in the fact that the relations individuals produce when interacting with one other often produce a "new reality" that, so to speak, "traps" individuals. This "new reality" is called an *emergent* outcome of the system. As in the case of undergraduate students living in a residence hall who, in a few days, develop a stable system of informal rules concerning the use of the common kitchen, once a given distribution of rights and resources is established in any realm of society it will likely show a self-perpetuating trend, since agents are now forced to mutually adjust their behavior under the new conditions, eventually reaching an equilibrium (though possibly "unfair"). These complex adaptive systems, where the emergent outcome feeds back on the original system of action, are known as "cumulative systems" (Boudon, 1979) and are a common object of sociological analysis.

2.3. Research methods and simulation experiments

Both quantitative and qualitative data are essential to produce simulation models insofar as these models are not built to reflect someone's fantasies about society. The representation of reality – the model – has to be grounded in empirical knowledge of the world if it is to provide an explanation of that world at all. However, in any given theory (whatever the theoretical style), we will find concepts (e.g., properties of individuals such as "sensitivity to the influence of others") with difficult, or even impossible, empirical measure. When agent-based modelers are faced with this problem, some solution must be found in order to make the simulation run. The solution consists of substituting unknown empirical data for random numbers which are extracted from a theoretical distribution (this is why the results of ABMs must be presented as averages of a sufficiently large number of simulation runs). While this procedure may be considered an artifact, notice that it is quite honest: the modeler explicitly recognizes the lack of knowledge that in a narrative style of theorizing often goes unnoticed (and sometimes hidden under a prose whose eventual literary beauty is not an essential element of a proper explanation).

This "artificial" way of proceeding has a further advantage, which is key for the analytical agenda in social sciences: the possibility of carrying out more complex "thought experiments" than those performed in the absence of this tool (such as Boudon's lottery thought experiment). When conducting field social research, it is almost impossible to answer "what if" questions that may be relevant for increasing the understanding of a social phenomenon. When there are competing theoretical understandings of an issue, relevant questions arise, such as, "What if the topology of the social network were different?" "What if people were not sensitive to others' expectations?", etc. Nevertheless, by artificially manipulating parameters, it is possible to show whether a given prerequisite (e.g., network closure) is actually a necessary condition to "grow up" the social pattern.

3. Contagion and homophily in a social network

3.1. A simple model of contagion

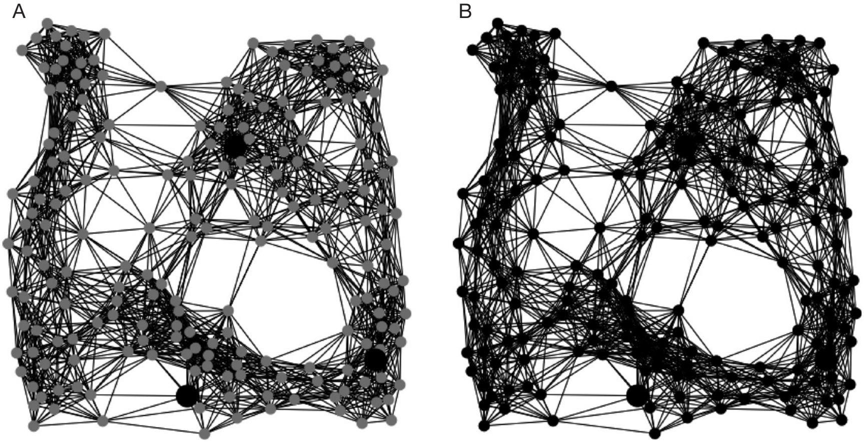
"Contagion" is a fairly well known social phenomenon. Since the celebrated study of Coleman, Katz and Menzel (1957) on the diffusion of the use of "gammanym" among doctors, it is widely accepted that the influence of peers on individuals' decision to accept or refuse a given socio-cultural trait produces a kind of "snow-ball process" that can usually be represented with a typical S-shaped diffusion curve, where the speed of the process depends on certain characteristics such as the "critical mass" or the network topology. This process usually ends with a fairly large proportion of the population adopting the new trait (Rolfe, 2009).

An example of this is Stonedhal and Wilensky's (2008) computational model of "virus on a network". The model represents a network of agents, a few of which are "infected by a virus" or, using a more general example, hold a certain trait with a dichotomic value. A number of parameters, such as the probability of virus-spreading or the probability of recovering after being infected, control the diffusion process in this model. However, the result is almost always the same. Even under the worst conditions for the spreading of infection allowed by the model, a very small number of initially infected nodes will foster the spreading of the virus through a large part of the population, as soon as there is a small probability of infection (5%) and the network is dense enough.

In Stonedhal and Wilensky's model, the artificial network is produced by a very simple algorithm which asks the potential nodes to create links with randomly selected partners until a certain number of connections are created. This final number of links is the result of multiplying the number of nodes by the average node degree (the average number of links per node), which is a parameter that can be manipulated in the model. This algorithm almost grants that, in every simulation, the topology of the network fits a "small world" topology, that is, a network with an average low distance between any two random nodes.

Figures 3.1.A and 3.1.B show a typical network generated by this model with 200 nodes and an average node degree equal to 20. Parameters that refrain diffusion are set to the minimum, and the virus spread chance is set to 5%. In figure 3.1.A, the large black circles represent the initial "infected" nodes.

Figure 3.1. A) Virus on a network (initial state). B) Virus on a network (final state)

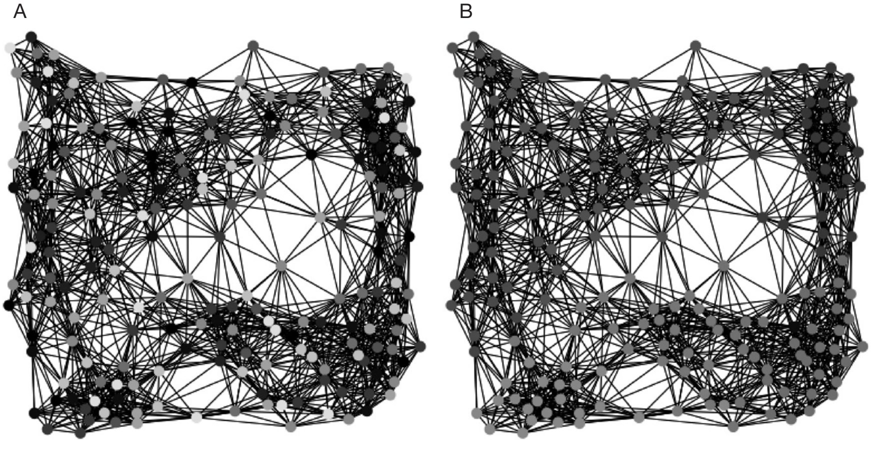


But what if the trait to be diffused is not a discrete-dichotomous variable, but a continuous one? The question can be explored by modifying the initial "virus on a network" model in the following way:

