Chapter 2

SOCIAL NETWORK ANALYSIS AND GAME THEORY: BASIC CONCEPTS AND ASSUMPTIONS

I use social network analysis and game theory in the models developed in this book. It is useful to discuss the basic concepts associated with these two "tools" before I begin. Readers familiar with social network analysis should check the formal definitions of a social network and the related network parameters in Section 2.1, because these definitions are somewhat different than those used in most network studies. Section 2.2 has nothing new for readers familiar with game theory, although some readers might be interested in the specific results for Trust Games presented in this section.

Social networks analysis is rich in conceptualization. Wasserman and Faust (1994) offer an extensive overview of concepts and operationalizations. However, social network analysis lacks testable implications (see Granovetter 1979). There are few theoretical predictions about what positions in a network can be expected to have what effect on a dependent variable. And, if phenomena can be explained by the structural aspects of a network, the arguments underlying the explanation are often rather informal and open to many theoretical counterarguments. For example, in the theory of structural holes (Burt 1992), the concept of a structural hole is defined in mathematical detail. Furthermore, the association between an actor who is "rich in structural holes" and the actor's performance is empirically convincing. However, the theoretical arguments underlying the association remain informal. As a consequence, it is hard to argue that network properties rather than personal characteristics that correlate with structural holes drive the performance (see Burt, Jannotta, and Mahoney 1998). In Section 2.1, I introduce the network parameters used in this book. These parameters summarize core structural properties of social networks. After introducing the network parameters, I present informal arguments about the effects they have on information communication processes. As discussed in Chapter 1, the effects of social networks on trust depend, at least in my theory, primarily on how fast trustors in a network can transmit information to other trustors and how quickly they receive information. Expectations on network effects based on my informal arguments will be referred to as "conjectures." In Chapters 3 and 4, I develop theoretical models from which I formally derive the effects of network parameters on trust. Conjectures that do turn out to be indeed logical consequences of the models will then be referred to as "hypotheses."

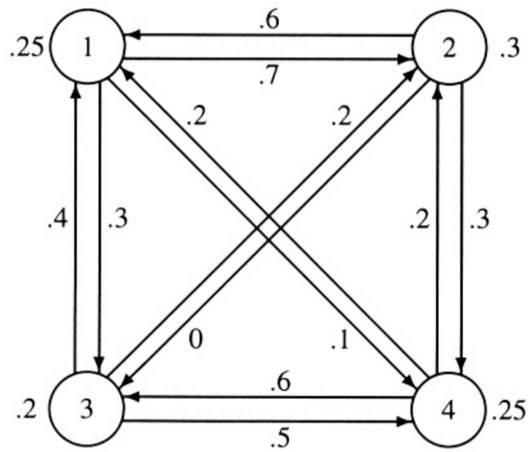
Game theory is a behavioral theory that is rich in implications. The theory assumes that actors are utility maximizers, and that actors decide upon their behavior taking into account that other actors are utility maximizers as well. Therefore, using actors' utilities for different outcomes in a game, it is possible to make predictions about actors' behavior. Chapter 3 contributes to an integration of social network analysis and game theory by providing a game-theoretic model of trust in a network of actors. The core concepts and basic assumptions of game theory needed to develop this model are introduced in Section 2.2 and models on trust are used as examples. Obviously, the emergence and maintenance of trust among rational actors will not only depend on network parameters but also on payoffs associated with transactions and on the temporal embeddedness of transactions between a trustor and trustee, for example. Section 2.2 provides on overview of related hypotheses that follow from game-theoretic analyses.

2.1 SOCIAL NETWORK ANALYSIS

I wish to explain the effects of social networks on the extent to which trustors can trust a trustee. Because I try to explain trust through learning and control effects, it is essential to know how fast trustors receive and transmit information in the network. Another element I want to include is the "importance" of a given trustor for the trustee. For instance, when a trustor who is involved in half of the trustee's transactions no longer trusts the trustee, this will be more problematic for the trustee than when a trustor who is involved in a much smaller proportion of transactions no longer trusts him. Consequently, the sanctions of a more important trustor can be more severe for the trustee than the sanctions of a less important trustor.[1]

To include these elements, social networks are conceived as *valued directed graphs with weighted nodes.* Figure 2.1 gives an example of such a network. Valued directed graphs are commonly used to represent (finite) networks (see, for example, Harary, Norman, and Cartwright 1965; Wasserman and Faust 1994). A *graph* is a set of nodes and ties between these nodes. If the set of nodes consists of n elements, the ties can be represented by an $n \times n$ matrix

[1] The generalization that trustors may be of varying importance is included in the definitions of network parameters and in the model in Chapter 3. The implications of Chapters 3 and 4 are based on networks in which all trustors are equally important.



$$\boldsymbol{\pi} = \begin{pmatrix} .25 \\ .3 \\ .2 \\ .25 \end{pmatrix}$$

$$\boldsymbol{A} = \begin{pmatrix} 1 & .7 & .3 & .1 \\ .6 & 1 & 0 & .3 \\ .4 & .2 & 1 & .5 \\ .2 & .2 & .6 & 1 \end{pmatrix}$$

Figure 2.1. Example of a network of trustors.

$\boldsymbol{A}$. In a *discrete* graph, the matrix $\boldsymbol{A}$ indicates only whether certain ties exist, i.e., the elements α_{ij} of $\boldsymbol{A}$ are 0 or 1. In a *directed* graph, the ties are directed from one node to another. This implies that the matrix $\boldsymbol{A}$ is not necessarily symmetric. In a *valued* graph, values are added to the ties to indicate, for example, the importance of a tie. If $\alpha_{ij} > 0$, a tie exists from actor i to actor j. I add *weights* to the nodes (actors) of the network, indicating the importance of a node (actor). I use the term "discrete network" for a network described by a discrete graph. Similarly, I use the terms "valued networks" and "valued networks with weighted nodes."

