Decision Support Systems 47 (2009) 169–184

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Online social networks in economics

Adalbert Mayer

Department of Economics, Texas A&M University, United States

ARTICLE INFO

Available online 3 March 2009

Keywords:
Economics
Social networks
Information transmission

ABSTRACT

This paper describes how economists study social networks. While economists borrow from other fields like sociology or computer science, their approach of modeling of social networks is distinguished by the emphasis on the role of choices under constraints. Economists investigate how socioeconomic background and economic incentives affect the structure and composition of social networks. The characteristics of social networks are important for economic outcomes like the matching of workers to jobs and educational attainment. I review the theoretical and empirical literature that investigates these relationships and discuss possible implications of new, Internet based, forms of social interactions.

1. Introduction

The economic problem states that there is scarcity and individuals face tradeoffs when making choices. The finite resources available are insufficient to satisfy all human wants. We have to decide for which purpose we want to use a scarce resource, such as oil, potatoes, or time. Choices made by individuals determine the allocation of these scarce resources. Therefore, economists are interested in the processes that lead to these choices. Social interactions and social networks are an important determinant of these choices.

New communication technologies are changing social contacts and lead to new forms of interaction like online communities. In this paper, I try to give an overview over how economists study social interactions and especially social networks. I illustrate effects on choices – and hence the allocation of scarce resources – and I analyze the implications of the increasing importance of online interactions. I discuss theoretical contributions that help to understand how social networks are formed. Moreover, I present empirical evidence on the characteristics of social networks and their effects on outcomes.

The interest of economists in social networks is motivated by the realization that social networks are an important channel for information transmission. Social networks have been studied in the sociology literature for a long time [25]. Economists are becoming more aware of this literature and are also increasingly utilizing insights on social interactions and social networks obtained by psychologists. Moreover, the use of formal, mathematical tools favored by economist leads to an overlap with other fields – such as computer science, mathematics, or physics [54]. Despite this overlap with other sciences there are distinct differences between the methodologies in these fields and economics. The approach of economists is distinguished by two main features. First, economists are ultimately interested in how social interactions relate to choices that affect the allocation of scarce resources. Therefore, I emphasize results of empirical research that tries to quantify these effects. Second, economics is characterized by a specific way to conceptualize how decisions are made [47]. Economists employ this approach to model the decisions that lead to the formation of social networks. Decisions are made by agents. Agents can be individuals, as well as firms or governments. Agents make their decisions based on preferences. This is usually formalized by stating that agents maximize a utility function – a function that ranks the desirability of outcomes associated with the different choices. Due to constraints only a certain set of choices is feasible. Agents maximize their utility subject to these constraints. For example a budget constraint states that a consumer can only purchase certain combinations of goods, while she cannot afford others.

The choices of agents can be interdependent because the decisions of one agent can affect preferences (e.g. through jealousy) and constraints (e.g. by using up resources or by providing information about the availability of a previously unknown option) of other agents. Economists use formal models of behavior to characterize likely results of the interaction between individuals and their choices. They employ the concept of equilibria. In an equilibrium the choices made by all agents involved are mutually consistent. Economists also ask whether the allocation of scarce resources that results from these choices is efficient or if a different allocation would be preferable according to some criterion.

This approach of studying choices of agents can be used to understand why individuals interact and form social connections. Therefore, it can be used to explain how social networks are formed and how they would change if some of the conditions that lead to their formation change. One such change is the rapid development of new communication technology. The introductions of the telephone and more recently the Internet have changed the environment in which interactions are determined. The cost of communicating over long distance has dropped dramatically; hence the constraints of

E-mail address: amayer@econmail.tamu.edu.

0167-9236/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.dss.2009.02.009

individuals have changed. New patterns of social interactions have arisen and we are only beginning to understand the implications for a number of outcomes.

In recent years the analysis of social networks has emerged as a distinct subfield within economics. However, networks play a role in a number of different areas of economics. Economists in the area of international trade investigate the effects of trade networks [see 58,24, 64,42]. Some industries, such as railroads or the telecommunication industry are based on networks. Economists in the field of industrial organization investigate the effects of these networks on market structure, pricing or investment decisions. Some of the insights and terminology from these other subfields within economics can be also used to study social networks. For example, network industries exhibit positive consumption or production externalities, as seen below externalities arise in social networks as well.

This paper is organized as follows. In Section 2, I describe the tools that economists use to study social networks. I introduce the formal, and mathematical concepts that economists employ to describe networks. Then, I survey the empirical literature that uses these tools to describe common properties of (online) social networks. Finally, I summarize the empirical literature that suggests that similar characteristics and physical proximity are important determinants of networks. In Section 3, I present two approaches of modeling the formation of social networks. One approach, related to the mathematics and computer science literature, is more mechanical and reveals algorithms that can explain how some of the observed features of social networks arise. As mentioned above, the approach of economists is different. The formation of networks is modeled as the result of choices by utility maximizing agents. In Section 4, I discuss the effects of social networks on outcomes that are of interest to economist. For example, many workers rely on social networks when searching for jobs. Hence, social networks can affect success of individuals in the labor market, as well as aggregate outcomes like the unemployment rate. Other outcomes affected by social networks are student performance or purchasing decisions. Finally, in Section 5, I explore the consequences of the recent changes in communication technology. Online social networks differ in some important dimensions from traditional "real world" networks. The cost of communication is very low, especially since physical distance is not important. I discuss implications for the structure and composition of social networks and examine possible effects on outcomes.

2. Describing social networks

In this section, I explain how economists describe social networks and present some of the empirical regularities that characterize social networks. In Section 2.1, I present the formal tools – based on graph theory – used to analyze social networks. In Section 2.2, I proceed to summarize the results of the empirical literature. In Section 2.2.1, I use the tools introduced in Section 2.1 to describe characteristics of the structure of social networks and online social networks. A common feature of social networks is homophily (the love of the same) or the tendency of individuals to associate with individuals with whom they share common characteristics. This tendency is especially important when considering the implications of social networks for economic outcomes (see Section 4). I discuss empirical evidence on the importance of homophile in Section 2.2.2.

2.1. Language and notation

To study social networks it is necessary to describe them. Economics is a quantitative science; therefore, it relies heavily on mathematics and tries to define quantifiable network properties. Economists borrow some of their modeling approaches from the vast literatures in sociology, mathematics, and computer science [25,54]). My discussion here is based on the more detailed and rigorous treatment of the topic by Jackson [37], I follow his notational conventions wherever possible.

A graph describes pair-wise relationships between the elements of a set. The elements of the set considered here are agents who can be potentially part of a social network. For example, these could be all students on a university campus or all individuals with Internet access. Let n be the number of agents who might be connected – in the terminology of network analysis a network with n nodes. The nodes of a network are connected by links, sometimes also called edges or vertices. A graph is characterized by a list of nodes and a list of all the connections between these nodes. The list of all the connections between nodes is also referred to as an adjacency list. Another way to represent the connections of a network is an adjacency matrix. Let G be an $(n \times n)$ matrix. If $g_{if}>0$ agent i is connected to agent j. If $g_{if}=0$ agents i and j are not connected. The elements of G can take on a range of values that reflect the strength (or other features) of the connection between two nodes. Unless otherwise stated, I will focus on the simplest case where two agents are either connected or not. In this case, the elements of this matrix are either equal to zero or equal to one.

The relationship between individuals is called "non-directed" if the fact that individual i is connected to individual j implies that individual j is also connected to individual i. The matrix G that characterizes a non-directed network is symmetric, $g(i,\ j)=1=>g(j,\ i)=1$. In a directed network it is possible that i is connected to j, while j is not connected to i.

Fig. 1a displays a simple non-directed network. Nodes A and B, and nodes B and C are connected to each other. Nodes A and C are not connected. This network is characterized by the following adjacency matrix:

$$G = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Fig. 1b displays a directed network. Nodes A and C are not connected. Node A links to node B, but node B does not link to node A.

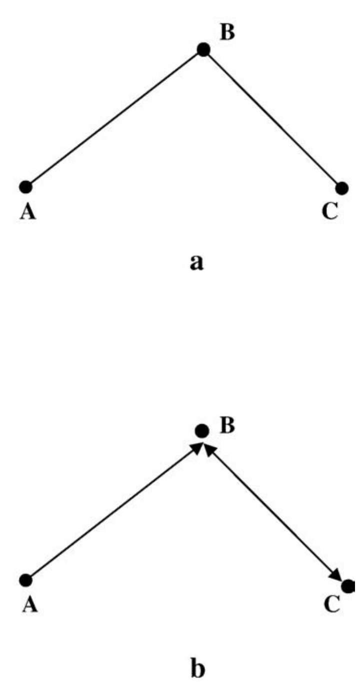


Fig. 1. (a) Non-directed network. (b) directed network.

Nodes B and C link to each other. The adjacency matrix of this network is given by:

$$G = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

The number of connections of an individual is referred to as the *degrees* of a node. The degree distribution describes how common different degrees of nodes are. A *path* between node i and node j is a sequence of nodes $i_1,\dots, i_K$ such that $g(i_k,i_{k+1})=1$ for all $k\in\{1,\dots, K-1\}$ and $i_1=i$ and $i_K=j$. The length of the path is given by K. The *distance* between i and j is the minimum path length between i and j. The *diameter* of a network is defined as maximum distance between any two nodes. In the network depicted in Fig. 1a all nodes are connected to each other. Nodes A and C each have one connection, node B has two connections. The distance between nodes A and C is two. The distance between nodes A and B, and B and C is one. The diameter of this network is two.

Consider a subset of nodes. If every node in this subset is connected through some path to each other node in the subset, the subset and its connections is called a connected sub-graph. A maximal connected sub-graph (a connected sub-graph that is not connected to any other nodes) is called a component.

One feature of social networks is that they are cliquish. That means, if i is connected to j and j is connected to k, there is a relatively high probability that i is connected to k, as well. One measure of the cliquishness of a network is the *cluster coefficient*. It captures the fraction of the friends of a given individual who are friends with each other. The literature considers different ways of calculating this measure. The total cluster coefficient is defined as [39]:

$$C = \frac{\sum\limits_{i;j\neq i,k\neq j,i} g_{ij}g_{jk}g_{ik}}{\sum\limits_{i;j\neq i,k\neq j,i} g_{ij}g_{jk}} \tag{1}$$

The cluster coefficient of an individual is defined by:

$$C(i) = \frac{\sum\limits_{j\neq i,k\neq j,i} g_{ij}g_{jk}g_{ik}}{\sum\limits_{j\neq i,k\neq j,i} g_{ij}g_{jk}}. \tag{2}$$

The average individual cluster coefficient is not necessarily equal to the total cluster coefficient.

Besides these simple measures presented above, a vast number of sometime complex ways to describe the structure of networks has been devised. For example, one way to define the role of individuals in a network is to calculate their Bonacich network centrality [6]. The Bonacich network centrality measures the total number of direct and indirect paths of any length stemming from a particular agent. Such paths are weighted by a factor that decreases with path length. Hence, Bonacich centrality depends both on the network structure and on the rate at which more distant paths are discounted. A more detailed discussion of this and other more complex network properties is beyond the scope of this paper and I refer the reader to [37,38] and [54] and their references. Nevertheless, the quantifiable network features presented here make it is possible to examine how a typical social networks looks like. In the next subsection, I use these measures to present common features of social networks.