a. All nodes have a state variable (representing a social trait) which can adopt values between 0 and 9; with the values being randomly assigned.
b. Nodes are sensitive to contagion from others with a probability which is controlled by a "social influence" parameter.
c. With probability fixed at point b), the nodes change the value of their traits to the median (the use of the median instead of the mean is recommended since individuals are not likely to be influenced by extreme options, but the main results of the analysis do not depend on the use of this statistical measure) of the nodes with which they have a directed link (i.e., nodes within a path distance of 1).

Figures 3.2.A and 3.2.B show the same network before and after the contagion process is finished. The color of the nodes indicates the "intensity" of their trait, and is darker the lower the values of the trait. It is important to note two main differences between these networks: first, a clustered distribution substitutes the initial random distribution, so nodes with the same color are close to each other. And secondly, the variability of the initial distribution has been lowered; which is shown by the fact that, in this network, the initial coefficient of variation of the trait decreases from 0.6 to 0.3. This process has little sensitivity to the "social influence" parameter (i.e., the end of the process is the same even for low chances of social influence).

Figure 3.2. A) Virus on a network with a continuous trait (initial state). B) Virus on a network with a continuous trait (final state)



3.2. A simple model of homophily in partner choice

Homophily is a basic principle of the structuring of social relations, meaning that similar individuals are connected among themselves more often than dissimilar ones. This tendency may be the product of the distribution of population over relevant social attributes (Blau, 1977), the structuring impact of social foci of interaction on individuals' networks (Feld, 1981), or the preferences of individuals for similar others (Kossinets and Watts, 2009). Whatever its cause, homophilic patterns imply larger homogeneity in social relations than would otherwise be expected. It is a well documented pattern in many realms of social life (see McPherson et al., 2001 and Cruz, 2013). In the realm of partner choice, Bearman et al.'s (2004) analysis of the structure of the romantic and sexual relations of 832 students at "Jefferson High" (a high school located in "Jefferson City") provides an excellent illustration. By conducting a series of simulation experiments, the authors conclude that homophily is a necessary mechanism (even if not sufficient) to account for the topology of this romantic network.

We may replicate a simplified version of that model with the same artificial network as in the previous section in the following way:

For every node:

a) Look for a partner among your directly linked neighbors, following the rules:
    a. The partner must be single (at the initial state of the simulation all of them are).
    b. The partner must be of the opposite sex (at the initial stage of the simulation sex is assigned with 50% chance).
    c. If your sex is male, the partner must be younger than you. If your sex is female, the partner must be older than you (age of the nodes ranges from 14 to 17).
    d. The difference between the values of the trait must be equal to, or lower than, 10%
b) With a certain likelihood, which is controlled by a parameter, a relationship may be broken.
c) The decisions iterate until no new relationships emerge.

The main results of this model are shown in Table 3.1. The figures (averages and standard deviations over 50 simulation runs) show that when partner choices are not homophilic, the number of relationships created through the simulation is higher the higher the value of the probability that a relationship will be broken, β. However, when partner choices are homophilic, assuming a tolerance to partner's difference of 10%, the number of relationships created is not only lower but also less dependent on the likelihood of breaking relationships (β).

Table 3.1. Number of relationships

	No homophily		homophily	
	μ	σ	μ	σ
β = 0.25	68.3	5.4	25.8	4.0
β = 0.50	92.2	9.0	30.4	5.6
β = 0.75	158.2	18.6	30.3	6.0

4. A social contagion and partner choice model

In order to analyze how these two basic and simple dynamics interact, I have built an agent-based simulation model (which I call SCPCM) where contagious and partner choice co-evolve at the same time within a population of 200 agents embedded in the same network as figures 3.2.A and 3.2.B. The model is fully described in the appendix following the ODD protocol designed by Railsback and Grimm (2012). A brief description of SCPCM is provided in this section. Afterwards some hypotheses on the expected relationships among variables are suggested.