The formal notation of a valued network with weighted nodes representing n actors is a pair $(\boldsymbol{\pi}, \boldsymbol{A})$, where $\boldsymbol{\pi}$ is an n-vector of weights indicating the importance of each actor with $\sum_{i=1}^{n} \pi_i = 1$, $\pi_i > 0$ for all i, and $\boldsymbol{A}$ is an $n \times n$-matrix where α_{ij}, $0 \leq \alpha_{ij} \leq 1$, is the importance of the tie from actor i to actor j.[2] By definition, $\alpha_{ii} = 1$ for $i = 1, \ldots, n$.[3]

A network is called *homogeneous* if $\alpha_{ij} = \alpha$ for all $i \neq j$. In this study, I am interested in heterogeneous networks and in the effects of heterogeneity on trust. In heterogeneous networks, some ties are stronger than others. Ties can be strong within subgroups of a network and weak between subgroups of a network. Networks might resemble a chain, a circle, or a star. All these

[2] I use boldface for vectors (lowercase) and matrices (uppercase).

[3] The same representation is directly applicable to stochastic blockmodels (see, for example, Wasserman and Faust 1994), where $\boldsymbol{\pi}$ indicates the percentage of actors in a block, while the matrix $\boldsymbol{A}$ represents the percentage of ties present between the blocks. Here, the diagonal elements should represent the percentage of ties within blocks and, hence, need not be constrained to 1.

properties might hinder or facilitate information diffusion. Network parameters such as the centrality of an actor in a network (see Freeman 1979) measure certain aspects of heterogeneity. Network parameters are usually defined for discrete networks in the literature, although sometimes informal indications are given about generalizations of network parameters for valued networks. In this book, network parameters are formally defined for valued networks with weighted nodes.

Although the list of network parameters below might seem a rather ad hoc selection, they cover the main structural properties that are frequently discussed in social network literature. In Chapter 1, I explained that I want to study network properties at the individual and global level. Therefore, I have differentiated between individual and global network parameters. Individual network parameters measure the properties of an actor within a network and can explain "within-network" effects. Global network parameters measure the properties of a network as a whole and can explain "between-network" effects. On the individual level, I focus on three properties:

- the extent to which an individual actor is connected to other actors in the network (outdegree and indegree),
- the extent to which the neighbors of an actor are connected to other actors (degree quality), and
- the extent to which neighbors of the focal actors are mutually connected (local density).

On the global level, four properties are distinguished:

- the density of the network,
- the extent to which the network is centralized around one or a few actors (outdegree variance, indegree variance, outdegree-indegree covariance),
- the transitivity of the network, i.e., a measure for the existence of dense subgroups with limited connectivity among these subgroups, and
- the number of actors in the network (network size).

I also describe intuitive expectations about the effects of the network parameters on information *transmission* and *reception rates,* i.e., how fast information is transmitted and received by actors in the network. I start with the individual network parameters.

OUTDEGREE

Outdegree is a parameter for the extent to which an actor in a network communicates information to other actors. For discrete networks, outdegree is defined as the number of ties an actor has divided by the maximal number of ties that are possible for an actor (Freeman 1979). For example, if there are five actors in a network and a focal actor has ties to three of the four other actors, the focal actor's outdegree is $\frac{3}{4}$. For valued networks, outdegree is defined as the sum of the values of the ties starting from the focal actor, divided by the highest possible value. Generalizing outdegree for valued networks with weighted nodes, the values of the ties are weighted by the importance of the actors at the other end of these ties. In this way, an actor who is connected to more important other actors has a higher outdegree than an actor who is connected to the same extent, but with less important others. Because actors cannot have ties to themselves, standardization is used such that the outdegree is 1 if an actor is "perfectly" connected to all other actors.[4] Formally, the outdegree for actor i is defined as

$$D_{out}(i) = \frac{1}{1-\pi_i} \sum_{j \neq i} \pi_j \alpha_{ij}. \tag{2.1}$$

Because an actor with a higher outdegree transmits information more often to other actors, I conjecture that an actor with a higher outdegree will transmit information faster to other actors than an actor with a lower outdegree. On the other hand, outdegree does not have an influence on the extent to which an actor receives information. Therefore, I do not expect an effect of outdegree on the time needed for an actor to receive information.

INDEGREE

Indegree is a parameter for the extent to which an actor receives information from other actors. The indegree is the number of incoming ties to the focal actor relative to its maximal value. Indegree is also called *degree prestige* (Freeman 1979). Incorporating weighted nodes, indegree can be defined here analogously to outdegree as

$$D_{in}(i) = \frac{1}{1-\pi_i} \sum_{j \neq i} \pi_j \alpha_{ji}. \tag{2.2}$$

Indegree and outdegree are identical for *symmetric* networks, i.e., if $\alpha_{ij} = \alpha_{ji}$ for all i and j.

If the indegree of an actor is higher, she will receive information more frequently from other actors and, therefore, I expect she will receive information sooner. The indegree of an actor does not influence the transmission of infor-

[4]The weights are chosen in such a way for all network parameters that these parameters are directly generalizable for stochastic blockmodels. The necessary adjustments for stochastic blockmodels involve the inclusion of ties *within* blocks (see Buskens 1998).

mation to other actors. Therefore, no effect is expected from indegree on how fast an actor transmits information.

DEGREE QUALITY

Outdegree and indegree are strictly local parameters of network position. These parameters could be made "less local" by studying the (weighted) proportion of actors an actor can reach in two, three, or more steps. *Outdegree quality* and *indegree quality* measure the extent to which actors can reach others or can be reached by others in two steps. Thus, these parameters indicate the extent to which an actor is linked to actors who have high degrees themselves. For example, if an actor communicates with only one actor who, however, communicates with all actors in the network, the first actor might still be able to transmit information relatively fast. An actor who receives information from only one other actor will receive information slower if this other actor receives information from only one actor. This is in contrast with the situation where this other actor is informed by a large number of third actors. The parameters outdegree quality $Q_{out}(i)$ and indegree quality $Q_{in}(i)$ are formally defined as the weighted covariance between the value of a tie to (from) an actor and the outdegree (indegree) of that actor.[5] Ties are weighted such that ties to (from) intermediate actors who are more important obtain a larger weight. Formally,

$$Q_{out}(i)=\frac{1}{1-\pi_i}\sum_{j\neq i}\pi_j\left(\alpha_{ij}-D_{out}(i)\right)\left(D_{out}(j)-\frac{1}{1-\pi_i}\sum_{k\neq i}\pi_k D_{out}(k)\right)$$

and

$$Q_{in}(i)=\frac{1}{1-\pi_i}\sum_{j\neq i}\pi_j\left(\alpha_{ji}-D_{in}(i)\right)\left(D_{in}(j)-\frac{1}{1-\pi_i}\sum_{k\neq i}\pi_k D_{in}(k)\right). \quad (2.3)$$

Thus, an actor has high outdegree quality if outgoing ties are toward actors with high outdegrees. Rogers (1995: 289) finds that actors search information from opinion leaders or those with high status who are expected to have higher outdegrees or indegrees. A positive effect of indegree quality on the diffusion of information might indicate that these actors use a "rational" strategy when choosing their ties.