2.2. Common characteristics of social networks

The concepts defined above make it possible to look at data on social networks and identify common features. In Section 2.2.1, I present the common patterns of network structure. Then, in Section 2.2.2, I discuss some empirical evidence on the role of individual characteristics in formation of social interactions.

2.2.1. Network structure

A number of common features that characterize social networks have been well established ([37] and [54]). Most members of social networks are connected to each other through a small number of intermediate nodes. In other words, social networks tend to exhibit a small diameter and a short average path length. This is also referred to as the small world effect. Empirical evidence for the existence of small world effects is presented by [30] who analyze the co-authorship network of academic economists.

The degree distribution of social networks is usually different from the degree distribution that would be generated by random network formation. The distribution is right skewed and has fat tails. Some links have a disproportionately high number of connections. Many social networks exhibit scale free degree distributions that follow a power law.

Social networks tend to be cliquish and exhibit a cluster coefficient that cannot be explained by random formation of links. In a randomly generated network with many nodes and few connections, the cluster coefficient basically equals the probability that two nodes are connected and is close to zero. Observed social networks exhibit higher cluster coefficients. The cluster coefficients for co-authorship networks in different academic disciplines range from 0.09 to 0.45 [37,54]. A network of actors, where a link is established when 2 actors co-star in the same movie, exhibits a cluster coefficient of 0.2 [54].

Furthermore, social networks exhibit positive degree correlations – nodes with a lot of links are connected to other nodes with a lot of links. Social networks also tend to exhibit a negative correlation between the individual cluster coefficient and the number of links of a node [54].

Gathering information on social networks is difficult. While some networks like co-author networks are clearly defined, many other definitions of social connections are more vague. It is not always clear what the relevant social network of an individual is. Moreover, traditionally researchers had to rely on surveys (for example, the National Longitudinal Study of Adolescent Health) to obtain information about social connections. The appearance of online social networking presents an opportunity for researchers to obtain access to information on vast networks. These networks may be of interest by themselves. Moreover, as discussed in Section 5, these networks are frequently related to underlying pre-existing real world social connections.

2.2.1.1. Online social networks. In an online social network the nodes are the users who are connected through some form of online communication. For example, a group of researchers who e-mail each other, or are 'friends' on pure social networking websites like MySpace.com or Facebook.com.

Mayer and Puller [48] analyze friendship networks on Facebook.com. They look at a snapshot of all student profiles from Facebook.com websites on January 17, 2005, at 10 Texas universities, Rice, the University of Texas at Austin, Texas A&M University, Baylor, Texas Tech, Texas Christian University, Southern Methodist University, the University of North Texas, the University of Texas at Arlington, and Texas State University. In 2005 Facebook.com was only available to college students, who needed to sign up with their official university e-mail address. On Facebook.com students set up a profile page that includes a picture, name, gender, high school, major, current classes, political orientation, music tastes, hobbies and other interests. The students' profiles also contain a list of 'friends'. A Facebook friendship is formed if student A sends a friendship request via the website to student B and student B accepts A's friendship invitation. Student A appears as a friend on B's Facebook profile and vice versa. Hence the friendship network on Facebook.com is a non-directed graph.

Table 1 displays the network features for these 10 networks. The average number of friends ranges from 17.2 at the UT-Arlington to 62.9 at SMU. This can be partially explained by the date that Facebook.com started on each campus. The variance of the number of friends is

Table 1
Characteristics of on-campus Facebook.com networks.

	Rice	U Texas	Texas A&M	SMU	Baylor	Texas Tech	Texas Christian	U North Texas	UT-Arlington	Texas State
Number of students	1300	8467	9299	2223	4295	4648	2342	2607	820	2922
Facebook uptake rate	0.80	0.40	0.44	0.57	0.61	0.31	0.52	0.18	0.08	0.22
Average number of friends	50.8	39.5	41.1	62.9	59.8	40.5	49.8	23.8	17.2	25.6
Variance of number of friends	31.9	36.5	38.4	48.3	50.8	35.6	36.0	23.9	17.7	23.8
Skewness of number of friends	1.06	2.01	2.06	1.75	1.74	1.50	1.11	2.28	1.52	1.69
Total cluster coefficient	0.24	0.20	0.17	0.23	0.19	0.21	0.23	0.21	0.27	0.23
Avg. individual cluster coefficient	0.30	0.22	0.19	0.27	0.21	0.23	0.25	0.22	0.25	0.23
Degree correlation	0.22	0.57	0.57	0.49	0.58	0.57	0.54	0.35	0.53	0.55
Degree–cluster correlation	−0.47	−0.04	−0.10	−0.17	−0.14	−0.08	−0.09	0.06	0.16	0.02

Note: This table is based on the data presented in Mayer and Puller [48]. Degree, degree correlation, total cluster coefficient, and individual cluster coefficient are defined in Section 2.2.1. The degree-cluster correlation is the correlation between students' number of friends (degree) and individual cluster coefficient.

closely associated with the mean; it ranges from 17.7 at UT-Arlington to 50.8 at Baylor. The distribution of the number of friends is clearly right skewed at all 10 universities. The skewness ranges from 1.06 at Rice to 2.28 at U North Texas. The cluster coefficients range from 0.17 at Texas A&M to 0.27 at UT-Arlington. The average individual cluster coefficient is in general slightly higher than the total cluster coefficient. Nodes with a lot of links are connected to other nodes with a lot of links, the degree correlation is always positive. It ranges from 0.22 at Rice to 0.58 at Baylor. The degree-cluster correlation is negative for all schools except the last three which were the youngest Facebook.com networks. Table 2 displays the degrees of separation between students in these 10 Facebook.com networks. Most students are connected to each other by no more than one or two intermediate connections. All of these features are similar to those reported for off-line social networks above.

Mislove et al. [50] use web-crawler software to assemble data on four popular social networking sites: Flickr, YouTube, LiveJournal, and Orkut. Their total data set contains 11.3 million users (nodes) and 328 million connections. They observe that these networks are characterized by a densely connected core of well connected nodes. The nodes in this core link to highly clustered groups of nodes with fewer connections. The nodes in networks with directed links (Flickr, YouTube, and LiveJournal) exhibit a strong correlation between the number of incoming links (in-degree) and outgoing links (out-degree). Moreover, they confirm that these online networks share some of the typical properties of off-line social networks. The cluster coefficients for the networks range from 0.14, for YouTube, to 0.33, for LiveJournal. The networks are characterized by scale free degree distributions and small world effects. Mislove et al. also investigate topic specific groups within the different social network sites. The members of these groups are highly clustered. They report within group cluster coefficients ranging from 0.34, for YouTube, to 0.81, for LiveJournal. These numbers are substantially higher than for the networks in their entirety.

Ahn et al. [2] study a data set that contains all nodes and links from a large South Korean social networking site (Cyworld). In addition they use smaller samples from web-crawls of MySpace and Orkut. They find a cluster coefficient of 0.16 for the Cyworld network and coefficients of 0.26 and 0.31 for their sample of the MySpace and Orkut networks.

Overall the features of online social network are very similar to those of off-line networks discussed above. This suggests that the mechanisms that lead to the formation of these networks are similar as well. In fact there is evidence that suggests that online social networks are related to underlying off-line social networks. There is empirical evidence that geography is reflected in online social networks. For example, physical distance does matter for the likelihood that two people are connected through LiveJournal.com [45]; and two students at Dartmouth College are more likely to have e-mail contact if they live close to each other on campus [63].

Mayer and Puller [48] find that the probability that two students form a friendship on Facebook.com increases substantially if they share common classes, lived in the same dorm or attended the same high school. They also present self-reported descriptions of how students meet their online friends. Using a sample of this information for Texas A&M, they report that the main channels of meeting friends are: being co-members of a school organization (26%), meeting through another friend (16%), attending the same high school (14%), and taking a course together (12%). Very few friendships are reported to be merely online interactions (0.4%).

The next step after describing social networks is to try to understand how they form and explain why they exhibit the above mentioned features. In Section 3, I present an overview of theoretical models of network formation. Before that, I discuss evidence that suggests that similarities in individual characteristics make social interaction more likely.

2.2.2. Empirical patterns of (online) social interaction

In this subsection, I focus on the simple – frequently atheoretical – description of who interacts with whom. The discussion above is focused on the network structure and uses relatively complex concepts to characterize networks. As seen below, there is a direct connection between network structure and interaction probability. One of the important stylized facts underlying the theoretical models of network formation presented in Section 3 is that individuals with similar characteristics associate with each other. This pattern is also referred to as homophily. The preference for interaction with individuals who are similar to one-self has been studied extensively by sociologists. I do not survey this literature here – for a summary and further references see [49]. Rather, I focus on recent evidence based on online social interactions. As discussed in Section 4, the consequences of the segmentation of social networks along socioeconomic and racial lines are of special interest to policy makers. Therefore, I emphasize the influence of race and socioeconomic background on social interaction.

The National Longitudinal Study of Adolescent Health (Add Health) includes self-reported friends of high-school students. Echenique and Fryer [22] propose a measure of segregation, the Social Segregation Index (SSI). Using the SSI and Add Health data they calculate the extent of within school racial segregation. They find that the social networks of African American students are fairly integrated in schools with a small number of black students. As the share of African Americans at a school increases, their social networks become more segregated. However, once the fraction of blacks at a school reaches 25% additional increases in the black student populations does not lead to a further segregation of their social network. Echenique and Fryer report similar but less pronounced patterns for students of other races.

Indirect evidence for the existences of social interactions can be obtained by observing interaction between individuals. Online interaction provides a source for such evidence. For example, Sacerdote and Mararmos [63] use e-mail traffic between students at Dartmouth College to identify the social networks of these students. They find that common characteristics facilitate the formation of friendships. While common interests, common major, or a similar family background all make social interactions more likely, the two most important determinants of social interaction are geographic

Table 2
Degrees of separation – Facebook.com within campus connections.

Fraction of all pairs with:	Rice	U Texas	Texas A&M	SMU	Baylor	Texas Tech	Texas Christian	U North Texas	UT-Arlington	Texas State
One degree of separation	0.04	0.00	0.00	0.03	0.01	0.01	0.02	0.01	0.02	0.01
Two degrees of separation	0.59	0.13	0.15	0.51	0.37	0.22	0.41	0.17	0.22	0.16
Three degrees of separation	0.36	0.62	0.66	0.43	0.57	0.62	0.52	0.54	0.43	0.56
Four degrees of separation	0.00	0.20	0.16	0.02	0.04	0.12	0.04	0.18	0.17	0.20
Five degrees of separation	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.03	0.02
Six degrees of separation	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Connected pairs	0.99	0.97	0.98	0.98	0.99	0.97	0.98	0.93	0.87	0.94
Average distance	2.30	3.00	2.95	2.40	2.62	2.81	2.54	2.81	2.57	2.88

Note: This table is based on the data presented in Mayer and Puller [48]. At six of the schools, a small number (less than .5%) of pairs have 7 degrees of separation. This table includes undergraduates in our F0acebook.com sample for whom we could identify race based upon the picture. To illustrate degrees of separation, if students *i* and *j* are friends, they have 1 degree of separation. If *i* and *j* are not friends but both are friends with student *k*, then *i* and *j* have 2 degrees of separation. The same logic applies to higher degrees of separation. Distance is defined in Section 2.2. A pair is connected if some path exists that connects a given pair of students.

proximity (e.g. being roommates or living in the same dorm) and race. They report that “Two randomly chosen white students interact three times more often than do a black student and a white student.” And that “ … placing the black and white student in the same freshman dorm increases their frequency of interaction by a factor of three.”

