University of Groningen

Trust and Exchange

Gautschi, Thomas

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2002

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Gautschi, T. (2002). *Trust and Exchange: Effects of Temporal Embeddedness and Network Embeddedness on Providing and Dividing a Surplus*. University of Utrecht.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 06-06-2022

Trust and Exchange

Effects of Temporal Embeddedness and Network Embeddedness
on Providing and Dividing a Surplus

Trust and Exchange

Effects of Temporal Embeddedness and Network Embeddedness
on Providing and Dividing a Surplus

Vertrouwen en ruil
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus
prof. dr. W.H. Gispen
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen op
vrijdag 7 juni 2002 des namiddags te 14.30 uur

door

Thomas Gautschi
geboren op 22 april 1969, te Thun (Zwitserland)

Promotor

Prof. dr. W. Raub

Co–promotoren

Dr. C.C.P. Snijders
Dr. H.M. Weesie

© Thomas Gautschi

ISBN 90–517–0617–0

This project was subsidized by the Netherlands Organization for Scientific Research (NWO, PGS 50–370)

Preface

In *A Course of Six Lectures on the Chemical History of a Candle* (1861), the printed version of his Christmas Lectures given at the Royal Institution in London, Michael Faraday constantly encourages his audience to be researchers on their own: "We come here to be philosophers; and I hope you will always remember that whenever a result happens, especially if it be new, you should say, "What is the cause? Why does it occur?" and you will in the course of time find out the reason." I too came to Utrecht to be a 'philosopher', eager to answer the 'whys' of a project which was then entitled *The Dynamics of Trust*. In the 'course of time', I was able to find out the reason for some of these 'whys' while leaving others unanswered. As a consequence of me trying to be a 'philosopher', the project slightly changed to what has eventually become this book on *Trust and Exchange*. A book on whether one should invest resources in the face of uncertainty and how the product of such investments is split between actors if they can negotiate its division.

The main chapters of this dissertation were presented at various occasions inside and outside the ICS. A slightly different version of Chapter 2 has been published in *Rationality and Society* (vol. 12, 2000). An adapted German version of Chapter 3 is forthcoming in *Modelle Sozialer Evolution*, edited by Andreas Diekmann and Ben Jann. Braun and Gautschi (2001), on which Chapter 5 is partly based, is currently under review.

My supervisors Werner Raub, Chris Snijders, and Jeroen Weesie have invested quite some resources in the course of this project. Even though the focus of the project has changed, I hope they find the surplus I produced worth their investments. I may have not always appreciated their comments and suggestions. This might thus be the place to apologize for being stubborn and not always following their advice. However, their input has not gone unnoticed and has significantly improved the quality of this book. I would like to thank them for spending more time on my supervision than I could have hoped for and for showing confidence in my work even in those times when progress was slow.

Regardless of his busy schedule, Werner always freed up time to read and comment on anything I handed in. Even though I did not always agree with his remarks on first sight, I had to realize that his comments were generally rooted on solid grounds. Werner's input has been very important for this project. He was also always able to structure my thinking and writing, telling apart the important from the unimportant.

From all I have learned from Werner, I will especially remember to think things over thoroughly before writing them down.

Chris always had time to listen to my questions. Whenever I dropped by, he was willing to help out right away, unless he was playing chess with Gerrit Rooks. In retrospect, I unfortunately made way too little use of Chris' willingness to share his knowledge and thoughts. Whenever I showed up with a problem, however, he was able to provide a solution or at least put me (back) on the right track. Also I would like to thank Chris for writing Chapter 4 with me. Besides his contribution to the text, his advice and comments on the statistical analyses were very helpful.

Jeroen I owe my thanks for his relaxed and humorous way of giving crucial advice. My feeling was always that he was looking at my ideas and writings (or the world in general?) from the 'Weesie point of view'. Do not ask me to define it—I only know it is a view on things different from the one the rest of us has, but which resulted in Jeroen seeing mistakes that otherwise would have remained unnoticed. Especially, however, I thank Jeroen for his advice and support on statistical matters. For some time, he even supplied 'long–distance advice and solutions' on the statistics of Chapter 3, which was written while Jeroen was working at Stata Corporation in Texas.

Many ICS colleagues have provided social and academic advice. I therefore owe thanks to a lot of the ICS members and especially the Utrecht members. I experienced the ICS as a stimulating and vivid surrounding to conduct my graduate studies. First of all, I am indebted to my yeargroup '97: Miranda, Johan, Marc and René, who has been my roommate over the last four and a half years. They made moving from Switzerland to the Netherlands a positive experience and have become good friends. Among other things, they introduced me to the Dutch culture. This included things so different as 'kroketten', de Efteling, and walks on the windy beaches of Schiermonnikoog. Despite all their efforts, some parts of the Dutch culture still remain a mystery to me (e.g. grown–ups with Mickey Mouse socks) while to others I show a remarkable resistance (e.g. 'erwtensoep'). Nevertheless, I appreciated and enjoyed all their efforts and I hope you will never give up trying to turn me into a 'real Dutchman'. I hope we can continue our regular 'jaargroep uitjes'.

My thanks also go to Vincent for discussions on trust and exchange and for untiringly solving all the LaTeX problems I could not master myself. Vincent and Johan have been my dart partners over the last couple of years. Our games provided the necessary breaks from work. I just hope the university will never hold us responsible for all the nice little holes in the wall of René's and my (former) office. And, I hope that my (one time thrown) 'one hundred and eighty' will remain unreachable for you! My roommate René, of course, I would like to thank for never loosing temper when confronted with my ups and, especially, downs. I very much appreciated our discussions on both our research topics. Finally, I thank all the members of the Pionier research program "The Management of Matches" who provided critical comments on my work. I will always remember our Friday afternoon Pionier seminars where sometimes the title of a paper already triggered discussions lasting several minutes. I would also like to thank the

Netherlands Organization for Scientific Research (NWO) for financial support through the Pionier program "The Management of Matches". Mariëlle I want to thank for helping to solve problems inside and outside the ICS and Arnout for translating my English draft of the Dutch summary.

Andreas Diekmann offered me my first scientific position at the University of Bern without which I may not have been here now. In Bern, I also met Norman Braun with whom I still share a (scientific) friendship. I thank both of them. To Norman I owe many more thanks for our discussions and our joint work during my ICS time of which not only Chapter 5 is proof but also our phone bills. Norman's untiring scientific devotion provided lots of new ideas and triggered thinking about various serious and less serious sociological problems. His enthusiasm for scientific work is 'infectious'. Many thanks also for being a generous host during my stays in München and for providing the facilities to run the experiment reported in Chapter 3.

Despite my early plans of just being a 'philosopher', life in the Netherlands has turned out to be more than writing a dissertation. Social get–togethers so different as playing soccer or hanging out in cafes turned out to be the necessary balance to work, thanks to Anne–Rigt, Miranda, Frank, Gijs, Johan, Marc, Mattijs, René, Rudi, Willem–Jan, and especially Wilfred. He has become a good friend and biking partner. He always allows me to at least beat him to the tops of the Swiss mountains. Thanks for leaving my 'Heimvorteil' untouched!

A couple of weeks a year my social life also takes place in Switzerland. Chrigu, Golda, Jürg, Martin, Markus and Rachel, Susi, Peschä, Viviana, and especially David, Mike, Noah, Pam and Tim are always making 'coming home' a pleasant experience. Others have been important to me for different reasons: Jill, who unfortunately was on the wrong side of the Atlantic most of the time; Amika and Kim, who always reminded me of my social life when my sofa looked too cosy; and Sonja for keeping her fingers crossed during the last few months of my work on this book.

Last but not least I thank my parents Margrith and Samuel for always and without hesitation supporting me in various ways and in everything I have or have not done. I owe you a great many thanks for continuously providing resources for the production of many surpluses. I hope I will finally be able to give you a large piece of it back.

Contents

List of Tables

List of Figures

Chapter 1

Introduction: Providing and Dividing a Surplus

1.1 Introduction

This book is about the provision and division of a surplus. How a surplus is provided and divided often rests upon bilateral relations. Two examples may help make this clear. Suppose two firms are working separately on the development of a new product. The complexity of the product requires investments into research and development, for example, efforts of engineers as well as production capacity. These investments are beyond the possibilities of each firm separately. One firm might have the technical know how but not the production capacity to implement it, the other might have the necessary production facilities but not the crucial knowledge. The two firms could join forces in an R&D alliance to combine both their knowledge, work force, and production capacity. It would be the only way for them to have a good chance of successfully developing the product. The cooperation gain earned by selling or licensing the new product constitutes the surplus. In situations of this kind, prior agreements often regulate the fair division of the surplus. However, how sure can either firm be that the surplus will indeed be divided according to the agreement once the surplus has materialized? Is it wise to join forces if problems on the division of the surplus are to be expected? As another example, take two hunters. In the first example, joining forces is the only way to produce a surplus. In the second example, each of the two hunters could kill some game on their own. Neither of their kills (e.g. a rabbit) would be enough to feed their families. If they hunt together though, the kill, for instance a mammoth, will be large enough to feed both their families. The 'additional' meat a mammoth provides, compared to two rabbits, constitutes a surplus. Now faced with the dead mammoth, how should the hunters divide it? Who gets the larger share and why?

These two scenarios on providing and dividing a surplus differ in various respects. As we discuss below, this leads to different research questions this book addresses. Common to both scenarios, however, is the provision and division of a surplus in *interactions between two actors*. They can be individual actors (i.e. natural persons) such as the hunters, or 'corporate actors' (Coleman 1990: 540) such as the two business firms. In this book, we analyze the provision and division of a surplus as abstract games between two actors. In these games, the provision and division of a surplus is conceptualized as *exchange between two actors*. Consequently, a surplus is not generally referred to as a cooperation gain, a mammoth, or the like. Throughout, *a surplus is an 'additional something' which, if produced, can be divided between two actors in some way with the division leaving both actors better off than if they had not produced and divided the surplus*. In the literature on division problems, the surplus is often referred to as 'cake' or 'pie'. We therefore use surplus, cake, and pie as synonyms.

The provision and division of a surplus as presented in this book is based on *two different basic scenarios*. In one part, we analyze how a surplus is provided and divided in the context of *delayed exchange*. In short, this means one actor needs to make a one–sided advance concession, not or not fully secured by legally enforceable contracts, without which it is impossible to provide the surplus. An example of this kind of one–sided advance concession is the knowledge transfer from one firm to the other in our R&D example. If the second actor can then divide the surplus, this kind of one–sided concession entails a certain risk for the first actor. The surplus could be divided in such a way that only the second actor profits. We will argue that providing and dividing a surplus through delayed exchange can be seen as a *trust problem* between two actors. Our focus is on 'placing trust', the move by the first actor making the *provision of a surplus* possible, rather than on 'honoring trust', the division of the surplus by the second actor later on. In the second part of the book, we neglect questions related to the provision of a surplus. In this part, we assume the production of a surplus does not pose any problems since, for instance, no one–sided concessions are necessary. The basic scenario in the second part of this book assumes that an encounter between two actors 'produces' a surplus. For instance, two hunters successfully kill a mammoth. The focus is then on the *division of the surplus* rather than its provision. Unlike the delayed exchange scenario, the second scenario gives both actors a right to co–determine the division of the surplus. It is supposed that they determine their shares by bargaining where a *delay in agreement* on how to divide the surplus entails cost for both actors. These costs determine the actors' bargaining powers, which crucially affect the bargaining outcome.[1] Besides the focus on the interaction between two actors, there is another aspect common to both scenarios: a focus on the effects of

[1] Since the whole book is about provision and division problems in a dyad, we always focus on two actors. We use a female pronoun to refer to one of them and a male pronoun to refer to the other. In the scenario of delayed exchange and trust, the provider of the surplus is called 'she' or 'trustor' and the divider of the surplus is called 'he' or 'trustee'. In the scenario of delayed agreement and bargaining, we can no longer distinguish between a provider and a divider. We tell the actors apart by calling the actor who starts the bargaining process 'ego' or 'he'. The second actor is consequently called 'alter' or 'she'.

delay. Either a delay in exchange as in the trust problem or a delay in agreement as in the bargaining problem.

We now *introduce the research questions addressed in this book*. For this, it is helpful to briefly discuss some assumptions on the scenarios addressed in this book. First, we consider three assumptions regarding the provision and division of a surplus as delayed exchange, that is, as a trust problem between two actors. Assume (1) that an actor's utility is given by the actor's share of the cake and that an actor prefers a larger share of the cake to a smaller one.[2] We model delayed exchange using an abstract game describing the interdependent situation between the two actors. As we will see in Section 1.2, this game is a Trust Game (Dasgupta 1988; Kreps 1990a). Assume (2) that actors behave according to game theoretic rationality or follow some other principle of incentive guided behavior. Game theoretic rationality is basically conceptualized in this book as Nash equilibrium behavior (Nash 1951). This means roughly that each actor maximizes utility, given the other actor's behavior (see for instance Rasmusen 1994: 22–28 for technical details). Last, assume (3) an 'isolated encounter' between two actors in the sense that their delayed exchange is not connected in any way with previous or future exchanges or any interaction between the two actors or between other actors. That is, we consider a one–shot Trust Game. Then, as will be shown below, the prediction would be that trust is not placed and, if nevertheless placed, would be abused. Thus, the surplus would not be provided. If it had been provided, it would have been divided in such a way that only the dividing actor would profit. Empirical results from systematic research and everyday observations, however, show that sometimes trust is placed, and sometimes it is also honored: surpluses are, at least sometimes, provided and more or less fairly divided in the framework of delayed exchange. In this book, we investigate whether and how this can be explained if we relax assumption (3) above but adhere as far as possible to assumptions (1) and (2). We consider situations where two actors repeatedly meet over time and exchange is embedded (Granovetter 1985). More precisely, we assume temporal embeddedness (Raub and Weesie 2000b) of exchange in a sequence of exchanges between the two actors. Thus, the original exchange between the two actors is connected with other exchanges between them. In terms of the abstract Trust Game, this means the game is repeated a finite or infinite number of times. As the discussion in Section 1.2 will show, such a temporal embeddedness of actors is a sociologically relevant extension of the one–shot situation (e.g. Blau 1964; Emerson 1972, 1981; Granovetter 1985; Voss 1982, 1985). The first part of the book (i.e. Chapter 2 through Chapter 4) then addresses the general research question:[3] *Given that the provision and division of a surplus is conceptualized as a situation of trust between two actors, which aspects of temporal*

[2] If the cake to be divided solely consists of money, then 'utility is own share of the cake' translates into the familiar 'utility is own money'.

[3] The research questions introduced in this chapter are formulated as 'descriptive' questions. The accompanying explanatory questions ('why questions') are implicitly contained.

embeddedness affect the placement of trust and how (i.e. positively or negatively) do these aspects affect the placement of trust?

A similar line of reasoning leads to the research questions on the division of a surplus as a delayed agreement. Consider again three assumptions. The first pertains to providing the surplus and assumes (1) that the provision of the surplus does not pose problems for the two actors. For example, the surplus is available to them beforehand.[4] As in the previous scenario, assume (2) that an actor's utility is given by the actor's share of the cake and the actor prefers a larger share of the cake to a smaller one. Since the scenario of delayed agreement is also based on an abstract, game theoretic representation of alternating offers bargaining, assume again (3) that actors behave according to game theoretic rationality. As we will discuss more extensively in Section 1.3, it follows from assumptions (1) to (3) that the division of the surplus depends on the relation between the bargaining powers of the two actors such that the actor with the higher bargaining power receives the larger share of the surplus. We assume that the actor who bears lower costs when the agreement on the division of a surplus is delayed has an advantage. The scenario of delayed agreement, however, does not contain any information on the costs to the actors if the agreement on the division of the surplus is delayed. We know the division of the surplus is a function of the actors' bargaining powers, but we do not know the bargaining powers. Nor does the scenario of delayed agreement give any indication of where differences in bargaining powers could come from. However, a sociologically relevant source of bargaining power has been suggested by Emerson (1972, 1981). It rests on the realization that an exchange between two actors should not be viewed as independent of their relations to other actors. We therefore embed our two actors in a network that also includes other actors. Following Emerson, this network embeddedness can be seen as an important source of the actors' bargaining powers. From an actor's point of view, networks either determine or prevent dyadic exchange alternatives. Consequently, networks define differences in opportunity structures between actors that are reflected in the actors' bargaining powers. The second part of the book (Chapter 5) addresses two research questions. First, we ask: *How does an actor's embeddedness in a network of negotiation partners determine his or her share of the surplus in relations with his or her partners?* Being embedded in a network means that actors with more than one partner have access to different bargaining partners with whom a division of a surplus is possible. In some networks, actual exchange may be restricted to one relation due to exogenously stipulated rules, such as for example monogamy. The research question stated above then leads to a second one: *With whom of his or her connected partners does an actor divide a surplus if restrictions in a network prescribe that the actor can only divide one surplus at a time?*

[4] The example of the two hunters is a situation where the provision of the surplus, that is, killing the mammoth, poses no problems. Another situation where a surplus is available is in laboratory experiments where the experimenter provides the subjects with the necessary surplus and they in turn can divide it.

