GETTING CLOSER OR DRIFTING APART?*

Tanya S. Rosenblat and Markus M. Mobius

October 30, 2003

Abstract

Advances in communication and transportation technologies have the potential to bring people closer together and create a 'global village'. However, they also allow heterogenous agents to segregate along special interests which gives rise to communities fragmented by type rather than geography. We show that lower communication costs should always decrease separation between *individual* agents even as *group-based* separation increases. Each measure of separation is pertinent for distinct types of social interaction: a group-based measure captures the diversity of group preferences which can affect the provision of public goods while an individual measure correlates with the speed of information transmission through the social network which affects, for example, learning about job opportunities and new technologies. We test the model by looking at coauthoring between academic economists before and during the rise of the Internet in the 1990s.

JEL Classification: D43, D46, K23, L12, L96
Keywords: social networks, global village, decreasing communication costs

---

*We would like to thank John Bonin, Yann Bramoulle, Ed Glaeser, Joyce Jacobsen, Larry Katz, Michael Schwarz, two anonymous referees, one editor and seminar participants at Wesleyan and Harvard Universities and the 2001 North American Summer Meetings of the Econometric Society for helpful comments. We are particularly grateful to Glenn Ellison for graciously sharing his data on the academic publishing process and to Ariel Pakes for detailed comments on our empirical strategy. Tony Saudek was an excellent research assistant.

1 Introduction

Do new communication technologies on balance bring us closer together, or do they push us apart? Observers greeted the introduction of new transportation technologies such as the railroad and the automobile on the one hand, and the spread of electronic communication such as the telephone and electronic mail on the other hand with the expectation that they would help overcome geographic boundaries and therefore draw communities closer together.[1] However, advanced communication technologies can create new divisions by making heterogenous agents more selective. If agents prefer to communicate with agents of their own type, communities will fragment along types rather than geographic location. The automobile and the telephone strengthened social interactions based on common interests and along generational lines.[2] Furthermore, the very development of the technological underpinnings of the Internet, emailing and the World Wide Web, were driven by the desire to facilitate cooperation between scattered groups of specialized researchers across the globe.[3]

We construct a simple theoretical model to explain how decreasing communication costs can simultaneously decrease separation between agents in one dimension and increase separation in another dimension. Each type of agents in our model belongs to some group such as a political party, an ethnic community or an aca-

[1]The telephone census of 1902 discusses the importance of both the telephone and the automobile in overcoming the isolation of rural life (Bureau of the Census 1902). Amongst the futurists who believe that advances in telecommunication will eventually make space obsolete are Toffler (1980), Negroponte (1995) andMcLuhan (1994) who coined the term *global village*.

[2]Sproull and Kiesler (1991) describe how the spread of the telephone strengthened affiliation among teenage peer groups. Lynd and Lynd (1929, chapter XIX, footnote 8) report in their *Middletown* study the tendency of young people to mingle with peers in neighboring cities: "The young people go miles away, but fail to get well acquainted with those near by.". Social life in the town became increasingly fragmented and centered around shared interests: club groups became prominent and an increasing number of friends were recruited in these organized environments. Lynd and Lynd (1929) interviewed a group of working class and business class wives. In the first group, ten out of 173 friends were recruited in clubs, compared to two out of 116 friends of their mothers. In the business class group 26 out of 75 friends were first met in clubs, compared to 6 out of 71 friends of their mothers. A similar trend holds for the husbands.

[3]Emailing was a by-product of the ARPANET program which was funded by the US defence department. The HTML markup language of the World Wide Web was invented by Tim Berners-Lee at CERN in an effort to make the information sharing between particle physicists easier (Hafner and Lyon 1996).

demic specialization or sub-fields. Agents prefer to collaborate with own-type neighbors but they face a tradeoff between conducting costly collaboration with distant own-type neighbors or starting less profitable projects with close distinct-type neighbors. We then study two measures of separation between agents. *Group separation* captures the separation between types of agents by looking at the share of messages which are exchanged within groups rather than between groups. In our model group separation always increases as communication becomes less expensive because agents become selective: the desire to segregate into type-based groups is the very reason agents communicate more with distant agents as costs fall. In contrast, *individual* separation describes the distance between two randomly chosen individual agents measured by the time it takes for news to travel between these agents. A priori, lower communication costs have an ambiguous effect on individual separation. While the distance between agents of the same type is always reduced, the increase in group separation makes it more difficult for news to reach members of a different group. However, we show that for sufficiently large social networks this latter effect is small and individual separation *always* decreases. This result holds both for lattice social networks such as the circle and for *small-world networks* that were popularized by Watts and Strogatz (1998). Small-world networks resemble real-world social networks much better than lattice graphs and can be easily constructed from lattice graphs by adding a small number of 'short-cuts'. We prove that if individual separation in the underlying lattice graph decreases by a factor a then it decreases in the corresponding small world graph by a factor $\sqrt{a}$.

In our model an agent's choice of communication partners affects the flow of information through the social network. Due to this externality, we should not expect the degree of separation to be socially optimal. In fact, our concepts of group and individual separation turn out to be useful measures for decomposing the welfare effects of changes in communication costs. Increased group separation gives rise to divergent group preferences because agents spend more time talking to like-minded neighbors. Such differences tend to be further amplified through *group polarization*. According to this robust finding in social psychology the beliefs and preferences of moderately opinionated agents tend to shift towards those of the most opinionated group members(Brown 1986).[4] Increased preference heterogeneity reduces mutual understanding between groups and makes coordination across groups more difficult because of divergent social norms. Taste heterogeneity has been associated with a decrease in public goods provision and increased conflict between groups (Alesina, Baqir, and Easterley 1999). Interventions which reduce group separation have been shown to align agents' preferences and promote empathy and cooperation: for example, Duncan, Boisjoly, Levy, Kremer, and Eccles (2003) show that white students with randomly assigned African-American roommates are more likely to support redistribution to the poor and affirmative action.[5] Decreased individual separation also tends to improve welfare because it increases the speed and efficiency of information transmission between agents. Finding out about better job opportunities and the spread of technological innovations will depend on how quickly news can travel from the sender to other agents (Granovetter 1973, Udry and Conley 2002).[6]

We demonstrate the differential impact of decreasing communication costs on group and individual separation empirically by looking at changes in patterns of coauthoring between academic economists. We believe that the results are interesting in their own right and have implications for academic knowledge production in a more connected world. Our data includes all coauthored papers in top economic journals between the years 1969 and 1999. This time period covers the rise of the Internet after the invention of the world wide web in 1991. It is a well documented fact that coauthoring, in particular coauthoring with distant collaborators, increased strongly during this time period. We find that the relative probability of

[4]Polarization appears to be particularly prevalent if communication is computer-mediated (Hightower and Sayeed 1995) as content on the internet can be easily searched for websites and newsgroups which support one's opinion (Sunstein 2001).

[5]Gurin, Peng, Lopez, and Nagda (1999) also finds a positive correlation between the degree of interaction and declining racial stereotypes. Experiments in social psychology found that cooperative activities between members of distinct groups tend to promote tolerance (et. al. 1961, Aronson 1975).

[6]Granovetter (1973) was the first to emphasize the importance of friends and relatives as sources of employment information. Montgomery (1991) reviews the case study evidence on job-finding methods used by workers which suggests that approximately 50 percent of all workers currently employed found their jobs through friends and relatives. Topa (2001) estimates a careful structural model of the interaction effects in the Chicago labor market. Udry and Conley (2002) illustrates the role of social networks in the spread of pineapple farming in Ghana.

realizing a potential project with a distant US collaborator increased by 30 percent in the 1990s compared to the 1980s. We also show that the increased attractiveness of long-distance collaborations made researchers more selective just as our model predicts: they were 20 percent less likely to realize a project with a dissimilar collaborator in the 1990s.

The remainder of the paper is organized as follows. Section 2 introduces a simple formal model. Section 3 defines our two distinct measures of separation which capture the social distance between groups of people and between individuals. Section 4 introduces our main result for lattice graphs which we extend to small-world networks in section 5. In section 6 we measure separation of researchers in academia using coauthoring and confirm the usefulness of our two measures. Section 7 concludes.

2 The Basic Model

We build a very stylized model of communication with two different types of agents and preference for communication with one's own type. There are $2n$ agents ($n > 3$) who are located along a circle (see figure 1). One half of all agents are of type A and the other half are of type B. Agents' types alternate along the circle: every type A agent has precisely two type B agents as direct neighbors and vice versa.

We will refer to agent's four neighbors who are located at most a distance two away from her as her *close neighbors* and the remaining neighbors as her *distant neighbors*. All other agents are non-neighbors.

2.1 Projects and Communication

Time is discrete and in each time period every agent can initiate *projects* with her neighbors. A project can be, for example, coauthoring a research paper, a lunch or dinner engagement, or merely a conversation. For simplicity, we assume that the benefits of a project accrue only to the initiator of a project. However, this assumption can be easily relaxed.

Each agent can start exactly four projects in each time period and can do at most one project with each of her neighbors. Collaborating on a project requires communication between both agents. We assume that the initiator has to send precisely one message to his partner. Moreover, communication with a close neighbor is costless while communicating with a distant neighbor has an (additive) cost C.

By choosing a lattice graph, we rely on the Euclidean notion of distance. Therefore, the types of communication technologies that best fit our basic model are those for which usage cost increases steeply with distance. Examples include the automobile and telephony before the dramatic decrease in long-distance rates during the second half of the 20th century.

An alternative notion of 'close' and 'distant' neighbors labels any agent who is not close to be distant. This notion of distance better fits communication technologies such as modern long-distance telephony, the World Wide Web and emailing with usage costs depending only weakly or not at all on geographical distance. We will be able to analyze both types of communication technologies together when we introduce small-world networks in section 5.

2.2 Preferences

Collaborating on a project with a distinct type neighbor gives utility $\underline{U}$ while collaborating with an own type gives utility $\tilde{U}$ which is distributed over $[\underline{U}, \infty)$ with cumulative distribution function $F\left(\tilde{U}\right)$. The utility which can be derived from each potential project is observable by agents before they initiate collaboration.