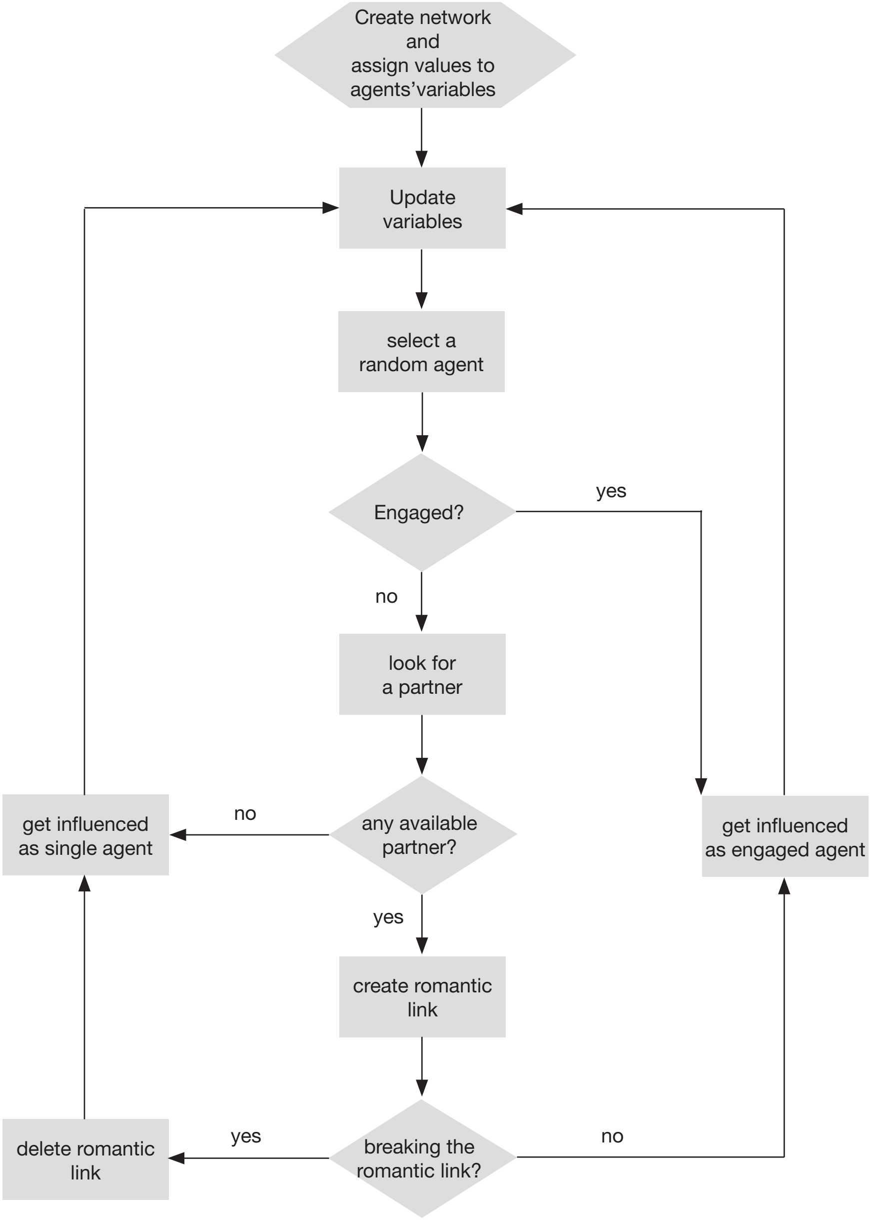
4.1. Brief description of SCPCM

The program was implemented in the Netlogo platform (Wilenski, 1998) and reproduces the following steps:

a) One of the agents is randomly chosen.
b) If the agent does not have a partner, the agent is asked to look for someone according to the following rules:
    a. The partner must be found among linked neighbors
    b. The partner must be of a different sex.
    c. If the agent is male, the partner must be younger; and the other way around if the agent is female.
    d. The difference between the values of the trait must be within a range of tolerance, which is set by a tolerance-parameter (τ).
c) If a partner is found, both agents engage. This relationship may be broken with a probability which is set by a breaking-probability-parameter (β).
d) Whether a partner is found or not, the agent is influenced by its linked neighbors according to the following rules:
    a. If the agent is not engaged, the value of the agent's trait becomes the median of its linked neighbors.
    b. If the agent is engaged, the value of the agent's trait is determined by both the value of the trait of the agent's partner, weighted by a weight-parameter (ω), and the median of the agent's linked neighbors, weighted by 1 - ω .

In summary, the model contains two different mechanisms of social interaction: on the one hand, agents select their partners following a homophilic

Figure 4.1. Flow chart



rule. The homophilic strength of the choice is determined by parameter τ, which ranges from 0 to 1.

On the other hand, agents are influenced by the other agents they are tied to, so the values of their traits converge to a central value of the local environment. This contagious process is, nevertheless, affected by the previous partner selection, since the value of the partner's trait has a special weight. The strength of the partner's influence, relative to other agents' influence, is determined by parameter ω, which also ranges from 0 to 1.

In the end there is a feedback process between partner choice and trait contagion: the distribution of trait values influences the agents' partners' pool; and, at the same time, agents' choices of partner influence the distribution of trait values. These dynamics are summarized in the flow chart above. The model attempts to show the outcomes of these reinforcing flows, paying special attention to the fact that the variation among agents' trait values within the network is determined by parameters τ, β and ω.

4.2. Hypothesis

Concerning the process of contagion, the variation in the trait distribution should be positively associated with tolerance, since tolerant individuals will be "comfortable" in a world with high diversity. It should also be negatively associated with the weight of the partner's influence, since if my partner has a strong influence on me, overall diversity is reduced. Nevertheless, there is not an obvious way to relate it to the probability of breaking a relationship. Therefore it can be expected that:

— The higher τ, the higher the coefficient of variation of the trait (H1).
— The coefficient of variation of the trait will not be sensitive to β (H2).
— The higher ω, the lower the coefficient of variation of the trait (H3).

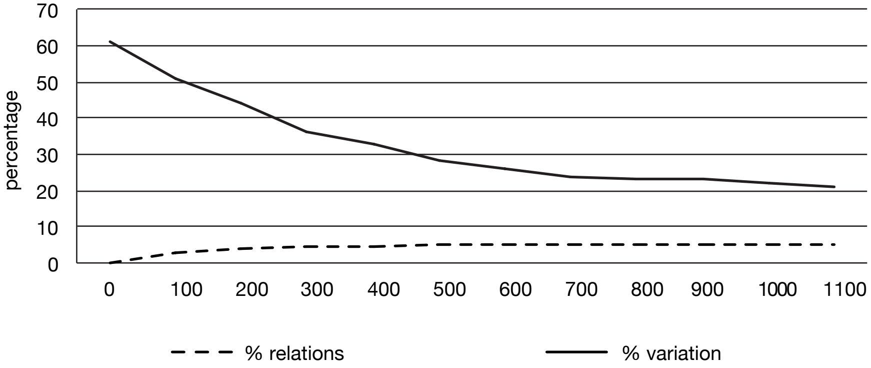
On the other hand, concerning the partner choice process, it is straightforward that as the probability of breaking romantic relationships increases, the final number of relationships must also increase. It would also seem quite obvious that the higher the tolerance to difference, the number of relationships should also increase. Therefore it can be expected that:

— The higher τ, the higher the number of relations (H4).
— The higher β, the higher the number of relations (H5).
— The number of relations will not be sensitive to ω (H6).

5. Results

A series of simulation experiments exploring the parameter space of τ, ω and β were conducted; reiterating the simulation 50 times for every experimental condition, which amounts to 66,500 simulation runs. The results of these experiments show the emergence of patterns which are quite different from the

Graph 5.0. Typical simulation run ($\beta = \omega = \tau = 0.5$)



simple models, where every dynamic operated independently. Graph 5.0 shows the evolution of the coefficient of variation (%) and number of relations (as a percentage of total friendship links) in a typical simulation run. The trends are quite clear: the trait variation continuously decreases from roughly 60% to roughly 20% as the simulation progresses, while a number of romantic relationships emerge in the early stages of the simulation. Although some of them disappear and new ones appear, the rate to total relations remains practically constant throughout the simulation run at a value of roughly 5%.