Embeddedness effects thus play a prominent role throughout this book. In addition to the focus on dyadic exchange and on effects of a delay in exchange or agreement, respectively, there is now a third common feature in the analyses of the two scenarios: the focus on effects of embeddedness. On the one hand, we ask whether we can explain the provision of a surplus, that is, the provision of the necessary resources for a production of the surplus, in terms of effects of temporal embeddedness. On the other hand, we seek to explain how the division of a surplus between two actors is affected by their network embeddedness. Even if we do address effects of embeddedness, the book still focuses on the outcome within a dyad. In this sense, we are interested in *micro–level* outcomes. This is immediately clear with regard to the scenario in the first part of the book. Beyond the dyad, there are no other actors involved. In the second scenario, however, networks play a prominent role. They remain relatively small though and, as Chapter 5 will show, only the immediate vicinity of the dyad under investigation is important. We therefore see the 'micro–level' as including two or at most a few actors, such as a dyad and third parties or small networks affecting the behavior of the actors in the dyad. Consequently, we view social settings beyond dyads or small networks as the *macro–level*, with society as a whole being one example. This book neglects macro–level consequences of micro–level behavior and questions on 'aggregated individual behavior' are beyond its scope.

There is literature that examines the effects (dyadic) trust has on society as a whole. Cook (2001), Fukuyama (1995), Knack and Keefer (1997), Misztal (1996), or Putnam (1993) present arguments on the beneficial effects of trust on the micro–level for outcomes on the society level. In short, the argument is that society vastly profits from trust on the micro–level because economic institutions such as laws or contracts do not provide a "[...] sufficient basis for both the stability and prosperity of postindustrial societies" (Fukuyama 1995: 11). However, trust on the micro–level can also have harmful effects on the macro–level. Think of the 'bads' inflicted on society by organizations that base their success on trustful relations between their members like the Cosa Nostra or organized crime in general (e.g. Hess 1986; Raith 1994, 1995), terrorist associations, or cartels. Not only can the bilateral provision of surpluses exhibit harmful effects on the macro–level, so can the micro–level division of these surpluses. One of sociology's main questions is concerned with these consequences: inequality and the unequal distribution of resources in society (e.g. Ultee, Arts, and Flap 1996: chs. 2 and 3). Engels (1845) or Marx (1867), witnesses to the Industrial Revolution, noticed that individual behavior related to the division of surpluses will increase inequality in society. Lenski (1966) highlights the distributive system throughout the history of mankind and seeks to explain who gets what and why based on the specific developmental stage of a society and its individuals.

In the remainder of this chapter, we give further details on the two scenarios introduced above. Section 1.2 discusses the provision and division of a surplus as delayed exchange and addresses the importance of temporal embeddedness of the actors. The main focus lies on the question whether or not the surplus is provided. Division is less prominent in the scenario of delayed exchange. Chapters 2 to 4 are about the provision and division of a surplus as delayed exchange, that is, as a trust problem. In Section 1.3, we introduce the division of a surplus as a series of offers and counteroffers, and thus as a consequence of delayed agreement. This scenario is expanded to include the sociologically relevant idea that network structure affects the division of the surplus. Here, the focus lies on the division rather than the provision aspect. Chapter 5 discusses in depth the division of a surplus as bargaining between two actors.

1.2 Delayed Exchange and Trust

Assume there is a surplus—say, a chocolate cake—to be divided between the trustor and the trustee and that one of them—in our case the trustee—has been given the right to divide the chocolate cake. His options are either to keep the whole cake himself or divide it more equally among the two of them. So as a result, either both actors profit or only the trustee profits.[5] This game is known in the literature as the *Dictator Game*.[6]

Let us now expand this scenario and incorporate how the surplus materializes. Assume the trustor has all the necessary ingredients for a chocolate cake but misses the recipe. She can keep the ingredients and nothing else will happen.[7] Alternatively, she can give the ingredients to the second actor, thereby *providing the necessary resources to produce a surplus*. In this case, the trustee bakes the cake and consequently decides on how to divide it in a Dictator Game–like manner. If he splits the cake more or less equally, both actors profit. Thus, the trustor who provides the ingredients prefers a fair share of the resulting cake to keeping her ingredients. If the trustee keeps the chocolate cake to himself, the trustor *loses her investments* (i.e. chocolate, milk, flour etc.) and does not profit from the surplus.[8]

[5]Assume that both actors assign a utility of 30 to their respective shares of the cake if it is split equally. If the trustee keeps the cake himself, he values eating the whole cake with 50. That is less than two times 30 since the last few pieces of the cake are less enjoyed than the first ones due to diminishing marginal utility. Since the trustor gets nothing of the cake, she values this outcome with 0.

[6]Normally, a Dictator Game allows the trustee in the role of the dictator to divide the cake as he pleases. We can simplify it and confine ourselves in the first part of the book to a situation where the dictator can only keep the whole cake or split it more evenly between himself and the trustor.

[7]An outcome giving each of the actors a utility of 10. Since, for instance, the trustor might use her ingredients to make pancakes which, however, she values less than a share of a chocolate cake. The trustee, on the other hand, has to have an old muffin that he also values less than a piece of chocolate cake.

[8]We could assume that losing the ingredients and not profiting from the surplus makes the trustor's utility negative. For the sake of argument, we need only assume that the utility of losing the investment is smaller than the utility of keeping the investment.

The same underlying structure emerges in more complex situations. Remember that a surplus can for example be a cooperation gain that two firms can realize by joining forces. Assume Airbus Industries and Boeing plan to jointly develop a supersonic airplane. Neither Airbus nor Boeing can build the plane on their own. However, with both companies' know how and Boeing's production capacity, it is feasible. At the start of the project, Airbus sends five hundred of its most skilled engineers to Boeing's Commercial Airplane Headquarters in Seattle. In other words, Airbus invests in the provision of a surplus. If the project succeeds and the airplane is built and sold, a gain (surplus) can be expected and can be equally divided. However, Boeing might abandon the joint project after accumulating enough of Airbus' know how on supersonic planes and produce the plane on its own. Even if the cooperation between Boeing and Airbus has been secured by legal contracts, Boeing knows a monetary fine for breaking the contract is less than the expected gains if it produces and sells the plane. If this happens, Airbus is giving away valuable know how without getting its share of the surplus.

Each of the two situations, baking a chocolate cake and producing a supersonic airplane, has a similar structure. The trustor who has important resources for producing the surplus makes a one–sided advance concession. This investment leaves the trustee with an opportunity for opportunistic behavior since the investment is either not secured by a legal contract (chocolate cake) or it is incompletely secured by a legal contract (supersonic airplane). The production and division of the surplus is only feasible after the trustor's investment. Consequently, there is a time asymmetry between providing the resources and dividing the surplus (a delay in exchange). This time asymmetry creates a crucial power difference in favor of the trustee. He is at an advantage since he can decide how the surplus is divided, as in the Dictator Game. Therefore, the trustor who provides the resources faces the incentive problem of whether to make them available to a partner who can use them for his own benefit. The trustor cannot be sure of the future size of her share of the surplus and thus has to anticipate the possible division. If the actors are each only interested in their own share of the cake, and if they each prefer more to less, the outcome could be that no surplus is produced. The trustor might anticipate not receiving a 'reasonable' share of the surplus, and hence keep the resources to herself.

This scenario of delayed exchange is a *situation of trust*. Coleman (1990: 97–99) specifies four characteristics of this type of situation. First, the trustor placing trust (i.e. providing the resources) makes an action by the trustee possible that otherwise would not have been possible. He can divide the surplus once it is produced. Second, the trustor benefits if trust is placed and honored. She values a fair share of the cake more than she values her resources. However, she regrets it if trust is placed and then abused. Third, placing trust is a voluntary action by the trustor. And fourth, the trustor and the trustee move sequentially thus creating the crucial time–lag and the incentive problem of the trustor.

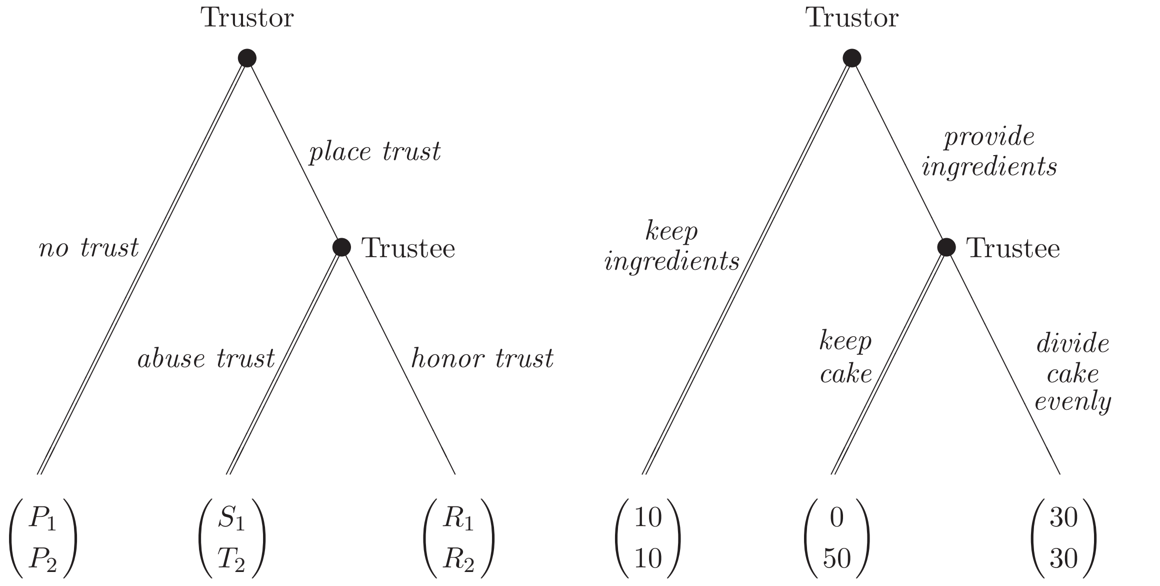


FIGURE 1.1 *Extensive form of the Trust Game ($R_1 > P_1 > S_1$ and $T_2 > R_2 > P_2$; double lines indicate equilibrium path of play) and a numerical example which refers to the chocolate cake example (see page 6).*

To describe this type of situation of trust, we use a simple game theoretic formalization introduced by Dasgupta (1988) and Kreps (1990a). The game is known as the *Trust Game* or one–sided Prisoner's Dilemma Game (Rasmusen 1994: 129–131). FIGURE 1.1 depicts the game with general payoffs and a numerical example that refers to the provision and division of a chocolate cake (the example on page 6). The ordinal ranking of the payoffs for the trustor is $R_1 > P_1 > S_1$ and for the trustee $T_2 > R_2 > P_2$.[9] Once the resources are provided—the trustor has placed trust—the trustee would prefer a division of the surplus in his favor to a more equal division since $T_2 > R_2$. As a result, he has an incentive to choose to abuse trust over honoring trust. However, the trustor should be able to anticipate this kind of behavior. She thus has an incentive not to provide the resources, that is, not to place trust since $S_1 < P_1$. The outcome of a one–shot Trust Game is then the *Nash equilibrium* where the trustor withholds trust and the trustee abuses trust if it has been placed. But this outcome is Pareto inferior to an outcome where the surplus is produced and divided in such a way that both actors profit. The Trust Game is thus an example of a *social dilemma*—a situation where individual and collective interests clash and rational and selfish actors end up in an inefficient (Pareto inferior) equilibrium.

A great deal of research has been conducted on situations where individual rationality in the sense of equilibrium behavior clashes with a socially desirable outcome. Predictions and experimental tests often focus on the (one–shot) Prisoner's Dilemma

[9] We refer to a specific combination of the trustor's and trustee's payoffs as the payoff structure of the game.

Game (for an overview, see for instance Dugatkin 1997: chs. 1 and 2).[10] Some recent research also uses the Trust Game to more adequately address the time asymmetry between the trustor's and the trustee's moves. Even though the results (i.e. rates of placed trust, cooperation rates) differ across experiments, one finding seems robust. Numerous experimental subjects do not behave in accordance with the game theoretic prediction as explained above, that is, trust is withheld but would be abused if placed in the Trust Game, or mutual defection should be observed in the Prisoner's Dilemma Game. Ahn et al. (1998), for example, report cooperation rates in a Prisoner's Dilemma of around 20%. The exact number varies with various features, such as the payoff structure of the game. In a series of Prisoner's Dilemma Games by Ahn et al., an increased difference between P and S (labelled 'fear': $P - S$), and between T and R (labelled 'greed': $T - R$) had negative effects on the cooperation rates. Findings reported in the literature about placing and honoring trust are often similar. For instance, in experiments with 36 one–shot Trust Games with different payoffs, Snijders (1996: Table 4.1) finds that on average trust is placed in 37% of the cases and honored in 36% of the cases. As in Ahn et al.'s experiment, trust rates vary with the payoff structure of the game. Snijders notes that the probability of placing trust decreases in the monetary risk to the trustor and the monetary temptation to the trustee.[11] On average, the cooperation rates he found are somewhat higher than in Ahn et al.'s experiment on the Prisoner's Dilemma Game. More results from experimental play of simple dilemma situations (e.g. Prisoner's Dilemma, Trust Game etc.) can, for instance, be found in the books by Davis and Holt (1993), Dugatkin (1997), and Kagel and Roth (1995) or in a paper by Hayashi et al. (1999) and especially in working papers of Bruno S. Frey's and Ernst Fehr's *Institute for Empirical Research in Economics* of the University of Zürich.[12]

Thus, experimental results often contradict game theoretic predictions that follow from the three assumptions on delayed exchange as introduced on page 3. The literature contains various suggestions as to how this could be explained. Mostly, the assumption

[10] In a Prisoner's Dilemma Game, two actors simultaneously choose one of two strategies. An actor can choose to defect, which in terms of the Trust Game means that trust is not placed and abused if the other does. On the other hand, an actor can choose to cooperate, which means placing trust and honoring trust if the other actor places trust. If both actors cooperate, they each earn R. If they both defect, they both receive P. If one cooperates and the other defects, the former receives S and the latter earns T. The ranking of the payoffs is $T > R > P > S$ for both actors. Individual rationality thus prescribes defection (defection maximizes an actor's payoff irrespective of the other actor's behavior) even though mutual cooperation would yield a Pareto superior outcome in terms of the payoffs.

[11] Snijders shows on the basis of a model of non–selfish utility that the payoff ratios RISK $:= \frac{P_1 - S_1}{R_1 - S_1}$ and TEMPTATION $:= \frac{T_2 - R_2}{T_2 - S_1}$ represent the influence of the payoffs on the subjects' behavior.

[12] Working papers by Frey, Fehr, and colleagues can be downloaded from http://www.iew.unizh.ch.

that 'utility is own share of the cake'[13] is relaxed in favor of considerations about fairness (e.g. Rabin 1993; Nelson 2001), relative payoffs (e.g. Bolton and Ockenfels 1998, 2000) or other, 'psychological' aspects (see for instance Camerer 1997; Kelley and Thibaut 1978; Rabin 1998; Snijders 1996). These approaches often generate better corroborated predictions on subjects' behavior in laboratory experiments. However, the probability of either cooperating or defecting is then generally seen as a result of individual traits. People are assumed to have different social orientations (e.g. egoistic, altruistic) that guide their behavior. In this book, however, we are not primarily interested in differences between individuals but rather in effects of different social situations on individual behavior. From a sociological point of view, an explanation of individual behavior in terms of the situation (e.g. one–shot game versus repeated game, different payoffs, third–party information etc.) itself would be more desirable, possibly in addition to social orientations. In other words, we would not like to challenge assumptions about utility but rather investigate the role of social conditions. In this book, this means studying the effects of temporal embeddedness on the probability that trust is placed, that is, that the necessary resources for the production of a surplus are provided. The research question addressed in the first part of the book therefore reads: *Given that the provision and division of a surplus is conceptualized as a situation of trust between two actors, which aspects of temporal embeddedness affect the placement of trust and how (i.e. positively or negatively) do these aspects affect the placement of trust?*

1.2.1 Temporally Embedded Trust

In a one–shot Trust Game, explaining individual behavior in terms of 'the situation' means explaining behavior specifically in terms of payoffs.[14] Analyses of this kind are limited in one important aspect: human interaction hardly ever takes place in isolated, one–shot encounters but is embedded temporally as well as socially. For one thing, people meet repeatedly, for another, they interact within a social network (e.g. Burt and Knez 1995; Buskens 1999; Raub and Weesie 1990). We start our analysis of embeddedness effects by bypassing the network aspect and focus first of all on the effects of repeated interactions on placing trust.

To conceptualize temporal embeddedness, we consider the situation where the one–shot Trust Game is repeatedly played by the same trustor and trustee. The question we would like to answer is: Which aspects of temporal embeddedness affect whether or not trust is placed in a dyadic relation between the trustor and trustee? That temporal

[13] Since in experiments the payoffs of a game typically correspond to monetary outcomes, 'own share of the cake' should be read here to mean 'money'.

[14] Individual (e.g. age, sex, education) or contextual (e.g. cultural background) variables may help explain behavior as well. Theoretically, we nevertheless overlook them. In addition, subjects participating in laboratory experiments usually constitute a homogenous population (e.g. Dutch freshmen in sociology) in such a way that effects of individual and contextual variables are less expected.

embeddedness can help facilitate cooperative behavior in dilemma situations has been noted by numerous researchers (e.g. Axelrod 1984; Coleman 1990; Granovetter 1985; Schelling 1960; Voss 1982, 1985). As the research question indicates, the focus of the following three chapters is on 'placing trust' and hence on the trustor's rather than the trustee's behavior. This means we address the provision of the surplus and thus the trustor's move to provide the necessary resources. Nevertheless, assuming incentive guided behavior, it is obvious that the trustor's behavior depends on her anticipation of the trustee's behavior. So, incentives for the trustee enter our explanation for the behavior of the trustor as well.