Because our model is symmetric in both types, we can restrict attention to the decision making process of a type A agent. Clearly, a type A agent will always collaborate with her two close own-type neighbors. The only tradeoff she faces is to collaborate on the remaining two projects with her two close type B neighbors at zero cost, or start a more profitable project with her two distant own-type neighbors and pay a communication cost C.

A type A agent will pay for costly communication with a distant type A neighbor if the project has sufficiently high potential:

$$\tilde{U} - C > \underline{U} \tag{1}$$

This will be the case with probability $\gamma(C) = 1 - F(\underline{U} + C)$. Note, that the probability $\gamma(C)$ is decreasing in C: new means of communication which decrease the cost C of sending messages make more projects with distant neighbors profitable.

In each time period our type A agent will make one of three decisions:

1. With probability $(1 - \gamma(C))^2$ communicating with her two distant type A neighbors is not profitable enough to justify the higher cost of communication and she will instead collaborate only with her four close neighbors. Hence our type A agent will send half her messages to own-type neighbors.

2. With probability $2\gamma(C)(1 - \gamma(C))$ exactly one of the two projects with distant type A neighbors is sufficiently promising to drop collaboration with a close type B neighbor. She will send 75 percent of her messages to own-type neighbors.

3. With probability $\gamma(C)^2$ collaboration with both distant type A neighbors is valuable enough to drop projects with both close type B neighbors. In this case type A agent will communicate only with own-type neighbors.

In case (2) there is a small indeterminacy because the type A agent can stop to collaborate with either of her two close type B neighbors in favor of the more profitable project with a distant own-type neighbor. As a tie-breaking rule we assume that every agent of type A drops the project with her left (right) type B neighbor if she wants to collaborate instead with her distant left (right) type A neighbor.

Note that in our model the total number of projects started by an agent (and hence the amount of communication she conducts in each period) is the same for all communication costs C. In a richer environment, the effects of lower communication costs on the total volume of communication is ambiguous. On one hand, agents substitute away from local projects towards less expensive long-distance projects (the substitution effect). On the other hand, the lower overall cost of communicating allows agents to start more projects (the income effect). The total amount of communication might therefore increase or decrease as a result. We choose to abstract away from these effects.

We also do not consider endogenous location choice of agents. If agents could costlessly choose their location before playing the communication game we would expect agents of similar types to move together and maximize utility through local communication alone. We acknowledge that the desire to live close to similar agents gives rise to some degree of clustering, and that complete mixing is an analytically convenient rather than a realistic assumption. However, complete segregation is unlikely because the choice of location is influenced by many factors other than the desire to be close to friends, such as career concerns, choice of school for children, or an idiosyncratic preferences for a certain location or apartment to name a few.

3 Measures of Separation

In this section we formally define group and individual separation and discuss the welfare implications of changes in each measure. Our measures of group and individual separation are closely related to the indices of 'balkanized affiliation' and 'balkanized communication' introduced by van Alstyne and Brynjolfsson (1997).

3.1 Group Separation

Assume that agent i sends and expected number x_{ij} of messages to agent $j \neq i$ in every time period. We can then define the degree of *group separation* Π between type A and type B as the share of total messages which are exchanged between agents of the same type:

$$\Pi = \frac{\sum_i \sum_{j \neq i} J\left(i,j\right) x_{ij}}{\sum_i \sum_{j \neq i} x_{ij}} \tag{2}$$

The indicator function $J\left(i,j\right)$ is defined as:

$$J\left(i,j\right) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are of the same type} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Since our model is symmetric in types and agents, this group measure collapses to the share of messages some single agent i sends to agents of different type. Larger values of Π indicate a greater degree of group separation. Society is completely segregated into non-communicating communities if $\Pi = 1$. This case is excluded as long as the cost of communication does not become zero.

Why do we care about group separation? We argue that group preferences or 'cultures' are affected by the relative degree of within and between-group interaction which is captured by our measure of group separation Π. Greater group separation gives rise to more diverse group preferences which makes collective decision-making more difficult. The public finance literature has identified several possible channels. First, group separation can affect the tastes of the median voter and, more generally, will increase the median distance from the median voter. Alesina, Baqir, and Easterley (1999) show how such an increase in the heterogeneity of preferences can reduce the provision of public goods in a community. Second, Alesina and la Ferrara (2000) build a model of group formation to explain the empirical fact that participation in social activities and hence social capital is lower in more heterogenous communities.

We present a very simple model of belief and taste formation in groups to illustrate the effects of changes in group separation. Each agent in our society has a preference η_i for the type of public good which is provided in their community. Each agent of type A at time t has taste $\eta_i^t = \alpha_A^t + \theta_i^t$ with real support and each agent of type B has taste $\eta_i^t = \alpha_B^t + \theta_i^t$. It consists of two components: a type dependent component α_A^t (α_B^t) and an idiosyncratic component θ_i^t. We assume that the idiosyncratic component is identically and independently distributed amongst agents and has mean 0. The type dependent component captures the idea that the median preferences of voters in each group differ a priori. For example, the young might prefer to spend money on bicycle lanes and playgrounds while the old prefer to spend money on public transportation and making building accessible for the disabled. Another example would be the preferences of researchers in different subfield for what type of research to fund: labor economists might prefer funding of large scale natural experiments while industrial economists prefer to collect better industry data. Similarly, astronomers would like NASA to build bigger and better space telescopes while particle physicists prefer to invest in accelerators.

Communication with neighbors affects preferences: we simply assume that the final preference of an agent is a weighted average of her own taste η_i and those of her neighbors with communication shares as the weights on her neighbors preferences. This captures the idea that an agent will be more heavily influenced by the preferences of neighbors with whom she communicates more frequently. We can calculate the final preference $\hat{\eta}_i^t$ of a type A agent as follows:

$$\hat{\eta}_i^t = \frac{(1+\Pi)\,\alpha_A^t + (1-\Pi)\,\alpha_B^t}{2} + \tilde{\theta}_i^t \tag{4}$$

The random variable $\tilde{\theta}_i$ is a weighted average of the agent's idiosyncratic component and those of her neighbors. We can calculate the 'average' or median preference for each type by summing over all individuals of the same type.[7] If society is sufficiently large, the idiosyncratic components cancels out by the law of large numbers, and we obtain the group preferences E_A^t:

$$E_A^t = \frac{(1+\Pi)\,\alpha_A^t + (1-\Pi)\,\alpha_B^t}{2} \tag{5}$$

Analogously, we obtain an expression for the group preference of type B agents after they update their initial preferences:

$$E_B^t = \frac{(1+\Pi)\,\alpha_B^t + (1-\Pi)\,\alpha_A^t}{2} \tag{6}$$

The difference in median group preferences can be calculated as:

$$\Delta E^t = E_A^t - E_B^t = \Pi\left(\alpha_A^t - \alpha_B^t\right) \tag{7}$$

This formula illustrates how group separation can preserve initial group-specific differences. The more separated agents are, the less they take the opinions of other types into account which tends to increase the 'median distance' to the median voter.

Differences in the preferences of group members can be further amplified by group polarization which is a robust finding in the experimental social psychology

[7] This would be the opinion observed by a Gallup poll over a large sample of individuals.

literature Brown (1986). Agents tend to weigh the views of more strongly opinionated peers more heavily than those of less opinionated ones when forming their own preferences.

3.2 Individual Separation

Our second measure of separation describes the degree of *individual separation* between two random agents in our society. Individual separation matters for welfare because it determines how quickly innovative technologies (or information about job opportunities) spread through the community.

Our concept is closely related to the game called*Six Degrees of Separation* that was popular on American campuses in the 1980s. The aim of the game is to find the shortest path of acquaintances that connects two randomly chosen players.[8] In the context of our model, we can provide a more realistic measure of individual separation which takes into account the strength of links along the path.

Our measure is defined with the following simple model of information diffusion in mind. Assume at time $t = 0$ agent 1 has some brilliant idea about a new technology which she starts to share with all of her four collaborators. We assume that the agent derives no utility from other agents using the technology and also cannot demand payment from any other agent for relaying the information.

Note that the share of messages sent to neighbors of her own type is exactly $\Pi(C)$, the degree of group separation. At the end of the first period, five agents will know about the news: herself, two neighbors to her right and two neighbors to her left. In the second period, each of these five agents will send two more messages to her right and left neighbors, e.g. agents who already know about the superior technology will continue to transmit to their neighbors. As long as communication with distant neighbors is at least somewhat costly (i.e. $\Pi < 1$) every agent j will hear almost surely about the new technology.

This will take a random number of time periods $\tilde{T}_j$. We then define the degree of individual separation S_j between agents 1 (the originator of the idea) and some

[8]The game was originally invented by a group of mathematicians who defined two agents to be linked if they had a co-authored paper. The aim of the game was to find the shortest path which linked the agent to the famous graph theorist and mathematician Paul Erdös.

agent $j \neq 1$ as the expected time it takes to communicate the news between those two agents:

$$S_j = \mathrm{E}\left[\tilde{T}_j\right] \tag{8}$$

The degree of separation $\hat{S}$ is defined as the average expected waiting time to reach a random agent j:[9]

$$\hat{S} = \frac{\sum_{j=2}^{n} S_j}{n-1}, \tag{9}$$

Note that our measure of individual separation only takes communication shares between agents into account. In particular, there is no scale effect: if all agents start to send double as many messages in each time period while keeping the share of messages sent to any neighbor constant, then our measure $\hat{S}$ will remain unaffected even though information should travel at double the speed. This specification is sufficient for the analysis of our model, however, because we have abstracted away from scale effects and the volume of messages sent/projects started does not depend on the cost of communication.[10]

4 The Effects of Lower Communication Costs on Group and Individual Separation

In this section we analyze how a decrease in the cost of communication affects group and individual separation respectively.

4.1 Group Separation

A decrease in communication costs will always increase group separation. Lower communication costs make agents more selective in their choice of collaborators,

[9]Due to the symmetry of our model the initial agent 1 (including her type) can be chosen randomly on the network.