Mayer and Puller [48] report similar patterns for friendship networks on Facebook.com at 10 Texas universities. They find that at all 10 universities, similar characteristics of two students make the formation of a friendship more likely. The patterns in social segmentation are similar across schools – despite the fact these schools are very different in size and other characteristics. They use two measures of segmentation. First, the relative probability of friendship, which compares the probability that two members of a subgroup are friends, to the probability that two random students are friends. For example, the relative probability of friendship of blacks is given by:

$$\text{Relative Probability of Friendship}(\text{black \& black}) = \frac{\frac{\text{Number of pairs of blacks who are friends}}{\text{Total number of pairs of blacks}}}{\frac{\text{Number of pairs of any students who are friends}}{\text{Total number of any pairs}}}.$$

This measure of relative segmentation is independent of the size of different groups. Their second measure of segmentation, reflects the actual social environment of an individual. It is determined by the likelihood of forming a friendship with a particular race *and* by the racial composition of the student body. The fraction of black friends of a black student depends on their relative probability of friendship formation and the share of blacks in the entire student population:

$$\text{Fraction black friends of black student} = \text{Relative Probability of Friendship}(\text{black\&black}) * (\text{share of blacks in population}).$$

Table 3 from [48] shows the results of both measure of segmentation for 10 Facebook.com networks at Texas Universities. The top part of Table 3 documents the absolute segmentation by race. If friendships were formed randomly, the distribution of characteristics among the friends of any subset of students should equal the distribution in the population. At all universities and for all races, students have a higher fraction of friends from their own race than implied by random assignment. For example, 13% of the students at Rice are Asian, but 30% of the friends of Asian students are Asian. 25% of the friends of blacks at Rice are Black while blacks comprise only 5% of the student population. It can be seen that the racial composition of the student body is related to the diversity of social networks.

Table 3 also reports the relative probability of friendship. The probability that two white/Hispanic students form a friendship is similar to friendship formation of any two random students (unity). Given that most students are white/Hispanic this is not surprising. Two Asian students are 1.59 (at UT-Arlington) to 7.42 (at A&M) times more likely to be friends than any two random students. For pairs of blacks, this ratio ranges from 5 (at U of North Texas) to 16.5 (at A&M). The relative probability of friendship is smaller than one for cross-race pairs. There seems to be more segmentation along racial lines in bigger universities than in smaller ones – a pattern similar to the one reported for off-line social networks among high-school students [22].

Moreover, Table 3 documents segmentation by major, cohort, and political orientation. Even though these features matter, they seem to be less important for the formation of social connections than race. Students have more friends with similar political views and in the same major or cohort than random friend assignment would predict.

For one of the universities – Texas A&M – Mayer and Puller [48] match the social network data to student-level administrative data. Using this more detailed data they estimate a linear probability model of any two students being friends. Again, common race is found to be strongly related to the existence of a friendship. Relative to the baseline rate that any two students chosen at random are friends, two black students are 17 times more likely to be friends, two Asian students are 5 times more likely to be friends, and two Hispanic students are about twice as likely to be friends. Living in the same dorm is also strongly related to friendship formation (by factor 13 relative to a random student). Socioeconomic background and academic achievement affect the probability of a friendship formation to a smaller but statistically significant degree. Surprisingly, these different predictors of friendship formation seem to be largely independent. In particular, the coefficients for race are robust to controlling for demographics, ability, dorm, major and activities.

The pattern that social ties tend to form among individuals with similar characteristics is present in both off-line and online social networks. Moreover, the papers discussed here suggest a strong connection between off-line social encounters and online social interaction.

3. Models of network formation

Models of network formation help us to understand how and why networks form. This knowledge enables us to predict how networks will evolve over time or adjust to changes in their environment. Furthermore, models make it possible to analyze whether networks that result from a specific formation process are efficient according to some criterion, or whether it would be possible to construct a preferable network. Jackson [37] distinguishes two approaches for the modeling of the formation of social networks. One is based on the random graph literature. It is influenced by the literature in computer science and statistical physics (see [54]). Networks are formed mechanically following some algorithm or stochastic process. The other approach originates in the economics literature and is founded on game-theoretic tools. Networks arise as the result of decisions by utility maximizing agents. In this section, I first describe the more mechanical models of network formation. Then, I explain the economic approach. Finally, I compare the two approaches.

Table 3
Segmentation in on-campus Facebook.com networks.

	Rice	U Texas	Texas A&M	SMU	Baylor	Texas Tech	Texas Christian	U North Texas	UT-Arlington	Texas State
Segmentation by race										
Fraction of students White/Hisp	0.82	0.85	0.96	0.94	0.91	0.97	0.95	0.92	0.82	0.96
Fraction of friends of Wh/H who are Wh/H	0.85	0.93	0.97	0.96	0.96	0.98	0.97	0.94	0.92	0.97
Fraction of students Asian	0.13	0.13	0.02	0.03	0.03	0.01	0.01	0.02	0.06	0.01
Fraction of friends of Asians who are Asian	0.30	0.58	0.16	0.22	0.25	0.07	0.05	0.10	0.23	0.02
Fraction of students Black	0.05	0.02	0.02	0.04	0.06	0.02	0.04	0.06	0.12	0.03
Fraction of friends of Blacks who are Black	0.25	0.38	0.27	0.32	0.47	0.17	0.25	0.33	0.58	0.18
Pair of:	*Relative probability of friendship*									
White/Hisp and White/Hisp	1.03	1.12	1.01	1.05	1.10	1.02	1.03	1.04	1.14	1.01
White/Hisp and Asian	0.79	0.42	0.74	0.61	0.43	0.52	0.55	0.77	0.32	0.84
White/Hisp and Black	0.87	0.56	0.77	0.53	0.41	0.70	0.65	0.66	0.55	0.75
Asian and Asian	2.41	4.13	7.42	6.24	4.23	3.85	2.45	3.58	1.59	1.78
Asian and Black	0.92	0.54	1.01	0.86	0.52	0.80	0.77	0.69	0.36	1.00
Black and Black	5.12	13.13	16.54	6.92	5.99	7.35	5.59	5.03	5.71	6.33
Segmentation by major										
Fraction of friends in same major if friendships were formed randomly	0.04	0.02	0.02	0.01	0.02	0.03	0.01	0.01	0.05	0.01
Actual fraction of friends in same major	0.08	0.08	0.07	0.08	0.06	0.06	0.07	0.08	0.10	0.08
Segmentation by cohort										
Pair of:	*Relative probability of friendship*									
Freshman and freshman	2.14	2.24	2.10	2.10	2.10	2.01	1.95	1.85	1.72	2.07
Freshman and sophomore	0.64	0.74	0.72	0.64	0.60	0.82	0.74	0.84	1.00	0.79
Freshman and junior	0.46	0.40	0.45	0.38	0.33	0.52	0.45	0.62	0.73	0.46
Freshman and senior	0.35	0.25	0.31	0.20	0.18	0.43	0.25	0.58	0.61	0.31
Sophomore and sophomore	2.18	2.28	2.04	2.42	2.62	1.80	2.19	1.74	1.29	2.01
Junior and junior	2.17	2.13	2.14	2.21	2.29	1.46	2.17	1.55	1.27	1.77
Senior and senior	1.80	2.05	2.43	2.08	2.06	1.71	1.92	2.38	1.95	1.93
Segmentation by political orientation										
Pair of:	*Relative probability of friendship*									
Liberal and liberal	1.22	1.06	1.28	1.00	1.13	1.07	1.09	1.18	1.24	1.05
Liberal and conservative	0.86	0.75	0.69	0.66	0.59	0.70	0.85	0.76	0.79	0.81
Conservative and conservative	1.35	2.17	1.28	1.36	1.41	1.44	1.30	1.45	1.84	1.53

Note: This table is based on Table 2 of Mayer and Puller [48]. It includes undergraduates in our Facebook.com. Students were classified as either White/Hispanic, Black, Asian, or Don't Know, as described based on their profile pictures. The fraction of pairs of students of race *X* and *Y* who are friends is the fraction of all possible pairs of students of race *X* and *Y* who report being friends (reported in percentage points).

3.1. Mechanical models of network formation

One approach of modeling the formation of social networks is based on mechanical models used in computer science or statistical physics. These models do not explicitly consider decisions of individuals. Networks are formed as the result of a stochastic process. Individuals form a link with others with some given probability or according to an algorithm.

A simple example for this approach is a model with n identical individuals (nodes). Each pair of individuals forms a link with some exogenously given probability, p. A graph generated by such a process is a so called random graph. The relatively simple process of network formation makes it easy to analyze the resulting network structure. It is possible to derive some interesting properties of random graphs [see 37]. If p is small relative to n then the network consist of a number of small disjoint components. If p is relatively large ($p>1/n$) then the network contains a "giant component" – a connected subgraph that contains most nodes of the network. The number of connections each node has follows a binominal distribution, the probability that a node has exactly k links is given by:

$$P(d=k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}. \tag{3}$$

While random graphs have the advantage of being tractable, they are not able to generate some of the common features of social networks discussed in Section 2.2.1. Real world social networks tend to have a degree distribution that is more right skewed – they contain some nodes with a very high number of connections. Moreover, for large n and a small p random graphs exhibit almost no clustering, the cluster coefficient is simply equal to p. This is a feature that is not consistent with observed real world social networks.

Jackson [37] discusses alternative models of mechanical network formation that can generate these properties [see also 54]. One approach is to allow for dependencies in the process of network formation. Jackson and Rogers [40] present such a model. They analyze a process of network formation in which a network grows as more and more nodes subsequently are added to the existing network. These new nodes form random links with existing nodes. In addition, they form links with nodes that are connected to these initial connections. Jackson and Rogers refer to these two link formation (meeting) mechanisms as random meeting and search based meeting. They fit their model to real world social networks and show that such model can generate the observed characteristics of social networks like clusteredness and degree distribution. In addition, their model makes it possible to infer the relative importance of the random and search based meeting mechanisms for each network.

While more complex mechanical models of network formation help us to understand what kind of stochastic processes are able to generate the observed network features, they do not tell us why humans make decisions that lead to these processes.

3.2. Economic models of network formation

Economic or strategic models of network formation explain how decisions of individuals lead to the social networks we observe. Ultimately those decisions are the driving force behind mechanisms captured by simple algorithms of network formation.

When economists study the decisions of individuals they explicitly model their behavior based on a set of assumptions [47]. Decisions are made by agents based on their preferences. Agents can be individuals, as well as firms or governments. Frequently, these preferences are frequently described by utility functions. (A utility function is a function that ranks the desirability of outcomes associated with the different choices.) Agents choose their actions to achieve the highest possible level of utility for a given set of constraints. For example, they purchase the combination of goods they can afford with a given budget.

This methodology can be applied to the decisions that lead to the formation of social networks. The agents who form a network want to maximize an objective function – their utility function (or profits). The connections of an agent affect her utility. As discussed below, agents benefit from social networks in a number of ways. Social connections can provide direct benefits, such as a pleasant conversation or the feeling of security. Generally, economists consider access to information as the most important benefit of connections in a social network. For example, information obtained in social networks can help us understand characteristics of products or allows us to learn about the availability of jobs. Establishing or maintaining connections can also generate a cost for the agents involved. Time invested in forming or maintaining connections is such a cost. Other costs may include expenses for communication or travel. Utility maximization implies that agents choose their connections based on the ensuing costs and benefits.