In a one–shot Trust Game where trust is placed, an abuse of trust bears no further consequences for the trustee since the relationship ends after one encounter. A rational and selfish trustee would see no need to honor trust that has been placed. If there are repeated interactions, the situation changes. The repeated Trust Game is similar to social exchange in the sense of Blau (1964; see also Emerson 1972, 1981; Thibaut and Kelley 1959). Basically, social exchange exhibits a simple pattern: by helping another person, a person creates an obligation the other person ought to discharge by doing something for the person who initiated the exchange (*reciprocity*). This in turn leaves the first person indebted to the second one and obligates her to reciprocate the favor, and so forth. Each of the actors pays the other back in kind. Partly because each reciprocation of a favor creates a new obligation, Blau (1964: ch. 4) and Emerson (1981) argue that social exchange can only be properly understood if the focus is on repeated interactions. Even though the situation of a repeated Trust Game is similar rather than identical to this conceptualization of social exchange, the mechanism of reciprocity that keeps social exchange going can also facilitate a repeated provision and division of a surplus in Trust Games.[15] Temporal embeddedness will weaken the trustee's advantageous position of having the second move. The trustor will most likely make her placement of trust dependent on the trustee's previous behavior. Gouldner (1960) assumes a universal and generalized norm of reciprocity, prescribing that in social exchange "people should help those who have helped them" and "people should not injure those who have helped them" (p. 171). Consequently, failure to behave in accordance with this norm is subject to sanctions by the trustor herself as well as by others. The question in the first part of the book is not whether there is such a norm of reciprocity. The question will rather be whether reciprocal behavior will evolve even under incentive guided behavior by the actors. Voss (1985: especially chs. 2 and 4),

[15]There is a specific reason why the repeated Trust Game "is only similar" but not identical to social exchange. In social exchange, reciprocating a favor creates a new obligation that requires a reciprocal move by the indebted person. In a repeated Trust Game, honoring placed trust can be viewed as a reciprocal move. By the same token, abusing placed trust means failing to reciprocate. However, by honoring trust if placed, the trustee does not create a new obligation towards the trustor. The trustor need not place trust in a new game, as would be the case if a repeated Trust Game was identical to social exchange. Nevertheless, the trustee's reciprocal move may increase the chance of the trustor placing trust again in the next game.

for instance, discusses whether or not reciprocity can be endogenously established and stabilized.

We agree with Blau (1964: 97) that failing to discharge an obligation "[...] has a number of disadvantageous consequences, several of which do not depend on the existence of a norm of reciprocity." Failure to reciprocate a favor can simply hamper future social exchange. Applied to the scenario of a repeated Trust Game, this means a trustee who fails to reciprocate by honoring placed trust eventually runs the risk of impeding any future provision of a surplus linked to 'his' trustor. In other words, a trustor whose trust has been abused is less likely to place trust in a future game. She sanctions the trustee by withholding trust, no further surpluses will be produced, and neither actor can profit in the long run. In the words of Blau, the "[...] disinclination to do favors [serves] as a punitive reaction against a violator" (p. 97) which helps people protect themselves from being abused.[16] We use Blau's analysis of 'punishment' in social exchange to our advantage in studying the repeated Trust Game.

The trustor's option for punishing deviant behavior and, similarly important, the trustee's anticipation of this punishment is one important aspect in the study of effects of temporal embeddedness on trust. Schelling (1960: 45) claims that mutual agreements based on trust are only enforceable because of the recognition of future opportunities. However, the future not only makes it possible to sanction deviant behavior. It also makes it possible to reward cooperative behavior. If the trustee honors placed trust, he increases the chances of a trustor once again providing resources for a new surplus. In this book, we refer to this aspect of the future as the *control effect* (Buskens 1999; Raub 1997) of temporal embeddedness (Blau 1964; Emerson 1981; Gouldner 1960; Homans 1961). In the literature, it is also referred to as conditional cooperation (Taylor 1987) or 'tit–for–tat' (Axelrod 1984). If a trustee recognizes that the future offers possibilities for obtaining shares of (many) more surpluses, which will not be produced if he falls prey to the temptation of the short–term gain, the trustor can effectively steer the outcome in a current situation.

We not only focus on this control mechanism but we also argue that control effects may be more or less severe or effective depending on the 'kind of future'. Temporal embeddedness is conceptualized in terms of a repeated Trust Game. The future, with respect to the current Trust Game, is therefore the number of Trust Games still to be played. To study the effect of the 'kind of future', we use Trust Games with different payoff structures. Let us assume that control effects increase or decrease in the earnings the trustee foregoes if the trustor withholds trust. Punishment, for example, can then

[16]Even though Blau's view of punishment exists independently of a norm of reciprocity, he sees this type of norm as reinforcing punishment. And, this norm makes the trustee feel guilty and thus subjects him to additional 'internal' sanctions independent of the actions of others. However, little or no attention is devoted in this book to the trustee's possible bad conscience. In addition, Blau argues that failure to discharge an obligation produces distrust and makes social exchange become economic exchange since now the actors will only help each other again for payment in advance. We will note in Chapter 3 that this is not completely true.

be seen as being less severe or effective if the trustee's payoff in the event that trust is withheld (i.e. P_2) is close to R_2 or even T_2, the payoffs the trustee may realize if trust is placed.[17] In sum, we assume that *two aspects characterize the future* in a repeated Trust Game. First, the expected length of the mutual future between the trustor and the trustee. Length is conceptualized as the number of possible games to punish deviant behavior or to reward cooperative outcomes. And second, the kind of future conceptualized by the payoff structure of the future games.

The trustor need not solely rely on the control effect of the future. She can base her decisions on the past as well. In a stable dyad, a shared past allows for *learning* about the trustee (Raub 1997).[18] A mutual past of the trustor and trustee can provide different kinds of information. Learning as presented in the literature reflects different approaches for dealing with this kind of information. Many of these approaches can be summarized as 'inductive learning' (Holland et al. 1986) in which subjects construct and update representations of their world as they discover regularities in the course of their interactions. Examples of models like this can be found in Lazaric and Lorenz (1998a, 1998b), Rivaud–Danset (1998), Sako (1998), or Schelling (1960).

We will, however, use a concept of learning more related to rational learning of game theoretic actors. In the beginning of their mutual relation, a trustor does not know the 'type' of the trustee. Put simply, a trustor does not know whether or not the trustee is trustworthy.[19] A trustee may be a person who always abuses placed trust: a 'bad trustee', so to say. Or, he may be the type who always honors trust if placed, that is, a 'good trustee'. However, he might also be a 'bad trustee' in disguise: a trustee who gives the impression of being a 'good trustee' by honoring placed trust in the first instance, so as to be able to abuse trust as soon as the trustor provides more and more valuable resources for the production of ever larger surpluses. As long as trust has not yet been abused, how can a trustor tell a 'bad trustee in disguise' apart from a 'good one'? In other words, game theory assumes that the trustor holds prior beliefs about the type of the trustee. After the trustor sees the trustee taking action, she updates her beliefs by applying Bayes' Rule (e.g. Rasmusen 1994: 52–56). It says that the trustor, depending on the outcome she observed, increases or decreases the probability with which she up to now assumed to play a certain type of trustee, for instance, a 'good trustee'.

---

[17] Punishment can be costly for the trustor as well since she might forego a possible gain (i.e. $R_1 - P_1$) by withholding trust.

[18] Neglecting the trustor and trustee's network embeddedness prevents other learning effects. For instance, reputation effects are an important source since they serve as a generalized commodity that can spread in a network and help trustors to learn something about a specific trustee (e.g. Burt and Knez 1995; Buskens 1999; Lahno 1995a, 1995b; Raub and Weesie 1990). In a network, a trustee should remember that "[...] one incentive not to cheat is the cost of damage to one's reputation" (Granovetter 1985: 490).

[19] For a more precise game theoretic definition of 'a player's type' and further discussions, see for instance Rasmusen (1994: 48–62).

We use a concept of belief learning, especially in Chapter 3. Beliefs, however, are not updated strictly rationally as is prescribed by Bayes' Rule.[20] A trustor in this book is a belief learning entity but not a strictly rational one such as game theory prescribes. She also holds beliefs on who the trustee is and how he is apt to behave in a given situation. In Chapter 3, we argue that these beliefs constitute the basis of the trustor's social capital in relation to the trustee. Initially, these beliefs reflect experiences from other relationships which are generalized to the current setting. Belief learning then simply refers to the trustor adjusting her social capital stock with each past outcome she observes *and* with each change in the characteristics of the future. Beliefs in our approach are not, as in game theory, purely history dependent.

1.2.2 Overview

We have presented arguments showing why the provision of a surplus, conceptualized as a one–shot Trust Game, is problematic. To facilitate the provision of necessary resources, we then suggested temporally embedding the trustor and trustee in a sequence of similar exchanges. However, we have not yet provided empirical evidence that this kind of temporal embeddedness does indeed facilitate the provision of resources by the trustor. **Chapter 2**, by using evidence from simple dilemma situations such as the Prisoner's Dilemma Game or the Trust Game, sheds some light on the issue of whether temporal embeddedness affects behavior. We do know of some models that typically emphasize the importance of the 'shadow of the future' in affecting behavior in social dilemma situations (e.g. Axelrod 1984; Maynard Smith 1982). However, less is known about how the past, also referred to as 'history' or 'shadow of the past', influences a subject's current behavior in social dilemma situations. To see whether temporal embeddedness can accelerate the provision of resources for the production of a surplus, Chapter 2 first of all investigates whether the past affects current behavior in social dilemma situations. We therefore compare dilemma situations that are equal in all but one aspect: one situation has a history while the other has none. A simple theory is developed that predicts no effects of the past on current behavior in these dilemma situations. We present experimental evidence on behavior in social dilemma situations that is in conflict with this theory, demonstrating that history effects do indeed influence a subject's behavior. Two main elaborations on the theory are suggested, namely, more complex utility arguments and more complex information structures. Since we would not like to challenge assumptions about utility any further, we concentrate in the following chapter on the 'information structure' as a way to learn from the past.

---

[20]Boyd and Richerson (1985) argue that whatever rules subjects follow if they learn, they do not use Bayes' Rule and learn in a completely rational way. The only living beings that have been observed to follow Bayes' Rule are birds and statisticians. However, there are also learning models that propose simple stimulus–response mechanisms (e.g. Bush and Mosteller 1955). For a related approach to learning in games that does not assume fully rational, but still 'sophisticated' players, see Fudenberg and Levine (1998).

Having established that the 'shadow of the past' and the 'shadow of the future' both matter, **Chapter 3** addresses how the provision and division of the surplus are affected in a temporally embedded relation. We conceptualize temporally embedded trust situations as a repeated Trust Game played by two actors without exit options. We simultaneously study the effects of the actors' mutual past and common future on the trustor's behavior. We discuss a simple model of trust depending on a dyadic social capital stock which, in turn, is based on learning and control by the trustor. It is assumed that past outcomes and the expected future are translated into social capital the trustor uses to make her decisions about whether or not resources for the provision of a surplus should be invested. This model makes it possible to derive hypotheses about the behavior of the trustor under different conditions of temporal embeddedness. Laboratory experiments are used to test these hypotheses.

While Chapter 2 and Chapter 3 present analyses from laboratory experiments, **Chapter 4** uses survey data to investigate the effects of temporal embeddedness on the provision of resources for the production of a surplus. We analyze data of buyer–supplier relations in the Dutch IT–business (for the data set, see Buskens and Batenburg 2000). A transaction between the buyer and the supplier is interpreted as a variant of a Trust Game, with the buyer being the trustor and the supplier the trustee. The buyer's investment in the provision of a surplus is mirrored in her investment in managing the transaction: the contract. The more complete a contract is, the more expensive it is for the buyer and, consequently, the smaller the possible surplus, that is, the utility of a working product minus the investment in management. A complete contract leaves little room for opportunistic behavior by the supplier and thus leaves him with a smaller share of the surplus as well.[21] On the other hand, an incomplete contract would save costs on the buyer's side—thus increasing the possible surplus—but also increase the risk of the supplier behaving opportunistically, for instance, by supplying inferior products since this would increase his share of the surplus. The buyer's investment in securing a transaction with the supplier can therefore be seen as a measure of the lack of trust in the buyer–supplier relation. The main question is thus whether or not the buyer should substitute written and legally enforceable but costly contracts for trust and how this decision is influenced by the buyer's and supplier's temporal embeddedness. Predictions on the buyer's behavior in relation to the supplier are based on the theory formulated in Chapter 3. The survey data used in Chapter 4 also make it possible to test whether the results from these data correspond with the experimental results presented in Chapter 3.

[21] A complete contract can be seen as a situation where the buyer places no trust in the supplier. As such, it would correspond to the (P_1, P_2)–outcome in the Trust Game. However, one could see contracts as hostages in the sense of Schelling (1960). The buyer–supplier relation could then be seen as a Hostage Trust Game rather than as a simple Trust Game. See, for instance, Snijders (1990) for theory and analyses of Hostage Trust Games.

1.3 Delayed Agreement and Bargaining

Unlike the first part of the book, the second part (Chapter 5) implicitly assumes that an encounter between two actors provides a surplus and addresses the division of the surplus. In Chapter 5, we discuss a situation in which both actors have a right to determine how the surplus is to be divided between them. An important aspect is that neither of the actors can make a claim on any part of the surplus based on an exogenously given right or on physical power. For instance, neither actor has a claim on the larger part of the pie because he happens to be older, more in need, or in a superior position, for example, being the boss or the trustee in a Trust Game, and thus favored by the structure of the game. We again assume that both actors care about nothing but how much of the surplus they get and that they prefer larger shares to smaller ones.

A feasible and simple formal description of how a surplus can be divided was proposed more than half a century ago. Von Neumann and Morgenstern (1944) introduced a game now known as the *Ultimatum Game* or 'splitting the dollar' (e.g. Binmore 1992). Actor 1, let us call him ego, and actor 2, we call her alter, can divide a dollar between them.[22] The game runs as follows. Ego makes a proposal to alter on how to divide the dollar in whole cents. Alter can then either accept or refuse. If she accepts, the dollar is divided according to ego's proposal. If she refuses, the game ends with alter and ego earning nothing. Assuming the dollar can only be split in whole cents, the game has two *subgame–perfect Nash equilibria*. Note that alter is indifferent between accepting or refusing a proposal of zero cents since she gets nothing either way. The first equilibrium thus prescribes that ego offers nothing and that alter accepts. In the second equilibrium, ego offers 1 cent and alter accepts.[23]

The Ultimatum Game is a simple *bargaining game* which mirrors the kind of real world behavior that can be observed millions of times a day. Whenever someone intends to buy something in a Western store, she plays an Ultimatum Game. The seller writes down a price and the buyer either accepts it or walks away.[24] The game has been studied

[22]Of course, any other surplus is also possible. Ego and alter could divide 100 dollars, 15 euros, or just 24 points. As long as they only care about their share of the surplus (i.e. they are risk neutral and their von Neumann–Morgenstern utility of x dollars is worth x utils), nothing changes.

[23]Note that any split of the dollar in whole cents is supported by a Nash equilibrium. This, however, leaves the question open which of these equilibria should be implemented by two rational and selfish actors? To rule out equilibria based upon *incredible threats*, we demand that an equilibrium is subgame–perfect (Selten 1965). Assume that alter threatens ego to deny any offer smaller than 60 cents. If ego now offers an amount smaller than 60 cents, alter should refuse. Neglecting this offer, however, would be irrational on the part of alter since it would leave her with nothing. Refusing any offer smaller than 60 cents is therefore an incredible threat: a threat a rational alter does not carry out once it is her turn to accept or deny ego's proposal.

[24]Think of the buyer's share of the surplus as the difference between what she is willing to pay and the actual price. The seller's share is then the difference between the price he sells the good for and the minimum price he would or could have sold it for.

in laboratory experiments. The first Ultimatum Game experiment was conducted by Güth, Schmittberger, and Schwarze (1982). They found the modal demand of ego to be half the cake. On average, ego kept a little more than 50% of the cake and no demands for 100% were observed. Other experiments also demonstrated ego's tendency to give away significant fractions of the cake. On average, ego's observed offers were between 40% and 50% of the cake (e.g. Binmore, Gale, and Samuelson 1995; Forsythe et al. 1994; Kahneman, Knetsch, and Thaler 1986; Roth 1995; Roth and Erev 1995).[25] As in the case of the Trust Game or Prisoner's Dilemma Game, scholars seek to explain the subjects' deviation from equilibrium play by means of, for instance, fairness (e.g. Bolton and Ockenfels 2000; Forsythe et al. 1994; Güth, Schmittberger, and Schwarze 1982) or sometimes by means of unintended procedural features in the experiment (e.g. Hoffman et al. 1993).