As outdegree quality and indegree quality are extensions of outdegree and indegree, respectively, the conjectures for these parameters are the same as for

[5]In an earlier paper (Buskens 1998), these parameters are called *individual outdegree centralization* and *individual indegree centralization,* but this terminology is confusing because centralization usually refers to global network parameters.

outdegree and indegree. I conjecture that higher outdegree quality facilitates a rapid transmission of information through the network and that outdegree quality does not have an effect on how quickly an actor receives information. Moreover, I expect that the time information needs to reach an actor will decrease with indegree quality, and indegree quality will not have an effect on how fast an actor transmits information in a network.

LOCAL DENSITY

The following individual network parameters are *local outdegree density* and *local indegree density.* Local density measures the extent to which an actor's contacts have contacts among themselves. It is expected that information is transmitted more slowly through the network if an actor informs two actors who have also frequent contacts among themselves than if the actor informs two actors who never have contacts with each other, because the probability is relatively high that these actors obtain the same information from multiple sources (Granovetter 1973).[6] Again, outdegree and indegree versions of this parameter are distinguished. The local outdegree density $LD_{out}(i)$ of actor i measures the extent to which actor i transmits information to connected neighbors. Local indegree density $LD_{in}(i)$ of actor i measures the extent to which actor i obtains information from connected neighbors. A tie is weighted with the product of the importance of the receiver and the transmitter, because the contribution of such a tie to the local density depends on the importance of the receiver as well as the importance of the transmitter.[7] Formally, the parameters are defined as

$$LD_{out}(i) = \frac{\sum_{j \neq i} \sum_{k \neq i, k \neq j} \pi_j \pi_k \alpha_{ij} \alpha_{ik} \alpha_{jk}}{\sum_{j \neq i} \sum_{k \neq i, k \neq j} \pi_j \pi_k \alpha_{ij} \alpha_{ik}}, \tag{2.4}$$

and

$$LD_{in}(i) = \frac{\sum_{j \neq i} \sum_{k \neq i, k \neq j} \pi_j \pi_k \alpha_{ji} \alpha_{ki} \alpha_{jk}}{\sum_{j \neq i} \sum_{k \neq i, k \neq j} \pi_j \pi_k \alpha_{ji} \alpha_{ki}}. \tag{2.5}$$

The local density parameters described above are only defined if actor i has at least two ties to other actors. Local density is defined as 0 if an actor has at most one tie.

Because a higher local outdegree density makes redundant information transmission more likely, I expect that a high local density will inhibit fast trans-

[6]This argument is similar to the redundancy argument in Burt's theory on structural holes. See Burt (1992: Chapter 1) for precise definitions and examples. Local degree density resembles the opposite of what Burt defines as the "effective size" of a network for an actor (1992:52), but it is not a straightforward generalization.

[7]Other forms such as the sum of the importance of the two actors connected by a tie are possible. One argument in favor of the product is that for stochastic blockmodels the proportion of actual ties from one block to another equals the product of the proportions of the actors in these blocks times the density of ties between the two blocks.

mission of information in a social network when density has been controlled. Therefore, an actor with a higher local outdegree density will transmit information more slowly through a network than an actor with a lower local outdegree density. No effect of local outdegree density is expected on the time information needs to reach a focal actor. For local indegree density the argument is somewhat different. If an actor receives information from other actors who frequently receive information from each other, information that has reached the neighborhood of the focal actor is also quickly received by many neighbors and, therefore, can be expected to be known fairly soon by the focal actor. Consequently, I conjecture that an actor with a higher local indegree density will receive information sooner than an actor with a lower local indegree density. I expect the local indegree density will not have an effect on the speed with which an actor transmits information through the network.

DENSITY

Density is the first global network parameter to be discussed. In discrete networks, the density of a network is defined as the number of ties divided by the number of possible ties. In valued networks, the density (Δ) is the sum of all values of the ties in the network divided by the sum of all maximal values of the ties. Ties are weighted with the product of the importance of the two actors adjacent to the tie. Formally,

$$\Delta = \frac{\sum_{i=1}^{n} \sum_{j \neq i} \pi_i \pi_j \alpha_{ij}}{\sum_{i=1}^{n} \sum_{j \neq i} \pi_i \pi_j}. \tag{2.6}$$

This corresponds with the average of all outdegrees or indegrees only if $\pi_i = \frac{1}{n}$ for all i. Then,

$$\Delta = \frac{1}{n} \sum_{i=1}^{n} D_{out}(i) = \frac{1}{n} \sum_{i=1}^{n} D_{in}(i). \tag{2.7}$$

I expect that information will be communicated faster through networks with a higher density than through networks with a lower density. Therefore, I conjecture that, on average, actors will transmit and receive information faster in a dense network than in a sparse network.

CENTRALIZATION

Global centralization parameters are always related to individual centrality parameters. Since I used outdegree and indegree as centrality parameters earlier, I discuss centralization parameters based on outdegree and indegree. *Centralization* of a network measures the differences in outdegrees or indegrees of the actors in a network. A suitable parameter for measuring this difference is the variance of actor degrees: degree variance (Snijders 1981). I define *outdegree* *variance* and *indegree variance* as the average variance in the outdegrees and indegrees of the actors in the network, with each term weighted by the importance of the actor. Note that the average outdegrees and indegrees could have been replaced by density if all actors are equally important. Formally, the outdegree variance is defined as

$$V_{out} = \sum_{i=1}^{n} \pi_i \left(D_{out}(i) - \sum_{j=1}^{n} \pi_j D_{out}(j) \right)^2 , \tag{2.8}$$

and, similarly, the indegree variance is defined as

$$V_{in} = \sum_{i=1}^{n} \pi_i \left(D_{in}(i) - \sum_{j=1}^{n} \pi_j D_{in}(j) \right)^2 . \tag{2.9}$$