On the one hand, the number of relations is higher than in the simple homophilic partner choice (the data are actually similar to the simple partner choice model when homophilic choice is *not* allowed). On the other hand, the standard result in the simple model of contagion is just a special case of SCPCM. Thus, the almost deterministic result found in the simple model (above), in which the coefficient of variation always decreases from 0.6 to 0.3, is no longer valid. In the new model, the coefficient of variation drops below that level (as low as 0.13) for most of the combination of values of the parameter space, but increases above that value (as high as 0.43) when the values of ω are very high.

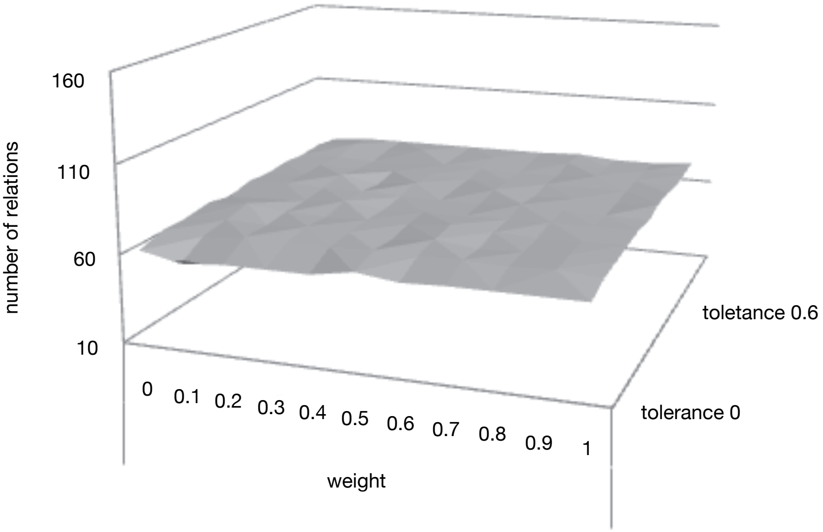
5.1. Number of romantic relationships

Concerning the number of romantic relationships, the simulation provides clear support for hypothesis H5 and H6, as can be easily seen in graphs 5.1, 5.2 and 5.3, which represent the number of final relationships for every combination of the spectrum parameter of τ and ω, when β equals 0.25, 0.50 and 0.75, respectively[1]. It is quite obvious that the variation in parameter β

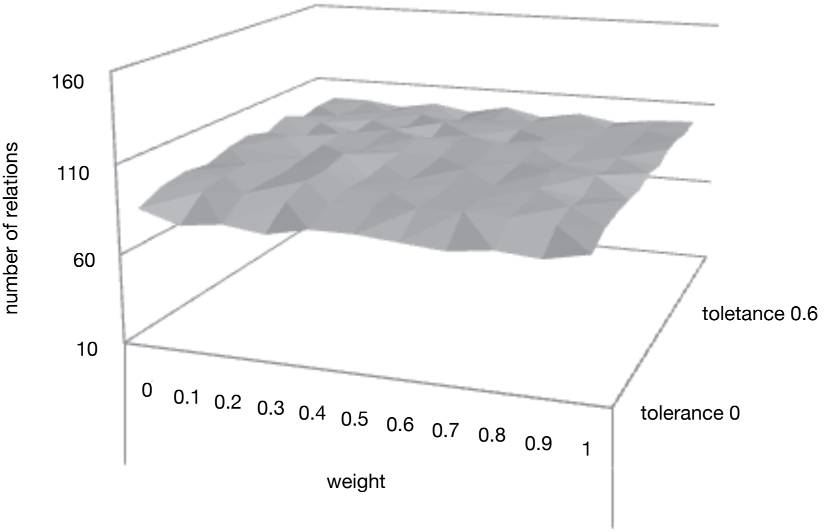
---

[1] The results shown in graphs are the mean values of 50 repetitions for every combination of parameters.

Graph 5.1. (β = 0.25)



Graph 5.2. (β = 0.50)



Graph 5.3. (β = 0.75)

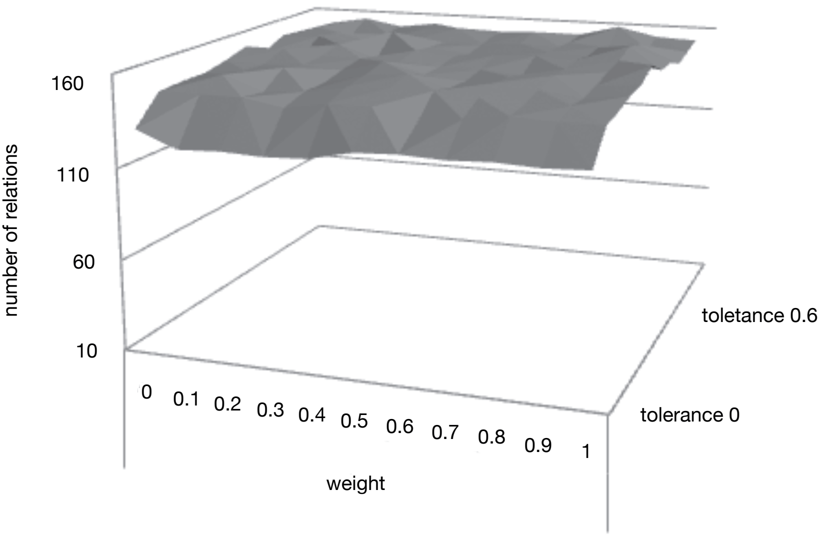


Table 5.1. Dependent variable: Number of relations

		Unstandardized Coefficients		Standardized Coefficients			99% Confidence Interval for B		Collinearity Statistics	
	Model	B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	30.075	.171		176.289	.000	29.635	30.514		
	probability-of-breaking-up	109.753	.203	.738	540.454	.000	109.230	110.276	1.000	1.000
	tolerance	8.801	.185	.065	47.606	.000	8.324	9.277	1.000	1.000
	weight	-.248	.185	-.002	-1.338	.181	-.725	.229	1.000	1.000

has the expected effect: the higher the probability of breaking a relationship, the higher the number of final relationships. It is not only the expected effect but also the greatest effect, since parameters ω and τ do not seem to have any influence. This result is clearly counterintuitive, since one would expect the number of relations to increase with tolerance to the partner's trait, as suggested by H4.