The resulting decision process can be fairly complex. The benefit of forming a link with an agent might depend on the other links of this agent and the links of these links. Moreover, forming a new link might change the value of existing connections. If an environment that is not altered by the response of others to a decision, the decision problem of an agent can be solved by simple maximization subject to constraints. However, in the case of network formation strategic interaction between agents can affect the decision process. Economists use game theory to study decision problems where the decisions of multiple agents affect each other. Game-theoretic modeling accounts for the fact that agents consider the reaction of others when making a decision.

One way to characterize likely results of strategic interaction is to find the Nash equilibria. A Nash equilibrium is defined as a set of strategies chosen by agents, for which it is true that no agent wants to change their behavior given the behavior of the other agents. This concept is used to characterize outcomes in many areas of economics and other sciences. For example, it can be used to describe the price setting strategies of oligopolies that face the trade-off between lowering prices to gain market share and raising prices to raise profits per unit sold. However, if agents' strategies consist out of the announcing the links she would like to form – and only links that are desired by both sides are formed – a network without any connections is always a Nash equilibrium. Each player anticipates that all other players announce no wish to form a link, and it makes no sense for the individual to deviate from this strategy. Therefore, Jackson [37] points out that the concept of “Nash equilibrium” is only suitable for directed networks, where mutual consent is not necessary. He refers to an alternative equilibrium approach for networks that require mutual consent to form links, a concept of pair-wise stability [41].

A network is pair-wise stable if none of the agents wants to sever an existing link and no pair of agents wants to form a new link. Let $u_1(g)$ denotes the utility of agent i under the network g. Then a network g is pair-wise stable if for all combinations of i and j:

$$u_i(g) \geq u_i(g') \text{ where } g_{ij} = 1 \text{ and } g'_{ij} = 1 \tag{4}$$

and

$$\text{if } u_i(g) > u_i(g') \text{ with } g_{ij} = 0 \text{ and } g'_{ij} = 1 \text{ then } u_i(g) < u_i(g'). \tag{5}$$

In a stable network no player wants to sever an existing link. Moreover, if an additional link is desirable for a player it has to be true that the player at the other end of this link must be worse off if this link is indeed established. Otherwise, this additional link will be added and the existing network is not stable. Different environments will lead to different network equilibriums as defined by pair-wise stability.

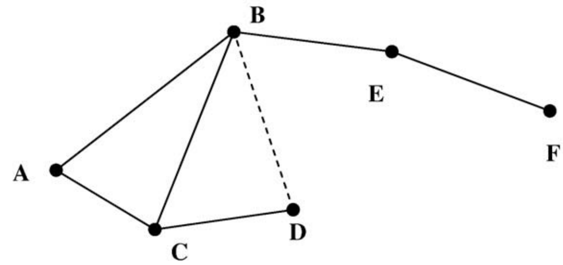
An equilibrium concept, like pair-wise stability, makes it possible to describe how stable networks look like. It is also possible to investigate whether those ‘equilibrium networks’ are efficient, or if other networks would be preferred according to some criterion. To do this some way to rank or compare different networks is necessary. One way is to compare the utility the agents who make up the nodes of the network. For example one can sum up the utilities of all the agents in a network and compare this ‘total’ utility for different network structures. However, this approach imposes strong assumptions about interpersonal comparability of utility. Another way to compare networks is the concept of Pareto efficiency. A network G is Pareto efficient if no other network exists that makes none of the agents worse of than under G and at least one person better off than under G.

Based on the exposition in [37] I present a simplified version of a connections model (see [41]). Such a model illustrates the benefits and costs derived from networks. Connections give benefits in the form of access to information or ability to trade. The variable δ_{ij} – defined to lie between 0 and 1 – captures the benefit of agent i from connection with agent j. There are also indirect benefits from friends of friends, here captured by δ_{ij}^2 if i and j are connected through a common friend. The benefit from more remotely connected agents is captured by δ_{ij}^3, for a connection through two friends,δ_{ij}^4 for a connection through three friends, and so on. Since $0<\delta_{ij}<1$the bigger the distance between two individuals the smaller the benefit they provide for each other. The costs of maintaining a friendship accrue only for direct friends. The cost of a connection to agent j accrued by agent i is given by c_{ij}. The utility that agent i derives from his direct and indirect connections in network g is given by:

$$u_i(g) = \sum_{(j \neq i)\&(p_{ij}<\infty)} \delta_{ij}^{P_{ij}} - \sum_{(j \neq i)\&g_{ij}=1} c_{ij}. \tag{6}$$

Where p_{ij} (g) gives the shortest path (or the distance) between agents i and j. Fig. 2 illustrates the utilities of different agents in such a connections model. In this figure it is assumed that the cost and benefit parameters are identical for all agents, $\delta_{ij}=\delta$ and $c_{ij}=c$. The table below the figure gives the utilities for each node. I compare two different network structures. In one case there is no direct connection between nodes B and D. In the other case such a connection exists. It can be seen that this additional connection affects not only the utility of nodes B and D, but also the utility of nodes E and F. In fact, while B and D have both a benefit and a cost of this additional connection nodes E and F accrue no cost for this additional link but benefit from closer connections to node D.

Given the assumptions about the cost and benefits of links in the connections model it is possible to derive properties of efficient networks. It can be shown that depending on the relationship between the costs of links and the benefit of connections there are only three simple networks structures that are efficient [41]. They show that if $c<\delta-\delta^2$, the complete network (every agent is connected to each other agent) maximizes the sum of the agents utilities. The costs of maintaining a link are so low that it makes sense to include all links in the network. The gain of forming a direct link to a friend of a friend outweighs its cost. The benefits of adding a link to an agent who is not a friend of a friend are even higher. If $\delta-\delta^2<c<\delta+\frac{n-2}{2}\delta^2$, total utility is maximized by a star shaped network, where one node is connected to all other nodes but no further links exist. The costs of adding a direct link to a friend of a friend outweigh its benefit. The star



Nodes	Utility without dashed connection	Utility with dashed connection
A	$2\delta+2\delta^2+\delta^3-2c$	$2\delta+2\delta^2+\delta^3-2c$
B	$3\delta+2\delta^2-3c$	$4\delta+\delta^2-4c$
C	$3\delta+\delta^2+\delta^3-3c$	$3\delta+\delta^2+\delta^3-3c$
D	$\delta+2\delta^2+\delta^3+\delta^4-c$	$2\delta+2\delta^2+\delta^3-2c$
E	$2\delta+2\delta^2+\delta^3-2c$	$2\delta+3\delta^2-2c$
F	$\delta+\delta^2+2\delta^3+\delta^4-c$	$\delta+\delta^2+3\delta^3-c$

Fig. 2. Connections model.

structure is the most cost effective way to connect all agents. It requires only $n-1$ links. A circle would require the same number of links but the average distance between different agents would be higher and the benefits from the connections lower (see [41] for a formal proof). Finally, if $\delta+\frac{n-2}{2}\delta^2<c$, an empty network where all agents are isolated maximizes total utility. If the cost are high enough even the most efficient structure (the star) does not generate benefits that are high enough to justify the costs of maintaining the required links.

The network structure that arises from the choices of self-interested agents sometimes coincides with the most efficient network structure. But for some combinations of cost and benefits pair-wise stable networks are not efficient and the most efficient network is not pair-wise stable.

If $c<\delta-\delta^2$ the complete network, where all nodes are connected to all other nodes, is pair-wise stable. As pointed out above, in this case the costs are so low that it always makes sense (for the individual or for the entire network) to form a direct link to a friend of a friend, outweighs its cost. Hence, there is only one pair-wise stable network and it is efficient.

If $\delta-\delta^2<c<\delta$ the star shaped network is pair-wise stable. However, there are other networks that are pair-wise stable, as well.

If $\delta<c<\delta+\frac{n-2}{2}\delta^2$ the star network described above is not pair-wise stable and all pair-wise stable networks are inefficient. This follows from the fact that if $\delta<c$ the center agent in the star wants to sever links with agents who have no other links. Her benefit from such a link is δ and hence lower than the costs c of this link. In fact, a pair-wise stable network will not contain agents with only one link. A link to an agent with no further connections generates a benefit that is smaller than the cost to maintain it. Each agent has either no link at all or at least two links.

If $\delta+\frac{n-2}{2}\delta^2<c$ there are no pair-wise stable connections and choices of self-interested agents result in the empty network. As seen above with such high costs the empty network is also the most efficient network.

Situations where actions of an individual do not only affect their own utility and the utility of their trading partners but also have indirect effects on others are referred to as externalities. Externalities are present in social networks if links formed or severed by one agent affect utility of agents that are not part of the link in question. Therefore, the networks that arise from individual decisions about the desirability of links are not necessarily the most efficient networks. For example, the additional link between nodes B and D in Fig. 2 results in a positive externality. The creation of this link increases the utility of nodes E and F. Depending on the values of δ and c this link may not be desirable for agents B and D.

If compensating payments between agents are possible more efficient networks might be obtained. In such a setting agent i who is linked to agent j can pay agent j to maintain links to other individuals – these links provide indirect benefit for agent i. Agents realize they can demand payment if they are more valuable – invest in links. The exact results of this process depend on the mechanism that leads to the determination of these payments. It is not always possible to internalize such externalities through transfer payments [41]. An example for improvement due to transfer payment can be seen in Fig. 2. Agents E and F might be willing to make a transfer payment to agents B and D if they agree to form a link. In a setting where agents move in a specified order and bargain over transfer payments associated with the formation of links, agents can benefit from their own direct links that provide indirect benefits to others [20].

The network structure in the models discussed in this subsection is very simple and does not necessarily exhibit the common characteristics of social networks. The reason is that in the networks considered here all agents are identical. This assumption makes the models tractable, but it is of course not realistic. When allowing for heterogeneity among agents or other generalizations more lifelike network structures arise – the downside is that these complex models are harder to analyze.

Clusteredness arises if individuals prefer to form links with individuals that are similar to themselves. As discussed in Section 2 empirical evidence suggest that homophily is a common feature in both online and off-line social networks. Similarity can refer to physical location or personal characteristics of agents. The costs of forming or maintaining a link are lower if it is easier to communicate with another person. The potential benefits are less uncertain if the

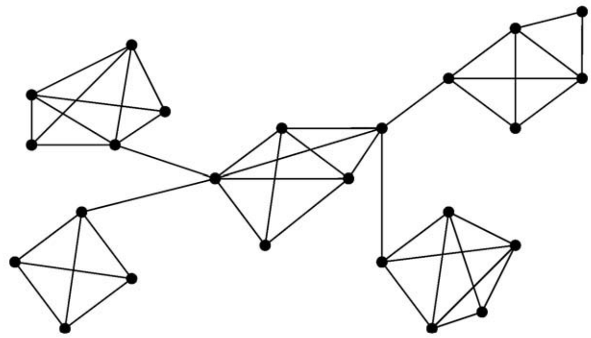


Fig. 3. Islands model.

other person is more familiar hence a link is more valuable if agents are risk avers.

The "small world effect" – the small distance between individuals in a network is generated by that high value of connection that decreases the average distance in the network significantly.

Both of these mechanisms are illustrated in the "island model" [39]. Agents live on separate islands. These islands can reflect a common taste in music, geographic location, a certain profession, or other characteristics. The cost of forming a link with another agent on the same island is relatively low and agents form connections to most other agents on their island. The cost of forming a connection to an agent on another island is high. Isolated agents would not form a connection across islands. However, a inter-island connection creates the additional benefit of the indirect connection to the other agents on the respective island. This pattern is illustrated in Fig. 3. It can be seen that this – still artificially simple – network structure exhibits clusteredness (agents on each island are connected) and short distances between agents (the islands are connected).