It can be expected that centralization accelerates information diffusion if the central actors are also the important actors in the diffusion process. In an earlier paper (Buskens and Weesie 2000a) on the model proposed in Chapter 3, I derived the result that trustors trust a trustee to a relatively large extent in a network where trustors who receive a large amount of information (have a high indegree) are the same trustors as those who transmit a large amount of information (have a high outdegree). I try to cover this aspect with a third network centralization measure: *outdegree-indegree covariance.* This network parameter is defined as

$$V_{out,in} = \sum_{i=1}^{n} \pi_i \left(D_{out}(i) - \sum_{j=1}^{n} \pi_j D_{out}(j) \right) \left(D_{in}(i) - \sum_{j=1}^{n} \pi_j D_{in}(j) \right) . \tag{2.10}$$

Other centralization parameters are proposed in the literature. Freeman (1979), for example, compares degrees in a network with the maximal degree in the network to obtain a centralization parameter. Freeman's parameter correlates highly with the centralization parameters defined above. I use degree variance instead of Freeman's parameter, because Freeman's parameter is too strongly related to the actor with the maximal degree. Yamaguchi (1994a, 1994b; Buskens and Yamaguchi 1999) proposes the "coefficient of variation" in degrees as a centralization parameter. This parameter is defined as the square-root of the variance in degrees divided by the average degree or density of the network and could be used for outdegrees as well as indegrees. Because the coefficient of variation correlates more with density, outdegree, and indegree than outdegree variance and indegree variance, it is not used in the following chapters. The higher the correlations among network parameters, the more problematic it becomes to disentangle the effects of the different parameters.

There are no clear conjectures for centralization parameters. If actors with a high indegree also have a high outdegree, central actors will receive *and* transmit information at a high rate. In such a network, I expect information will be transmitted and received faster by all trustors. Thus, actors in a networks with higher outdegree-indegree covariance will transmit and receive information on average more quickly than actors in a network with a lower outdegree-indegree covariance. Controlling for outdegree-indegree covariance, I expect that outdegree variance and indegree variance will inhibit the transmission of information, as it seems to be inefficient if actors who obtain information frequently, do not communicate this information to others or if actors who are able to transmit information to many others receive very little information. Thus, I conjecture that the information transmission rate and the information reception rate decrease with outdegree variance and indegree variance.

Transitivity

To define transitivity, I recall some definitions. A *triad* consists of three actors (i, j, k) in a network. A discrete network is called *transitive* if the existence of a tie from actor i to actor j and from actor j to actor k implies that there is a tie from actor i to actor k. A discrete network is called *intransitive* if it is not transitive; thus, intransitivity implies that there is a triad (i, j, k) such that ties from actor i to actor j and from actor j to actor k exist, while a tie from actor i to actor k is absent. I define the extent to which a triad (i, j, k) is transitive as $\alpha_{ij}\alpha_{ik}\alpha_{jk}$ for a valued network. To obtain network *transitivity,* the transitivity of each triad is summed over all triads and weighted by the product of importance of the three actors involved, divided by the maximal possible value on the basis of all pairs of ties of the actors:

$$Tr = \frac{\sum_{i=1}^{n}\sum_{j\neq i}\sum_{k\neq i,k\neq j}\pi_i\pi_j\pi_k\alpha_{ij}\alpha_{ik}\alpha_{jk}}{\sum_{i=1}^{n}\sum_{j\neq i}\sum_{k\neq i,k\neq j}\pi_i\pi_j\pi_k\alpha_{ij}\alpha_{ik}}. \tag{2.11}$$

Transitivity as described above is defined only if at least one actor has two or more ties. Transitivity is defined as 0 if all actor have at most one tie.

Transitivity is related to density. For example, if all α_{ij} are equal to 1, transitivity as well as density are equal to 1. If density is low, transitivity measures the extent to which ties are concentrated in subgroups of actors within the network. I expect that high transitivity slows down information diffusion through the network, because information can be caught within subgroups of actors. Information will need a relatively long time to reach other actors in the network. This argument corresponds with Granovetter's (1973:1374) statement that trusting a leader is more problematic for the followers if the network of followers is fragmented. Therefore, I conjecture that transitivity has a negative effect on the time needed for information to reach actors in the network as well as on the time actors need to spread information in the network.

Table 2.1. Conjectures on the effects of network parameters on information diffusion rates.

	Transmission rate	Reception rate
INDIVIDUAL NETWORK PARAMETERS		
Outdegree	+	0
Indegree	0	+
Outdegree quality	+	0
Indegree quality	0	+
Local outdegree density	−	0
Local indegree density	0	+
GLOBAL NETWORK PARAMETERS		
Density	+	+
Outdegree variance	−	−
Indegree variance	−	−
Outdegree-indegree covariance	+	+
Transitivity	−	−
Network size	?	?

NETWORK SIZE

The last network parameter, which does not need extensive explanation, is *network size* n. This is simply the number of actors in the network. Because I generally assume that $\pi_i = \frac{1}{n}$ for all i, the number of actors in the network can certainly have effects on information diffusion. In larger networks, more time may be needed to inform the same proportion of actors than in smaller networks. The effects of network size are important as far as generalizations of results to large networks are concerned. If network size has a considerable additive effect on information diffusion, generalizing results from small networks to large networks will not be straightforward. In the simulations, I only consider networks with between two and ten actors. Therefore, if the effects of network parameters depend on the size of the network, it is more difficult to predict what the effects of these networks will be for networks with more than ten actors.

THE CONJECTURES SUMMARIZED

In conclusion, Table 2.1 gives an overview of the conjectures relating to the effects of network parameters on the rate of information transmission and reception amongst actors in a network. The relationship with conjectures relating to trust is straightforward. Trustors who transmit information more quickly are expected to place more trust in the trustee because they have more extensive control opportunities. Trustors who receive information more quickly place more trust in the trustee if the information is positive and place less trust if the information is negative.

2.2 GAME THEORY

[R]epeated games may be a good approximation of some long-term relationships ... — particularly those where "trust" and "social pressure" are important, such as when informal agreements are used to enforce mutually beneficial trades without legally enforced contracts.

—Fudenberg and Tirole (1991: 145)

Game theory allows predictions to be made about behavior of actors in interdependent choice situations. In Subsection 1.2.1, I argued that the trustor would not place trust in the Trust Game. This, in fact, is the game-theoretic solution of a one-shot Trust Game that is played without any connection being made to past or future transactions. In general, I define the *solution* of a game as the outcome that results from a combination of strategies played by rational actors.[8] In this section, I will explain how such a solution can be found.