The linear multivariable regression model estimated for this dependent variable clearly confirms the impression produced by the graphs. "Probability of breaking a relationship" has the strongest significant effect on the dependent variable, while "weight" has no significant effect at all and "tolerance" has a very weak (although significant) effect as shown in Table 5.1, which displays the results of the model. The value of R squared for the model is 0.547[2].

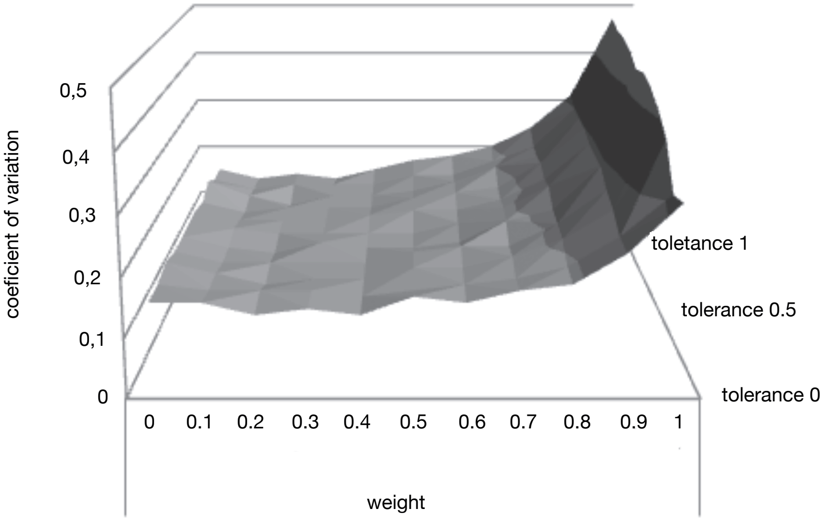
5.2. Coefficient of variation

The results concerning the coefficient of variation of the trait are even more counterintuitive. The trait variation among agents increases as β decreases, but only with high values of both ω and τ. If the weight of the partner's trait is high but tolerance to partner's difference is low or the other way round (i.e., tolerance is high but weight is low), the probability of breaking a relationship does not seem to have an effect on the coefficient of variation.

Tolerance to partner's trait and weight of partner's influence have very different effects. On the one hand, the coefficient of variation *does not seem to be very sensitive to the values of parameter* τ. On the other hand, *parameter ω seems to have a strong influence*, as the coefficient of variation of the trait clearly increases the higher the values of ω.

The linear multivariable regression model estimated for this dependent variable shows (see Table 5.2) that "weight" has the strongest significant effect on the dependent variable, but it is positive rather than negative. The effect of "tolerance" has the expected direction, although it is rather small. Contrary

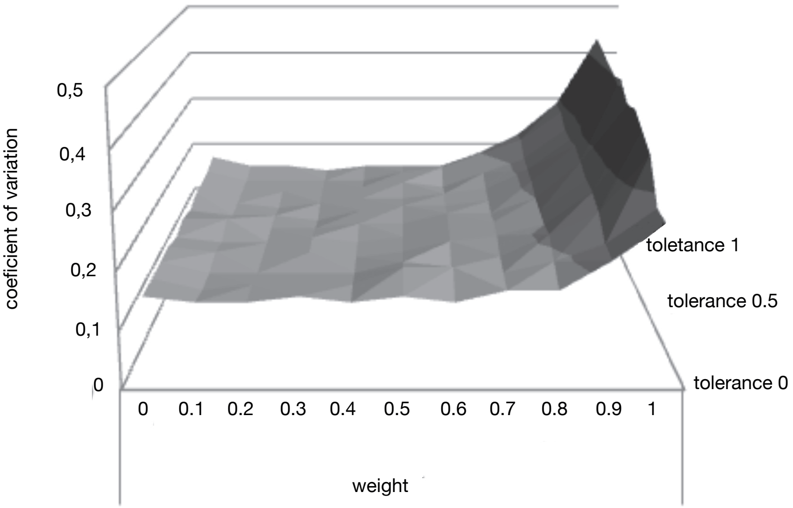
Graph 5.4. (β = 0.25)



---

[2] It cannot be taken for granted that the distribution of residuals is fairly homoscedastic in this statistical model or in the models presented below.

Graph 5.5. (β = 0.50)



Graph 5.6. (β = 0.75)