[10]A more general definition of individual separation could include a scaling factor for the total amount of communication by each agent. In our definition, one time unit corresponds to the average time interval between two consecutive messages sent by the same agent. An increase of total communication from X to X' would therefore imply a down-scaling of the time unit by a factor $\frac{X}{X'} < 1$.

and allows them to collaborate on more projects with own-type agents. Formally, we can derive the degree of group separation as follows:

$$\Pi(C) = \frac{1}{2}\left[1 + 2\gamma(C) - \gamma(C)^2\right] \quad (10)$$

Note that group separation $\Pi(C)$ is decreasing in the cost of communication C, i.e. communication increasingly focuses on own-type neighbors. In particular, we have $\Pi(C) = \frac{1}{2}$ for $C = \infty$ and $\lim_{C \to 0} \Pi(C) = 1$.

4.2 Individual Separation

The effect of a decrease in the cost of communication from C_H to $C_L < C_H$ on individual separation is a priori ambiguous. Agents will send more messages to distant own-type neighbors which will tend to increase the speed of within-type diffusion. This *connectivity effect* on its own would work to decrease the degree of individual separation. However, higher within-type communication increases group separation and makes it harder for messages to travel between types. In particular, if communication costs become very small ($C_L \to 0$) and news from a type A agent will almost never reach type B agents. This would make the degree of individual separation infinite. Therefore, the net effect of lower communication costs on the degree of individual separation is ambiguous.

However, for large networks sizes n we can show that the first effect dominates the group separation effect. The intuition is that connectivity is a global property while group separation is a local one. As the size of agents' neighborhoods increases due to better communication technologies the number of time periods necessary for news to travel the distance between two randomly drawn same-type agents decreases proportionally. Nevertheless, the expected number of time periods to bridge this distance will be of order n. In contrast, the waiting time for news to travel between some agent and her close distinct-type neighbor is $\frac{1}{F(\underline{U}+C)}$:[11] news will at some point 'cross over' the type barrier as they spread through the (large) social network. Therefore, the presence of distinct types does not greatly affect

[11] The probability of sending a message to this neighbor is $F(\underline{U} + C)$.

average individual separation.

The next theorem formalizes this first result: while group separation increases as a result of lower communication costs individual separation will always decrease for sufficiently large societies. We normalize individual separation $\hat{S}$ by dividing through by n to compare separation for different communication costs.

Theorem 1 *Average individual separation $\hat{S}(C)$ is*

$$\lim_{n\to\infty}\frac{\hat{S}(C)}{n}=\frac{2+5\gamma(C)+\gamma(C)^2}{2(1+\gamma(C))}=\frac{1}{2}a(C).$$

It is increasing in C.

Proof: see appendix A

An heuristic proof of theorem 1 proceeds as follows. Two random agents live on average a distance $\frac{n}{2}$ apart (since the circle has length $2n$). The cluster of agents who heard about the news expands on both ends an expected distance of $\Delta d(C)=\frac{1}{a(C)}>2(1-\gamma(C))+4\gamma(C)$.[12] Hence, it will take on average $\frac{n}{2}/\Delta d(C)$ time periods for news to travel between two randomly selected agents.

Theorem 1 has the following immediate corollary.

Corollary 1 *The relative degree $a(C_H,C_L)$ of individual separation in a regime with a high cost C_H of communication versus a regime with a low cost $C_L<C_H$ of communication satisfies:*

$$\lim_{n\to\infty}\frac{\hat{S}(C_H)}{\hat{S}(C_L)}=\frac{a(C_H)}{a(C_L)} \tag{11}$$

Theorem 1 and corollary 1 can be extended to more general lattice graphs. In particular, if we look at circular graphs with a larger radius of interaction both results will hold but the function $a(C)$ will change. On two and more-dimensional graphs we have to additionally normalize by the degree of individual separation by n^d where d is the dimension of the lattice.

[12]The outermost agent on the right boundary of a cluster can expand the cluster by a length 2 or 4. However, if the agent to the direct left of him send a message to a distant agent the cluster expands by a length of 3.

Our results differ from van Alstyne and Brynjolfsson (1997) who focus on group separation and argue that their measures of group and individual separation co-move. However, they derive this conclusion from simulations on small networks where lowering communication costs has ambiguous effect on individual separation.

5 Small-World Networks

Recently, various researchers have observed that real-world social networks exhibit *small-world* features (see, for example, Watts and Strogatz (1998)). Small-world networks are characterized by (a) a high degree of *clustering* and (b) *small characteristic path length*. The coefficient of clustering $C(G)$ of some graph G measures the degree to which neighboring agents' individual neighborhoods overlap, and therefore captures the degree of 'cliquishness' of the network.[13] Regular lattice graphs such as our circle are the prototypes of highly clustered networks.

The characteristic path length $L(G)$ of a graph G measures the average 'length' of the shortest chain connecting two random agents. The length of a chain is defined as the time it takes to transmit a message along the chain.[14] The average path length is closely related to our measure of individual separation.[15] In partic-

[13]Formally, assume each agent has m neighbors and starts p projects in each time period (in our model we have $m = 8$ and $p = m/2 = 4$). She starts a project with some neighbor j with probability x_{ij} in each time period. Agents i and j conduct an expected number $Y_{ij} \leq \min(p - x_{ij}, \tilde{m})$ of projects with the same set of agents in each time period where $\tilde{m} \leq m - 1$ is the number of common neighbors of both agents which can be at most $m-1$. Then the coefficient of clustering $C(G)$ is defined as the weighted average of the share of messages neighboring agents send to the same destinations:

$$C(G) = \frac{1}{n(m-1)} \sum_{i,j \neq i} x_{ij} \frac{Y_{ij}}{p - x_{ij}} \tag{12}$$

The coefficient of clustering always lies between 0 and 1. It tends to zero for large random graphs and is 1 for complete graphs (where all agents communicate equally with each other).

[14]Formally, consider the chain of agents $C = (A_0, A_1, .., A_m, A_{m+1}, .., A_{\overline{m}})$ which connects agents i and j (i.e. $A_0 = i$ and $A_{\overline{m}} = j$) and denote the volume of messages which are sent between agents A_m and A_{m+1} with x_m. The length of the chain is then defined as $L_C = \sum_{m=0}^{\overline{m}-1} \frac{1}{x_m}$.

[15]However, it is not the same. When constructing our measure $\hat{S}$ we allowed agents to continue sending measure even after they have heard about news in order to model the spread of news through society. Each agent j can therefore be reached by agent i through more than one path. Hence agent j will not always find out about the news through the shortest possible path.

ular, as we increase the size n of a circle graph both individual separation and the average path length increase at rate n.

5.1 A Simple Small-World Network

We adapt a model of small-worlds developed by Watts and Strogatz (1998): the 'skeletal' network of our small world is the same circular network we introduced in section 2. However, we allow for additional shortcuts between agents on the circle. Each type A (type B) agent on the circle has a random link with *some* other own-type agent in the network.[16] For simplicity, we assume that each agent has precisely one of those 'weak' links.

We call a shortcut between two agents i and j a *weak link* because their individual neighborhoods do not overlap and they communicate rarely. On the other hand, agents have *strong links* to neighbors on the circle because they both share many common neighbors and communicate often with each other. This distinction between weak and strong links was first made by Granovetter (1973) and Granovetter (1995) in his analysis of job search.

In each period there is a small probability δ that the agent can conduct one additional project. This project can only be conducted (a) either with one of the two direct neighbors to the right or left which yields utility $\underline{U}$ and involves costless communication; or (b) with the weak-link neighbor which yields random utility $\hat{U}$ distributed according to the distribution function F over $[\underline{U}, \infty)$.[17] In the latter case the cost of communication is assumed to be $\hat{C}$. Therefore, the probability of conducting this additional project with the weak-link neighbor is

$$1 - F\left(\underline{U} + \tilde{C}\right) = \frac{1}{A\left(\tilde{C}\right)} \tag{13}$$

[16]Our results do not depend on agents having links only to own-type agents. In fact, they stay unchanged for any kind of type-correlation along weak links.

[17]Collaboration with distant weak-link neighbors seems contrived. A more natural extension of the model would have agents choose collaborators amongst their three distant own-type neighbors and their two close distinct-type neighbors. However, the advantage of extending the model as suggested here is that it is easier to compare the new model with the original set-up.

Note that the function $A\left(\tilde{C}\right)$ is the inverse probability of conducting the additional project with the weak-link neighbor and is hence increasing. While the previously defined function $a\left(C\right)$ captures the rate of expansion of news along 'strong' links, the new function $A\left(\tilde{C}\right)$ describes the rate of expansion of news along 'weak' links.

An improvement in communication and transportation technology weakly decreases both the short-range communication cost C and our new long-range communication cost $\tilde{C}$. However, they can improve at different rates: early telephony and automobiles decreased short-range communication costs but had little effect on the long-range cost. Vice versa, modern advances in electronic communication decreased the long-range cost at a faster rate than the short-range cost.

Our definitions for group separation Π^S and individual separation $\hat{S}^S$ carry over to small-world networks with one caveat. Our definition of a small-world network does not define a unique network but a probability distribution over a class of small networks because the identity of the weak-link neighbor is chosen randomly. With a probability that is exponentially declining in the network size n some of these networks look just like lattice networks.[18] To focus attention on 'typical' small-world networks we calculate the expected time S^j for news to travel from agent 1 to agent j by taking the expectation over all possible small-world networks.

5.2 Group and Individual Separation in Small Worlds

Unsurprisingly, the presence of weak links does not greatly affect group separation. In fact, the degree of group separation $\Pi^S\left(C,\tilde{C}\right)$ of the small world model is simply a linear transformation of the degree of group separation $\Pi\left(C\right)$ of the original model:

$$\Pi^S\left(C,\tilde{C}\right) = \frac{\Pi\left(C\right) + \frac{1}{A\left(\tilde{C}\right)}\frac{\delta}{4}}{1+\frac{\delta}{4}} \tag{14}$$

Therefore, group separation increases as communication costs decrease just as it did in the original model.