Some concepts are needed for this explanation.[9] The Trust Game in Figure 1.1 is a representation of the game in extensive form. The *extensive form* is a configuration of nodes and branches, without any loops and originating from a single node. This is often referred to as the *game tree.* At each node there is a description as to which actor has to *move.* In my case, actors are the trustor or the trustee. A *move* made by an actor is a choice he or she makes at a given node. In the Trust Game, the trustor can choose between the moves "placing trust" and "not placing trust." An *end node* is a node with no outgoing move. At every end node, the payoffs for the actors are specified. I added labels at the branches of the game tree to indicate the interpretation of the moves.

Before or during a game, events might occur that are not under the control of the actors in the game, but which nevertheless influence the course of the game. These events can be included in the extensive form of the game by using "chance" moves. A chance move is said to be played by *Nature.* For example, the value of T_2 in the Trust Game may be unknown before the game begins, but might be chosen from a given distribution at the start of the game. This would be indicated by an additional node at the top of the game tree (see Figure 2.2). If Nature chooses between two or three values, each of these choices will be indicated by a different branch giving the probabilities of each choice at each branch, respectively. This case is illustrated in Figure 2.2, where Nature chooses $T_{2,1}$ with probability 0.6 and $T_{2,2}$ with probability 0.4. In Chapter 3, a game is analyzed in which payoffs are chosen from a continuous distribution. This will be indicated by specifying the distribution from which Nature "chooses."

---

[8]See Harsanyi's solution concept that consists of solution payoffs and a set of joint strategies that realize these payoffs (1977: 135-137).

[9]For a general introduction to game theory see Rasmusen (1994), for example.

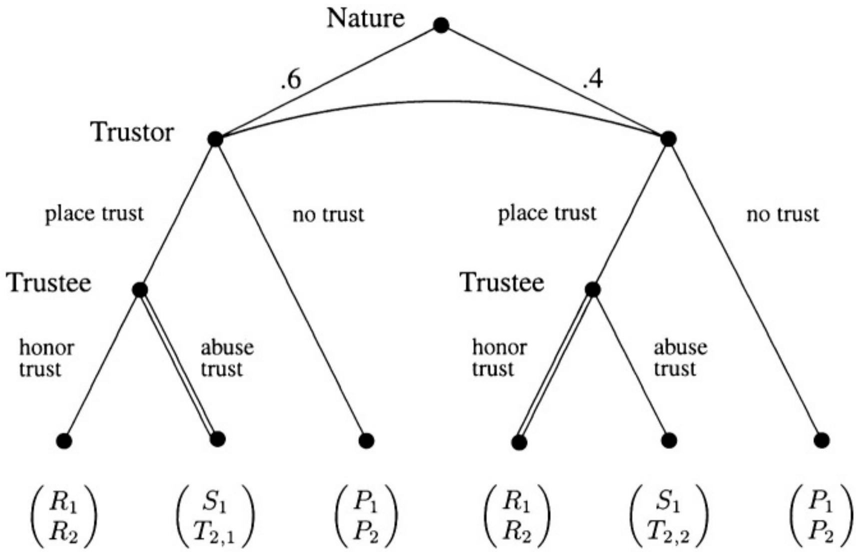


Figure 2.2. Extensive form of a Trust Game with incomplete information Γ^I, where $R_i > P_i, (i = 1, 2), P_1 > S_1, T_{2,1} > R_2$, and $T_{2,2} < R_2$.

Moves in a game may or may not be observed by actors in a game before the actors concerned have to move themselves. Information sets in the extensive form of a game indicate whether an actor is able to observe a certain move. The actor's *information set* is a set of nodes among which the actor is not able to distinguish this by direct observation (see Rasmusen 1994: 40). In Figure 2.2, the line between the two nodes in which the trustor has to move indicates that they belong to the same information set for the trustor. This line means that the trustor could not observe the initial move by Nature, and she does not know which of the two nodes the game reached at the moment she has to move. An information set that consists of one single node is called a *singleton.* In analyzing games, there is an implicit assumption that the actors know the game tree. It is also assumed that the actors know that other actors know the game tree. The term *common knowledge* is used for information of which everybody is aware, and everybody knows that everybody knows and so on.

A *strategy* of an actor is a rule that prescribes the moves of an actor for each of his or her information set in the game. For example, a strategy for the trustee can specify that he will abuse trust if trust is placed by the trustor.

A *Nash equilibrium* is a combination of actor strategies such that no actor can obtain a higher payoff by changing to another strategy under the assumption that the other actors do not change their strategies (Nash 1951). In the Trust Game, the strategy for the trustor "never place trust" and the strategy for the trustee "always abuse trust if trust is placed" form a Nash equilibrium. The reason is that given that the trustee will abuse trust, the trustor prefers not to place trust because $S_1 < P_1$. And, given that the trustor never places trust, the trustee will obtain P_2 whatever his strategy might be. Thus, he cannot improve his payoff by changing to another strategy. A Nash equilibrium is the minimal requirement for a combination of strategies to be a candidate for the solution in a game. A combination of strategies that does not form a Nash equilibrium cannot be part of the solution, because if one player can obtain a higher payoff by changing to another strategy, he or she is expected to do so.

In many games, however, there are multiple Nash equilibria that lead to different outcomes of the game. In such cases, I need arguments as to why rational actors prefer some equilibria to others. Ideally, I would like to have one outcome that could be selected as the solution. Assume that $\frac{R_1+S_1}{2} < P_1$ in the Trust Game. Then, the strategies "never place trust" of the trustor and "choose with equal probability between abuse and honor trust if trust is placed" of the trustee also form a Nash equilibrium. This is a Nash equilibrium because the trustor is still better off not placing trust and the trustee cannot improve his payoff by changing to another strategy. The reason why this equilibrium is not considered realistic is that if trust were in fact placed, the trustee's best option would be to abuse trust. Below, it will be argued that the strategies "not placing trust" by the trustor and "always abusing trust if trust is placed" form the most "reasonable" Nash equilibrium for the Trust Game, and therefore the outcome "trust is not placed by the trustor" can be considered as the solution. Before I formalize some equilibrium selection arguments, I will introduce some additional concepts concerning information in a game.