If a given surplus is divided according to the rules of the Ultimatum Game, how come ego has the privilege of making the proposal on how to split the surplus? Why cannot alter make the proposal and ego accepts or refuses?[26] Of course, ego and alter could toss a fair coin and let chance decide who should have the privilege to submit the proposal. Instead of arguing about who should make the one and only proposal, we could slightly relax the strict procedure of the Ultimatum Game and allow for a counteroffer by alter if she turns down ego's proposal on how to divide the surplus.[27] Two–stage Ultimatum Games of this type have been studied by, for instance, Binmore, Shaked, and Sutton (1985) or Güth and Tietz (1988). Moreover, experimental tests also employed three or five–stage Ultimatum Games (e.g. Neelin, Sonnenschein, and Spiegel 1988) and ten–stage games (e.g. Bolton 1991; Ochs and Roth 1989). In later rounds offers and counteroffers have been found to tend towards a fifty–fifty split of the surplus. For further information and experimental evidence on ultimatum bargaining, I refer the reader to the overview article by Roth (1995) and chapter 5 in Davis and Holt (1993).

Adding a finite number of stages to the Ultimatum Game may reduce some of ego's advantage of making the first and only proposal in the original game. However, why should ego and alter stop making offers and counteroffers after a given number of stages? In the second part of the book, we explain the division of the surplus if ego and alter can negotiate about its distribution as long as they please. In terms of our

[25] Again, the working paper series of the *Institute for Empirical Research in Economics* of the University of Zürich (Bruno S. Frey and Ernst Fehr, see http://www.iew.unizh.ch) gives many more experimental results for the Ultimatum Game.

[26] Why cannot the buyer propose a price she is willing to pay and the seller either accepts and sells, or refuses and keeps his goods? At least among my friends, this is how matters are sometimes settled.

[27] One could argue that it would be possible to alternate the role of the trustor and trustee in the repeated Trust Game as well. It would probably even help to establish mutual trust more easily. Although this is worth investing, it is beyond the scope of this book. A game somewhat similar to a repeated Trust Game where the trustor and the trustee switch roles is the Centipede Game (e.g. Hollis 1998: 14–25; McKelvey and Palfrey 1992).

example, we are not interested in the price of a good in the Western store. Instead, we are now interested in the price of the same good, say a carpet, on sale at an oriental bazaar. The buyer knows the maximum price she is willing to pay. The seller, on the other hand, has a clear conception of the minimum price he will still accept and sell the carpet for. If the price the buyer is willing to pay exceeds the seller's minimum price, a surplus is created and can be divided. In general, the seller and buyer then make offers and counteroffers until the bargaining process eventually leads to an agreement on the price. They do not have to stop bargaining after an exogenously given number of offers. However, bargaining may break down if they cannot agree on a price. Their roads might part without a surplus being divided. Mostly, however, bargaining leads to an agreement with the buyer paying less than she would have been willing to pay and the seller getting more than the minimum price he would have still sold the carpet for. Both get a share of the surplus. The buyer and seller have taken turns making *offers and counteroffers until an agreement was reached.*

Chapter 5 examines the division of a surplus in a game of offers and counteroffers without a fixed number of rounds: roughly speaking, an Ultimatum Game with an infinite number of stages known as the *Alternating Offers Game* (Rubinstein 1982). The game goes as follows. In the first time period, ego starts by making an offer on how to split the dollar. If alter refuses, she makes a counteroffer in the second time period which can then be accepted or refused by ego. If he refuses, it is his turn to propose a division of the dollar in the third time period, which alter can accept or refuse, and so on. Delaying an agreement on the division of the surplus, however, is not to the advantage of ego and alter. They would rather divide the cake now than later. A delay in agreement inflicts costs upon the bargaining actors. These costs are introduced in the model via *discount factors*. The higher an actor's costs, the less patient this actor can be in delaying the agreement and the smaller the actor's discount factor.[28] As in the Ultimatum Game, any split of the surplus is supported by a Nash equilibrium and we can again ask which of these equilibria will be implemented by two rational and selfish actors. Common sense says that 'good things come to those who can wait' and one would thus intuitively expect the division of the surplus to depend on the actors' patience in delaying an agreement. An impatient player can and will not argue long about the division of the surplus and, compared to a more patient actor, is thus content with less. As Rubinstein (1982) shows, this is exactly what happens: the Alternating Offers Game has one unique subgame–perfect equilibrium where ego and alter split the dollar in the first round and the equilibrium split only depends on ego's and alter's discount factors. Binmore (1985: 273) explains that being patient can be

[28]The longer ego and alter argue about the division of their surplus, the less they value their respective shares once they agree on an division. An actor's utility of getting the share x of the surplus at time $t = 0, 1, 2, \ldots$ is discounted using a discount factor $0 < \delta < 1$ and thus taken to be $x\,\delta^t$. Actors who are very patient have a discount factor close to one and rather impatient actors have a discount factor close to zero.

translated into *bargaining power*. In equilibrium, the actor with the higher bargaining power thus receives the larger share of the surplus.

The Alternating Offers Game captures the essentials of bargaining and we use the game in Chapter 5 to capture ego's and alter's negotiation process over a given surplus. The outcome of the negotiation process is reflected in the relation of ego's and alter's bargaining powers, which rest in turn, on their respective discount factors. Concrete predictions on the division of a given surplus require information about ego's and alter's bargaining power. Bargaining powers can of course be assumed by the researcher. This is not really satisfactory. We therefore follow Emerson (1972, 1981) whose idea, roughly speaking, was that an actor's success in exchange depends on his or her embeddedness in a larger network of other actors. Consequently, we use network embeddedness to determine an actor's patience in delaying an agreement. This leads to the first research question of Chapter 5: *How does an actor's embeddedness in a larger network of negotiation partners determine his or her share of the surplus in relations with his or her partners?*

1.3.1 Embedded Bargaining

A network structure is a set of bargaining relations between a set of actors. Say, there are four actors: A, B, C, and D. A network is established if, for example, A is linked to B, B is linked to C and D, and C and D are linked to each other. This network is known as STEM.[29] FIGURE 1.2 depicts this structure. Embedded bargaining simply means that connected actors can engage in exchanges with each other. While Thibaut and Kelley (1959) and Blau (1964) studied social exchange almost exclusively within two–party systems, Emerson (1972, 1981) shifted the focus from dyadic exchange to more complex exchange systems, such as the example above. He stressed that a network structure gives actors an opportunity to participate in several exchange relations. In the STEM structure, for instance, B has three opportunities to engage in exchange, namely, with A, C, or D. Actor A, however, is at the periphery of the network and can only exchange with B. From A's point of view, opportunities are scarce and the network restricts his choices to exchanging with B. Eventually, A is even worse off than in a simple dyad with B since B now has alternatives to an exchange with A. As a result, A's expectation about a lengthy and successful negotiation process with B should be small. A delay in agreement will be very costly for A and A thus tends to be impatient and his bargaining power can be expected to be small.

Emerson (1972, 1981) proposed a theory of social exchange centered around the concept that actors' success in their exchanges depends on the network structure they are embedded in. His point of departure is a simple one. Actors are involved simultaneously in many exchange relations but the actors differ as regards the number of

[29] Actors C and D are structurally equivalent and thus generally referred to as C_1 and C_2; Chapter 5 will adopt this labelling scheme. For the sake of clearness of the example, we nevertheless adhere to the labels C and D here.

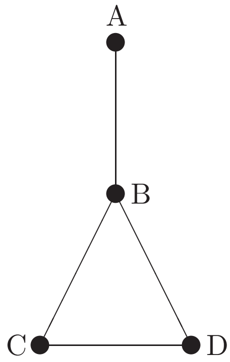


FIGURE 1.2 STEM *network structure where lines indicate possible exchange relations.*

exchange relations they are in. The power of one actor over the other is determined by the dependency of the latter on the former. In Emerson's theory of exchange, opportunities to substitute between partners are the central source of power. Our approach to bargaining power rests on Emerson's idea of dependency. Actors' positions in a given negotiation network can be defined by the number of their possible exchange relations. These exchange relations determine the actors' patience in negotiating the division of a surplus, which translates into bargaining power. In the STEM example given above, A only has B to negotiate the division of a surplus with. B, however, has two other possible exchange partners, namely C and D. Assume B can obtain the same good in all his relations. A is fully dependent on B for exchange but C and D can at least exchange with each other. B is in a very advantageous position and A is in a very unfavorable one. C and D are somewhere in between. B can therefore be very patient in negotiating a surplus. A must be very impatient and C and D can be more patient than A but less patient than B.

However, A is only highly dependent on B for exchange if all of B's partners, from B's point of view, are alternatives. Emerson (1972: 70–71) refers to this type of relations as *negatively connected*. An example is a sports network where connected actors can and want to play squash with each other. A only has B to play squash with. If B and C choose to play each other, because they agree on the 'prize money', A and D have no choice but to stay home. If B plays D, A and C stay home. A is not in an enviable position. C and D, in turn, are in a somewhat better position. They do not depend fully on B for exchange. If B and A decide to play squash, C and D can still play each other. Whatever couples form to play squash, B will always be in the game. B is thus in the most favorable position if the STEM network of FIGURE 1.2 is negatively connected. However, Emerson also argues that relations can be *positively connected*. This means exchange in one relation tends to stimulate exchange in another one. Say B can only exchange with C or D if he obtains a specific good from A first. For instance, A has two squash rackets and B has none. If B wants to play C or D, he first has to exchange with A. Even though A is at the periphery of the network, he is important for exchange in the whole structure. This puts a positively connected A in a better position than a negatively connected A in the same structure.

Summing up, Emerson (1972, 1981) claims that networks are important even if the focus is on dyadic exchange. Networks determine the alternatives or restrictions that produce differences in power. In **Chapter 5**, we use Emerson's logic of dependency in network exchange to define the bargaining power of actors in a negotiation network. The difference in bargaining powers between actors represents differences in the actors' positions in the network and whether the actors are negatively or positively connected (relational assessment). The equilibrium outcome of the Alternating Offers Game informally described in the previous section depends on these bargaining powers. So, the answer to the research question 'who gets how much of the surplus', is the result of a negotiation procedure where offers and counteroffers are sequentially made in bilateral sessions. The outcome of this negotiation process is in accordance with the solution of Rubinstein's Alternating Offers Game.

1.3.2 Overview

Emerson's initial work on network exchange was further elaborated in Cook and Emerson (1978). Eventually, their ideas evolved into the research program of *sociological exchange theories* (for an overview see Willer 1999). These theories adopt the idea that position and relational assessment contribute towards determining an actor's bargaining success. Research on exchange theories thus seeks to explain how structure affects negotiated outcomes between rational and selfish egoists. The focus is generally on negatively connected networks where a resource transfer in one relation prevents transfers in others. To test theoretical predictions, experimental situations are used where actors are placed in an exogenously given network that limits feasible matches between bargaining partners. An example is the STEM structure introduced above. In this type of experiments, connected actors typically negotiate—generally via computer terminals—the division of a surplus of 24 points. The experimental protocol prescribes negotiation rules that closely match alternating offers bargaining. Since experimental subjects are stimulated to maximize their earned points, network exchange experiments match the theoretical approach of embedded bargaining as presented in Chapter 5. To answer our research question and determine the empirical relevance of our approach, Chapter 5 thus uses published experimental data from network exchange theory.

Following Thibaut and Kelley (1959), Gouldner (1960), and Blau (1964), Emerson's primary focus was on social exchange and therefore on reciprocity as the driving force behind exchange. According to Emerson, actors seek to discharge obligations created by the reciprocal behavior of others. They do not want to remain indebted. Moreover, actors try to decrease power differences between each other by initiating further exchange, including exchange with new partners if it diminishes their interdependence. Finally, initially more powerful actors become dependent on less powerful ones if they engage in repeated exchange with the latter. In the long run, dyadic exchange relations thus tend to become balanced. In other words, the actors become equal in power. Emerson claims that if the focus is on exchange networks, this balancing process consequently changes network structures over time. For Emerson, social structure is therefore a dependent variable (see also Cook and Whitmeyer 1992): negotiation outcomes affect the structure.

We start from the assumption that the structure affects negotiation outcomes. Our contribution to sociological exchange theory, however, does not regard exchange as a consequence of reciprocal behavior but merely as a rational and selfish way of negotiating the division of a surplus. After all, such utility maximizing behavior is stimulated in experimental subjects. If a given network structure changes, it is not due to Emerson's balancing process that seeks to minimize power differences between actors. Instead, rational and selfish actors strive to maximize their power—and hence, the power differences between them—by eventually permanently breaking some of their existing negotiation links if it improves their bargaining power. As a consequence, some realized exchange structures may deviate from the given negotiation structures. The first research question pertains to who gets how much of the surplus depending on his or her network embeddedness. In negatively connected networks where exchange in one relation prevents exchange in other relations, a second research question becomes relevant: *With whom of his or her connected partners does an actor divide a surplus if restrictions in a network prescribe that the actor can only divide one surplus at a time?* The network exchange model in Chapter 5 thus also addresses the change of a given network structure due to breaks in certain negotiation links.[30]

1.4 Structure of the Book

The following three chapters all focus on the provision of a surplus and temporal embeddedness. Chapter 2 first of all addresses the question whether the past in any way affects behavior. Experimental evidence from social dilemma situations is used to elaborate on this question. Chapter 3 then uses these experimental findings, combined with the better known idea that the future does affect behavior in current situations, to study how temporal embeddedness influences the provision of a surplus. We conceptualize the provision of a surplus as a Trust Game. Experimental evidence from repeated Trust Games is used to test the hypothesized effects of temporal embeddedness on the provision of a surplus: that is, whether or not trust is placed. Chapter 4 builds on the theory formulated in Chapter 3 but tests the predictions using survey data on buyer–supplier relations in the Dutch IT–industry. The amount of trust the buyer places in the supplier is measured as his investment in managing the current transaction with the supplier. We study how the buyer's and supplier's temporal embeddedness affect this investment in management. Chapter 5 finally addresses the division of a surplus as a game of alternating offers. We focus on networks of bargaining actors and predict

[30]Unlike Emerson (1972, 1981), however, we do not consider situations where actors initiate new exchange relations beyond the ones given in the initial negotiation structure. We thus do not address the growth of network structures.

who gets how much of the surplus in each relation. To test these predictions we use existing experimental evidence from network exchange theory.

All chapters of this book are self–contained and can be read without knowing the content of the other chapters. Consequently, a certain amount of overlap between Chapter 2, Chapter 3, and Chapter 4, all focusing on the provision of a surplus, could not be prevented.

Chapter 2

History Effects in Social Dilemma Situations*

> "For no man giveth, but with intention of Good to himselfe; because Gift is Voluntary; and of all Voluntary Acts, the Object is to every man his own Good; of which if men see they shall be frustrated, there will be no beginning of benevolence, or trust; nor consequently of mutuall help; nor of reconciliation of one man to another; and therefore they are to remain still in the condition of *War*; which is contrary to the first and Fundamentall Law of Nature, which commandeth men to *Seek Peace*."
>
> —— Thomas Hobbes (1651: 105)

2.1 Introduction

Some scholars have argued that a dynamical analysis of trust typically reveals an asymmetry between a slow emergence and a swift decline of trust (e.g. Blau 1964; Coleman 1990; Dasgupta 1988). Intuitively, it seems reasonable that building up a trust relation is more time consuming than ruining such a relation. This is based on the understanding of trust relations as a risky situation for at least one person, namely, the trustor, the person who decides to place or withhold trust (Coleman 1990: 97–99). 'To trust someone' inheres the possibility to become disappointed. Once trust is placed, it can either be honored or abused by the trustee. Therefore, as long as the trustor sees a chance that the trustee abuses trust if placed, she should be 'careful'. Since we assume that trust is a binary decision (i.e. trust is either placed or withheld), 'to be careful' means therefore, that one keeps a probable loss due to wrongly placed trust small (i.e. only place trust if the outcome at stake is small). Taking this risk to trust can be worthwhile, however, since it is the only way to obtain information about the trustee's incentives. Further, placing and subsequently honoring trust secures a payoff to both players which is larger than their payoffs when trust is withheld (surplus). This suggests that the risk a trustor takes by placing trust should increase after the trustee

*A slightly different version of this chapter has been published in *Rationality and Society* 12 (2): 131–162, 2000.

has proved himself trustworthy by honoring placed trust. On the other hand, once betrayed, the trust relation normally ends swiftly. It then has become obvious that the trustee is, in fact, not trustworthy and it makes sense to try to find others, who are more worthy of our trust. Alternatively, one may make use of safeguards (cf. Schelling 1960) which make it unfavorable for the trustee to abuse trust. Unfortunately, we do not dispose of any hard evidence that this theoretically plausible asymmetry does indeed exist empirically.

The assertion about the asymmetry in the emergence and decline of trust is mainly based on two straightforward foundations: the importance of both the past and the future in a trust relation. That is, a relation where an investment of resources bears the risk of an unfavorable outcome for the investor. While the past assumably shapes the experience of the trustor about the efforts and incentives of the trustee and possibly reduces the investments into the trustor's current 'risk management' (e.g. complex and lengthy contracts in business relations or extensive and costly searches for new trustworthy partners), the future will allow the trustor to control and sanction the actions of the trustee (e.g. Buskens 1999: ch. 1). Such future–effects are normally referred to as *reciprocal behavior* (e.g. Blau 1964) or *conditional cooperation* (e.g. Taylor 1987).[1] While these potential effects of the 'shadow of the future' are relatively well known (e.g. Axelrod 1984; Maynard Smith 1982; Trivers 1971), this does not hold for the 'shadow of the past'. Whether subjects take into account past occurrences or not and how these occurrences influence their behavior is not very well understood. An example of an exception, showing that past effects may affect current behavior, can be found in Rapoport (1988). He compared, among other things, the level of cooperation in a Prisoner's Dilemma Game. He assigned the subjects to two groups, one group playing a Prisoner's Dilemma after they participated in another social dilemma situation, the other group playing the Prisoner's Dilemma before they were confronted with the other social dilemma situation.[2] The latter group showed a smaller percentage of cooperation in the Prisoner's Dilemma Game. This finding could, to some degree,

[1]That is, if the past indeed shapes expectations about the (future) behavior of the partner, then ongoing cooperative exchange needs a way to react and/or sanction any violation of these expectations. However, in doing so, the system also needs a method of absolution (Deutsch 1958).