The effects of weak links on individual separation are dramatic. Weak links provide shortcuts through which distant parts of the circle graph can get 'infected'

[18]For example, in one realization each agent has a random link to his right own-type neighbor.

by news. Therefore, the individual degree of separation no longer increases linearly in the size n of the circle but only at the rate $\ln(n)$ as the next theorem shows. It is in this sense that weak links make the world 'small'.

Theorem 2 *In the small world average individual separation satisfies*

$$\lim_{n\to\infty} \frac{\hat{S}^S(C)}{\ln(n)} = \sqrt{\frac{a(C)\,A\left(\tilde{C}\right)}{2\delta + o(\delta)}}.$$

Proof: see appendix B

In fact, when we compare the relative degree of separation in the high and low cost regime we find that our insights from the basic model continue to hold. Interestingly, improvements in short-range and long-range communication technologies affect the rate of diffusion in exactly the same way.

Corollary 2 *The relative degree of individual separation in a regime with a high cost* $\left(C_H, \tilde{C}_H\right)$ *of communication versus a regime with a low cost* $\left(C_L, \tilde{C}_L\right) \leq \left(C_H, \tilde{C}_H\right)$ *satisfies:*

$$\lim_{\delta\to 0} \lim_{n\to\infty} \frac{\hat{S}^S\left(C_H, \tilde{C}_H\right)}{\hat{S}^S\left(C_L, \tilde{C}_L\right)} = \sqrt{\frac{a(C_H)\,A\left(\tilde{C}_H\right)}{a(C_L)\,A\left(\tilde{C}_L\right)}} \tag{15}$$

Note that in small worlds a doubling of the local rate of diffusion $a(C)$ decreases the individual degree of separation only by a factor $\sqrt{2}$ instead of a factor 2 as in the basic model.

An interesting special case are *homogenous* improvements in communication technologies which increase the rate of short-range and long-range diffusion (i.e. the parameters $a(C)$ and $A(C)$ by the same factor f. In this case, the degree of individual separation will also decrease by the same factor f.

Our small-world results easily generalize to circular graphs with larger individual neighborhoods. If we consider different skeletal networks (i.e., a square lattice rather than a circle) our proofs can be adapted to derive precise results on a case by case basis. However, qualitatively, the results remain the same: lower communication costs decrease individual separation in both small worlds and on lattice graphs, but the relative decrease is less pronounced for small worlds.

6 Collaboration of Academic Economists between 1969 and 1999

We test our model by looking at the evolution of academic coauthoring between 1969 and 1999. Several new technologies decreased the cost of communication substantially starting around 1980. First, fax technology became ubiquitous in the 1980s: by 1985 already more than 100,000 machines were shipped annually and by 1990 this number had increased 20 fold (Economides and Himmelberg 1995). Second, emailing and file transfer through FTP was common by the beginning of the 1990s at US universities (Arfman and Roden 1992, Walsh 1997).[19] Third and perhaps most importantly, the rise of of the Internet in the 1990s made it dramatically easier to publish and search for working papers using the HTML markup language and browser software. Moreover, deregulation of the US airline and telephone industries in the 1980s drastically decreased the cost of travelling and making long distance telephone calls. The calling rates for state to state calls, for example, fell almost by half between 1984 and 1989 during the price wars which followed the break-up of AT&T in 1984 (FCC 1999).

The period 1980-1999 therefore provides a natural testing ground for our theory. We would also expect the effects of decreasing communication costs on group and individual separation to be particularly strong within the academic community. Research departments were early adopters of fax machines and academics were the first users of both email and the Internet because the original Arpanet was specifically designed as a research tool.

Our model predicts that decreasing communication costs should lead to more collaboration between 'similar' researchers but at the same time decrease individual

[19] In the US 24 percent of physicists and 34 percent of mathematicians had email addresses in 1991 (Walsh 1997).

separation of all researchers. Increased group separation can have undesirable welfare consequences: if subfields develop divergent methodologies such as 'natural experimentalists' versus 'empirical labor economists' it can complicate resource allocation and affect teaching of the discipline. On the other hand, we would expect lower individual separation to be unambiguously positive because it accelerates the transmission of useful and yet unpublished word-of-mouth information such as the availability of new data sources or preliminary results of other researchers.[20]

We use a dataset which contains all articles published between 1969 and 1999 in eight top economics journals.[21] We measure collaboration and communication between researchers by looking at their coauthored publications. Our dataset contains 8,838 authors of whom 6,201 authors published at least one coauthored paper.

6.1 Changes in Group Separation

In order to measure changes in group separation over time we first have to define metrics for measuring geographic distance between coauthors and for measuring the similarity of their types which can be easily mapped into our model.

We measure the type similarity of coauthors by the field overlap of their publication records prior to publication of their coauthored article. The distance between coauthors is coded in two different ways. First of all, we simply distinguish between coauthor relationships where both coauthors are affiliated with US institutions (US/US coauthors) and US/foreign collaborators. Second, we restrict attention to US/US coauthors and distinguish between coauthors who work less than 200 kilometers (125 miles) apart and those who live further apart. Although the precise cutoff distance is somewhat arbitrary we had several reasons to choose 200 kilometers. First of all, it is close to the median distance between US coauthors. Second, it implies a total commute time by car of about 4-5 hours for one coauthor to visit her collaborator. We consider this close to the maximum distance

[20] Such word-of-mouth transmission is particularly important in economics where publication of new results typically takes 2 years and more.

[21] Glenn Ellison generously shared his data with us. The data has been collected from the CD version of EconLit.

which would allow regular face-to-face contact between two collaborators without having to travel for more than one day or use an airplane.

We test our predictions on group separation by embedding our model into a simple discrete choice framework. There is a stream of potential projects y_i with characteristics (D_i, S_i, X_i) where D_i and S_i are dummy variables which are set to 1 if both coauthors are distant and similar respectively. The vector X_i captures other attributes of the potential project such as the field of study and other coauthor attributes such as their degree of specialization and the number of previously published papers. The probability that a potential project y_i will be realized is

$$Prob(y_i = 1 | D_i, S_i, X_i) = g(\alpha_D D_i + \alpha_S S_i + \beta X_i) \tag{16}$$

where g is an increasing function. With probability $1 - g(D_i, S_i, X_i)$ the project will not be realized. We estimate the empirical model separately for the periods 1980-1989 and 1990-1998 and make the following predictions:

H1: Improved means of communication decrease the cost of coauthoring with a distant author such that

$$\alpha_D^{90} > \alpha_D^{80}. \tag{17}$$

H2: The opportunity cost of coauthoring with a distinct type coauthor increases because it becomes more profitable to wait for a project with an own-type coauthor. Agents become more selective which implies that

$$\alpha_S^{90} > \alpha_S^{80}. \tag{18}$$

6.1.1 Description of the Data

We extract all coauthored papers between 1980 and 1999. To simplify our analysis we extract just the first two coauthors of each paper - more than 80 percent of all coauthored papers during this period have exactly two coauthors. We only include papers where at least one coauthor is affiliated with a US research institute and where each coauthor has at least one prior publication during the preceding 10 years.[22] The latter restriction is necessary because we use an author's publication record to determine her type and to measure the degree of type similarity between two coauthors. The resulting subsample contains 1,772 coauthored articles. Summary statistics are provided in table 1.

We measure type similarity of coauthors by the degree of overlap of their publication records. Thus we do not define an author's type directly but only relative to her coauthor: they are of more similar type if their publication records overlap to a greater degree. Formally, for each paper i and author j we construct a vector $v_{ij}(c)$ of size 17 which summarizes the share of publications in field c. We then define our basic type similarity measure AUSIMIL as follows:

$$AUSIMIL = \sum_{c=1}^{17} \min\left(v_{i1}(c), v_{i2}(c)\right) \tag{19}$$

This index also takes values between 0 and 1: larger values indicate greater similarity. A value of 1 implies that both authors allocated their research time equally across the same fields. Note that AUSIMIL always takes the value 0 if both authors work in distinct fields. To map the data more closely into our model we construct a discrete measure of type similarity. We define a dummy variable AUSIMIL50 which is 0 if the similarity between the two coauthors is below the median value of AUSIMIL50 and which is 1 otherwise.

NATDIFF is an indicator variable which is 1 if one of the coauthors is foreign. DISTANCE measures the geographic distance between academic institutions of both coauthors in kilometers for US/US coauthors. LONGDIST is an indicator variable which is 1 iff the distance between two US-based coauthors is more than 200 kilometers.

We collect information about the 10 year prior publication record of each coauthor j by counting her total number of publications AUPREVj ($j = 1, 2$) and her degree of specialization AUSPECj. We use a simple Herfindahl-type index of

[22] Econlit provides affiliation only after 1988. For 1969-1988 affiliations were manually added to the data set by searching through paper copies of the eight journals in our sample.

specialization defined as follows:

$$AUSPECj = \sum_{c=1}^{17} v_{ij}(c)^2 \tag{20}$$

This index of specialization is a real number between 0 and 1 and takes the value 1 if the author is completely specialized, i.e. all her publications are in a single field.

6.1.2 Analysis

The patterns of coauthoring with foreign authors (NATDIFF) and coauthoring with long-distance US authors (LONGDIST) are consistent with hypothesis H1. Between 1969 and 1979 and 1980 to 1989 the share of US/foreign papers was about 16 percent and increased to 19 percent thereafter. Amongst US/US coauthor relationships long-distance collaborations increased from 43 percent before 1980 to 50 percent between 1980 and 1989 and 55 percent thereafter.

When we regress both distance measures on a year trend and 16 field controls an interesting trend emerges: US/foreign coauthoring increased mainly in the 1990s while US/US long-distance coauthoring already accelerated in the 1980s. US/foreign coauthoring increased at an annualized rate of 1.4 percent in the 1990s after decreasing slightly in the 1980s (see table 2). In contrast, long-distance collaborations within the US increased at an annualized rate of 1.4 percent in the 1980s (see table 3). This is consistent with the fact that the US deregulated their airline and telecommunications markets earlier than most other countries and was also a leader in introducing electronic means of communication.