3.3. Comparison of the two approaches and applications

The mechanical models of network formation show us which processes are able to generate networks that exhibit features or real world networks, like clusteredness or a right skewed degree distribution. The mechanical models of network formation help us to understand 'how' networks are formed. However, they do not tell us 'why' networks are formed. The game-theoretic approach is able to address this question. It is based on utility maximizing agents who face a trade-off between the cost of maintaining links and the benefits derived from these links. For example, these models explain why there is an increased likelihood of link formation in a specific neighborhood and why links that lead to the small world effect are formed.

The two approaches do not contradict each other and can be combined. For example, Mayer and Puller [48] add preferences for connections to individuals with similar characteristics (homophile) to a mechanical model of network formation (see [40]). They show that such a model can generate features of online friendship networks on university campuses. They calibrate their model to fit these characteristics and use the parameterized model to simulate the results of counterfactual policies. They predict the impact of university policies, such as affirmative action, on social networks on university campuses. This illustrates how an understanding of the process of network formation makes it possible to infer features of unobservable networks. Networks may be unobservable because it is not possible to collect the appropriate data. Networks that do not exist (yet) are also unobservable. A model of network formation makes it possible to predict features of networks that arise due to changes in the network environment. In Section 5, I examine how the increasing importance of online interaction can change the structure of social networks. But first, in Section 4, I discuss the effects of social networks on individual decisions and the allocation of scarce resources.

4. Effect of social networks on outcomes

Economists are interested in social networks because there is ample empirical evidence that social networks affect our behavior. The structure and composition of social networks are reflected in "economic" outcomes. These outcomes include the allocation of time, the decision to purchases certain goods, or the ability to find a job. Moreover, information obtained through social networks may affect investment in human capital (education), occupational choice, or the decision to engage in criminal activities [13,14]. In this section, I discuss how social networks lead to differences in these outcomes and summarize the empirical evidence on these effects. I first discuss how information transmission can affect economic outcomes in general and present empirical evidence from different areas of economics. Then, in Sections 4.2 and 4.3, I focus on two areas where social networks are considered to be of special importance, peer effects in education and the labor market.

4.1. Choices and imperfect information

The transmission of information is the most important mechanism through which social networks affect the behavior of agents in an economy. Information is an important element of any decision process and social contacts are an important source of information. Individuals connected through a social network can obtain information from each other. Information can be shared actively or it can be obtained through the observation of others. Obtaining information can affect the behavior of individuals in a number of ways. Agents may change their behavior if they become aware of additional options, or they may reconsider a choice due to information about its costs and benefits.

In economics, the exchange of goods or services in a market is the most common way of modeling interactions. In the simplest case, all potential sellers and buyers of a good or service meet on a centralized market. Demand and supply are equated through the adjustment of prices. Frequently, it is assumed that all market participants are informed about the properties of the goods and services traded. All potential buyers and sellers are aware of each other. If there are many buyers and sellers exchange on such perfect markets can lead to an optimal allocation of goods. Economists speak of 'Pareto Optimality' or 'Pareto Efficiency', none of the market participants can be made better off without making someone else worse off.

Centralized markets for uniform, frequently traded goods – like commodity markets – are described relatively well by this simple model. However, there are some markets where access to information is not universal. Imperfect information can lead to market frictions and to an inefficient allocation of goods and services – it would be possible to reallocate goods in a way that does not make anyone worse off and at least one agent better off.

Lack of information can interfere with trade in two ways. First, attributes of goods or services are unknown. This is most important in markets for goods that are not standardized. The goods exchanged have unique – or hard to comprehend – properties. Examples include the real estate market or markets for professional services. Anecdotal evidence suggests that referrals are relatively important in these markets. Second, not all the potential trading partners are known. Agents are only aware of a limited number of other market participants and the goods they supply or demand. For example, a job seeker is not necessarily aware of all available vacancies.

Below, I give brief examples on how these two kinds of information problems are related to information transmission through social networks. Models of learning address how individuals use social interactions to update information about attributes associated with choices or actions.

Models of trade options analyze how social networks can provide information about potential trading partners. The research on the effects of social networks and their effect has been especially active in the areas of labor markets and peer effects in education. I discuss these two topics separately and in more details in Sections 4.2 and 4.3.

4.1.1. Models of learning

In learning models agents observe the actions of other members of their social networks and the results of these actions [31]. They compare these observations to their own actions and adapt their own behavior if the actions of others lead to more desirable results. Bala and Goyal [4] analyze a theoretical model of learning in social networks. They illustrate how learning from social contacts leads to the spread of behavior through networks. They assume a given social network. The agents in the network have identical tastes and all have to make the same choice. The outcome associated with this choice is uncertain and depends on unknown exogenous factors. Agents observe members of their social network. They observe their actions and the resulting outcomes. Agents will adopt the more successful strategies if they can observe them. This will lead to the spread of successful actions through the network. Bala and Goyal show that eventually all agents will adopt the same behavior. This behavior does not necessarily have to be the optimal action but if the network is large enough the probability that all agents converge to the optimal behavior is very high.

Empirical evidence for the importance of learning though social networks is found in a number of different areas. These include the adoption of agricultural technologies in developing countries, the purchase of consumer electronics, retirement savings decisions, and investment decisions of portfolio managers.

Farmers in Ghana use social networks to obtain information about technology and adopt successful technologies from their neighbors [17–19]. Farmers experiment with different technologies (for example fertilizer use) and share their experiences with members of their social networks. Farmers modify their behavior based on this information. For example, neighbors of a farmer who obtained higher than expected profit from extensive fertilizer use increase their fertilizer use.

There is also evidence of learning spillovers in the adoption of agricultural technology from India [28]. Household-level data to investigate technology adoption during the “green revolution” in India reveals that imperfect knowledge about new technologies delays their adoption. Profits of farmers are higher if they have neighbors with experience in new technologies. The presence of experienced neighbors also increases the amount of land devoted to new technologies. Farmers rely on their own experiences as well as on the experience of their neighbors to learn about the implementation of new technologies [28].

Mobius, Niehaus, and Rosenblatt [51] investigate how university students use social networks to update their perceptions about consumer products, such as MP3 players. They present a model that distinguishes between weak social learning (forming an opinion based on the observation of friends' behavior) and strong social learning (the active sharing of information). Furthermore, they differentiate between weak social learning and social persuasion (the influence of a fad or fashion) “ …where a consumer's valuation is directly affected by the valuation of his friends' for a product.” They conduct a field experiment and find that social learning is more important for consumer demand than traditional advertising channels. A result that supports the need to understand the structure of social networks is that as social distance decreases the number of acquaintances increases faster than the decrease in their influence. Therefore, they conclude that “distant acquaintances might have the largest effects on social learning.”

Devising a retirement savings plan involves decisions that are more complex and more important than the purchases of consumer electronics. Duflo and Saez [21] study retirement savings behavior of employees at a large university. They find that co-workers influence the decision to enroll in a Tax Deferred Account plan, as well as the choice of the mutual fund vender who manages the plan. They conclude that “…peer effects may be an important determinant of savings decisions.”

There is also evidence that social connections affect the choices of highly sophisticated decision makers when the problem they face is complex enough. For example, Cohen et al. [16] study the impact of shared education networks between mutual fund managers and corporate board members on investment choices. They find that portfolio managers invest more in companies that are managed by members of their education network. Moreover, fund managers tend to perform better with these investments than with investments in companies that are not managed by members of their education networks. Cohen et al. interpret this as an evidence for the hypothesis that social networks are an important source of information that leads to trading decisions and hence influences the determination of asset prices.

Mislove et al. [50] present evidence that suggests that learning takes place through online networks. Based on evidence from Flickr.com they argue that online social networks are used to locate content.

4.1.2. Models of trade options

Agents are frequently not aware of all the potential partners that are available for interaction. This issue is of importance in the labor market. The related literature is extensive and I discuss it in a separate Section 4.3. Here, I focus on the more general implications for the formation of prices. One way to model limited interaction possibilities between market participants is to assume that agents can only trade with agents with whom they are connected in a network. For example, Kakade et al. [43] study the effect of trading opportunities that result from connections in a network on price dispersion (Price dispersion is present if buyers pay different prices for identical products or services). Buyers can only purchase a consumption good from a seller with whom they have connections and vice versa. Prices are determined to achieve market clearance (every seller sells her products) but they may vary between sellers. In such a framework better connected agents are able to obtain better prices. Hence, the price that an agent pays or receives for a good is related to their position in the network, and the overall network structure has an effect on price dispersion. They use data on international trade networks to calculate the resulting price dispersion.

4.2. Peer effects in education

The choice of a particular school (district) or university is frequently influenced by concerns about suitable peers. Students can affect their classmates' behavior and academic achievement in a variety of ways. Knowledge spillovers of talented classmates can improve learning. At the same time disruptive behavior of some students can severely affect the learning environment of an entire classroom. A number of undesirable behaviors, such as smoking, drinking or drug consumption, are often induced by peers.

Empirical evidence of peer effects can be found for all age groups. For example, Hoxby [34] uses fluctuations in the composition of cohorts at Texas elementary schools to estimate peer effects on academic performance. She finds that a 1 point increase in the reading scores of peers increases a student's own reading score by 0.15 to 0.4 points. These effects are strongest for peers of the same race. She also finds that a student is not only affected by their peers' academic achievement. A higher share of female classmates improves the math performance of students, even though females do not perform better on math tests than male students. Sacerdote [62] examines peer effects by exploiting that at Dartmouth College roommates and dorm-mates are randomly assigned. He finds that “peers have an impact on grade point average and on decisions to join social groups such as fraternities.” However, he cannot find evidence that suggest that residential peers affects “ …other major life decisions such as choice of college major.”

Besides academic outcomes educators are also concerned about with ‘social outcomes’, such as drinking or smoking. In fact, Foster [27] points out that in student populations peer effects tend to be more important with respect to social outcomes than with respect to performance.

In the papers mentioned above, random assignment makes it possible to investigate causal effect of peers. It is harder to isolate a causal effect of social connections when they do not arise as the result of some exogenous, random assignment process. The majority of social networks – online and off-line – are the result of an endogenous process which makes it almost impossible to isolate causal effects.

Calvó-Armengol et al. [12] point out that most empirical studies of peer effects rely (frequently implicitly) on two assumptions. "First, peer effects are conceived as an average intra-group externality that affects identically all the members of a given group. Second, the group boundaries for such a homogeneous effect are often arbitrary, and at a quite aggregate level, in part due to the constraints imposed by the available disaggregated data." Calvó-Armengol et al. explore the connection between the position of students within their social network and two outcomes: delinquency and academic performance. They develop a theoretical model of peer effects. The model predicts that in equilibrium, a common group externality that affects all the group's members to the same degree does not exist. The position of an individual within a social network determines the strength of peer influences. This prediction is tested with data from the National Longitudinal Survey of Adolescent Health (AddHealth). They find that the position of an individual in a network as measured by her Bonacich centrality explains differences in behavior. A one standard deviation increase Bonacich centrality is associated with a 44% standard deviation increase in delinquency, and a 32% standard deviation increase in academic performance. Hence the authors conclude that "the position of each individual in a network plays an important role in shaping individual behavior."