INFORMATION IN GAMES

Rasmusen (1994: Section 2.3) presents a suitable classification for information in a game. A game with *perfect information* is a game in which all nodes are singletons. Where this is not the case, a game is one with *imperfect information.* Thus, in a game with perfect information, actors observe all moves by Nature and by other actors that occur before they have to move. Moreover, actors never move simultaneously in a game with perfect information. The Trust Game is an example of a game with perfect information. A game of *certainty* is a game in which Nature never moves after an actor has moved. A game of uncertainty can still be a game with perfect information if the moves by Nature are observed immediately by all the actors. An example of such a game is analyzed in Chapter 3.

Two special types of games with imperfect information are games with asymmetric and games with incomplete information. In a game with *incomplete information,* Nature moves first and this move cannot be observed by at least one of the actors. In a game with *symmetric information,* an actor's information set at any node where an actor chooses an action or at an end node contains at least the same elements as the information sets of every other actor. An example of a game with asymmetric information is a game in which Nature moves first and this move is not observed by all actors. The Trust Game with incomplete information Γ^I in Figure 2.2 is an example of a game with incomplete and asymmetric information. If the trustee would also be unable to observe the move by Nature, the game would be one with symmetric information.

EQUILIBRIUM SELECTION

The *equilibrium path* is the path through the game tree that is followed in equilibrium. However, also the responses at the nodes that are never reached in equilibrium have to be specified in the (equilibrium) strategies. Above, I have mentioned two possible strategies followed by the trustee that prescribe what the trustee does after the trustor places trust although that move of the trustee is not part of equilibrium play. The *perfectness* of an equilibrium is related to whether a strategy of an actor is still optimal on paths away from the equilibrium path (Selten 1965).

In order to provide a formal definition of a *subgame-perfect equilibrium,* I first define a *subgame.* A *subgame* is a game that starts at a node that is a singleton for every actor and includes all the branches and nodes that follow this singleton. In the Trust Game, the game starting at the node where the trustee moves is a subgame; the whole Trust Game is a subgame of itself. A combination of strategies form a *subgame-perfect equilibrium* if it is a Nash equilibrium for the entire game and the induced strategies form a Nash equilibrium for all subgames. The double lines in Figure 1.1 indicate equilibrium play in the two relevant subgames. In general, one can find a subgame-perfect equilibrium in a finite game with perfect information by *backward induction:* start at the end of the tree, and find the equilibrium paths from the last moves in the tree; given these moves at the end of the tree, find the equilibrium paths from the last but one moves, and continue this procedure up to the top of the tree. Note that this procedures does not work in Figure 2.2 because the trustor's move does not start a subgame. The information set in which the trustor has to move consists of two nodes.

Because the only Nash equilibrium for the subgame that starts at the trustee's node prescribes abusing trust by the trustee, the only subgame-perfect equilibrium is the equilibrium in which the trustor never places trust and the trustee always abuses trust if trust is placed. I will use the subgame-perfect equilibrium concept to analyze the game in Chapter 3. In games with imperfect information, the concept of a subgame-perfect equilibrium is not always useful because there are no subgames with the exception of the whole game tree. Adjusted concepts for perfect equilibria are *perfect Bayesian equilibria* and *sequential equilibria* (see, for example, Rasmusen 1994: Chapter 6). I will not use these concepts in this book.

Another important equilibrium selection criterion is *payoff dominance.* In dynamic games that are more complex than the Trust Game, such as the repeated Trust Game, there will usually be a large number of subgame-perfect equilibria. In such situations, the set of equilibria can often be restricted by only considering payoff dominant equilibria. A payoff dominant equilibrium (Harsanyi and Selten 1988: 80–81) is an equilibrium for which no other equilibrium exist that is a Pareto improvement, i.e., an equilibrium for which at least one actor is better off and the others are not worse off. Thus, if there is an equilibrium for which the actors obtain a payoff 2, and another for which they obtain 4, the latter equilibrium is a Pareto improvement of the first. The outcome for which all players obtain a payoff 4 is selected as the solution on the basis of the payoff dominance criterion. If there would be two equilibria for which the payoffs are 2 for actor 1 and 4 for actor 2 in one equilibrium and the payoffs are reversed in the second equilibrium, I cannot select one of the equilibria with the payoff dominance criterion, because these equilibria cannot be compared with the Pareto criterion: in one equilibrium one actor obtains more and in the other equilibrium the other actor obtains more.

Although equilibrium selection is a rapidly expanding research field, a generally accepted selection theory that guarantees a unique solution in the games analyzed below is still lacking.[10] In this book, I will apply subgame perfectness and payoff dominance for equilibrium selection.

Repeated Games

"Repeated" or "iterated" games are games in which actors have to make the same choices repeatedly. This allows them to take into account what has happened in previous periods of play and to anticipate on future play (Fudenberg and Tirole 1991: Chapter 5; Gibbons 2001). In Chapter 1, I argued that actors are hardly ever involved in isolated transactions, but that they are embedded in a social context that evolves over time. Temporal embeddedness can be modeled by assuming that a trustor and trustee do not only play one Trust Game. Instead, after each period, they play another Trust Game with probability $1 - \delta$. The parameter δ, $0 < \delta \leq 1$, is the *drop-out rate* of the trustor. In this book, it is assumed that a trustee continues to have transactions for ever, but that the trustor has an exogenously given probability of stopping a series of transactions with the trustee.[11] In this repeated game, the Trust Game Γ shown in Figure 1.1

[10]Other equilibrium refinements, such as renegotiation proofness and properness, are proposed and criticized (Van Damme 1987; Harsanyi 1995; Norde, Potters, Reijnierse, and Vermeulen 1996), but it is not necessary to discuss them here.

[11]Note that this probability does not depend on the behavior of the trustee.

is called the "constituent game." Before the repeated game can be analyzed, the payoffs of the game have to be defined. In every period, actor i obtains a payoff u_{it} related to the outcome of that period of play. Here, $t = 0, 1, 2, \ldots$ is a discrete time parameter, and $i = 1, 2$ with 1 indicating actor 1 or the trustor and 2 indicating actor 2 or the trustee. In all periods in which the actors do not play they obtain a payoff $u_{it} = 0$. I will use standard exponential discounting with discount factor w_i, $0 < w_i < 1$, to obtain the accumulated payoff over the whole game:

$$u_i = \sum_{t=0}^{\infty} w_i^t u_{it}, \tag{2.12}$$

where w_i denotes the pure time preferences of actor i and u_i the utility for actor i.