In table 2 we decompose the changes in the patterns of coauthoring between the 1980s and the 1990s. Two related trends emerge from 'eye-balling' the US data: (1) coauthoring between distinct-type and close coauthors has declined strongly; while (2) coauthoring with distant own-type collaborators has increased by roughly the same amount. These trends also show up in the US/foreign coauthoring data but less strongly so. Both phenomena are exactly consistent with the predictions of our model: agents become more selective when communication costs decrease and substitute low-value projects with close but dissimilar coauthors with high-value projects with distant but similar collaborators.

Ideally, we would like to formally test our joint hypothesis H1 and H2 by separately estimating the probability p_i that a potential project y_i with characteristics (D_i, S_i, X_i) is implemented in the 1980s and 1990s. Unfortunately, we lack the data to fully estimate such a discrete model because we only observe successful projects ($y_i = 1$). However, under some mild assumptions we can estimate the change in coefficients between the periods 1980-1989 and 1990-1998.

We choose the following functional form for estimating our discrete choice model:

$$Prob(y_i = 1|D_i, S_i, X_i) = \exp(\alpha_D D_i + \alpha_S S_i + \beta X_i) \tag{21}$$

Note that we can interpret α_D (and similarly α_S) as the relative percentage increase in probability that a potential project will be realized if both coauthors are distant (or of similar type). Using Bayes' rule we obtain:

$$p(y_i = 1|D_i, S_i, X_i) = p(D_i, S_i, X_i|y_i = 1)\frac{p(y_i = 1)}{p(D_i, S_i, X_i)} \tag{22}$$

The additional project characteristics X_i include field controls and dummy variables capturing the degree of specialization of each coauthor and the experience of each coauthor measured by the number of articles which he or she has published previously. The cutoff values for our two specialization dummies (one for each coauthor) and our two experience dummies are simply the median values of AUSPEC1, AUSPEC2, AUPREV1 and AUPREV2. The joint project characteristics (D_i, S_i, X_i) therefore divide the dataset into discrete *cells*. To simplify notation we will use the subindex i both to denote an individual observation and a cell with characteristics (D_i, S_i, X_i).

In order to ferret out the effect of distance on coauthoring we can simply compare two cells with the same characteristics except distance:

$$\frac{p(y_i = 1|D_i = 1, S_i, X_i)}{p(y_i = 1|D_i = 1, S_i, X_i)} = \underbrace{\frac{p(D_i = 1, S_i, X_i|y_i = 1)}{p(D_i = 0, S_i, X_i|y_i = 1)}}_{TermI}\underbrace{\frac{p(D_i = 0, S_i, X_i)}{p(D_i = 1, S_i, X_i)}}_{TermII} \tag{23}$$

Using our functional form assumption the left hand side of this equation simplifies to $\exp(\alpha_D)$. Term I on the right-hand side can be easily estimated from the data. Only term II presents a problem because we do not know the distribution of coauthor characteristics in the universe of *potential* (as opposed to actual) projects.

However, if we assume that this distribution did not change between the periods 1980-1989 and 1990-1998 then we can obtain a formula for the *change* in the distance coefficient α_D when estimated separately for both periods:

$$\Delta\alpha = \alpha_D^{90} - \alpha_D^{80} = \ln\left(\frac{p^{90}(D_i=1, S_i, X_i|y_i=1)}{p^{90}(D_i=0, S_i, X_i|y_i=1)}\right) - \ln\left(\frac{p^{80}(D_i=1, S_i, X_i|y_i=1)}{p^{80}(D_i=0, S_i, X_i|y_i=1)}\right) \tag{24}$$

For each pair of cells with characteristics $(D_i = 0, S_i, X_i)$ and $(D_i = 0, S_i, X_i)$ we thus get a different estimate $\widehat{\Delta\alpha}_{(S_i,X_i)}$ with precision $h_{(S_i,X_i)}$.[23] By summing over all cell pairs we can thus get an improved estimate of $\widehat{\Delta\alpha_D}$ and its standard error σ^2:

$$\begin{aligned} \widehat{\Delta\alpha_D} &= \frac{\sum_{(S_i,X_i)} h_{(S_i,X_i)} \widehat{\Delta\alpha}_{D(S_i,X_i)}}{\sum_{(S_i,X_i)} h_{(S_i,X_i)}} \\ \sigma^2 &= \frac{1}{\sum_{(S_i,X_i)} h_{(S_i,X_i)}} \end{aligned} \tag{26}$$

We derive an estimator for the change $\widehat{\Delta\alpha_S}$ in the preference for coauthoring with a similar author in an exactly analogous way.

The assumption that the distribution of coauthor characteristics for the universe of potential projects did not change between 1980-1989 and 1990-1998 is

[23] For each (S_i, X_i) the estimate $\widehat{\Delta\alpha}_{(S_i,X_i)}$ and precision $h_{(S_i,X_i)}$ are calculated as follows. From the data we can estimate for each $p^j(D_i, S_i, X_i)$ $(j = 80, 90)$ the sample mean $\hat{p}^j_{(D_i,S_i,X_i)}$ and variance $\sigma^{2,j}_{(D_i,S_i,X_i)}$. We then obtain:

$$\begin{aligned} \widehat{\Delta\alpha}_{(S_i,X_i)} &= \ln\left(\frac{\hat{p}^{90}_{(D_i=1,S_i,X_i)}}{\hat{p}^{90}_{(D_i=0,S_i,X_i)}}\right) - \ln\left(\frac{\hat{p}^{80}_{(D_i=1,S_i,X_i)}}{\hat{p}^{80}_{(D_i=0,S_i,X_i)}}\right) \\ \frac{1}{h_{(S_i,X_i)}} &= \frac{\sigma^{2,90}_{(D_i=1,S_i,X_i)}}{\left(\hat{p}^{90}_{(D_i=1,S_i,X_i)}\right)^2} + \frac{\sigma^{2,90}_{(D_i=0,S_i,X_i)}}{\left(\hat{p}^{90}_{(D_i=0,S_i,X_i)}\right)^2} + \frac{\sigma^{2,80}_{(D_i=1,S_i,X_i)}}{\left(\hat{p}^{80}_{(D_i=1,S_i,X_i)}\right)^2} + \frac{\sigma^{2,80}_{(D_i=0,S_i,X_i)}}{\left(\hat{p}^{80}_{(D_i=0,S_i,X_i)}\right)^2} \end{aligned} \tag{25}$$

important for the derivation of this estimator. It implies that the geographic distribution of economists across US universities according to fields and degree of specialization has not changed very much during the last 20 years. We do not have data to verify this assumption: but to the extent that economics departments tend to replicate themselves when replacing vacant positions with new researchers in order to preserve the balance of the various subfields within the department we believe that the assumption can be justified.

Table 4 reports our estimates for $\widehat{\Delta\alpha_D}$ and $\widehat{\Delta\alpha_S}$ using the data on US/foreign coauthoring and table 5 repeats the exercise for US/US coauthoring data. In each table we estimate four different specifications. In the first column we characterize cells only by the similarity dummy S_i and the distance dummy D_i which gives us four distinct cells. In the second column of both tables we add controls for 17 fields giving us $4 \times 17 = 68$ cells. In the third column we also control for each coauthor's degree of specialization giving us $4 \times 68 = 272$ cells and in the fourth column we control for each coauthor's experience. Increasing the number of controls any further is problematic given our sample size: while each new control dummy allows us to better control for project heterogeneity it also doubles the number of cell pairs and cuts the observations per cell on average by a half. Eventually, an increasing number of cells contain no observations.

For our US/US coauthoring data our estimates of both $\Delta\alpha_D$ and $\Delta\alpha_S$ are positive and significant which confirms hypothesis H1 and H2. The effects are quite large: in the 1990s a potential project with a distant author is 30 percent more likely to be realized than in the 1980s. The increased choice set makes researchers about 20 percent less likely to realize a project with a dissimilar author compared to the 1980s. The estimated coefficients are remarkably stable across all four specifications.

For US/foreign coauthoring we estimate a slightly bigger increased preference for coauthoring with own-type coauthors compared to the estimated coefficients for US/US data in table 4. However, we do not obtain estimated of $\Delta\alpha_D$ which are significantly different from zero. Unfortunately, the share of US/foreign coauthored papers only lies around 17 percent from 1980-1998. This makes it hard to apply our estimation technique while at the same time controlling for sources of heterogeneity such as fields and experience.

6.2 Changes in Individual Separation

To demonstrate changes in individual separation we simply calculate the average number of coauthors which separate two randomly chosen researchers i and j. We calculate this distance using 1990 and 1999 as base years and by considering all papers published during a 15 or 20 year time frame prior to those base years. We say that two researchers are linked if they have coauthored a paper during the respective 15 or 20 years time frame.

We then compare the network distance between two random researchers in the 1974-1989 network with the same measure in the 1984-1999 network. We repeat the same exercise with a comparison based on 20 year time frames (i.e. 1969-1989 compared to 1979-1999).

One problem with this simple approach is that the resulting graphs are not always connected (and our measure of individual separation is hence not well defined): some researchers never coauthor or coauthor with an exclusive clique of colleagues. Fortunately, almost all researchers belong to the same *giant connected cluster*.[24] For example, the coauthoring graph of all economists who have published at least one paper between 1969 and 1998 consists of a giant component consisting of 3,443 distinct authors while the next largest component of the graph only contains 15 authors. The same giant components exists within all the subgraphs defined for the pre-1990 and pre-1999 networks. To keeps our analysis simple we ignore the small number of author-nodes which do not belong to the giant component.

Another complication arises due to the fact that coauthoring has become increasingly common. Between 1969 and 1989 every author in the giant cluster coauthored on average $C_1 = 2.54$ papers. Between 1979 and 1999 that number increased to about $C_2 = 2.73$ papers, an increase of about 7.5 percent. We want to abstract away from this increase in the density of the network when comparing

[24]The presence of a giant connected cluster is typical for real world social networks. Watts and Strogatz (1998) analyze the network of film actors who are linked if they acted in a film together and find that about 90 percent of all actors belong to the giant connected cluster.

average individual separation. We achieve this by deleting links in the pre-1999 coauthoring network with probability $1 - \frac{C1}{C2}$. This 'weeding out' of links preserves the structure of the network in terms of the share of 'close' and 'distant' links and makes the pre-1989 networks comparable to the pre-1999 networks.