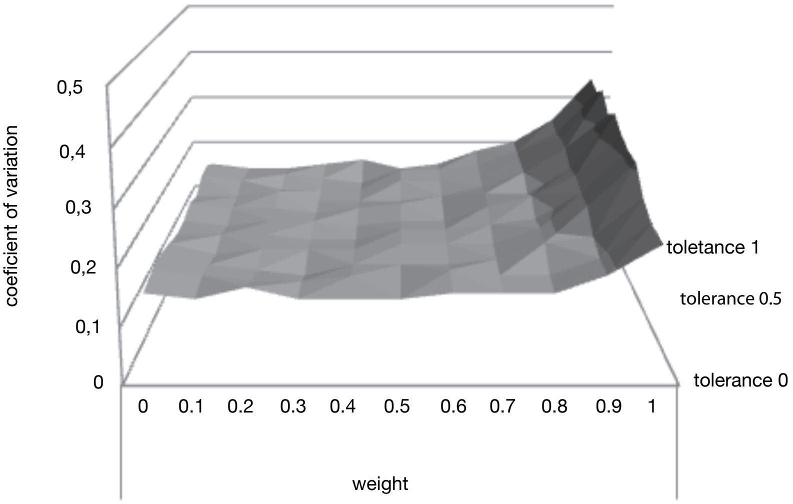


Table 5.2. Dependent variable: Coefficient of variation of trait

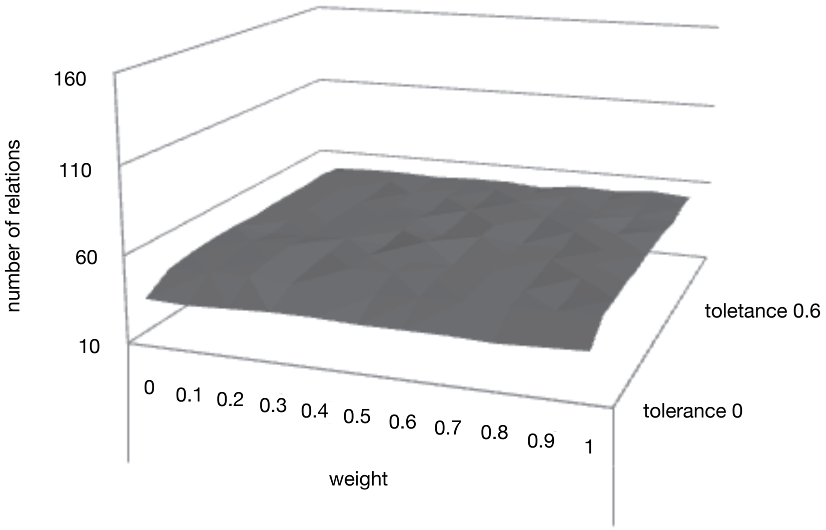
	Model	Unstandardized Coefficients		Standardized Coefficients			99% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	.181	.001		353.862	.000	.180	.183		
	probability-of-breaking-up	-.030	.001	-.090	-49.124	.000	-.032	-.028	1.000	1.000
	tolerance	.018	.001	.059	32.170	.000	.016	.019	1.000	1.000
	weight	.129	.001	.424	232.622	.000	.128	.131	1.000	1.000

to our expectations, “probability of breaking a relationship” has a significant, although small, negative effect. The value of R squared for the model is only 0.192, implying that the model poorly captures the logic behind the variation of the dependent variable.

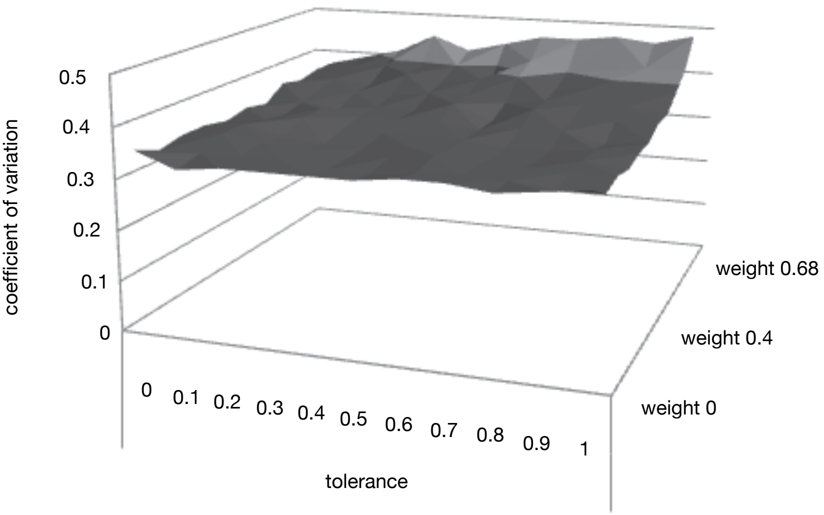
5.3. Sensitivity to average number of links

A straightforward question regarding these results is whether they are dependent (and if so, to what extent) on the topology of the network. As explained above (see the appendix for a more detailed account), the network is created by means of an algorithm which randomly assigns links to agents until the number of links per agent fits a certain average node degree parameter, which has been set to 20 throughout the whole range of simulations.

Graph 5.7.



Graph 5.8.



In order to answer this question, new simulations were conducted varying the average node degree. Graphs 5.6 and 5.7 show the number of relations and coefficient of variation when the average node degree equals 4 (i.e., agents have 4 links on average) and the probability of breaking a relationship equals 0.5 (so they can be compared to graphs 5.2 and 5.5). The influence of average node degree on the number of relations is obviously a deterministic outcome of the model: since agents choose their partners from among their linked neighbors, the lower the number of ties, the lower the number of romantic relationships. The estimated regression model (see Table 5.3) shows a strong significant effect for this variable. The R squared for this model increases to 0.743.

Table 5.3. Dependent variable: number of relations

		Unstandardized Coefficients		Standardized Coefficients			99% Confidence Interval for B		Collinearity Statistics	
	Model	B	Std. Error	Beta	t	Sig	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	-12.346	.162		-76.092	.000	-12.764	-11.928		
	probability-of-breaking-up	109.790	.153	.738	716.837	.000	109.396	110.185	1.000	1.000
	tolerance	8.773	.139	.065	62.921	.000	8.413	9.132	1.000	1.000
	weight	-.253	.140	-.002	-1.811	.070	-.613	.107	1.000	1.000
	average-node-degree	3.394	.008	.442	429.131	.000	3.374	3.415	1.000	1.000

Table 5.4. Dependent variable: coeficient of variation of trait

		Unstandardized Coefficients		Standardized Coefficients			99% Confidence Interval for B		Collinearity Statistics	
	Model	B	Std. Error	Beta	t	Sig	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	.302	.001		599.056	.000	.301	.304		
	probability-of-breaking-up	-.030	.000	-.090	-63.098	.000	-.031	-.029	1.000	1.000
	tolerance	.018	.000	.059	41.360	.000	.017	.019	1.000	1.000
	weight	.129	.000	.424	297.775	.000	.128	.131	1.000	1.000
	average-node-degree	-.010	.000	-.561	-393.762	.000	-.010	-.010	1.000	1.000

On the other hand, the influence of average node degree on the coefficient of variation is less straightforward. The linear multivariable regression model estimated for the coefficient of variation (see Table 5.4) shows that "average node degree" has the strongest significant effect, which is negative. The effects of "weight", "probability of breaking a relationship" and "tolerance" are similar to the model shown in Table 5.2 above. The value of R squared increases to 0.507 in this model.

6. Discussion

Summarizing, in SCPCM the probability of breaking a relationship has a very strong positive effect on the number of relations, and a weak (but still significant) negative effect on trait variation. Tolerance to others has a positive significant effect on both variables, but in every case this effect is rather small. The strongest effects on trait variation comes from the average node degree and from the weight of partner influence (which has no effect at all on the number of relations).

There are a number of counterintuitive results that should be stressed. One would expect higher levels of trait variation the higher the tolerance; however, *tolerance to others has no strong effect on trait variation*. In the model, the coefficient of variation invariably falls, mainly driven by the number of available ties per agent and the weight of partners' influence. These effects are also counterintuitive insofar as one would expect the influence of these variables to operate in the opposite direction than they actually do. The trait variation should be expected to increase as the number of different neighbors also increases, and to decrease as the weight of partner's influence increases. But the statistical tests show that it actually happens the other way round. Why is this so?