[2]Briefly, this social dilemma situation represented a Tragedy of the Commons game where subjects could, in each round, simultaneously claim a share of a common resource (pool). The game lasted for seven rounds and subjects played in groups of four persons. Subjects were informed about the number of rounds to be played but, while the game lasted, not about the individual shares taken out of the pool. The remaining resource in the pool was doubled after each round and made common knowledge. However, the game only continued as long as the sum of the subjects' shares did not match or even exceed the resource in the pool. Consequently, an empty or overdrawn pool ended the game. A collectively rational strategy would therefore be to equally split the pool in the last round while it would be individually rational to claim the whole pool in the second to the last round, given everybody else cooperates (i.e. claims nothing).

be understood as an effect of the past (i.e. playing another social dilemma game) on the present (i.e. playing the Prisoner's Dilemma).[3]

Using three data sets on experimental social dilemmas, this chapter addresses the question whether we can find any proof that past effects exist. The remainder of the chapter is structured as follows. The next section presents a simple theory which allows to compare a subject's behavior in social dilemma situations with and without a past, where the present and the future are kept constant. The theory predicts similar behavior in both situations. Hence, a possible difference in behavior in the 'non–history' and the 'history' situation has to be due to the past. Section 2.3 presents results from three data sets on experimental social dilemmas. The results show that we indeed find evidence of past effects in all cases. Therefore, the game theoretic models underlying the experiments are ill suited to explain these effects. The last section offers intuitive theoretical reasons why these history effects occur. Moreover, we come up with some suggestions on how to model them.

2.2 History Effects in Simple Games

This section discusses behavior in two simple and similar situations and its implications in connection with past occurrences. These situations will further be referred to as Game A and Game B. Even though the games can be understood in terms of game theory, we abstain from a game theoretic analysis of the situations since it would only add more complexity without bringing about an additional insight into the effect of the past. We will, nevertheless, come back to game theory later in this section (and in the discussion section) and link our 'theory' to game theory under complete information.[4] Let us first in general describe Game A and Game B. In the third section, we will specify the characteristics of Game A and Game B chosen to execute the three experiments.

2.2.1 Situations With and Without a History in General

We define a Game A to be a game of $n_A \geq 1$ players, where every active player disposes of at least two actions, leading to a unique strategy for every player of the game, telling him which action to choose at every node of the game. The game ends with a specific

[3]Some of the early studies on trust also addressed, to some degree, the question whether or not the past affects the present. Boyle and Bonacich (1970), for example, found support for their assumption that subjects' current expectations about the outcome in a Prisoner's Dilemma Game will change in the direction in which experience is different from their previous expectations. In the same vein, Deutsch (1958), amongst other findings, shows that an experienced benevolent rather than malevolent treatment in the past leads subjects to be more likely to act benevolently when in charge of the outcome of a Prisoner's Dilemma Game. Finally, Swinth (1976), based on evidence from two–times repeated Prisoner's Dilemma type games, argues that a commitment period in which good intentions can be signalled is enough to establish trust.

[4]The data at hand is based on simple game theoretic models but can, for our purpose, be analyzed without applying the tools of game theory.

payoff for every player in the game (i.e. active as well as inactive players). We assume the game to be one of complete information (e.g. Rasmusen 1994: ch. 2).

Game B resembles the situation captured by Game A but is characterized by **preceding** actions of $k \leq n_B$ players, with $n_A \leq n_B$ allowing for the introduction of new players in Game B. This will extend Game A such that it becomes a subgame of the larger Game B, now referred to as Game A_B (read: "Game A in B").[5] We require that Game A_B and Game A are the same in terms of players, actions, and payoffs. Hence, the 'preceding actions' in Game B can extend Game A in two ways. First, Game B may be a repeated version of Game A. Second, for at least one player the 'preceding actions' are actions which were not available to him in Game A and which give him the possibility to form other strategies.[6] Therefore, Game B becomes a game where Game A_B forms one of several possible subgames.

What we assume throughout the whole chapter is that every player has all information necessary for a decision in every node of Game A and Game B and also applies it with perfect rationality.[7] Hence, a player's decision on what strategy to play in Game A as well as in Game B should only be based on his expected payoffs.

The 'extension' of Game A which makes up the larger Game B as discussed above is precisely what is considered to be the 'history' of the game. Loosely speaking, it denotes (i) anything which has occurred before subgame Game A_B was reached, and (ii) anything which could have happened if Game A_B had not been reached. Again, we assume that the history is known to all players.

What we are interested in is the comparison between the strategy chosen by a player i in Game A and the strategy chosen by (the same) player i in Game A_B. Note that this means that Game A and Game A_B are not only similar with regard to players, actions, and payoffs, but also with regard to i's decision problem. Since we are solely interested in the comparison of i's decision in Game A and Game A_B, we do not have to discuss the other possible subgames of Game B.[8] Following, we sketch a simple argument which generally describes what a player i should do in a situation which resembles a Game A and a Game B. As we explain in some detail below, behavior in Game A and Game A_B should be identical.

---

[5]For a formal definition of a subgame see, for example, Rasmusen (1994: 94). Loosely speaking, a subgame is the game which follows the history of the (whole) game played so far and requires that the subject who moves next precisely knows what has been 'going on' so far in the game.

[6]Note that we do not exclude the possibility that this can also be a new player, not having played Game A at all.

[7]It will be sufficient to define rationality as maximizing one's own interests, that is, choosing the alternative which yields the largest utility. When discussing the theory and the experiments, we assume that the utility for any actor increases (decreases) with a higher (lower) payoff to himself and is independent of the payoffs of the remaining actors (i.e. 'utility is own money').

[8]Even though under a game theoretic analysis the characteristics of these subgames are important in finding the equilibria of Game B.

2.2.2 A Simple 'Theory' on History Effects

As explained above, Game A is similar to the subgame Game A_B of a larger game Game B. Therefore, the only difference between Game A and Game A_B is that the latter represents a situation with a history. The important point to make is that this history brings about *no further information* necessary for a decision by i once he reaches subgame Game A_B. He ends up in the same situation we denoted as Game A, facing the same players, actions, and payoffs. Therefore, an actor i has precisely the same information to base his decision on in subgame Game A_B as he would have in Game A where no preceding actions take place. From this point of view, the incident which forms the history of Game B is irrelevant for a decision of i in Game A_B. Therefore, the behavior of an actor i in subgame Game A_B should be the same as in Game A. We find the same argumentation in Harsanyi and Selten (1988). In connection with subgame consistency, they remark: "After all, once the subgame has been reached all other parts of the game are strategically irrelevant." (p. 90).[9] We already put forward the arguments leading to this conclusion, but let us in short summarize them.

- Actor i actually faces the same alternatives, leading to the same payoffs whether he is in a Game A or in a Game A_B. Therefore, Game A and the subgame Game A_B are similar with regard to i's decision problem.
- Actor i is rational and selfish and, therefore, his behavior is only contingent on the payoffs of the game.
- The action(s) which extend Game A to become Game B form what we call the history of the subgame Game A_B.
- The information i gains from the history is irrelevant for his decision in subgame Game A_B since it is not reflected in the payoffs of Game A_B.
- Hence, i's behavior in Game A and Game A_B should be the same.

However, let us briefly mention one aspect of i's rationality. It is important that *all* actors are opportunistic in the sense that they seek to improve their own monetary

[9] A note on the difference between subgame perfectness and subgame consistency seems appropriate here. Especially since our argumentation about 'no difference between Game A and Game B should be found' actually hinges on the criterion of subgame perfectness (Selten 1965) but we have just cited Harsanyi and Selten (1988) on subgame consistency. However, subgame perfectness and subgame consistency differ to some extent. Even though related, subgame consistency applies whenever a game has no proper subgames and, in addition, Bayes' Rule cannot be applied since at least one information set is not reached with positive probability in the course of the game. For weaker notions of consistency (i.e. if Bayes' Rule can be applied), we speak of a perfect Bayesian equilibrium (e.g. Kreps 1990b). In short, every subgame–perfect equilibrium is also a Bayesian equilibrium, and every Bayesian equilibrium is also subgame consistent. Harsanyi and Selten (1988: 90ff, 344ff), however, do not make a difference between the definition of a subgame and a proper subgame. Together with the fact that Bayes' Rule can always be applied to the situations we focus on, there is no difference between subgame perfectness and consistency with respect to the simple games under consideration.

utility. Otherwise it is possible that the history might indeed disclose some relevant information to i (e.g. about some other actor(s) having positive sentiments towards the i–th actor).

What we are further going to do in this chapter is compare situations which resemble a Game A with situations which resemble a Game B and Game $\mathsf{A_B}$, respectively. That is, situations where a subject faces the same decision in both games but 'bears the burden' of an irrelevant (from a rational decision point of view) history in Game B. Hence, these situations are theoretically identical and should therefore lead to the same outcome.

But what if we indeed find different behavior in those two situations? Since the situations reflect a simple decision problem and the only difference between them is the history of Game B, we have to explain possibly different behavior in Game A and Game $\mathsf{A_B}$ by the existence of its history. If the experiments analyzed in Section 2.3 do show different behavior in Game A and Game $\mathsf{A_B}$, we need to take into account that people may not only base their decision on the payoffs alone. But what would a theory say about Game A and Game B which assumes that actors are fully rational in their behavior?

2.2.3 Game Theory and Its Prediction for Game A and Game B

We have so far considered a theory whose assumptions come close to game theory under complete information.[10] We will now lay out under what conditions (a variant of) game theory would predict the same behavior in Game A and Game $\mathsf{A_B}$. However, the aim of this chapter is not to make a contribution to game theory but to analyze simple experimental situations based on game theoretic understanding.

- If we assume that Game A and Game B are games with complete information such that there exists a unique (subgame perfect) equilibrium, Selten's (1965) theory of subgame perfection trivially predicts that the behavior in Game A and Game $\mathsf{A_B}$ is indeed the same (see, e.g. Rasmusen 1994: 93–95).

- If there are multiple (subgame perfect) equilibria in Game A, this prediction no longer holds. Predictions would then depend on the equilibrium selection arguments to be applied to the games. Consequently, it need not be true that the same equilibrium is predicted for Game A and Game $\mathsf{A_B}$ (for an example, see Rasmusen 1994: 95–96). The existence of multiple (subgame perfect) equilibria especially becomes a problem if Game A itself is a (in)finitely repeated game (Folk Theorem, see e.g. Rasmusen 1994: 123–126).

- Some authors have discussed the validity of subgame perfectness, which is especially relevant in the case where the subgame Game $\mathsf{A_B}$ is not on the equilibrium

[10] See for instance Rasmusen (1994: ch. 2) for a discussion of information requirements in games and for an introduction to game theory in general.

path of Game B (e.g. Kreps 1990b: 432–437). What should rational actors conclude when they arrive at Game A_B although they know they should not have? Subgame perfection is a strictly forward–looking refinement of Nash Equilibrium that assumes that players believe they arrived at Game A_B 'by coincidence' and do not take into account whether it is at all likely that they end up in that particular decision node (for a discussion on Nash refinements, see, e.g. Kreps 1990b: 417–443). The theory of forward induction (e.g. Kohlberg and Mertens 1986) which was developed in a context of signaling games (e.g. Rasmusen 1994: 165ff), argues that rational actors would in fact modify their beliefs when 'impossible' (i.e. probability zero) events indeed occur. Hence, this would very much question whether the predictions for Game A and Game A_B would be the same or not.

If we seek to model games like Game A and Game B using game theory, the question is of course what would be an appropriate model? If we would model the situations using games with incomplete information, game theory would not necessary predict any longer that players should behave the same in Game A and Game A_B. Under the assumption of incomplete information, players may, for example, learn something about the incentives of each other from playing the part of Game B that leads to Game A_B.

2.3 Empirical Analysis of History Effects

In this section, we present an analysis of three different data sets. Even though not specifically designed to test history effects, they are nevertheless well suited for our purposes. The setup of all three experiments is similar. Each employs a simple game theoretic representation of a social dilemma situation which allows for a Game A–Game A_B comparison. If history effects do not play a role in affecting the subjects' behavior, we should then not find different results in any of the Game A and Game A_B under study in this section.

2.3.1 Experiment 1 (Snijders and Keren 1997)

This experiment tests the most simple case we would regard as 'history effects'. It will be checked whether subjects play differently if confronted with a simple choice out of two alternatives (Game A) or the same choice, but with a history (Game A_B). Before we describe the games in more detail, we first lay out some characteristics of the experiment.

Background: The experiment was conducted by Snijders and Keren at Utrecht University. The results presented here are based on one out of five experiments executed. For an overview of all experiments see Snijders and Keren (1997).
Subjects: 172 students from Utrecht University participated in an experiment reflecting a Game A and a Game B. For their participation, students were paid NLG 11 (Dutch Guilders, approximately € 5 or US$ 6 at the time of the experiment). The experiment

lasted for about 45 minutes. The participants received written information as well as a graphical representation of the games.

Setup: Participants were randomly assigned to one of five conditions, from which we only consider the relevant four. Two games represent a Game B whereas the other two conditions represent a Game A, identical to the subgame Game A_B of Game B. We will further refer to these game–couples as Game A–1 and Game B–1 and Game A–2 and Game B–2, respectively. The games are depicted in FIGURE 2.1:[11]

Game A–1: This game represents a decision situation where actor j is asked to split an amount of money either equally between himself and actor i (choosing right) or in favor of himself (choosing left). Subjects were asked whether they would choose right or left as actor j. Note that this game it the Dictator Game discussed in Chapter 1.

Game B–1: This game represents the same choice situation for actor j as A–1 but this time, actor i is an 'active' player. He can choose between splitting a smaller amount of money equally between himself and j (choosing left) or let j split a larger amount of money (choosing right). Players i and j move sequentially, i before j. Again, subjects were asked whether they would choose right or left as actor j. Note that if i chooses right, j is confronted with Game A–1. Game B–1 is the Trust Game discussed in Chapter 1.

Game A–2: This game represents a decision situation for actor k where he is asked to split an amount of money either equally between himself and the actors i and j (choosing right) or in favor of himself (choosing left). Subjects were asked whether they would choose right or left as actor k. This is a Dictator Game with three actors.

Game B–2: B–2 represents the same choice situation for actor k as described by Game A–2 but this time, actors i and j are 'active' players. Actor i chooses to either split a smaller amount of money in favor of himself (choosing left) or lets j decide to split this amount of money (choosing right). Actor j can then choose between splitting the amount of money in favor of himself (choosing left) or let k split a larger amount of money (choosing right). Players i, j, and k move sequentially, i before j and j before k. Again, subjects were asked whether they would choose right or left as actor k. This is a Trust Game with three actors.

Design: The setup of the experiment, assigning the students randomly to one of five groups, leads to a 'between–subjects' design in which all students only played either a Game A or a Game B. This allows for two 'Game A–Game A_B' comparisons of the following kind:

- Situations described by Game A–1 and Game A–2 are simple decision situations for the subjects.
- Situations described by Game A_B–1 and Game A_B–2 (i.e. the subgames of Game B–1 and Game B–2, respectively) denote precisely the same decision situation for

[11] In FIGURE 2.1 and all following figures, payoffs following from a certain combination of actions are represented as vectors. Those vectors should be read top–down, first indicating player i's payoff, then player j's payoff, and, at the bottom, player k's payoff.

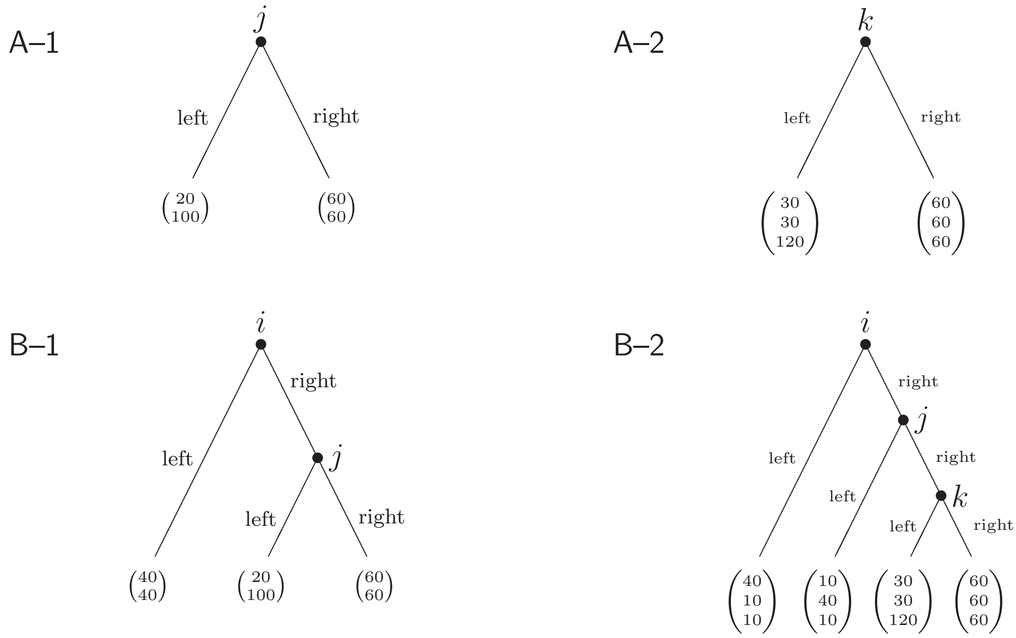


FIGURE 2.1 *Four games representing a choice situation for j in Game 1 and k in Game 2.*

the subjects as in Game A–1 and Game A–2 but this time with a history. The history consists of the moves by i (and j, respectively).