The left two columns of table 6 shows the evolution of average individual separation over time for 20 and 15 year time frames. In both cases individual separation decreases by 5 and 14 percent respectively between 1989 and 1999. These declines are amplified if we restrict attention to authors with more than 2 publications (see right two columns in table 6). This restriction excludes many 'peripheral authors' who have only weak connections to the giant cluster and tend to drive up the degree of separation. Now average individual separation decreases by 19 and 24 percent respectively.

7 Conclusion

Our model shows how advances in communication technology have the potential to at the same time bring us together and push us apart. We test this hypothesis by looking at collaborations between academic economists and find support for both lower average individual separation and greater group separation.

There are a number of possible directions for extending our empirical analysis. First of all, it would be interesting to see whether our dual observations of lower individual separation and greater group separation can be replicated for other datasets. Second, the link between different measures of separation and economic outcomes should be explored carefully. This would mean, for example, to carefully map a process of technological diffusion through different types of social networks.

A Proof of Theorem 1

We start by introducing some notation. We denote the cluster of agents who 'hear about' news sent by some agent i by time t with H_i^t. We adopt the convention $H_i^0 = \{i\}$. Since a cluster has two boundaries on a circular graph we focus on the expansion of the right boundary without loss of generality.

The leading agent $\overline{h}_i^t$ is the member of the set H_i^t who is furthest away from i on the right. We define the distance between agents i and $\overline{h}_i^t$ as d_i^t:

$$d_i^t = \left|\overline{h}_i^t - i\right| \tag{27}$$

The expansion of the right boundary of the cluster is determined by the leading agent and the agent next to him if she has heard the news already. Therefore, the right boundary can be in one of two states: in state 01 the agent next to the leading agent has not heard about the news. In state 11 the agent next to the leading agent has heard about it, too. From state 01 the process transits to state 11 with probability $1 - \gamma(C)$ (if the leading agent collaborates with his close neighbors only). From state 11 the process transits to state 01 with probability $(1 - \gamma(C))\gamma(C)$ (if the leading agent collaborates with his two own-type neighbors and the agent next to him only collaborates with his close neighbors). Since the probability flow between both states has to be the same in steady state we can deduce that the probability that the right boundary of the cluster is in state 01 converges to $\frac{\gamma}{1+\gamma}$.

In state 01 the leading agent sends a message with probability $1 - \gamma(C)$ to her two close neighbors and with probability $\gamma(C)$ to her two own-type neighbors (one close and one distant). We can therefore describe the evolution of d_i^t through the following transition matrix:

$$Prob\left(d_i^{t+1}|d_i^t;01\right) = \begin{cases} 1-\gamma(C) & \text{if } d_i^{t+1} = d_i^t + 2 \\ \gamma(C) & \text{if } d_i^{t+1} = d_i^t + 4 \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

In state 11 both the leading agent and the agent next to her can send messages with probability $1 - \gamma(C)$ to their two close neighbors and with probability $\gamma(C)$ to their two own-type neighbors. The boundary of the cluster can therefore expand by either a distance of 2, 3 or 4. The transition matrix becomes:

$$Prob\left(d_i^{t+1}|d_i^t;11\right) = \begin{cases} (1-\gamma(C))^2 & \text{if } d_i^{t+1} = d_i^t + 2 \\ \gamma(C)(1-\gamma(C)) & \text{if } d_i^{t+1} = d_i^t + 3 \\ \gamma(C) & \text{if } d_i^{t+1} = d_i^t + 4 \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

We can then calculate that the time-averaged spread of news converges in probability to

$$\Delta d = \text{plim}_{t\to\infty} \frac{d_i^{t+1} - d_i^t}{t} = 1 + \gamma(C) + \frac{1 + 3\gamma(C)}{1 + \gamma(C)} = \frac{1}{a(C)} \tag{30}$$

which is decreasing in C.

Two random agents i and j are on average a distance of $\frac{n}{2}$ apart (since the circle has length $2n$). Reaching j or a direct neighbor of j will therefore take on average $\frac{n}{2\Delta d}$ time periods. Reaching j will then take at most $\frac{1}{F(\underline{U}+C)}$ time periods. QED

B Proof of Theorem 2

The proof proceeds in two steps.

1. We show that the expected waiting time W_β for news to spread to at least a share β of the population of agents satisfies:

$$\lim_{n\to\infty} \frac{W_\beta}{\ln(n)} = \frac{1}{\sqrt{\frac{2\delta}{a(C)A(\tilde{C})}} + o(\delta) + O(\beta)} \tag{31}$$

2. We show that the average time it takes for news to spread from the share β of infected agents to the remaining $1-\beta$ non-infected agents is bounded above by a constant which is independent of n.

¿From these two steps we can deduce

$$\lim_{n\to\infty} \frac{\hat{S}^S}{\ln(n)} = \frac{1}{\sqrt{\frac{2\delta}{a(C)A(\tilde{C})}} + o(\delta) + O(\beta)} \tag{32}$$

Since we can choose β as small as we desire we immediately obtain the result stated in the theorem.

B.1 Step I

News spreads through two channels: (a) existing clusters of infected agents expand around their boundaries, and (b) new clusters form thanks to weak links. We start by analyzing a simplified stochastic process which provides an upper bound for the diffusion of news and hence a lower bound on the waiting W_β until a share β of agents have heard about the news. The simplifying assumptions are:

1. *Each* agent within the convex hull of an infected cluster can start a *new* cluster with probability $\frac{\delta}{A(\tilde{C})}$ through his weak link *in each period.*

2. Clusters evolve *without overlap.*

Both assumptions speed up diffusion. To derive the rate of diffusion of this simpler process we introduce some notation. At each point in time t the process generates a new (stochastic) set of clusters Ξ_t. The superset of all these clusters is denoted with $\Pi_t = \bigcup_{s \leq t} \Xi_t$. Each cluster $\xi \in \Xi_\tau$ is said to have a vintage τ. It grows over time at a stochastic rate and we call the size of its convex hull at time t the *span* $D(\xi, t)$. We use the convention $D(\xi, t) = 0$ if $t < \tau$. The number of clusters formed at time t is denoted with $X_t = |\Xi_t|$ and the total number of agents inside the span on all clusters which have formed up to time t is Y_t. Note, that for the coupled process we have:

$$Y_t = \sum_{\xi \in \Pi_t} D(\xi, t) \tag{33}$$

The number of infected agents Y_t increases over time because there are new infections and because existing clusters expand:

$$Y_{t+1} - Y_t = X_{t+1} + \sum_{\xi \in \Pi_t} \left[D(\xi, t+1) - D(\xi, t) \right] \tag{34}$$

We take expectations on both sides and define $y_t = E[Y_t]$, $x_t = E[X_t]$ and $z_t = E|\Pi_t|$. We also know from the proof of theorem 1:

$$\lim_{t \to \infty} \frac{E\left[\sum_{\xi \in \Pi_t} \left[D(\xi, t+1) - D(\xi, t) \right]\right]}{z_t} = 2\frac{1}{a(C)} \tag{35}$$

Therefore we can simplify equation 34 and obtain

$$y_{t+1} - y_t = x_{t+1} + \frac{2}{a(C)} u(t) z_t \tag{36}$$

where $|u(t) - 1| \leq A \exp(-\frac{\epsilon}{\delta})$ for some $A, \epsilon > 0$. We next note that

$$x_{t+1} = \frac{\delta}{A(\tilde{C})} y_t \tag{37}$$

We then get:

$$y_{t+1} - y_t = \frac{\delta}{A(\tilde{C})} y_t + \frac{2}{a(C)} u(t) z_t \tag{38}$$

Next we note that:

$$z_{t+1} - z_t = x_{t+1} = \frac{\delta}{A(\tilde{C})} y_t \tag{39}$$

We next take first differences of equation 38:

$$y_{t+2} - y_{t+1} - (y_{t+1} - y_t) = \frac{\delta}{A(\tilde{C})} y_{t+1} + \frac{2}{a(C)} u(t+1) z_{t+1} - \left[\frac{\delta}{A(\tilde{C})} y_t + \frac{2}{a(C)} u(t) z_t \right] \tag{40}$$

We assume that $u(t) = 1$ at first. We then have a simple difference equation of the following form:

$$y_{t+2} - 2y_{t+1} + y_t = \frac{\delta}{A(\tilde{C})} (y_{t+1} - y_t) + \frac{2\delta}{a(C)A(\tilde{C})} y_t \tag{41}$$

When we solve the characteristic equation we get a solution of the form:

$$y_t = B \exp\left(\sqrt{\frac{2\delta}{a(C)A(\tilde{C})}} t + o(\sqrt{\delta}) \right) \tag{42}$$

It can be shown that $u(t)$ is sufficiently close to 1 for small enough δ that it does not affect this solution to the difference equation. From this solution equation 31 follows.

Next, we have to relax the simplifying assumptions we made for the coupled process.

- *Agents can start a new cluster through their weak link only once.* Agents start new clusters at rate δ and become unavailable for starting a second cluster. However, we have just shown that the population of infected agents expands at a rate proportional to $\sqrt{\delta}$. Hence, we get again a solution as in equation 42.

- *Not every agents inside the convex hull is infected.* Whenever an agent at the boundary communicates with a distant agent a 'gap' is created which fills up with probability $F(\underline{U} + C)$ in each time period. For each τ the ratio $\frac{y_{t-\tau}}{y_t} \to 1$ as $\delta \to 0$: an arbitrarily large share of infected agents live in clusters of vintage τ for small δ. By choosing τ large enough we can ensure that the share of infected agents inside the convex hull of these clusters converges to 1. Hence we get again solution 42.