The dynamic of the simulation model allows us to understand this puzzle. Because the process of contagion is necessarily stronger the denser the network of ties, the diversity among agents is reduced (and, at a network level, the emergence of a number of trait-clusters is fostered); a result which has already been observed in previous models (Schelling, 1978; Axelrod, 1997). Furthermore, when the influence of the (similar) partner has a higher weight than the influence of other linked neighbors, homophilic choices of partners seem to reinforce the homogenization effect of social contagion. To the best of my knowledge there is no previous research accounting for this effect nor a clear explanation of why tolerance does not have a strong, statistically significant influence on trait variation.

Finally, an important result of the analysis is that since the effect of "weight" on the trait variation necessarily depends on the number of romantic relationships created through the simulations, and because this number is only a small proportion of the total amount of relations (unlikely to be higher than 5%), it follows that the behavior of a small number of agents has a strong impact on the evolution of the whole system, which is an usual feature of complex adaptive systems.

7. Conclusion

One of the most interesting features of Raymond Boudon's sociological theory is the use of simple "thought experiments" in order to ascertain the basic logic underlying the emergence of "social facts". This methodology can be widely extended nowadays by means of agent-based models, a computational tool for analyzing complex adaptive systems, such as those which are usually the object of sociological analysis. Following Boudon's approach, in this paper we have built a "thought experiment" whose goal is to ascertain the joint consequences of a contagious process and a homophilic partner choice in a population of artificial agents.

Agent-based models clearly reproduce the basic principles of Boudon's sociological analysis. In our model, agents are programmed to pursue certain goals (i.e., finding a partner) while being sensitive to other agents' characteristics. Due to the effect of agents' interactions in a given social environment (a social network), some properties of the system (i.e., total number of romantic relations and the variation of a trait among agents) change. It is worth noting that these system-level changes are driven by very simple mechanisms which operate at the level of agent's behavior (i.e., peer influence and homophilic choice).

Following this generative approach, system properties are shown to emerge from a cumulative process: when an agent finds a partner, it has a consequence for the agent and for the agent's partner; but it also has consequences for agents to which they are both linked. These consequences may drive other agents to change their behavior, and so forth. Of course these long chains of reactions cannot be disentangled in the absence of simulation experiments. Since the value of partner trait has a special weight, the influence of an engaged agent on others will not be the same as that of a "single" agent. Even if engaged agents are a minority in every simulation run, their presence has an effect which spreads over the whole network. In the end, it is possible to observe the emergence of patterns which were not "intended" by any of agents since they are the product of local responses that have global consequences.

Acknowledgments

I am grateful to Beatriz Abdul-Jalbar and to two anonymous reviewers from *Papers*, whose comments were useful to improve a previous version of the paper. I am also grateful to my excellent students of social theory at Universidad de La Laguna (year 2013-2014) whose faithful attendance to my lectures encouraged this work. Financial support is acknowledged from research projects CSO20012-31401 (MINECO) and CSD2010-0034 CONSOLIDER-INGENIO (MICINN).

Bibliographic references

Axelrod, Robert (1997). "The dissemination of culture: a model with local convergence and global polarization". *Journal of Conflict Resolution*, 41 (2), 203-226.

Bearman, Peter S.; Moody, James and Stovel, Katherine (2004). "Chains of Affection. The Structure of Adolescent Romantic and Sexual Networks". *American Journal of Sociology*, 110 (1), 44-91.

Blau, Peter M. (1977). "A Macrosociological Theory of Social Structure". *American Journal of Sociology*, 83 (1), 26-54.

Boudon, Raymond (1979). *La logique du social. Introduction à l'analyse sociologique*. Paris: Hachete.

— (1984). *Le place du désordre*. Paris: Presses Universitaires de France.

Coleman, James S. (1986). "Social Theory, Social Research and a Theory of Action." *American Journal of Sociology*, 91 (6), 1309-35.

Coleman, James S., Katz, Elihu and Menzel, Herbert (1957). "The Diffusion of an Innovation among Physicians". *Sociometry*, 20: 253-270.

Cruz Gómez, Irene (2013). "La homofilia como forma de estructuración de las relaciones sociales en Cataluña". Universitat Autònoma de Barcelona. PhD. Diss.

Elster, Jon (1989). *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

Epstein, Joshua M. and Axtell, Robert L. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. Washington (DC): The Brookings Institution.

Feld, Scott L. (1981). "The Focused Organization of Social Ties". *American Journal of Sociology*, 86 (5), 1015-1035.

Gilbert, Nigel (2008). *Agent-based Models*. London: Sage Publications.

González-Bailón, Sandra and Murphy, Tommy E. (2013). "The effects of social interactions on fertility decline in nineteenth-century France: An agent-based simulation experiment". *Population Studies: A Journal of Demography*. http://dx.doi.org/10.1080/00324728.2013.774435.

Hedström, Peter (2005). *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press.

Hedström, Peter and Bearman, Peter S. (2009). "What is analytical sociology all about? An introductory essay". In: Hedström, Peter and Bearman, Peter (eds.). *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford University Press.

Hedström, Peter and Swedberg, Richard (1998). *Social Mechanisms. An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.
Kossinets, Gueorgi and Watts, Duncan J. (2009). "Origins of Homophily in an Evolving Social Network". *American Journal of Sociology*, 115 (2), 405-450.
Manzo, Gianluca (2013). "Educational Choices and Social Interactions: A Formal Model and a Computational Test". *Comparative Social Research*, 30, 47-100.
McPherson, Miller; Smith-Lovin, Lynn and Cook, James M. (2001). "Birds of a Feather: Homophily in Social Networks". *Annual Review of Sociology*, 27, 415-444.
Miller, John H. and Page, Scott E. (2007). *Complex Adaptive Systems*. Princeton, NY: Princeton University Press.
Railsback, Steven F. and Grimm, Volker (2012). *Agent-based and Individual-based Modeling*. Princeton, NY: Princeton University Press.
Rolfe, Meredith (2009). "Conditional Choice" In: Hedström, Peter and Bearman, Peter (eds.). *The Oxford Handbook of Analytical Sociology.* Oxford: Oxford University Press.
Schelling, Thomas (1978). *Micromotives and macrobehavior*. New York: Norton.
Stonedhal, Forrest and Wilensky, Uri (2008). Netlogo Virus on a Network model. http://ccl.northwestern.edu/netlogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.
Wilenski, Uri (1998). NetLogo. http://ccl.northwestern.edu/netlogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Appendix: Odd protocol

##Overview

1) *Purpose*: The model aims to explore how two different social dynamics, diffusion of a cultural trait and romantic matching, influence each other. The specific problem the model addresses is how both of these processes are dependent on interaction based on three individual characteristics: sensitivity to others' similarity, influence of partner on one's own decision, and likelihood of breaking a romantic relationship. The model explores these dynamics in a fixed network of 200 agents which are intended to represent teenagers who have friendship relationships which may evolve, if the right partner is found, to romantic relationships.
2) *Entities, state variables, and scales*: The model has three kinds of entities: boys, girls and links. The environment consists of a torus of 81x81 patches which have no state variable. All agents, whether boys or girls, have the following state variables: sex (boolean), age (numerical), engaged? (boolean), partners-memory (list), trait (numerical), and influence-threshold (numerical). Links represent the type of relationship between two agents by means of a color code (see below).