- Subjects (playing as actor j in Game B–1 and as actor k in Game B–2) are able to conclude from the game played so far that i (and j) must have chosen 'right' to give them the possibility to actually perform an action (i.e. choose between 'left' and 'right'). This history is known to all actors. According to our theory, this 'shadow of the past' should neither influence j's (Game B–1) nor k's (Game B–2) decision to choose between right and left since their decisions should only be based on the payoffs of the game. And these payoffs are precisely the same as under conditions Game A–1 and Game A–2, respectively.

- Hence, based on our simple theory, we expect the same percentage of subjects choosing right or left in Game A–1 (Game A–2) and Game B–1 (Game B–2).

Before we discuss the results of the experiment, let us in short lay out the game theoretic solution to these games.

Game Theory: Game B–1 is a two–stage Trust Game (Dasgupta 1988) with a unique subgame perfect equilibrium where both i and j choose left and both earn 40. Game A–1 represents the only (proper) subgame of this Trust Game, starting with the node where j has to make a decision.[12] Game B–2 is a three–stage Trust Game which follows the same logic as the (two–stage) Trust Game (i.e. Game B–1). It also has a unique

[12]Note that Game A–1 is a Dictator Game. Normally, in a Dictator Game j would be given the possibility to split an amount of money as he pleases. Here, he is given the possibility to split the money equally between him and actor i or in favor of himself.

TABLE 2.1 *Choosing 'Right' in games representing situations with and without a history in experiment 1.*

	A–1—B–1 Comparison		A–2—B–2 Comparison	
	Game A–1	Game B–1	Game A–2	Game B–2
Right	35.5%	73.5%	25.8%	60.0%
N_{Right}	11	25	8	21
N_{total}	31	34	31	35

Notes: Comparison of behavior shows significant differences between Game A–1 and Game B–1 ($\chi^2(1) = 9.4986,\ p < .005$) and between Game A–2 and Game B–2 ($\chi^2(1) = 7.8630,\ p < .01$).

subgame perfect equilibrium where i, j, and k choose left. This leads to a payoff of 40 for i and 10 for j and k each. Again, Game A–2 presents the last subgame of the three–stage game, starting with the node where k gets to choose an action. We assume that the utility for an actor increases (decreases) with a higher (lower) payoff to himself and is independent of the payoffs of the remaining actors.

Even though game theory predicts that j in games A–1 and B–1, respectively k in games A–2 and B–2 should choose 'left', we do not focus on testing this game theoretic prediction.[13] We only want to stress that if someone chooses 'left' in the subgame–only version, he should—under mild conditions—also choose 'left' if confronted with the whole game tree. The same holds true for someone choosing 'right'. Put differently, we want to pinpoint the fact that the history should not influence the decision to choose between 'left' or 'right'.

Results: The relevant results of the experiment are summarized in TABLE 2.1. Contrary to what our theory predicts, subjects who were confronted with a game and its history (Game B–1 and Game B–2) showed a significantly higher percentage of choosing 'right' (73.5% and 60% versus 35.5% and 25.8%) than did the subjects who played Game A–1 and Game A–2, respectively. This finding is neither in line with our theory nor with game theory, according to which everybody should have chosen 'left'.

Discussion: Even though the results go against our theory, there seems to be a common sense explanation, stemming from the history of the game. It seems reasonable to assume that the history of the game contains some information which is not reflected in the payoffs of the B–games. Even though a rational actor should, according to our theory, base his decision on the monetary payoffs and neglect anything not represented

[13] If we deal with strictly rational players, no i in B–1 and B–2 (and, additionally, also no j in B–2) should ever choose 'right'. Any j in B–1 and k in B–2 should thus be very well aware that they will never have the chance to move. The reader may now argue that this resembles the situation of an 'impossible' event where Kohlberg and Mertens' (1986) forward induction theory applies. However, since the situation is one of complete information, forward induction is irrelevant and no j in B–1 (or j and k in B–2) would want to change his strategy of always choosing 'left' just because i in B–1 (or i and j in B–2) for whatever reason now played 'right'. Even though games A–1 and A–2 are not on the equilibrium path of the larger games B–1 and B–2, respectively, it does not affect our conclusions.

by them, this seems not to be what happened here. In the situation described by Game B–1 and Game B–2, subjects can easily see that they are dependent on the decision of one, respectively two, other actors. In Game B–1, the other actor could have chosen 'left' to secure himself a 'safe' payoff since he can foresee that a rational subject should always choose 'left' after he has chosen 'right', which would leave himself with the worst possible payoff. The same, of course, holds true for Game B–2, where i and j should choose 'left' if confronted with a rational k who goes for 'right' if the subgame is reached. Any subject in the role of j in Game B–1 and k in Game B–2, respectively, sees that if he indeed reaches the node where he is to decide how to split the money, the other actor(s) must have chosen 'right' and thereby running the risk of a possible 'loss', namely, the difference between what they could have gotten by playing 'left' in the first place and what they get if j in Game B–1 (k in Game B–2, respectively) plays 'left'.

What might have happened is that the subjects honor this outcome (viz. reaching 'their' subgame) by reciprocal behavior. Many subjects turned out to be fair enough to split the amount of money equally instead of grabbing the bigger piece of the cake in the B–games. The situations without a revealing history (exposing the fair character of the other players) leave no room for reciprocal behavior since the situation only asks to split the money (un)equally between oneself and an unknown other (or two unknown others, respectively). It seems that subjects are more selfish in situations where they do not know anything about the character of the other subject(s). Why should I share the pie equally if I do not have any evidence that the other is worth that gesture? Put differently, the results speak in favor of some history dependent altruism (or history driven reciprocity) but would question the existence of a general norm of reciprocity (Gouldner 1960). If the latter indeed exists, we would then be in need to explain why 26.5% and 40% of the subjects in Game B–1 and Game B–2, respectively, played uncooperatively?

As this simple experiment shows, the history of the game is not only capable of revealing some information (about the other players, in the situation at hand) not reflected in the payoffs, but also that the subjects do indeed take this additional 'piece of evidence' into account when choosing which alternative to perform: a first hint that our theory may need revision.

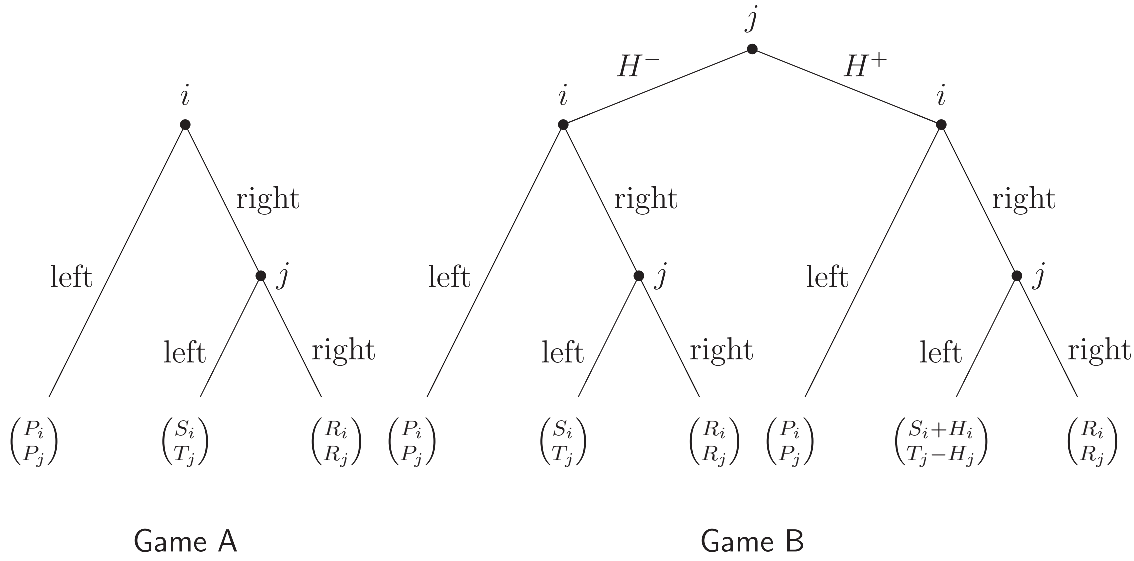
2.3.2 Experiment 2 (Snijders 1996)

Compared to Experiment 1, the history of the game presented here is a bit more complex. Nonetheless, from a rational point of view, it will again be irrelevant for a proper decision. Let us again first discuss the setup of the experiments.

Background: The experiments were conducted by Snijders at Groningen University and Amsterdam University. The results presented here are only a small fraction of several experiments by Snijders. With regard to a Game A–Game B comparison, only two experiments are of relevance. For more details, see Snijders (1996: 65–84).

Subjects: 106 students from Groningen University and 102 students from Amsterdam University participated in an experiment reflecting a Game A and a Game B. The subjects were promised a minimum of NLG 10 (approximately € 4.50 or US$ 6 at the time of the experiment) for participating. In addition, about 15% of the subjects would earn more (up to a maximum of NLG 120 or € 54.50) depending on their performance. All participants were orally instructed by the experimenter about the task to perform. In addition, they also received written instructions and a graphical representation of the games to be played.
Setup: Each subject in Groningen and Amsterdam played five times a Game A and later five times a Game B. The A– as well as the B–games only differed in terms of their absolute payoffs. This does not change the characteristics of the games since the ordinal ranking of the payoffs assigned to each combination of alternatives remained the same in each Game A as well as in each Game B. Subjects were asked to make a decision in each Game A and Game B as an actor i as well as an actor j. To be able to make an appropriate comparison with respect to the history effects, we are only interested in the participants' behavior as actor i. Game A and Game B are depicted in FIGURE 2.2:
Game A: This game represents a decision situation for actor i where he is asked to split an amount of money equally between himself and another actor j (choosing left) or handing over the decision to split a larger amount of money to j (choosing right). Actor i knows that actor j must then choose between two alternatives, namely, splitting this larger amount of money equally between himself and i (choosing right) or in favor of himself (choosing left). Subjects were asked whether they would choose right or left as actor i.
Game B: This game again represents the same choice situation as described by Game A, but preceded by an additional move by actor j. He has to decide whether or not to 'post a hostage', a good over which j voluntarily renounces disposal and which changes (due to its value) his incentive structure such that opportunistic behavior on his behalf becomes unattractive (or at least less attractive).[14] If posted by j, the hostage is lost (i.e. goes to a third party or the opponent) in case of opportunistic behavior by j. If j does not post a hostage, subgame Game A_B will be played. If j does post a hostage, a 'new' subgame Game C_B will be played which differs in payoffs compared to Game A_B if i chooses right and j consequently chooses left. The relevant question for our purpose is: "If you were actor i and actor j had not posted a hostage, would you then choose left or right?".
Design: The setup of the experiment, letting each subject play (five times) Game A as well as (five times) Game B leads to a 'within–subjects' design. A 'Game A–Game B' comparison of the following kind can be made:

[14] See, for example, Schelling (1960) for an intuitive discussion of hostages in social interactions, Raub and Keren (1993) for a theoretical analysis and prior experimental evidence, and Raub and Weesie (2000a) and Weesie and Raub (1996) for a general discussion on hostages and incentive problems.



Value of hostages: $H_i = 0$, $H_j > 0$ in Groningen, and $H_i = H_j > 0$ in Amsterdam.

FIGURE 2.2 *The two games representing a choice situation for i in experiment 2* $(T_j > R_i = R_j > P_i = P_j > S_i)$.

- Situations described by Game A are simple decision situations for the subjects playing as actor i.
- Situations described by Game B reflect the same decision situation for the subjects (playing as actor i) as in Game A if actor j does not post a hostage. The hostage–option for j forms the history of the subgame Game A_B.
- A subject playing as actor i knows that a posted hostage would reduce j's temptation to choose left (i.e. splitting the money in favor of himself) after he has chosen right, that is, letting j split the money. They could both benefit from the situation compared to the situation where no hostage is posted. If j does not post the hostage the choice situation for i remains the same as in the Game A.
- Hence, all subjects choosing right (left) as actor i in Game A should also choose right (left) in Game B where j does not post a hostage.

We present the game theoretic solution of the two games before we discuss the results of the experiments.

Game Theory: Game A is a (two stage) Trust Game such as discussed in the first experiment (there it was named Game B–1). Due to the ordinal ranking of the payoffs (see FIGURE 2.2), it has a unique subgame perfect equilibrium where i and j choose left and both earn the punishment–payoffs P_i and P_j. Game B is a Hostage Trust Game with a subgame perfect equilibrium depending on the hostage value. If the value of the hostage is smaller than the gain from defection, $H_j < T_j - R_j$, the subgame perfect equilibrium is such that j does not post a hostage and consequently plays left while i also plays left. The outcome is the same as in Game A, both earning P_i and P_j, respectively. On the other hand, if $H_j \geq T_j - R_j$, j will post a hostage and consequently play right, while i also chooses right. Both players will earn the reward payment R_i and R_j, respectively.

Again, as in Experiment 1, the relevant task here is not to test the game theoretic equilibrium. Since i was asked to make a decision in Game A_B under the assumption that j refrains from posting a hostage, the choice situation for i is the same as in the 'two–stage' Trust Game, that is, Game A. We want to know whether not posting a hostage in Game B leads the subjects playing as actor i to change their strategy compared to Game A. Our theory would predict that who chose left (right) in Game A should have also chosen left (right) in Game B as well.[15]

Results: The results of this experiment are summarized in TABLE 2.2. The predictor 'Game A_B' shows a negative effect, disclosing that less subjects (actors i) chose to play right in the subgame Game A_B when no hostage was posted by j compared to the otherwise similar Game A. The column 'Marginal effect' reveals that the estimated probability of playing right decreases by 0.10 to 0.03 (0.13 − 0.10) in Game B.[16] Put differently, the data unveils a tendency to trust less in a Hostage Trust Game where no hostage was posted than in the Trust Game. According to our theory, the coefficient of the predictor 'Game A_B' should actually be zero.[17] Since five Game A and five Game B with different payoffs were played, it might be relevant to check whether the magnitude of the payoffs might have influenced the behavior of the subjects playing the part of actor i. In short, the predictor TEMPTATION (roughly, monetary gain for j from playing 'left') says that the more j would earn by playing 'left', the less willing should i be to play 'right' in the first place. The regression coefficient for TEMPTATION, however, shows no effect of such a monetary gain for j on the behavior of the subjects in the role of i. The predictor RISK (roughly, monetary gain given up by i if playing 'right'), on the other hand, shows a statistically significant and negative effect on choosing 'right'. The higher a possible loss becomes, the less likely it is that i hands over the decision to split the money to j. However, in connection with our analysis, it is necessary to mention that the influence of both predictors (i.e. TEMPTATION and RISK) on i's behavior turns out to be quite similar in the analysis of Game A (cf. Snijders 1996:

[15] Under some of the conditions in Snijders' (1996) experiments (i.e. payoff matrices and hostage values such that $H_j \geq T_j - R_j$), it would indeed have been rational for an actor j to post a hostage. Note, however, that this is not the point of interest here. Moreover, since subjects in the role of actor i were literally asked to make a decision given that j has refrained from posting a hostage, a game theoretic argumentation about actor j's rational behavior regarding hostage posting is beyond the scope of the chapter and, thus, irrelevant for the further discussion.

[16] On how to calculate marginal effects in logistic or probit regression see, for instance, Long (1997).

[17] Due to the 'within–subjects' design of the experiments, we are able to compute the size of the history effect (i.e. the coefficient of the predictor Game A_B). In Experiment 1 and its 'between–subjects' design, we could only show that an effect stemming from a mutual history exists (see Snijders (1996: 72–74) for a discussion of the topic).

TABLE 2.2 *Probit analysis that actor i chooses right in* Game A_B *of* Game B.

Predictors†	Coefficient	p–Value	Marginal effect
Game A_B	-0.48**	0.00	-0.10
RISK	-2.08**	0.00	-0.44
TEMPTATION	-0.02	0.95	
Constant	-0.62	0.33	0.13
$N = 1883$, Pseudo $R^2 = .29$, $L = -668.3$			

Notes: ** = significant at 1%–level

† Only relevant predictors of TABLE 6.5 (Snijders 1996: 158) are reported here.