Weinberg [66] develops a model that is designed to study the interactions between friendship formation and behavior. He presents a model of network formation with endogenous associations. He presents evidence based self-reported friendships from Add Health data that is consistent with the predictions of the model. He assumes that individuals choose the amount of time they invest in social interactions to maximize their utility. This utility depends on their own behavior and the behavior and characteristics of the people they interact with. The model generates the feature that people with similar characteristics and patterns of behavior are more likely to interact. This reinforces underlying behavioral tendencies and the effects of groups on individual behavior will be nonlinear even when the underlying behavioral model is linear. The nonlinearity introduced by the endogenous association in social networks and leads to different predictions than a linear peer affects model. For example, an increase in the number of students that smokes will provide students who are inclined toward smoking an opportunity to form friendships with additional smokers and reinforcing their original tendencies.

Echenique and Fryer [22] and Echenique et. al., [23] also exploit AddHealth data. They calculate racial within school segregation of students and find some (small) correlations between the segregation of a student – as measured by their Social Segregation Index (SSI) – and, the likelihood of smoking, test scores and college attendance. Fryer and Torreli [29] study effects of peer groups on academic achievement of minority students, again using Add Health Data.

Mayer and Puller [48] document relationship between online interaction and outcomes of college students. They link information from Facebook.com networks of students to information from university records at Texas A&M University. Stressing that any associations should not be interpreted as causal, they investigate relationships between the characteristics of students' friends and outcomes. The outcomes are grades as measured by GPA, self-reported drinking behavior, and participation in student organizations. They report that after controlling for own 'pre-treatment' characteristics of students the 'pre-treatment' characteristics of friends are correlated with a students grades. When controlling for contemporaneous GPA friends pre-treatment characteristic no longer matter. The contemporaneous outcome GPA of friends does however matter. This is consistent with the notion that current behavioral patterns of friends are important to influence grades, while friends' characteristics are of lesser importance. They also report that students' own characteristics and to a lesser extent friends' characteristics are correlated with a self-reported inclination to drink alcohol. Again current behavior of friends – in this case self-reported drinking – shows a stronger effect than pre-treatment characteristics.

Mayer and Puller [48] also investigate the relationship between online social connections and the participation in volunteer, religious and political organizations. They find that membership in a political organization and students' Facebook.com networks are only weekly related. However, there are clear associations between students' Facebook.com networks and the membership in volunteering groups (like Habitat for Humanity) and religious organizations. They also find that volunteer groups seem to serve as a meeting channel. While in the case of religious groups common tastes seem to be responsible for the association between social networks and group membership.

4.3. Labor markets

Labor income is the most important source of income for the vast majority of individuals. Hence, economists pay special attention to this market. The labor market is also a market in which information transmission through social networks is especially relevant. Heterogeneous workers with unique characteristics are matched to diverse jobs with varying demands. Therefore, no centralized central market for labor exists and it is difficult for both sides, job seekers and employers, to find a suitable match. The process that matches workers and jobs is characterized by frictions that arise from imperfect access to information. The consequences of these frictions are analyzed in a vast literature on search and matching. See for example Pissarides [57] and Rogerson, Shimer, and Wright [60]. This literature explains why we observe at the same time both vacancies and unemployed job seekers suitable for these vacancies. In Section 4.3.1, I summarize the empirical evidence suggesting that social networks are important in the job-search process. In Section 4.3.2, I discuss some of the theoretical implications of information transmission through social network in a job-search framework.

4.3.1. Empirical evidence

The importance of information transmission through social contact for the labor market has been recognized for a long time. Rees [59] distinguishes formal and informal sources of information acquisition in the job-search process. Formal sources include public and private employment agencies, newspaper advertisements, school placement services, and more recently online advertisement or specialized websites. Informal sources are referrals from other employees or employers. Granovetter [33] stresses the importance of weak ties – more casual acquaintances outside the closest social circle of friends – for the job-search process. Bewley [5] and Montgomery [53] summarize the findings of various surveys and conclude that about half of all currently employed workers found their job through information provided by friends or family members.

Ioannides and Loury [35] survey the literature on the effects of information transmission though social networks in job search. They list seven stylized facts that emanate from this literature:

1) "There is widespread use of friends, relatives and other acquaintances to search for jobs and it has increased over time." Based on evidence from the Current Population Surveys they suggest that in the early 1970s about 15% of unemployed workers used friends and relatives in their job search. This number has increased to above 20% in the early 1990s.
2) The use of information from social networks in the job-search process varies by location and demographic characteristics. More educated workers and women are less likely to use informal job-

search channels. Workers in low income neighborhoods are more likely to use social connections in job search.
3) "… job search through friends and relatives is generally productive." Workers who use social contacts receive more offers (and accept more jobs) per application. This explains the finding that about half of all workers report that they found the current job through social contacts. Moreover, workers who found their jobs through referrals are less likely to quit and have longer job tenure. The evidence of the effect on wages is mixed.
4) The differences between different demographic groups in the effectiveness of their jobs search can partly be explained by their reliance on referrals. Women are less likely to use referrals and report less frequently that they found their current job through friends or relatives.
5) " .. many differences in productivity of job search by age, gender, race, and ethnic group cannot be completely accounted for by differences in usage." Conditional on using social referrals men and whites are more successful in finding a job than women or members of minority groups.
6) The use of the Internet as a job-search tool is increasing. There is some evidence for a digital divide. Members of minority groups are less likely to utilize the Internet to search for employment.
7) There are large differences across countries and industries in the use of social connections for job search.

Sacerdote and Marmaros [63] examine the role of social connections in the job search of graduates of Dartmouth College. They report that students believe that networking is an important tool in the job-search process. The students use different networking strategies depending on their gender and race and on the job they are seeking. Students who exploit connections to alumni members through fraternities or sororities are reported to be more likely to obtain high paying jobs. This observation does not imply that there is a causal effect of the connections provided by sororities or fraternities. An alternative explanation would be that students who have certain characteristics – such as high energy or likeability – tend to join sororities or fraternities and these same characteristic lead to success in the labor market, independent of the social connections. To address this selection issue Sacerdote and Marmaros exploit the fact that roommates and dorm-mates at Dartmouth College are assigned randomly. They find similarities in the career paths of students who where randomly chosen to reside in close proximity to each other during their freshmen year. This random assignment policy makes it possible to rule out self-selection. However, the reflection problem described by Manski [46] is still an issue. Students who where randomly assigned to be roommates might have been exposed to common influences that determined their career paths but had nothing to do with information transmission through social contacts.

Topa [65] presents a model where agents share information about job openings with members of their social neighborhood networks. In such a model the probability that an unemployed agent finds a job is increasing in the number of his friends or neighbors who are currently employed. He assumes that agents exchange information with agents in their own and the neighboring census tracts and that social networks form long ethnic lines. This leads to testable predictions about the relationship between the unemployment rates in adjacent neighborhoods. Using Census Tract data for Chicago he finds that the relationships in the unemployment rates of adjacent neighborhoods cannot be explained by similarities in the characteristics of the residents. This suggests that social interactions are responsible for part of these relationships. Moreover, these social interactions seem to be most important in minority neighborhoods and in neighborhoods with relatively low levels of education.

4.3.2. Theoretical models of job search through social networks

The empirical evidence suggests that information transmission through social networks plays an important role in the process that matches workers to jobs. Therefore, understanding how social networks influence this process is vital to the understanding of the labor market. As mentioned above, information transmission through social networks can affect the job matching process in two ways [44]. First, employers obtain information about worker characteristics from other employers or workers. Second, workers can obtain information about vacancies through their social networks.

4.3.2.1. Information flow to employers. While the second idea has received considerably more attention in the literature, a number of theoretical papers investigate the information flow from social networks to employers. One example is Montgomery [53] who presents a theoretical model of the effects of hiring due to referrals. In his model both workers and firms use formal and informal hiring channels. The informal hiring channel is based on the social ties of workers. The model illustrates why profit maximizing firms have an incentive to utilize referrals. He shows why social connections of workers can be associated with labor market outcomes. The model also suggests that the structure of the social network through which referrals take place affects the income distribution. Greater stratification with respect to productive worker characteristics increases the dispersion of wages.

4.3.2.2. Information flow to workers. The flow of information about vacancies through social network is usually modeled by augmenting a simple search or match framework with information transmission though social networks. For example, Calvo-Armengol and Jackson ([9] and [10]) consider a model where existing jobs are destroyed with some random probability and new – initially vacant – jobs arrive. Some but not all workers hear about these vacancies. There are n individuals who are connected by an exogenously given $n \times n$ matrix G. If $g_{ij} > 0$ individuals i and j are linked and share information about job openings. (The value of g_{ij} captures the intensity of the connection between i and j.) An unemployed worker does not share information about vacancies – she tries to secure the job for herself. An employed individual shares information about a vacancy with her social contacts. She passes it on to each of her social contacts with a probability that is proportional to the intensity of the contact captured by the elements of the matrix G. Once an unemployed worker hears about a vacancy it is filled and the worker is no longer unemployed.

Calvo-Armengol and Jackson assume that wages are a non-decreasing function of the past wage and the number of employment opportunities of an agent. The fact that friends of employed workers are more likely to find employment leads to a positive correlation of employment status and wages across agents and time. Heterogeneity in the probability of finding a job can lead to duration dependence. Consider a simple example with two workers. One worker has a 'good' social network, she has connections to many employed workers. She has a relatively high probability of finding employment. The other worker has a 'bad' social network – few connections to employed workers – and a lower probability of finding employment. If both workers become unemployed in the same period the worker with the 'good' social network is more likely to find new employment quickly. The worker with the bad network remains unemployed for longer. Therefore workers who have been unemployed for a longer period of time tend to be workers with bad social networks. These workers have a lower probability of finding employment in a given period. We observe duration dependence of unemployment. Note that, more conventional explanations of duration dependence are based on the same logic, but assume unobserved characteristics of workers as the reason for differences in the probability of finding employment. Calvo-Armengol and Jackson also show that the likelihood of dropping out of the labor force is higher for workers with 'bad' social networks. Initially small (random) differences between different groups can be magnified through this referral process and result in eventually large differences in the labor force participation.

Most other papers that study the role of information transmission through social networks in the job-search process use modeling approaches similar to Calvo-Armengol and Jackson ([9] and [10]). By simulating a calibrated version of their model Arrow and Borzekowski [3] provide some quantitative results for the effects of social networks on difference in labor market outcomes. Their findings suggest that some labor market patterns that have been previously explained by unobserved heterogeneity among workers might be the result of information transmission through social networks.

Calvo-Armengol and Zenou [15] and Ioannides and Soetevent [36] augment a standard labor market matching model [see 57] with information transmission though social networks. They model the spread of information in fashion similar to Calvo-Armengol and Jackson ([9] and [10]). In this framework an aggregate matching function provides the number of matches of unemployed workers to vacancies as a function of job seekers and vacancies posted by firms. Assuming a network structure where all workers have the same number of connections, Calvo-Armengol and Zenou suggest that the introduction of matching through social networks preserves some of the standard properties of the matching function. It is increasing and concave in both unemployment and vacancy rates. However, the matching function is no longer homogenous of degree one. The network size has a non-monotonic effect on the matching rate. A denser network increases the probability that unemployed workers hear about a vacancy from a social connection. However, a denser network also increases the chance that multiple vacancies reach the same unemployed worker. This leads to congestion effects and after a critical network size has been reached an increase in the network density leads to a decrease of the matching rate. Ioannides and Soetevent [36] relax the assumption that all workers have the same number of connections. In their model the number of connections of each worker is drawn from an arbitrary probability distribution. They use a Poisson distribution to calibrate their model to match US data. They find that a denser network leads to lower unemployment and higher mean wages.