Global variables are: number-of-romantic-relationships (numerical) and trait-variability (numerical), which are the main outputs of the model. Other global variables are set as parameters: likelihood-of-breaking-a-relationship (numerical) tolerance-to-cultural-difference (numerical) and weight-of-partner-influence (numerical). All three of these variables are key parameters to explore in the model. Besides these variables, average-node-degree (numerical) and mean-influence-threshold (numerical) are parameters that control the average number of ties per agent, and the average sensitivity of agents to social influence.

There are no temporal or spatial scales, since real time and/or real environment are not simulated.

3) *Process overview and scheduling*: The model includes the following actions executed every time-step in the same order:
   1. One agent is randomly chosen.
   2. If the agent is not engaged, the agent is asked to look for a partner.
   3. If a partner is found, the agent is asked to engage.
   4. Whether engaged or not, the agent is always asked to be culturally influenced (i.e., change the value of its cultural trait).
   5. Variables are updated.
   6. The simulation stops after 1200 time steps, which is enough for the model to reach an equilibrium point.

##Design concepts

4) *Design concepts*

Basic Principles: The model attempts to capture the interaction of two different mechanisms: homophily and contagion. Homophily is the principle by which people tend to engage in relations with other people similar to them in certain traits. For the sake of simplification, only one trait is represented. Contagion is a process which produces the spread of a certain trait among a population by means of social influence. In the model, agents look for a romantic partner similar to them in a certain cultural trait, which is measured on a quantitative scale. At the same time, agents are also influenced by their relationships, whether romantic or friendship, although these two different sources of influence do not have the same weight.

Emergence: The model shows how the dynamics of romantic-matching and social influence are interdependent so the rate of variation of the cultural trait among the population and the number of romantic relationships both differ from the scenario where these two processes are independent.

Adaptation: Agents perform two kinds of adaptive behavior. They become engaged if there is an agent in their local environments who meets the conditions to be chosen as a partner (details below). Second, agents change the value of their trait by means of a social influence process (details below).

Objectives: There is not a fitness or utility measure in the model to be optimized. However, agents behave as if they had the goal of finding a romantic partner.

Learning: Agents do not learn from past experience.

Prediction: Agents do not predict future conditions.

Sensing: All agents occupy a position in a network, which is assumed to not evolve as time progresses. The network represents the web of friendship relationships among teenagers. When searching for a partner and when updating the value of its cultural trait, every agent has access to state variables of its local environment (i.e., other agents it has a direct tie with).

Interaction: Boys and girls in the same local environment interact by making (and breaking) romantic relationships (see details below). All agents in the same local environment interact by influencing one another on the value of their cultural traits (see details below).

Stochasticity: Stochastic processes are used in the initialization in different ways. The social network is seeded with random number 1111 in order not to confound the effect of variation in the network topology with the effect of agents' behavior. State variables of agents are randomly initialized in every simulation run. The agent behaving in every simulation is also randomly chosen. Since there are 200 agents and the simulation lasts for 1200 ticks, every agent has on average 6 chances of engaging in a relationship and being influenced. Random numbers are also used in some sub-models (see details).

Collectives: There are two agent sets: boys, who may match with girls younger than them; and girls who may match with boys older than them. Both boys and girls are the subject of social influence in the same way.

Observation: At the end of every simulation run, the required outputs are: a) number of social relationships engaged in through the simulation; b) actual coefficient of variation of the cultural trait. Plots show the evolution of these indicators through time steps. In addition, it is also shown in the interface whether a certain link represents friendship (black links), a current romantic relationship (green links) or a past relationship (grey links). Agents are represented by means of circles whose color shows the value of the cultural trait (from light gray for low values to dark gray for high values).

##Details

5) *Initialization*: The simulation is initialized with 200 agents, whose state variables are randomly assigned. Sex is assigned with a 50% chance. The age of agents is picked from a uniform distribution within the range 14 to 17. The trait of agents is picked from a uniform distribution within the range 0 to 9. Influence-threshold is set by a parameter between 0 and 1 (currently set to 1, i.e., maximum sensitivity to influence). The variable engaged? is set to false for all agents. Memory of past partners is initially empty.

Links are then created with a random seed. The random assignment of links to agents ends when the condition of 20 links per agent on average is met. This produces a small-world type of network. The procedure is copied from Stonedhal and Wilensky (2008).

6) *Input data*: No input data are required.
7) *Submodels*:

Look for a partner:

If an agent is selected to look for a partner it will randomly pick, if any, one of its linked neighbors which meets the following three conditions:

a) opposite sex
b) if the agent is male, the partner must be younger. If the agent is female, the partner must be older.
c) the absolute difference between the two trait values divided by ten must be less than the value set by the parameter tolerance. This grants that agents will engage with agents with a very similar trait value when tolerance is low, but the pool of possible partners will be larger when tolerance is high.

Get engaged:

If a partner has been selected, the agent checks that the partner is not a member on the list of previous partners. Then it includes the partner in this list, changes the state of engaged? to true, and asks the partner to do both actions. However, if a random number extracted from a uniform distribution between 0 and 1 is below the value set for the parameter probability of breaking the relationship, the variable engaged? is again set to false for both agents.

Get influenced:

Regardless of whether the agent is engaged or not, it will be the subject of social influence. If the agent is engaged, the agent's trait will become equal to the value of the trait of its partner, weighted by the value of parameter weight, plus the median of the values of its local neighbors, weighted by one minus weight. When the agent is not engaged, the value of the agent's trait becomes the median of the value of the agent's local relationships.