99, Table 4.5). The smaller probability of choosing 'right' in Game A_B compared to Game A is, therefore, consistent across different levels of payoffs.[18]

Discussion: As already seen in Experiment 1, history affects behavior. Due to the 'within–subjects' design of the experiment, we can even compare the decisions of the same subjects. The theoretically unimportant effect of a hostage that 'could have but was not posted' negatively affects the choice to perform the action 'right'. By not posting a hostage, actor j seems to become less trustworthy in the eyes of an actor i. Subjects in the role of actor i may have had a reason to conclude from j's decision that such behavior eventually prevented a more positive outcome for both actors (i.e. both earn the reward payoff R). In other words, a given hostage would have been a signal of trustworthiness, but the signal 'no hostage' might work the other way round.[19] Moreover, not posting a hostage when one could have might even be a stronger and more distinct signal of untrustworthiness than posting a hostage can be a signal of trustworthiness. It seems again clear from this experiment that the payoffs alone are not the sole basis for the subjects' decision which alternative to choose. Not placing a hostage seems to contain important information not reflected in the payoffs.

[18] All other predictors from Snijders' (1996) analysis are omitted in TABLE 2.2. Their regression coefficients are also roughly similar in both games, revealing that they can neither be responsible for the fact that less subjects chose right in the subgame Game A_B (for a more detailed discussion of the results, see Snijders [1996: 96–100 and 157–159, especially Table 4.5 and Table 6.5]). One might, however, argue that the probability of actor i choosing 'right' in Game B decreases with an increasing value of the not posted hostage. Even though the probit–coefficient confirms this assumption, it is by no means statistically significant ($p = 0.34$). The fact that the i–th actor chose to play 'right' less often can, therefore, be explained solely by j's binary decision whether or not to post a hostage.

[19] Hostages are apparently only one possible form of commitments in mixed–motive games. There exists experimental evidence that *communication effects* serve a similar purpose. That is, verbal (and even non–verbal) communication without a binding and enforceable character can facilitate the development of cooperation and trust in a proceeding mixed–motive game. The effect of communication is stable in two–person as well as multi–person games. For a short overview see Colman (1995: 141–142, 218–221).

2.3.3 Experiment 3 (Mlicki 1996)

The last experiment we discuss runs along the same lines as Experiment 2. The main difference lies in the fact that another dilemma situation is reflected in the games. Let us again start by focusing on the setup of the experiment.

Background: The experiment was conducted by Mlicki, Raub, and Weesie at Utrecht University. The experiment was originally designed to test the influence of different factors (e.g. transaction costs) on posting a hostage, and hence, on the equilibrium outcome of the games (for a discussion and results, see Mlicki 1996). Manipulations were partly present in what we consider to be Game B but in none of the Game A's. By only taking into account those Game B's where the manipulations were omitted, we are able to use the experiment for a proper 'Game A–Game B' comparison.
Subjects: 216 students participated in an experiment reflecting a Game A and Game B situation. The subjects were promised at least NLG 11 (approximately € 5 or US$ 7 at the time of the experiment) for participation in the study. Depending on their performance they could make more than NLG 20. The experimental games were played on computers hooked up to a network and subjects were instructed that they would play against another subject on the network. In fact they unknowingly played a preprogrammed computer which always played the equilibrium strategy according to game theory. Instructions were partly displayed on the computer screen and partly given on paper. In addition, an experimenter was present in the computer lab for assistance. Due to the fact that we only consider subjects who played the Game B without any of the manipulations (e.g. transaction costs) present, we are restricted to $N = 23$ subjects, a fairly small number of subjects compared to the first two experiments.
Setup: Each subject played a Game A and later a Game B. Participants were only told that they would play a Game B against the same other player after Game A was played. As in the preceding experiments, Game A_B is a subgame of Game B reached by a specific history. The social dilemma situation represented by the games is completely symmetric with regard to the players' actions and payoffs. In addition, they also had to make their decision simultaneously. Therefore, actors i and j are actually interchangeable. The games are depicted in FIGURE 2.3:[20]
Game A: This game represents a decision situation where i has to choose between two alternatives, left and right. The same holds true for j. Since i and j move simultaneously, they do not know what the other chooses to do, but they do know what the opponent's options are. If both choose 'left' they split an amount of money equally. If both choose 'right' they split an amount of money equally which, however, is larger than the amount they split by mutually choosing 'left'. If i chooses 'right' ('left') and j chooses 'left' ('right'), i (j) receives the lowest payoff possible and j (i) earns the largest payoff possible. Subjects were now asked whether they would choose

[20] The dotted lines in FIGURE 2.3 represent 'information sets', the set of different nodes in the game tree that j knows might be the actual nodes but cannot with certainty distinguish between them. In the games of FIGURE 2.3, j's information sets are used to represent simultaneous moves by i and j.

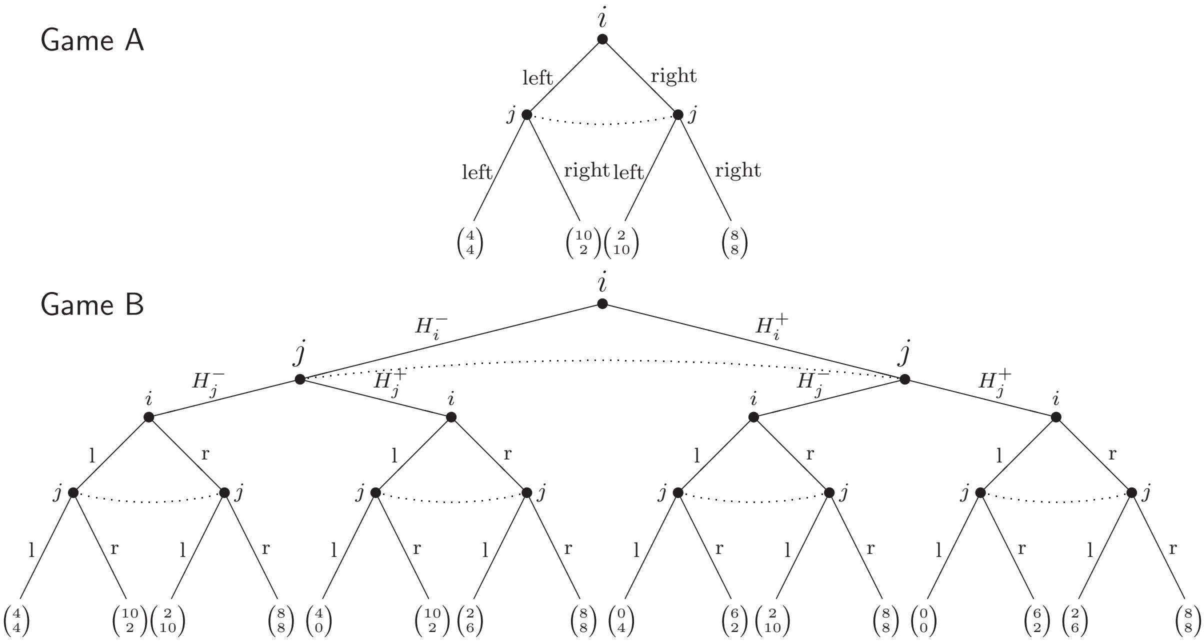


FIGURE 2.3 *The two games representing a choice situation for i in experiment 3.*

right or left if they play as an actor i.
Game B: This game represents the same decision situation for actor i as described by Game A but this time preceded by an additional move by i and j. Both actors can decide whether or not to post a hostage. If neither of them posts the hostage, the subgame Game A_B will be played. If at least one of the two actors decides to post a hostage, a different subgame will be reached.[21] After i and j have simultaneously decided about their hostage option, decisions will be made public and they again have to decide whether to play 'right' or 'left'. We are interested in the decision of those subjects i who themselves denied to post a hostage (the computer never posted a hostage under the condition we examine here).
Design: Letting each subject play a Game A and a Game B leads, as in Experiment 2, to a 'within–subjects' design. A comparison of the following kind can be made:

- The situation described by Game A is a simple choice situation for actor i (where his payoffs are also dependent on j's behavior).
- The situation described by Game B reflects the same decision situation for actor i as in Game A if neither j nor i post a hostage. The hostage option of both players forms (part of) the history of the game. To be precise, the history of Game A_B is twofold. On the one hand, it is given by the 'first' Game A, played

[21]Actually, the combination of posting and not posting a hostage leads to four different subgames, namely, Game A_B if neither i nor j post a hostage, Game C_{B1} and Game C_{B2} if only one actor posts a hostage and Game C_{B3} if both post a hostage. Again, we are only interested in those Game B's where the history leads to the subgame Game A_B.

immediately before Game B and, on the other hand, the possibility to post a hostage in Game B.[22]

- Actor j (viz. the computer program) never posts a hostage. Hence, each actor i who does not post a hostage makes sure that the subgame of interest, namely, Game A_B, is played.

- Hence, all subjects choosing right (left) in Game A should also choose right (left) in Game A_B after they denied to post a hostage.

Before we present and discuss the results, let us, in short, again focus on the game theoretic solution of Game A and Game B.

Game Theory: Game A reflects the well–known Prisoner's Dilemma with its unique Nash equilibrium where both actors choose 'left'. Game B is a Hostage Prisoner's Dilemma such that subgame Game A_B is reached when neither actor posts a hostage. The hostage mechanism in the experiment was such that the hostage (with a value of NLG 4) was lost if placed and subsequently followed by a 'left'–move (identical to 'defection' as it is normally referred to in the Prisoner's Dilemma).[23] Remember that the decision of i and j with regard to their choice about the hostage was made public before they had to choose between 'right' or 'left'. Posting a hostage makes 'right' the dominant strategy but, if the other player does not post a hostage and plays 'left', choosing 'right' leads to the worst possible outcome. A rational actor should, therefore, deny to post a hostage, leading to Game A_B and subsequently play 'left'. Hence, Game B has a *unique* subgame perfect equilibrium which leads to the same outcome as in Game A.

As pointed out throughout this whole section, we are not interested in analyzing those subjects who played in accordance with the game theoretic prediction. However, in this case, our interest coincides with the equilibrium–behavior of the subjects. For the analysis, we restrict ourselves to those subjects who indeed reached the subgame Game A_B in the hostage game Game B. Only then we can again compare their behavior in Game A with their behavior in Game A_B which, according to our theory, should be equal, that is, either 'right' or 'left' in both games.[24]

Results: While from the subset of those subjects who reached Game A_B, 73.90% chose 'left' in the prior Game A, 91.30% of them played 'left' in Game B. Put into the parlance

[22]Note that the history in this experiment slightly differs from the history in the two preceding experiments. Due to its twofold character, different behavior in Game B, if found, can hinge on either one of these two 'different' histories. Statistically, it is not possible to precisely tell apart up to which degree the outcome in Game A and/or the not posted hostage in Game B affect possible deviating behavior in Game A_B.

[23]For the impact of the hostage value on the payoffs of the three Game C_{Bx} subgames, see FIGURE 2.3.

[24]We do not have to worry about the actions of player j (i.e. the computer) since 'he' did not post a hostage and, what is irrelevant to our analysis, played 'left'. Therefore, the computer makes sure that subgame Game A_B can be reached.

of the Prisoner's Dilemma and considered from the perspective of cooperation, a little bit more than one fourth cooperated in the Prisoner's Dilemma (i.e. Game A) while less than one in ten cooperated in Game A_B as part of the larger Hostage Prisoner's Dilemma (i.e. Game B). Again, we find a difference (McNemar's $\chi^2(1) = 8.33$, $p < .005$, $N = 46$) in the two games, even though they are identical with respect to payoffs.

Discussion: Even though we found the same result as in the previous experiments, the explanation seems not to be as straightforward as in Experiment 1 and 2. This time, the history of Game B could be twofold. One the one hand, we have the history consisting of additional moves by the actors in Game B. This is what we so far referred to as 'the history of a game' and also examined in the two preceding experiments. It seems not too farfetched if we again assume that the hostage that 'could have but was not posted' by the other subject positively influenced, that is, increased the likelihood to choose 'left' in the subgame Game A_B.

On the other hand, we have another game played against the same opponent, namely, Game A.[25] It might very well be that the subjects took their experiences from Game A into account when deciding on an action in Game A_B. Hence, this would mean that Game A becomes part of the history. Since all actors had complete information, even this 'enlarged' history should theoretically not change the predictions based on our theory. But we must assume that the subjects were able to learn from the outcomes of Game A that only playing 'left' saves them from being ripped off by their opponent.[26] Moreover, having experienced actor j playing 'left' in the preceding Game A and now seeing him refraining from posting the hostage may just be the last hint needed by i to confirm that j cannot be trusted. Therefore, it might be that a decision by actor i in Game B is subject to three history effects. First, j's behavior in Game A, second, j refraining from posting the hostage, and third, an interaction between these two effects. Unfortunately, we are not able to (statistically) tell apart which effect (the outcome of Game A or 'not posting a hostage' in Game B) can be hold responsible for the different behavior in Game A_B. But, the important conclusion we draw from this experiment runs along the same line as earlier: history affects behavior.

As the three experiments clearly show, all the results are in contradiction to the theoretical prediction. We assumed that actors are rational with respect to their decision and only consider the information inherent in the payoffs to be relevant for them. This

---

[25]Of course one could argue that in Experiment 2 we also had a Game A played before a Game B was played. But there, the two games were not linked together as this is the case in Experiment 3, where the subjects were explicitly told that they now play another game against the same opponent. Here, Game A and Game B can be considered as being part of one larger situation. In Experiment 2, Game A and Game B were considered to be separated situations. Subjects were confronted, loosely speaking, with the following: "Let's assume you are playing a Game A. What would you do?" Later they were confronted with a new situation: "Now, let's assume that you are playing a Game B. What would you do?" There was no link between Game A and Game B, except that Game A was also identical with the subgame Game A_B of Game B.

[26]Whatever the other actor does, by playing 'left' an actor is better off than by playing 'right' in the Prisoner's Dilemma. Game theoretically, we say that 'left' is the dominant strategy.

theory was not supported by the experiments and it must be concluded that a theory which does not take the history of a game into account leads to inadequate predictions about the behavior of the subjects. Therefore, in one or the other way, the information which was disclosed by the history must be reflected in a theory (model) about behavior in Game B. The last section summarizes the findings and sketches a possible place to look for such an extension of the theory.

2.4 Discussion

In this chapter we studied whether the past affects behavior in situations where the present and the future were kept constant. For this purpose, we compared situations we referred to as Game A and Game A_B, a subgame of Game B. We argued that simple game theoretic models with complete information predict no difference in behavior between Game A and Game A_B. In other words, no 'history effects' are to be expected. However, in three analyses of experimental data we did find substantial history effects. For each of these effects we came up with, admittedly, ad hoc arguments to explain the divergent behavior in Game A and Game A_B. In the following, we summarize these arguments and discuss some suggestions on how to model them.

To begin with, let us take a look at the most simple game we studied, namely, the Trust Game and its subgame (see FIGURE 2.1, games A–1 and B–1). In Game B–1, any actor i who is choosing 'right' (as we assumed he did to put j into the same situation as described by Game A–1) runs the risk to end with less than he could have gotten by choosing 'left'. One could now argue that actor j honors this risk taken and plays 'right' because he wants to reciprocate i's cooperative behavior. But, to apply this sort of 'tit–for–tat' behavior, an actor j must first know that there is at least one other actor on whom he is dependent and, second, that this actor has been cooperatively. But reciprocity out of fairness is not explained easily if we stick with the assumption made in the beginning of this chapter, namely that 'utility is own money'. For an explanation of the subject's behavior in our experiments we need to amend this assumption. In addition, we need to introduce incomplete information (e.g. Rasmusen 1994: ch. 2) in explaining the results of Experiment 2 and 3.

One possible way to work around the 'utility is own money' assumption is by employing the so called *social orientation models*. In these models, how much someone cares for the well–being of another subject is expressed by his social orientation parameter. Therefore, a subject's utility is not only a function of his own (monetary) payoffs, but also takes into account the payoffs of his opponent, weighted by his social orientation. Experimental evidence for such behavior can, for example, be found in

Liebrand (1984) or McClintock (1972, 1978).[27] We assume that actor j's utility is defined as follows (with general payoffs $T_j > R_j > P_j$ and $R_i > P_i > S_i$):

$$U_j(\text{right}) = R_j + \theta_j R_i \quad (2.1)$$
$$U_j(\text{left}) = T_j + \theta_j S_i \quad (2.2)$$

where $-1 < \theta_j < 1$ denotes actor j's social orientation (i.e. how much he cares for actor i's well–being).[28] Hence, in explaining the results of Experiment 1, reciprocal behavior in Game B–1 would demand that $U_j(\text{right}) > U_j(\text{left})$ and thus $\theta_j > \frac{T_j - R_j}{R_i - S_i} =: \theta_j^*$. But as can be seen easily, the same inequality can also be derived from the payoffs of Game A–1 since the 'additional' punishment payoffs (P_i and P_j) are not even taken into account. Consequently, social orientation models do not explain the findings of Experiment 1. Put differently, the outcomes from Experiment 1 speak in favor of a history dependent fairness norm. New experiments to test whether or not a player's social orientation θ_k is also a function of this player's history—moreover, histories with different characters—and not only a 'personality trait', could easily be set up and would allow for some interesting conclusions.[29]

On the other hand, a simpler way to model divergent behavior in Game A and Game A_B is to assume that actor j gains some additional utility for reciprocal behavior. The difference between this approach and the social orientation model is that actor j only earns an additional payoff in case he cooperates out of reciprocity considerations. Hence, $U_j(\text{left})$ remains unchanged. A utility function for actor j in the Trust Game then looks as follows:

$$U_j(\text{right}) = R_j + \Delta_j \quad (2.3)$$
$$U_j(\text{left}) = T_j \quad (2.4)$$

where Δ_j denotes j's additional utility ('reciprocity revenue') in case of reciprocal behavior and is only defined in case of a preceding action by another actor i. That is,

[27] For a more detailed discussion of social orientation models, see for example Braun (1998) or Weesie (1994).