These advances rely on simplifications to reduce the complexity generated by the introduction of information transmission though social networks. One approach [for example 9,10, or 26] is to assume a very simple structure of the networks. An alternative is to assume that networks are constantly dissolved and reformed randomly [for example 15 or 36]. Given these assumptions it is difficult to obtain a sense for the magnitude of the effects generated by social networks. Mayer [67] uses numerical simulations to examine the implications of more realistic – and complex – networks and quantify the effects of information transmission through these social networks on labor market outcomes. He finds that differences in social networks of individuals can generate substantial variation in the length of unemployment spells and the probability of being employed. However, wage differences that arise from differences in the bargaining behavior of workers with different kinds of social networks are only minor.

4.3.2.3. Endogenous job Information networks. Ioannides and Datcher Loury [35] distinguish between exogenous job information networks, a given social structure affects the information flow about job openings, and endogenous job information networks, the actions of individuals that lead to the formation of social networks are based on their role in the information transmission process. The models presented above all assume exogenously given social networks. Models of endogenous job information networks are special cases of models of strategic network formation discussed in Section 2. Calvo-Armegnol [8] studies a model where workers hear about job openings either directly our through their social network [similar to 9,10]. Additional social ties increase the likelihood that a worker hears about a vacancy. At the same time more indirect links leads to increased competition for the information provided by the direct links. In such a setting the network structure affects the efficiency of information transmission. Two networks with the same number of total connections can lead to different unemployment rates. Therefore it is possible to rank these networks. Calvo-Armegnol investigates which network structures arise endogenously. He assumes that links are costly to maintain and the benefit of maintaining a link is the resulting increase in the probability of finding job. Links are only maintained if they are beneficial to both sides.

4.3.2.4. Occupational choice and intergenerational persistence. Information transmission through social networks can also have other more indirect effects on the labor market. For example workers might anticipate that their success in certain careers depends on their social network. Consequently they take this into account when investing in human capital (education) or choosing an occupation. Calvo-Armengol and Jackson [11] illustrate that if decisions of both parents and their children decisions depend on an overlapping social environment, strong parent–child correlations can be generated without any direct relationship between.

To summarize, there is ample empirical evidence that information transmission through social networks is an important part of the process that matches workers to jobs. The rapidly evolving theoretical literature helps to understand this process in more detail. In the next section, I discuss how the increasing importance of online communication affects the matching of workers to jobs and other outcomes.

5. Online social networks

In this section, I explore the consequences of the rise of online social networks. I first discuss the differences between traditionally formed social networks and online networks. Then, I explore the implications for the network structure and for economic outcomes.

5.1. Differences in network structure

As seen in Section 2.2 online social networks exhibit many of the features observed in real world networks, such as clusteredness or small world effects. However, the nature of online communication leads to a number of differences between online interactions and real world interactions. The characteristics of communication partners are perceived differently online and the relative importance of these characteristics may be different than in off-line interaction. It is easily possible to be part of a number of separate online social networks. Online networks make it easier to provide false or misleading information, and it is harder to verify information provided by others. Moreover, in online networks it is easy and basically costless to search a vast number of potential interaction partners. This makes it possible to find communication partners with very specific interests.

In terms of the models discussed in Section 3 the main difference between online networks and real world networks is in the cost of communication. Compared to traditional social networks, both, monetary cost and time costs of communication are lower and depend less (or not at all) on distance. It is not necessary to travel, pay for long distance phone calls or pay postage and wait for physical mail. In fact, the physical location of members of these networks does not affect their ability to communicate online. However, as mentioned above, online social networks are used to augment other forms of social interaction. Therefore, physical distance plays an indirect role. Related to the decrease in the cost of interaction is the frequency of interaction. While we are limited to a small number of face to face interactions every day, online networks make it possible to share information with (and update information on) a large number of individuals in a short period of time. Reduced cost of interaction can also lead to the formation of temporary or specialized networks.

Direct empirical evidence for changes in interaction due to changes in communication technology is provided by Goyal et al. [30] who analyze the co-authorship networks of academic economists. They find that over the time period from 1970 to 2000 the number of connections of each author has increased and even though the overall number of academic economists has increased the average distance between them has decreased.

5.2. Implications

The factors mentioned above influence the kind of social connections that are formed in online social networks. Costs of forming and maintaining connections are lower in online networks and the connections are less dependent on constraints like distance.

5.2.1. Improved information flow

A straight forward implication of lower costs of communication is more communication. This leads to potentially more social contacts and better flow of information. In learning models, discussed in Section 4, a denser network leads to a faster spread of information and learning [56]. Therefore, the increased network density that arises from the lower communication costs can be welfare improving. It can lead to a faster spread of information about new technologies. It can also facilitate the spread of information about product characteristics. For example, customers share feedback related to online purchases, and online merchants provide information on purchases of other customers with similar profiles. This can be viewed as a specialized and temporary social network. It is made possible by the low cost of interacting. Oestreicher-Singer and Sundararajan [55] explore the influence of visible co-purchase decisions by other costumers of Amazon.com. Amazon.com provides information about complementary products under the heading "Consumers who bought this item also bought. . . ". This leads to peer effects in purchasing decisions and Oestreicher-Singer and Sundararajan estimate "that the visibility of co-purchases more than triples the influence that complementary products have on each others demand." Possible implications of this improved information flow are higher price transparency and a decrease in search cost.

As mentioned in Section 4, the job-search process is one area where information flow through social networks is especially important. In general, more available information should improve the matching of workers to firms and increase efficiency. However, Calvo-Armengol and Zenou [15] point out that in their labor market matching model the networks size has a non-monotonic effect on the matching rate. Due to congestion effects an increase in the network density leads to a decrease of the matching rate after a critical network size has been reached. However, their results hinge on the assumption that all workers have the same number of connections. Ioannides and Soetevent [36] relax this assumption. They allow for a Poisson distribution for the number of connections and calibrate their matching model to match US data. They find that a denser network leads to lower unemployment and higher mean wages. This suggests that better understanding of the structure of social networks is needed to assess if a decrease in communication cost does in fact reduce mean unemployment.

5.2.2. Increased segmentation along some characteristics

One implication of lower communication cost is a decrease in separation along some dimensions – such as physical distance. However, as illustrated by Rosenblat and Mobius [61] this is accompanied by an increase in separation along other dimensions, like individual characteristics. Rosenblat and Mobius examine the effects of decreasing communication costs in a theoretical model. They consider a social network where agents have closer and more distant neighbors. Distance increases the cost of interaction. Individuals are part of two different groups. Distant and close neighbors can be members of either group. The benefit of interacting with members of the own group is higher than the benefit of interacting with members of the other group. If the costs of interacting with distant neighbors are higher, individuals tend to interact with close neighbors from both groups. A drop in the cost of interacting with distant neighbors leads to more interaction with members of the same group even if they are not close neighbors. Rosenblat and Mobius define two measures of separation, group separation and individual separation. Group separation measures the share of interactions within a group relative to in-between group interaction. Individual separation measures the average time a new piece of information needs to travel to a random agent. Lower communication costs lead to increase in group separation (agents become more selective) and in general to a decrease in individual separation.

In their model agents derive welfare from social interactions. Their welfare depends on "a communication externality that results from the collaboration decisions of all other agents." This externality consist of three components: the benefit of information received from other agents, costs generated by differences in group opinions, and the benefit of institutions serving the needs of specific groups. Lower individual separation increases the benefit from transmission of ideas and improves welfare. It improves access to valuable information for example, about jobs, products, or technologies. An increase in group separation has ambiguous welfare consequences. Higher within group differences make the coordination of public good provision more difficult. This may lead to a reduction in the provision of public goods. An increase in group separation facilitates coordination within a group, "… group separation gives rise to institutions complementing private gains." Rosenblat and Mobius use data on collaboration patterns between academic economists empirically to test their hypothesis. They find evidence that supports both lower average individual separation and greater group separation.

Currarini et al. [20] make a similar point. They build an economic model of friendship formation to study how the homophily can lead to segregation of minorities. Their findings suggest that: larger groups tend to form more same-type ties and fewer other-type ties than small groups; larger groups form more ties per capita; and all groups are biased towards same-type relative to demographics. Therefore, an overall increase of potential interaction partners would lead to increased segregation along demographic lines. Some evidence that confirms this prediction can be found in Mayer and Puller [48]. By looking across Facebook.com networks at different university campuses it can be seen that larger universities tend to be more racially segregated than smaller universities.

Hence it seems reasonable to conjecture that the decreased cost of online interaction can increase segmentation along some demographic dimensions. The question is which are the characteristics that become relatively more important? From a policy perspective, we should be most concerned about race and socioeconomic status.

Race and socioeconomic status are strongly related to physical location. Therefore, a decreasing importance of the geographical distance of the formation of ties in social networks makes it easier to overcome barriers like physical distance. Hence decreasing costs of communication can also lead to less segmentation along the dimensions race and socioeconomic status. Their overall effect is ambiguous. The rise of online social networks can lead to either an increase or decrease of segmentation along demographic lines. More segmentation is in general associated with undesirable outcomes. Montgomery [53] points out that more segmentation along productive abilities increases income dispersion. Calvo-Armengol and Jackson [9] illustrate that a subgroup that is isolated from the information flow from others can end up with a high rate of workers drop out of labor market. The findings of Calvo-Armengol and Jackson

[11] suggest that an increase in segmentation can also lead to more persistence in earnings across generations.

5.2.3. Network size and pro-social behavior

Due to the low costs of interacting the community sizes of online social networks are potentially very large relative to real world social networks. Putnam [68] reports that community size is frequently negatively related to pro-social behavior like helping strangers or volunteering. Networks in small communities are connected closer. They can generate more social pressure and better enforcement mechanisms to encourage behavior beneficial to society. Similarly, Mobius and Szeidl [52] relate the ability to trust others to the structure of social networks. In closely connected social networks there are more opportunities for sanctioning behavior and hence more trust. In a network that is characterized by dense interconnectedness of small groups (networks that exhibit a high degree of network closure) an individual can ask valuable favors of a small number of individuals. In more dispersed networks (a low value of network closure) a larger number of people can be asked for less valuable favors. Allcot et al. [1] use the concept of network closure to study the effect of community size on pro-social, trust related, behavior. They use data on the self-reported sizes of social networks of middle-school students and high-school students. They find a strong negative relationship between grade size and network closure. They consider the three outcome variables: whether a student reported “trouble” with another student since the beginning of the school year; whether students feel safe; and grade point average. The latter does not directly capture pro-social behavior but might be affected by it. They confirm the finding that grade size is negatively associated with pro-social behavior.

While low cost of communication improves information flow it may increase segmentation along some dimensions and reduces the ability to induce pro-social behavior. It can reduce trust. Trust is essential for many transactions. Due to their large size online social networks make it hard to generate trust through conventional mechanisms. The need for trust leads to alternative solutions – for example the use of reputation. Brown and Morgan [7] explain how eBay uses a feedback system to generate informative reputation for its users. This illustrates that new mechanisms may be developed that can overcome some of the disadvantages that result from the low communication cost and anonymity of online communication.

It is possible to make some predictions of the effects of the increasing importance of online social networks. However, many of the predictions are ambiguous and more empirical research is necessary to understand the magnitudes of any effects.