The game introduced above is the Iterated Trust Game ITG$(\Gamma, w_1, w_2, \delta)$. It is assumed that the constituent game Γ, the discount factors, and the drop-out rate are common knowledge. Folk theorems (see, for example, Abreu 1988; Kreps 1990b: Chapter 14; Fudenberg and Tirole 1991: Chapter 5; Rasmusen 1994: 124) show that in a repeated game such as the ITG, a very large number of subgame-perfect equilibria may exist. Unconditional play of the one-shot equilibrium is always an equilibrium in the repeated game. In the Trust Game, this means that the trustor never places trust and the trustee always abuses trust if trust is placed. This is an equilibrium because the trustor never has an incentive to place trust if trust is always abused, and the trustee cannot improve his payoff if the trustor never places trust. Equilibria in which trust is placed and honored can never be based on trust being placed unconditionally by the trustor. For, if there are no threats for the trustee that the trustor will change to "not placing trust," the trustee is always better off if he abuses trust. "No trust" is the only sanction the trustor has in the ITG and, therefore, equilibria in which trust is placed and honored have to be based on "conditionally cooperative" strategies in which the trustor will change to uncooperative behavior, i.e., not placing trust, if the trustee abuses trust.

Trigger strategies (Friedman 1971) are examples of such conditionally cooperative strategies. A trigger strategy for the ITG is defined as follows:

1. Trustor and trustee act cooperatively (place and honor trust) as long as the other actor acts cooperatively.

2. As soon as one actor deviates from cooperative behavior (withholds or abuses trust), the other actor changes to uncooperative behavior forever.

The ITG may have a cooperative equilibrium in *trigger strategies.* Trigger strategies are interesting because they are associated with the most severe punishment threat against the abuse of trust by the trustee (see Abreu 1988 on optimal punishment). This implies that if there exist equilibria in which trust emerges, trigger strategies are certainly in equilibrium. Furthermore, if cooperative strategies are in equilibrium the punishment threats of the trustors are credible and, therefore, the trustee will not abuse trust in equilibrium. Hence, the trustor *never* has to execute sanctions by withholding trust. Because trust is never abused in a trigger strategy equilibrium and sanctions for abusing trust need not be implemented, it is also a fact that there are no conflicts due to non-cooperative behavior in a trigger strategy equilibrium and, consequently, there is no need for conflict resolution. Trigger strategies, however, are not in equilibrium for all possible values of the parameters in the game. The following theorem states the necessary and sufficient condition for an equilibrium in trigger strategies and, hence, for the existence of a solution in which trust is always placed and never abused.

THEOREM 2.1 *Consider the* ITG $(\Gamma, w_1, w_2, \delta)$. *Then, trigger strategies are a subgame-perfect equilibrium if and only if*

$$w_2(1-\delta) \geq \frac{T_2 - R_2}{T_2 - P_2}. \tag{2.13}$$

Proof. For a formal proof see Friedman (1986: 88–89). □

An intuition for the proof is the following. First, the trustor never has an incentive to deviate from the equilibrium path because obtaining R_1 is the highest payoff she can receive in every period. The trustee has an incentive to deviate if his expected payoff from receiving T_2 now and P_2 in all the remaining periods is higher than the expected payoff from receiving R_2 now as well as in all the remaining periods. The condition where this is not the case is given in the theorem. This is the formal representation of the statement that placing trust is possible for the trustor if the trustee's long-term losses from trust withheld by the trustor is larger than the short-term gains obtained by abusing trust.

Thus, if the condition for equilibrium in trigger strategies is fulfilled, there exist conditionally cooperative strategies that are in equilibrium and in which trust is always placed and is never abused. These strategies might be trigger strategies. However, it is possible that other conditionally cooperative strategies exist that form an equilibrium and that imply the same solution, namely, trust is always placed and never abused on the equilibrium path. Such an equilibrium is certainly not payoff dominated by another equilibrium because the trustor cannot obtain more than R_1 in every period.[12] The maximal payoff for the

[12] I do not know any convincing argument why this equilibrium is more realistic than other Pareto incomparable equilibria where, for example, the trustee is allowed to abuse trust in every one out of ten periods of play without "triggering" the trustor to withhold trust. The reader can check that such equilibria exist for appropriate parameter values.

trustor is reached because the trustee will never abuse trust in this equilibrium. Therefore, this equilibrium describes the situation such that the control effects for the trustor are large enough to compensate for the short-term incentive of the trustee to abuse trust. Thus, for example, if the trustor expects "enough" periods of play with the trustee in the future, she can trust the trustee.

The following hypotheses follow from this theorem about parameters in the ITG that will recur frequently throughout this book. It follows from the theorem that

- trust increases with the discount factor of the trustee (w_2),
- trust decreases with the drop-out rate (δ),
- trust decreases with the incentive of the trustee for abusing trust ($T_2 - R_2$), and
- trust increases with the loss experienced by the trustee when the trustor does not place trust ($R_2 - P_2$).[13]

The payoffs and the discount factor of the trustor do not affect the equilibrium condition. This is due to the fact that trust is never abused in equilibrium. This issue will be addressed again in Chapter 3.

Two elements that are discussed in Section 1.2 have not yet been incorporated in this game-theoretic model. First, the actors are not yet embedded in a social network. Second, because all actors have perfect information over the incentives of the partners, they do not learn about their partners. Therefore, the model presented above does not imply predictions about learning effects.

NETWORKS AND GAMES

In most game-theoretic models of repeated games, the same (two) actors are playing in every period. Some repeated games with varying opponents have been studied (see, for example, Kreps 1990a; Milgrom et al. 1990; Fudenberg and Tirole 1991: Section 5.3), but these studies do not model social networks in any detail. Also, they employ simple assumptions such as the fact that new actors in the game know what happened in the past during all periods and with all actors. Therefore, these models lead more or less to the same results as models in which the same two actors are playing all the time.

One type of model that has found a large number of successors is similar to Axelrod's (1984) computer tournament in which actors are randomly matched together to play certain games (Heckathorn 1996; Lomborg 1996). Actors in

[13]Note that $\frac{T_2-R_2}{T_2-P_2} = \frac{T_2-R_2}{(T_2-R_2)+(R_2-P_2)}$.

these games have prescribed strategies such as always abusing trust or playing Tit-for-Tat, but there is no network structure that facilitates information transmission about past periods among actors.