- *An agent's weak link can become infected before the agent can infect that link herself.* At this point we use the fact that we only model the evolution of the system until a share β of agents has become infected. That implies that an agent can infect another agent through her weak link at least with probability $(1-\beta)\frac{\delta}{A(\tilde{C})}$ which gives us formula 31.

- *Clusters can overlap.* To deal with this contingency we use again the fact that the share of infected agents is at most β. Assume the y_t infected would be randomly distributed along the circle - in this case the average distance between them would be at least $\frac{1}{\beta}$. But since clusters grow around the boundaries the average distance between the boundaries of the z_t clusters is also at least $\frac{1}{\beta}$. Two neighboring clusters can grow together if their boundaries are less than eight agents apart. The probability for this event is $O(\beta)$. Hence a share $O(\beta)$ of non-overlapping clusters disappear in each time period which gives us again formula 31.

B.2 Step II

For the second step we create a coupled process which governs the evolution of the system *after* a share β of agents has become infected. We bias the evolution of this process against the spread the news - therefore the process provides a lower bound on the true process. Note that after a share β of agents has become infected a share $\eta = \beta \left(1 - O\left(\delta\right)\right)$ of agents can spread news through their weak links (see step I). We call this set of agents I and the set of agents they are linked to N_I. The following rules govern the coupled process:

1. Weak links generating from agents outside of the set I cannot spread news.
2. Each agent only sends news to her *direct* neighbor *to the right* if she collaborates with her and this agent does not belong to the set N_I.

The agents in the set N_I are on average a distance $\frac{1}{\eta}$ apart and subdivide the circle into fragments on which the coupled process develops independently. The expected waiting time to cover each fragment is bounded above by $\frac{1}{\delta} + \frac{1}{\eta}\frac{1}{F(\underline{U}+C)}$. By the law of large numbers the average time W' it takes for news to reach all agents on the circle in the coupled process is the same, and hence finite. But since the coupled process systematically discriminates *against* the spreading of news the average time for news to infect all agents is also finite. QED

Authors' Affiliations

Tanya Rosenblat
Department of Economics
Wesleyan University
PAC 123, 238 Church Street
Middletown, CT 06459
Telephone: (860) 685 5351
Fax: (860) 685 2781
E-Mail: trosenblat@wesleyan.edu

Markus Möbius
National Bureau of Economic Research and
Department of Economics
Harvard University
Littauer Center
Cambridge, MA 02138
Telephone: (617) 496 3419
Fax: (617) 495 8570
E-Mail: mobius@fas.harvard.edu

References

Alesina, A., R. Baqir, and W. Easterley (1999): "Public goods and ethnic divisions," *Quarterly Journal of Economics*.

Alesina, A., and E. la Ferrara (2000): "Participation in heterogenous communities," *Quarterly Journal of Economics*, pp. 847–904.

Arfman, J. M., and P. Roden (1992): "Project Athena: Supporting distributed computing at MIT," *IBM Systems Journal*, 31(3), 550–.

Aronson, E. (1975): "The Jigsaw Route to Learning and Liking," *Psychology Today*.

Brown, R. (1986): *Social Psychology: The Second Edition*. The Free Press, New York.

Bureau of the Census (1902): *Telephones and Telegraphs*. Government Printing Office, Washington.

Duncan, G. J., J. Boisjoly, D. M. Levy, M. Kremer, and J. Eccles (2003): "Empathy or Antipathy: The Consequences of Racially and Socially Diverse Peers on Attitudes and Behaviors," mimeo, Harvard University.

Economides, N., and C. Himmelberg (1995): "Critical Mass and Network Size with Application to the US Fax Market," Discussion paper no. ec-95-11, Stern School of Business, N.Y.U.

et. al., M. S. (1961): *Intergroup Conflict and Cooperation: The Robbers' Cave Experiment*. University of Oklahoma Institute of Intergroup Relations, Norman, Okla.

FCC (1999): "Statistics to Communication Common Carriers: Historical Rate Tables," report available from http://ftp.fcc.gov/bureaus/common_carrier/reports/fcc-state_link/socc/987-all.pdf, FCC.

Granovetter, M. (1973): "The Strength of Weak Ties," *American Journal of Sociology*, 78, 1360–1380.

Granovetter, M. S. (1995): *Getting a Job: A Study of Contacts and Careers*. The University of Chicago Press (2nd. Edition), Chicago.

Gurin, P., T. Peng, G. Lopez, and B. Nagda (1999): "Context, identity and intergroup relations," in *Cultural divides: Understanding and overcoming group conflict*, ed. by D. Prentice, and D. Miller. Russell Sage Foundation, New York.

HAFNER, K., AND M. LYON (1996): *Where Wizards Stay Up Late : The Origins of the Internet*. Simon & Schuster, New York.

HIGHTOWER, R., AND L. SAYEED (1995): "The Impact of Computer-Mediated Communication Systems on Biased Group Discussion," *Computers in Human Behavior*, 11, 33–44.

LYND, R. S., AND H. M. LYND (1929): *Middletown: A Study in Contemporary American Culture*. Constable and Co LTD, London.

MCLUHAN, M. (1994): *Understanding Media: The Extensions of Man*. MIT Press (2nd. Edition), Cambridge.

MONTGOMERY, J. D. (1991): "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," *American Economic Review*, 81, 1408–1418.

NEGROPONTE, N. (1995): *Being Digital*. Vintage Books, New York.

SPROULL, L., AND S. KIESLER (1991): *Connections - New Ways of Working in the Networked Organization*. MIT Press, Cambridge.

SUNSTEIN, C. (2001): *Republic.com*. Princeton University Press, Princeton, New Jersey.

TOFFLER, A. (1980): *The Third Wave*. Bantam Books, New York.

TOPA, G. (2001): "Social Interactions, Local Spillovers and Unemployment," *American Economic Review*, 68, 261–295.

UDRY, C., AND T. CONLEY (2002): "Learning About a New Technology: Pineapple in Ghana," working paper, Yale University.

VAN ALSTYNE, M., AND E. BRYNJOLFSSON (1997): "Electronic Communities: Global Village or Cyberbalkans?," working paper, MIT Sloan School.

WALSH, J. P. (1997): "Telescience: The Effects of Computer Networks on Scientific Work," consultancy paper for oecd, University of Illinois at Chicago.

WATTS, D. J., AND S. H. STROGATZ (1998): "Collective Dynamics of 'Small-World' Networks," *Nature*, 393, 440–442.

Figure 1: Society with $n = 5$ type A and 5 type B agents.

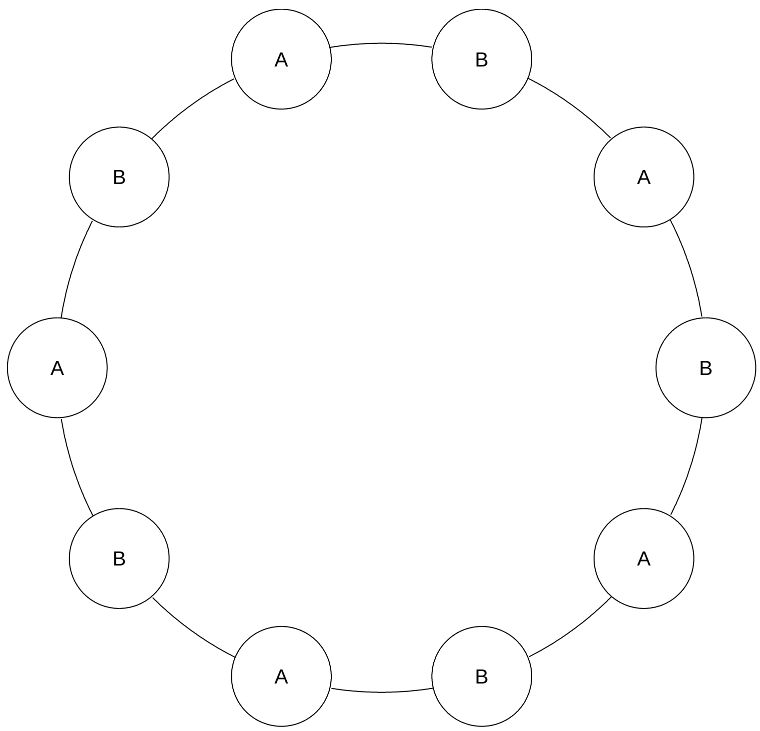

Table 1: Variable means and standard deviations for coauthored papers in eight economic journals 1980-1998

Variable	**Mean**	**Std. Dev.**	**Variable**	**Mean**	**Std. Dev.**
			MTHEORY	0.197	0.398
Project Characteristics			IO	0.145	0.352
			FINANCE	0.071	0.257
AUSIMIL	0.452	0.317	MACRO	0.188	0.391
AUSIMIL50	0.515	0.5	INTERNAT	0.058	0.234
DISTANCE	1068.261	1510.552	DEVELOP	0.012	0.111
NATDIFF	0.169	0.375	URBAN	0.009	0.095
LONGDIST	0.544	0.498	HISTORY	0.004	0.063
YEAR	1989.196	5.123	PF	0.07	0.255
			LABOR	0.116	0.321
Coauthor Characteristics			METRICS	0.055	0.228
			PRODUC	0.017	0.131
AUPREV1	6.812	7.217	ENVIRON	0.005	0.071
AUPREV2	7.218	9.502	POLITECO	0.011	0.106
AUSPEC1	0.613	0.286	LAWECON	0.008	0.092
AUSPEC2	0.617	0.287	OTHER	0.016	0.127
Fields					
EXP	0.015	0.123			

$N = 1772$

The dataset comprises all coauthored papers in eight economics journals between 1980 and 1998 with at least one US author: *Journal of Political Economy*, *American Economic Review*, *Quarterly Journal of Economics*, *Econometrica*, *Review of Economic Studies*, *Review of Economics and Statistics*, *Rand Journal of Economics*, *Brookings Papers on Economic Activity*. The variable AUSIMIL measures the similarity of coauthors based on their publication records up to 10 years prior to publication of their joint paper. AUSIMIL50 is an indicator variable and set to 1 if AUSIMIL is greater than its median value. AUSPEC indicates how specialized authors are. NATDIFF is an indicator variable which is 1 if one of the coauthors lives outside the US. DISTANCE is distance between US coauthors locations in kilometers and LONGDIST is 1 if the distance exceeds 200 kilometers (125 miles). YEAR indicates the year of publication and calendar year 1980 is set to 0. Each paper falls into one of 17 field categories, which are labor, econometrics, productivity, experimental, micro theory, industrial organization, finance, macro, international, development, history, public finance, environmental economics, political economy, law and economics, and other fields.