6. Conclusion

Changes in communication technology have dramatically altered the nature of social interaction. Online social networks are one new way of communication that changes how individuals interact. In this paper, I explore online social networks in economics from two angles. First, I explain how economists model the formation of social networks. Second, I discuss how social networks affect economic outcomes.

Economists use models of network formation based on stochastic algorithms but more importantly they model the formation of social networks as the result of optimization decisions. Social connections are associated with costs and benefits and decisions makers weigh these against each other. These models make it possible to analyze the consequences of technological changes that alter the cost of communication.

Social networks are an important mechanism of information transmission. Online social networks facilitate interaction between individuals. This improves information transmission. At the same time it alters the structure of social networks. Better information flow improves price transparency, facilitates learning, may advance technology adoption, and makes it easier to obtain information about product features or characteristics of trading partners. Larger social networks may reduce the ability to induce pro-social behavior. Online social networks also facilitate specialization which may increase or decrease segmentation along some demographic dimensions.

Online social networks are still evolving and research on the topic has only started. Research approaches that borrow from multiple disciplines may be especially fruitful to understand how online social networks form and what their consequences are.

References

[1] H. Allcott, D. Karlan, M. Mobius, T. Rosenblat, A. Szeidl, Community size and network closure, American Economic Review 97 (2) (2007).
[2] Y.-Y. Ahn, S. Han, H. Kwak, S. Moon, H. Jeong, Analysis of topological characteristics of huge online social networking services, Proceedings of the 16th international conference on World Wide Web (WWW'07), Banff, Canada, 2007.
[3] K.J. Arrow, R. Borzekowski, Limited network connections and the distribution of wages, FEDS Working Paper No. 2004-41, 2004.
[4] V. Bala, S. Goyal, Learning from neighbours, Review of Economic Studies 65 (3) (1998).
[5] T. Bewley, Why Wages Don't Fall during a Recession, Harvard University Press, Cambridge, MA, 1999.
[6] P. Bonacich, Power and centrality: a family of measures, American Journal of Sociology 92 (5) (1987).
[7] J. Brown and J. Morgan, Reputation in Online Markets: Some Negative Feedback, Mimeo, University of California, Berkeley (2006).
[8] A. Calvo-Armegnol, Job contact networks, Journal of Economic Theory 115 (1) (2004).
[9] A. Calvo-Armengol, M.O. Jackson, The effects of social networks on employment and Inequality, American Economic Review 94 (3) (2004).
[10] A. Calvo-Armengol, M.O. Jackson, Networks in labor markets: wage and employment dynamics and inequality, Journal of Economic Theory 132 (1) (2007).
[11] A. Calvo-Armengol and M. O. Jackson, Like father, like son: network externalities, parent–child correlation in behavior, and social mobility, Mimeo (2007).
[12] A. Calvó-Armengol , E. Patacchini, and Y. Zenou, Peer effects and social networks in education and crime, Mimeo (2005).
[13] A. Calvó-Armengol, T. Verdier, Y. Zenou, Strong ties and weak ties in employment and crime, Journal of Public Economics 91 (1–2) (2007).
[14] A. Calvó-Armengol, Y. Zenou, Social networks and crime decisions. The role of social structure in facilitating delinquent behavior, International Economic Review 45 (3) (2004).
[15] A. Calvo-Armengol, Y. Zenou, Job matching, social network and word-of-mouth communication, Journal of Urban Economics 57 (500–522) (2005).
[16] L.A. Cohen, A. Frazzini, C. Malloy, The small world of investing: board connections and mutual fund returns, NBER Working Paper, 2007.
[17] C. Conley, T. Urdy, Social learning through networks: the adoption of new agricultural technologies in Ghana, American Journal of Agricultural Economics 83 (3) (2001).
[18] C. Conley, T. Urdy, Social networks in Ghana, Yale Economic Growth Center Working Paper No. 888, 2002.
[19] C. Conley, T. Urdy, Learning About a New Technology: Pineapple in Ghana, Yale Economic Growth Center Working Paper no. 817, 2007.
[20] S. Currarini, M. O. Jackson, and P. Pin, An Economic Model of Friendship: Homophily, Minorities and Segregation, (with Matt Jackson and Paolo Pin), Econometrica (forthcoming).
[21] E. Dufl, E. Saez, Participation and investment decisions in a retirement plan: the influence of colleagues' choices, Journal of Public Economics 85 (1) (2002).
[22] F. Echenique, R. Fryer, A measure of segregation based on social interactions, Quarterly Journal of Economics 122 (2) (2007).
[23] F. Echenique, R. Fryer, A. Kaufman, Is school segregation good or bad? American Economic Review 96 (2) (2006).
[24] G. Fagiolo, J. Reyes, S. Schiavo, On the topological properties of the world trade web: a weighted network analysis, Physica A 387 (15) (2008).
[25] K. Faust, S. Wasserman, Social Networks Analysis: Methods and Applications, Cambridge University Press, Cambridge, UK, 1994.
[26] F. Fontaine, Why are similar workers paid differently? Journal of Economic Dynamics and Control 32 (12) (2008).
[27] G. Foster, It's not your peers, and it's not your friends: some progress toward understanding the educational peer effect mechanism, Journal of Public Economics 90 (8–9) (2006).
[28] A. Foster, M. Rosenzweig, Learning by doing and learning from others: human capital and technical change in agriculture, Journal of Political Economy 103 (6) (1995).
[29] R. Fryer and P. Torelli, An Empirical Analysis of ‘Acting White’, Mimeo, Harvard University (2006).
[30] S. Goyal, M.J. van der Leij, J.L. Moraga-Gonzalez, Economics: an emerging small world, Journal of Political Economy 114 (2) (2006).
[31] S. Goyal, Learning in networks, in: G. Demange, M. Wooders (Eds.), Group Formation in Economics: Networks, Clubs and Coalitions, Cambridge University Press, Cambridge, UK, 2005.
[33] M.S. Granovetter, The strength of weak ties, American Journal of Sociology 78 (6) (1973).
[34] C. Hoxby, Peer effects in the classroom: learning from gender and race variation, NBER Working Paper 7867, 2000.
[35] Y.M. Ioannides, L. Datcher Loury, Job information networks, neighborhood effects and inequality, Journal of Economic Literature 42 (4) (2004).

[36] Y.M. Ioannides, A. Soetevent, Wage and employment in a random social network with arbitrary degree distribution, American Economic Review 96 (2) (2006).
[37] M.O. Jackson, The economics of social networks, advances in economics and econometrics, theory and applications, in: Richard Blundell, Whitney Newey, Torsten Persson (Eds.), Ninth World Congress of the Econometric Society, vol. 1, Cambridge University Press, 2006, Chapter 1.
[38] M.O. Jackson Social Networks in Economics, (2008) forthcoming in the Handbook of Social Economics (edited by Benhabib, Bisin, Jackson).
[39] M.O. Jackson, B.W. Rogers, The economics of small worlds, Journal of the European Economic Association 3 (2–3) (2005).
[40] M.O. Jackson, B.W. Rogers, Meeting strangers and friends of friends: how random are social networks? American Economic Review 97 (3) (2007).
[41] M.O. Jackson, A. Wolinsky, A strategic model of social and economic networks, Journal of Economic Theory 71 (1) (1996).
[42] R. Kali, J. Reyes, F. Mendez, Trade structure and growth, Journal of International Trade and Economic Development 16 (2) (2007).
[43] S. Kakade, M. Kearns, L. Ortiz, R. Pemantle, and S. Suri, The economics of social networks, 2004. Preprint.
[44] A. Kugler, Employee referrals and efficiency wages, Labour Economics 10 (5) (2002).
[45] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, A. Tomkins, Geographic routing in social networks, Proceedings of the National Academy of Sciences 102 (33) (2005) 11623–11628.
[46] C.F. Manski, Identification and endogenous social effects: the reflection problem, Review of Economic Studies 60 (3) (1993).
[47] C.F. Manski, Economic analysis of social interactions, Journal of Economic Perspectives 14 (3) (2000).
[48] A. Mayer, S.L. Puller, The old boy (and girl) network: social network formation on university campuses, Journal of Public Economics 92 (1–2) (2008).
[49] M. McPherson, L. Smith-Lovin, J. Cook, Birds of a feather: homophily in social networks, Annual Review of Sociology 27 (2001).
[50] A. Mislove, M. Marcon, K.P. Gummadi, P. Druschel, B. Bhattacharjee, Measurement and analysis of online social networks, Proceedings of the 5th ACM/USENIX Internet Measurement Conference, (IMC'07), San Diego, CA, 2007.
[51] M. Mobius, P. Niehaus and T. Rosenblat, Social learning and consumer Demand, Mimeo, Harvard University (2005).
[52] M. Mobius, A. Szeidl, “Trust and social collateral”, NBER Working Paper No. W13126, 2007.
[53] J.D. Montgomery, Social networks and labor–market outcomes: toward an economic analysis, American Economic Review 81 (5) (1991).
[54] M. Newman, The structure and function of complex networks, Society for Industrial and Applied Mathematics Review 45 (2003).
[55] G. Oestreicher-Singer and A. Sundararajan, The Visible Hand of Social Networks in Electronic Markets, Mimeo, New York University (2008).
[56] J.H. Oh, A. Susarla, and Y. Tan, Examining the Diffusion of User-Generated Content in Online Social Networks, Mimeo, University of Washington (2008).
[57] C.A. Pissarides, Equilibrium Unemployment Theory. MIT Press, Cambridge, 2000.
[58] J.E. Rauch, Business and social networks in international trade, Journal of Economic Literature 39 (4) (2001).
[59] A. Rees, Labor economics: effects of more knowledge. Information in labor markets, American Economic Review Papers Proceedings 56 (2) (1966).
[60] R. Rogerson, R. Shimer, R. Wright, Search-theoretic models of the labor market: a survey, Journal of Economic Literature 48 (4) (2005).
[61] T. Rosenblat, M. Mobius, Getting closer or drifting apart? Quarterly Journal of Economics 119 (3) (2004).
[62] B. Sacerdote, Peer effects with random assignment: results for Dartmouth roommates, Quarterly Journal of Economics 116 (2) (2001).
[63] B. Sacerdote, Mararmos, How do friendships form? The Quarterly Journal of Economics 121 (1) (2006).
[64] M.A. Serrano, M. Boguna, A. Vespignani, Patterns of dominant flows in the world trade web, Journal of Economic Interaction and Coordination 2 (2) (2007).
[65] G. Topa, Social interactions, local spillovers and unemployment, Review of Economic Studies 68 (2) (2001).
[66] B. Weinberg, Social Interactions and Endogenous Association, Mimeo, Ohio State University (2007).
[67] A. Mayer, Quantifying the Effects of Job Matching through Social Networks, Mimeo, Texas A&M University (2008).
[68] R. Putnam, Bowling Alone: America's Declining Social Capital, Journal of Democracy (1995).

Adalbert Mayer is an Assistant Professor at the Department of Economics at Texas A&M University in College Station, TX. He graduated with a Diploma-degree from the University of Augsburg, Germany, in 1998. He received a Ph.D. degree in economics from the University of Rochester in 2003. His research interests are in the field of labor economics. The topics he has investigated in his research include the intergenerational persistence of income and education, the economics of education, and – more recently – the role of social networks in education and in the labor market. His research has been published in Applied Economics, the Journal of Public Economics, and the Journal of Human Capital.