Raub and Weesie (1990; Weesie 1988: Chapter 5) explicitly model a network of information relations between actors playing Prisoner's Dilemmas with each other. Only one global property of the network, namely, the minimal outdegree in the network, affected the trigger strategy equilibria. There is a growing literature on models in which actors are assumed to be placed on a grid ("cellular automata"), which can be interpreted as a social network, and actors play with neighbors on the grid. However, in these models all actors have the same number of neighbors and equal probabilities of playing with each of these neighbors (Nowak, Szamrej, and Latané 1990; Messick and Liebrand 1995; Hegselmann 1996). There are some analyses in which the number of neighbors of an actor is varied. This allows predictions about differences *between* networks but not *within* networks, because all positions in the network are equivalent. In most real-world networks there are, of course, individual differences between actors. Some actors have more contacts than others and these contacts may be more intense. The models in Chapter 3 and 4 seek to make predictions about such individual differences.

Incomplete Information

Game theorists have long realized that the complete information assumption used in many models is very problematic. Harsanyi (1967–68) laid the foundation of including incomplete information in game-theoretic models. Actors may be uncertain about the payoffs of other actors, for example. This kind of uncertainty can be modeled with a random move by Nature at the beginning of the game. Imagine that there are two trustees as illustrated in Figure 2.2: a "good" trustee who has no incentive to abuse trust in any period ($T_{2,2} < R_2$) and a "bad" trustee who has payoffs equal to the ordinary Trust Game ($T_{2,1} > R_2$). Nature chooses at the start of the game with given probabilities which of the two trustees has to play with the trustor. However, the trustor is not able to observe the outcome of this move, i.e., she does not know whether the trustee is good or bad. The trustee knows his own "type," i.e., he knows whether or not he has an incentive to abuse trust. The game that starts after the move by Nature can also be repeated. During the game the trustor can adjust her beliefs about the type of trustee and maybe she can deduce from his behavior with which trustee she plays. For example, if the trustee were ever to abuse trust, he would be the bad type. Thus, *learning* becomes an issue if incomplete information is introduced. Moreover, a bad trustee can try to behave as if he is the good type, because it is more profitable for him if the trustor believes he is the good type. In other words, trustees might be concerned to maintain the reputation of being a good type of trustee (see also Kreps, Milgrom, Roberts, and Wilson 1982).

Analyses of the finitely repeated Trust Game with incomplete information yield some appealing results with respect to control and learning effects (see Dasgupta 1988).[14] The problem with the finitely repeated ordinary Trust Game Γ is the following. If the trustee has an incentive to abuse trust, this implies that the trustee will certainly abuse trust in the last period. Therefore, the trustor will not place trust in the last period. Consequently, the trustee will abuse trust in the last period but one, because in the last period he will receive P_2 anyway. This implies that the trustor also cannot place trust in the last period but one. This argument continues up to the first period, which means that the trustor can never place trust. The situation changes if there is even a slight probability that the trustee does not have an incentive to abuse trust (see Camerer and Weigelt 1988; Neral and Ochs 1992). Then, the trustor might want to test whether the trustee has an incentive to abuse trust. Therefore, "some" trust is possible if there are "enough" periods to be played in the future. An equilibrium exists that consists of three phases. In the first phase, all types of trustees honor trust and, consequently, the trustor places trust. If the end of the game is approached (the exact timing depends on the parameters of the game), the bad trustee starts to randomize between abusing and honoring trust.[15] Of course, the good trustee continues to honor trust throughout the game. As long as the trustee continues to honor trust in this randomization period, the trustor is more and more convinced that she is playing with a good trustee and, therefore, continues to trust although the end of the game comes closer and closer. As soon as the trustee abuses trust for the first time, he reveals himself as being a bad trustee and the trustor will never place trust again.

Recently, games resembling the Trust Game have been analyzed (Cripps and Thomas 1997; Levine and Martinelli 1998). In these games, one actor cannot observe the type of another actor. Levine and Martinelli consider a buyer-seller relationship as a starting point of their game. The seller has the possibility of selling high-quality or low-quality products. The seller has to make that decision in advance because he has to make an extra investment, for example in production technology, committing himself to one of the two strategies. He cannot change that decision later in the game. The buyers cannot directly observe the choice of the seller. In Levine and Martinelli's model, the probability that the seller will make the investment depends on the time he expects to be in the market with that product and the extent to which buyers are able to evaluate whether he is selling high-quality or low-quality products. Interpreting the results in terms of social networks, sellers have a larger incentive

[14] In the finitely repeated Trust Game, the first move is a move by Nature similar to the first move in Figure 2.2 and the part of that game after the move by Nature is repeated.

[15] The term "randomize" indicates that the trustee chooses abusing trust with probability p and honoring trust with probability $1 - p$, $(0 < p < 1)$.

to sell high-quality products to buyers who obtain more information about the behavior of this seller from other buyers.

The two examples of games with incomplete information discussed above indicate that such games can shed some light on the relation between control and learning. I want to stress here that (rational) actors in these models have to be forward-looking as well as backward-looking to optimize their expected payoffs. They have to be backward-looking to learn from the information obtained from past periods of play and forward-looking to take into account potential sanctions and learning effects in the future. Some recent studies (Macy 1993; Flache 1996) suggest that actors in game-theoretic models are assumed to be forward-looking only and contrast these models with backward-looking learning models. This contrast makes sense only in game-theoretic models with complete information. However, game-theoretic models with incomplete information are themselves examples of learning models in which completely rational actors optimize their future payoffs using what they learned from past periods and even trying to exploit the learning efforts of other actors. However, modeling incomplete information and detailed social networks in a game-theoretic context at the same time is beyond the scope of this book and beyond the scope of current research in general. Therefore, I limit the analysis of learning effects to modeling how fast trustors can obtain information, assuming that trustors who obtain more information learn faster. Consequently, trustors who obtain more positive information can place more trust in the trustee than other trustors. In this book, I will not model the strategic use of and search for information. Developing and analyzing a model that combines social networks with a game-theoretic model including incomplete information may well be part of future research efforts. A first model in this direction consists of two trustors who play a finitely repeated Trust Game with one trustor and can inform each other between periods about the behavior of the trustee (see Buskens 2000).