Table 2: Testing for trends in coauthoring between US and foreign economists by regressing NATDIFF on YEAR and field controls

Variable	(80-98)	(80-89)	(90-98)
YEAR	0.002	-0.008†	0.014*
	(0.002)	(0.005)	(0.006)
Field controls	Yes	Yes	Yes
N	1772	857	792
R^2	0.042	0.045	0.063

Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%

The dependent variable is NATDIFF; standard errors are shown in paranthesis. The field controls include experimental economics, micro theory, industrial organization, finance, macro, international, development, urban economics, history, public finance, labor economics, econometrics, productivity, environmental economics, political economy, and law and economics. The first column includes all coauthored papers published between 1980 and 1998 while the next two columns restrict attention to the 1980s (1980-1989) and 1990s (1990-1998).

Table 3: Testing for trends in coauthoring between North American economists by regressing LONGDIST on YEAR and field controls

Variable	(80-98)	(80-89)	(90-98)
YEAR	0.009**	0.013†	0.003
	(0.003)	(0.007)	(0.009)
Field controls	Yes	Yes	Yes
N	1415	697	622
R^2	0.015	0.027	0.017

Significance levels : † : 10% ∗ : 5% ∗∗ : 1%

The dependent variable is LONGDIST; standard errors are shown in paranthesis. The field controls include experimental economics, micro theory, industrial organization, finance, macro, international, development, urban economics, history, public finance, labor economics, econometrics, productivity, environmental economics, political economy, and law and economics. The first column includes all coauthored papers published between 1980 and 1998 while the next two columns restrict attention to the 1980s (1980-1989) and 1990s (1990-1998).

Figure 2: Changes in pattern of coauthoring with similar and distant authors between the periods of 1980-1989 and 1990-1998

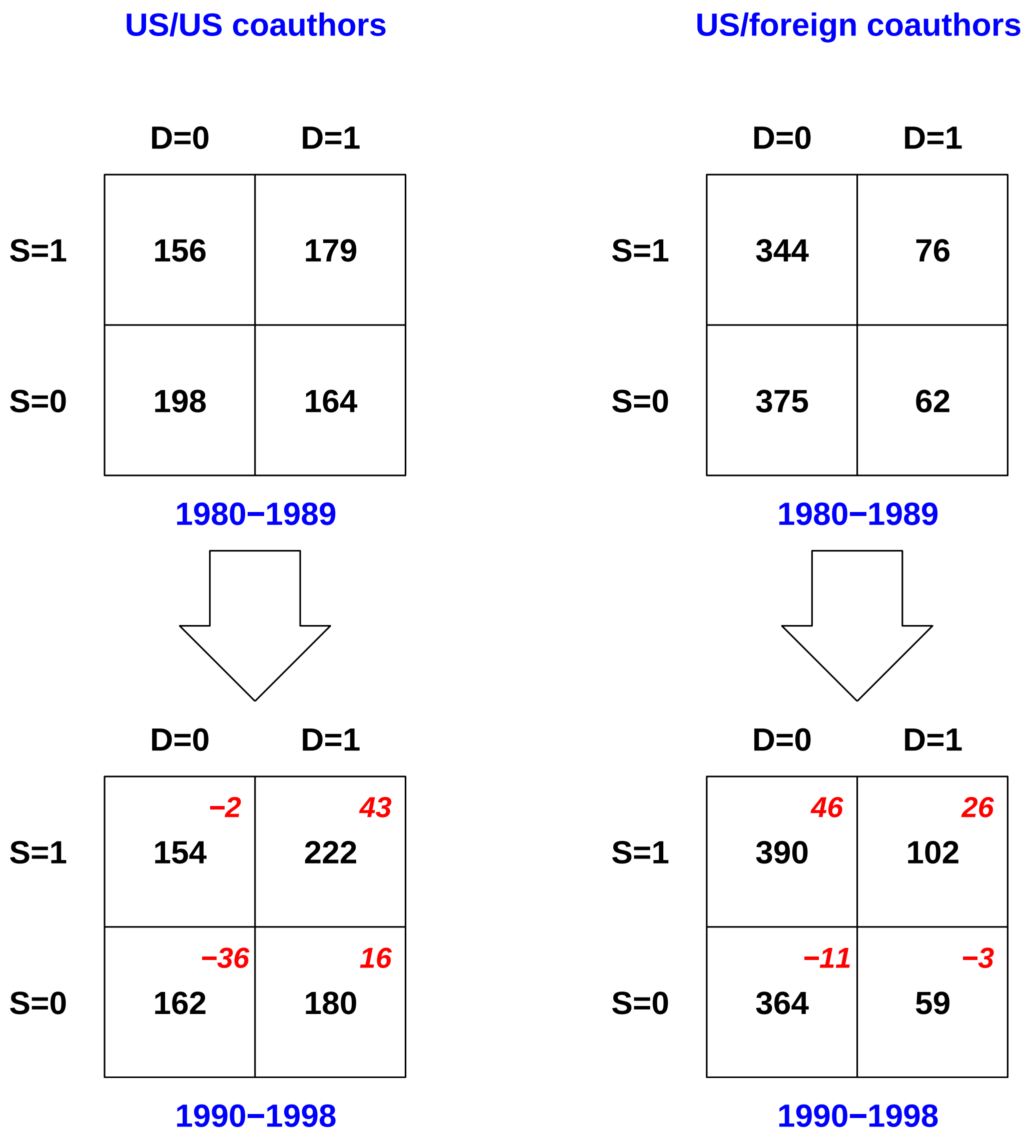

Table 4: Estimated increase in preference for coauthoring with distant/similar foreign coauthors between the periods 1980-1989 and 1990-1998

Variable	(1)	(2)	(3)	(4)
$\alpha_D^{90} - \alpha_D^{80}$	0.091	0.098	0.062	-0.029
	(0.118)	(0.137)	(0.159)	(0.151)
$\alpha_S^{90} - \alpha_S^{80}$	0.176*	0.220*	0.236*	0.237*
	(0.075)	(0.097)	(0.108)	(0.107)
Field Controls	No	Yes	Yes	Yes
AUSPEC Controls	No	No	Yes	No
AUPREV Controls	No	No	No	Yes
N=1772				

Significance levels: † : 10% ∗ : 5% ∗∗ : 1%

Standard errors are shown in paranthesis. Distance is measured by NATDIFF. The field controls create 17 cells and include experimental economics, micro theory, industrial organization, finance, macro, international, development, urban economics, history, public finance, labor economics, econometrics, productivity, environmental economics, political economy, and law and economics. The second column adds controls for coauthor specialization using the median values of AUSPEC1 and AUSPEC2 as cutoffs to distinguish between non-specialized and specialized authors. This control subdivides each field cell into four subcells. The third column adds controls for the number of previously published papers by each coauthor using the median values of AUPREV1 and AUPREV2 as cutoffs.

Table 5: Estimated increase in preference for coauthoring with distant/similar US coauthors between the periods 1980-1989 and 1990-1998

Variable	(1)	(2)	(3)	(4)
$\alpha_D^{90} - \alpha_D^{80}$	0.276^{**}	0.289^{**}	0.290^{*}	0.294^{*}
	(0.091)	(0.108)	(0.127)	(0.123)
$\alpha_S^{90} - \alpha_S^{80}$	$0.153^{\dagger}$	$0.189^{\dagger}$	$0.213^{\dagger}$	0.190
	(0.093)	(0.110)	(0.129)	(0.126)
Field Controls	No	Yes	Yes	Yes
AUSPEC Controls	No	No	Yes	No
AUPREV Controls	No	No	No	Yes
N=1415				

Significance levels: $\dagger$: 10% $*$: 5% $**$: 1%

Standard errors are shown in paranthesis. Distance is measured by LONGDIST. The field controls create 17 cells and include experimental economics, micro theory, industrial organization, finance, macro, international, development, urban economics, history, public finance, labor economics, econometrics, productivity, environmental economics, political economy, and law and economics. The third column adds controls for coauthor specialization using the median values of AUSPEC1 and AUSPEC2 as cutoffs to distinguish between non-specialized and specialized authors. This control subdivides each field cell into four subcells. The fourth column adds controls for the number of previously published papers by each coauthor using the median values of AUPREV1 and AUPREV2 as cutoffs.

Table 6: Average individual separation of coauthors in economic journals

		20 years	15 years	20 years > 2 publ.	15 years > 2 publ.
pre-1989	n	1813	1269	889	601
	L	2.54	2.52	2.58	2.54
	$\hat{S}$	**8.57**	**9.03**	**7.87**	**8.14**
pre-1999 (unadj.)	n	2629	1658	1177	738
	L	2.73	2.72	2.88	2.78
	$\hat{S}$	**8.39**	**8.07**	**7.13**	**6.89**
pre-1999 (adj.)	n	2458	1542	1055	677
	L	2.60	2.58	2.69	2.60
	$\hat{S}$	**8.18**	**7.90**	**6.63**	**6.55**

The size of the giant connected cluster is n and the average number of links of every author in this cluster is L. The average degree of separation is $\hat{S}$. Each column corresponds to a different time frame - the 20 year frame for example compares the period 1969-1989 (pre-1989) to 1979-1999 (pre-1999). The two right columns only include authors who had more than two publications during the time frame. The adjusted pre-1999 network is obtained by (a) calculating the average number of links in the pre-1989 network ($C1$) and the pre-1999 network ($C2$); and (b) deleting links in the pre-1989 network randomly with probability $1 - \frac{C1}{C2}$. For this table the values of $\frac{C1}{C2}$ were 0.930, 0.928, 0.896 and 0.915 (left to right).