[28] Usually, it is assumed that the own monetary payoffs are also weighted by some parameter $-1 \leq \vartheta_j \leq 1$, the individual orientation of j, that is, how much he cares for his own well–being. Different combinations of these two orientation parameters lead to different behavioral orientations, for example discussed by Herkner (1991). For several reasons discussed by Braun (1998), it suffices to restrict the orientation parameters to $\vartheta_j = 1$ and $-1 < \theta_j < 1$.

[29] A similar conclusion can be drawn by specifying a so–called *guilt model* with $U_j(\text{left}) = T_j - \gamma_j (T_j - S_j)$ and $U_j(\text{right}) = R_j$ (see also Snijders 1996). Actor j feels guilty if abusing placed trust. Even though the social orientation model cannot be transformed into a guilt model, the conclusions remain comparable. If the guilt parameter $\gamma_j > \frac{T_j - R_j}{T_j - S_j} =: \gamma_j^*$ is also supposed to be history dependent, comparing both models shows then that fair behavior is more easily found under the assumption of the guilt model (as specified in this note) since $\gamma_j^* < \theta_j^*$ always holds. On empirical grounds, this seems even plausible since 'guilt' is a stronger behavioral mechanism than 'fairness'. Loosely speaking, $\gamma_j (T_j - S_j)$ is equal to the amount an actor j is willing to pay to obtain the privilege of being abusive. Note, however, that different specifications of the guilt part of the utility function are possible and consequently lead to different conclusions.

Δ_j is a function of i's behavior and, therefore, history dependent. Following, we try, in short, to explain the empirical findings by applying such a Δ_j–parameter.

Experiment 1: The Trust Game

If we only consider the subgame A–1 of the two–stage Trust Game B–1, we see that an actor j can choose between left and right, earning T_j and R_j, respectively. Since the temptation payoff T_j is always larger than the reward payoff R_j, a rational actor j will play left. In Experiment 1, about two third chose to do so. Consider now an actor j who is confronted with the 'whole' two–stage Trust Game. Playing left still is the dominant strategy and secures him the payoff T_j. But if he chooses to reciprocate, therefore playing right, we will assume that he earns $R_j + \Delta_j$. Depending on the value of Δ_j, the temptation to play left becomes smaller or even vanishes completely. Something along those lines could have been happening in Experiment 1 since in Game B–1, after trust has been placed by actor i (he plays right) only about one fourth of the subjects keep playing left. The same results are found in Game A–2 and Game B–2.

Our argumentation about Δ_j would imply that the smaller the payoff difference between T_j and R_j becomes, the more likely it is that an actor j plays cooperatively in Game A_B. Such a statement could easily be tested by setting up Trust Games with different payoffs. Snijders' (1996: ch. 4) analyses support this hypothesis to some degree. Although the introduction of Δ_j may seem ad hoc from a rational point of view, it at least produces predictions which are in line with the experimental results presented in Section 2.3. Moreover, Δ_j can be linked to the guilt model (cf. note 29) if we assume that, for example, $\Delta_j \equiv \delta_j\,(T_j - R_j)$. In contrast to the guilt model, a subject receives some additional benefit from being reciprocative instead of paying the debts for being abusive. Even though it is, as it has already been stressed above, obvious from the results in the experiment that the parameter δ_j (or γ_j in the guilt model) are history dependent, it leaves open the question to which degree history affects the parameter. Moreover, since such a transformation of utilities takes place in the mind of the actors, it brings about the more psychological question about how players transform their given matrix into an effective matrix (Kelley and Thibaut 1978).

Experiment 2: The Hostage Trust Game

While we only employed pure utility arguments to explain the different outcomes in the games of Experiment 1, for Experiment 2 we need an additional argument in setting up an explanation for the behavior found there (since j moves before and after i moves). We assume that actor i has some belief about j's value of Δ_j, without being completely sure about it. Technically, we model Game A and Game B as games with incomplete information (e.g. Rasmusen 1994). Note that in Experiment 2 we are concerned with the behavior of actor i. In explaining Snijders' (1996) findings, we need to show that it is possible that actor i in the Hostage Trust Game where no hostage was posted by actor j (Game A_B) is more often inclined to play left compared to his behavior in the corresponding Trust Game (Game A).

Setting up the two games under incomplete information leads to a simple condition on when an actor i plays 'right' in Game A: $\pi^{\mathsf{A}} > \frac{P_i - S_i}{R_i - S_i} =: \pi^*$, where π^{A} denotes the probability that he faces an actor j who will reciprocate.[30] Assume that $P_i - S_i$ measures the potential loss of actor i due to unjustified trust and $R_i - P_i$, in turn, denotes i's gain due to justified trust. Hence, above inequality precisely resembles Coleman's (1990: 97–102) assertion about whether or not actor i should place trust (i.e. choose to play right): $\frac{p}{1-p} > \frac{\text{potential loss}}{\text{potential gain}}$, where p denotes the probability that a trustee is trustworthy (i.e. our π^{A}), reduces to π^* if we substitute the loss and risk terms into the righthand side of Coleman's inequality and solve for p. What needs to be shown in explaining Snijders' results is that by adding Δ_j to j's payoffs, we can find an equilibrium in Game B in which actor i's assessment about facing an actor j who will reciprocate even after not posting a hostage, is lower than π^*: actor i's assessment π^{A} of facing an actor j who reciprocates needs to be smaller in Game B, implying that i would then indeed choose to play 'left' more often than in Game A. Gautschi (1999) shows that there exists an equilibrium in Game B which supports the empirical findings.

Unfortunately, there is no straightforward equilibrium in this Hostage Trust Game with incomplete information where all game theoretic types of actors j refrain from placing a hostage and i's assessment about a reliable trustee would indeed fall short of π^*.[31] Nonetheless, Game B has an equilibrium in which all game theoretic types of actors j are mixing with the same positive probabilities in an open interval $(0, 1)$ over whether or not to post a hostage and actor i withholds trust: he plays 'left', as long as π_G and π_M (cf. note 31) together are not exceeding $\frac{P_i - S_i + H_i}{R_i - S_i + H_i}$.[32] Yet, we would need to consider a situation where no hostage is posted. However, without affecting the equilibrium, it can safely be assumed that the probabilities for each of the game theoretic types of actors j to provide a hostage are close to zero. Therefore, we have a situation where hostage posting by any of the trustees is possible but highly unlikely to occur. Due to the fact that there exists an actor j with probability π_M who reciprocates only after posting a hostage, π_G can, in equilibrium, indeed be smaller than the similar probability π^{A} in Game A. Hence, in equilibrium the condition $\pi_G < \pi^{\mathsf{A}} < \pi^*$ can be fulfilled (for more details and a numerical example, see Gautschi 1999).

---

[30] Due to the introduction of incomplete information, Nature moves first and decides on the probabilities with which actor i faces a specific game theoretic type of j, depending on their values of Δ_j. In Game A, i plays an actor j who reciprocates with probability π^{A} and an actor j who will not reciprocate with probability $1 - \pi^{\mathsf{A}}$.

[31] Due to the additional introduction of a hostage, Nature now distinguishes three different game theoretic types of actors j in Game B. With probability π_G actor j reciprocates whether or not he posts a hostage (this is the same game theoretic type as the one who occurred in Game A with probability π^{A}; therefore, call him 'good trustee'), with probability π_M actor j only reciprocates after he has posted a hostage (call this one 'mediocre trustee'), and with probability $\pi_B = 1 - \pi_G - \pi_M$ he never reciprocates (hence, call him 'bad trustee'). For more details, see Gautschi (1999).

[32] A second equilibrium, in which all game theoretic types of trustees are mixing with different probabilities in the open interval $(0, 1)$ over whether or not to post a hostage, also supports the empirical findings. However, the conditions for a trustor to withhold trust are less straightforward.

The introduction of the Δ_j parameter again leads to predictions which support the experimental findings. However, the theoretical solution is less straightforward than in Experiment 1. Due to the more complex nature of the history of the game in Experiment 3 (see the discussion of the results) we refrain here from discussing possible implications on the behavior of i in this experiment by also introducing a Δ_j parameter. However, we expect that a theoretical model, even though more cumbersome than in Experiment 1 and Experiment 2 (since Δ_j would be dependent on the outcome of the first Prisoner's Dilemma as well as on the hostage decision in the second Prisoner's Dilemma), would provide a possible explanation for the empirical findings.

Summing up, we have seen that the shadow of the past clearly influences an actor's current behavior. Unfortunately, we are not yet equipped with a model to convincingly explain these empirical findings. The introduction of a reciprocity parameter Δ_j leads, on the one hand, to predictions which are in line with the experimental results but, on the other hand, remains to some degree ad hoc and yet lacks a proper argumentation. As long as the considerations about such an additional 'fairness payoff' (or guilt payoff) cannot be put on a proper foundation (e.g. explaining j's moral disposition and where it comes form), it will always bear the negative connotation of just being an ad hoc argument which happened to fit the purpose best.

It seems, however, likely that such a reciprocity parameter must be history dependent. This would then be a hint that an analysis of such situations is probably more easily done in the context of simple learning games (e.g. Bandura 1977; Bush and Mosteller 1955; Fudenberg and Levine 1998; Macy 1989, 1993, 1996). The setup of the experiments is unfortunately limited with respect to a proper application of learning models (too few repetitions of the dilemma situation). And, in connection with learning models, one may have to abandon the assumption of rational Bayesian learners. Applying Bayes' Rule in finding the equilibria of Game B in Experiment 2 leads to outcomes which are far from simple 'common sense' understanding (see Gautschi 1999). Moreover, it has been argued by, for example, Boyd and Richerson (1985) that subjects in situations with incomplete information do not at all update their beliefs according to Bayes' Rule.

Even though this chapter has only scratched the surface of an interesting problem without yet giving a satisfactory solution, there have been some modest first steps done in the right direction. An elegant approach is presented by Rabin (1993). He sets up a model in which feelings of altruism or envy are triggered by actions (or intentions) of others. The idea is actually as simple as the underlying assumptions of our Δ_j parameter or the guilt parameter. Rabin also assumes that positive or negative feelings towards the opponent add to the own monetary utility. His model, however, has two desirable features. First, feelings of sympathy become less important as monetary payoffs increase. Second, preferences for being fair or not are explicitly based on beliefs about the opponent's behavior. Therefore, Rabin's (1993) fairness part of the utility function has a proper foundation while our assumptions with respect to the role of history remain vague. Rabin's model is, however, restricted to two–person (normal form) games of complete information.[33] Extending the approach to multi–person, sequential games under incomplete information might be problematic or even change the implications of the model (Rabin 1993: 1296). A model which would be applicable to (multi–person) games of incomplete information but employs a different approach of fairness is due to Bolton and Ockenfels (1998, 2000). They argue that subjects are motivated by their monetary payoffs as well as their own relative payoffs. The latter measures how an individual's monetary payoff compares to that of the rest of the players. The aspect of fairness is such that ego wants to be treated fairly and not that ego treats alter fairly due to some reason. Unfortunately, relative payoffs only refer to the current game and not the entire history which would be necessary in our context.[34]

It must be concluded that the game theoretic models highlighted in this last section have desirable features to explain behavior which seems irrational at first sight. Unfortunately, the shortcomings of all models (the one proposed here as well as those by Rabin (1993) and Bolton and Ockenfels (1998, 2000)) are yet too severe to be able to provide satisfactory answers about the impact of history effects on current behavior.

[33]An example of such a situation is given in Experiment 1 (i.e. the two–stage Trust Game B–1). Employing Rabin's (1993) model shows that there exists no fairness equilibrium for the payoffs specified. Its prediction is such that monetary incentives outperform any sympathy aspects and the subjects reach the subgame perfect equilibrium. However, this is not what we observed in Experiment 1. Rabin (1993: 1296–1297) reaches a similar conclusion and suggests the incorporation of additional emotions into such models.

[34]On grounds of experimental evidence stemming from four mini–ultimatum games with different cake divisions, Falk, Fehr, and Fischbacher (1999) conclude that responders (in either accepting or turning down a proposer's offer) do not only take into account the proposer's actual offer but also the set of available, but not chosen offers. They conclude that the models by Rabin (1993), and Bolton and Ockenfels (1998, 2000) are incomplete since the former does not capture distributional concerns (however intentions) while the latter does not take into account the fairness of the proposer's intentions (but only distributional concerns).

Chapter 3

Trust over Time: The Effects of Dyadic Social Capital

> "Again, one of the Contractors, may deliver the Thing contracted for on his part, and leave the other to perform his part at some determinate time after, and in the mean time be trusted; and then the Contract on his part, is called PACT, or COVENANT: [...] he that is to performe in time to come, being trusted, his performance is called *Keeping of Promise*, or Faith; and the fayling of performance (if it be voluntary) *Violation of Faith*.
> [...] For he that performeth first, has no assurance the other will performe after; because the bonds of words are too weak to bridle mens ambition, avarice, anger, and other Passions, without the feare of some coërcive Power; which in the condition of meer Nature, where all men are equall, and judges of the justnesse of their own fears, cannot possibly be supposed. And therfore he which performeth first, does but betray himselfe to his enemy."
> — Thomas Hobbes (1651: 94, 96)

3.1 Introduction

Many transactions are characterized by a sequence of moves with a time–lag between the 'delivery of' and the 'return on' an action. Such transactions carry the burden that, because of the time–lag, there is an obvious risk of actions being delivered, but not reciprocated. When enforceable contracts are not available or feasible, trust is often seen as an aid to this problematic aspect of 'social exchange' (Blau 1964). In a recent book, Misztal (1996) reviews and summarizes a large body of literature and identifies several functions of trust. One of these functions is 'trust as a lubricant for cooperation' (Arrow 1974) which has mainly been analyzed using a Rational Choice framework. This research can be summarized in three categories: trust in one–shot situations, in repeated situations, and within networks. Repeated situations and network embeddedness are often considered to provide the trustor[1] with helpful 'clues' as to

[1]The trustor is the individual who decides whether or not to place trust while the trustee needs to decide whether or not to honor given trust. Trustor and trustee are sometimes also referred to as principal and agent, respectively.

decide whether or not to place trust, knowing that the trustee has a certain temptation to abuse trust.

Snijders (1996) presents research on one–shot trust situations with and without the use of credible commitments in the form of hostages, as proposed by Schelling (1960).[2] Besides hostages, Coleman (1990), Blau (1964), Granovetter (1985), and especially Voss (1982, 1985) have analyzed how trust problems can be overcome by embedding the actors in a framework of repeated interactions. Most research on repeated games, however, employs the Prisoner's Dilemma Game. And, some of these studies explicitly interpret mutual cooperation in the Prisoner's Dilemma Game as trust (e.g. Ahn et al. 1998, 1999; Ahn, Ostrom and Walker 1998; McCabe, Rassenti, and Smith 1996). The literature on the (repeated) Prisoner's Dilemma Game and cooperation via reciprocity is vast. Dugatkin (1997: ch. 2) nicely summarizes many theoretical perspectives on the emergence of cooperation between subjects which have been published over the last twenty years. Repeated games reflecting true trust problems in the sense that they contain a time–lag between the actors' moves have been used to experimentally test predictions about placing and honoring trust by, for example, Kollock (1994) and Camerer and Weigelt (1988).

Networks offer additional possibilities to actors that need to overcome problems of whether or not to place trust. Third parties, for instance, provide the trustor with the possibility to gain control over the own trust relation(s). Actors can collect or spread information about the behavior of the partner or simply substitute an untrustworthy exchange partner for another, supposedly more trustworthy one (e.g. Burt and Knez 1995; Buskens 1999; Lahno 1995a, 1995b; Raub and Weesie 1990; Schüssler 1989; Boone and Macy 1999).

Although this literature offers insight into cooperation and trust under different conditions, it hardly ever addresses the question of what precisely induces actors to either act cooperatively or defectively in a given situation. Put differently, systematic explanations regarding an increase or decrease in mutual cooperation or trust are often not treated explicitly. By and large, the literature reports percentages of cooperative or defective moves without focusing on the question of what precisely leads to an increase or decrease of these observed or predicted percentages. Of course, there are notable exceptions which do not only focus on the outcomes but also on the reasons behind the observed behavior (e.g. Burt and Knez 1995; Buskens 1999; Kollock 1994). Lewicki and Benedict Bunker (1995) and Boon and Holmes (1991) present an elaborated description of the development of trust within personal relationships. However, both models are adapted to the very specific situation of close personal relationships and therefore offer explanations on the development of trust only in such a setting.[3]

[2]See, for example, also Keren and Raub (1993), Mlicki (1996), or Raub and Keren (1993) for hostage analyses in the Prisoner's Dilemma Game or Raub and Weesie (2000a) and Weesie and Raub (1996) for a general theoretical analysis of Hostage Games.

[3]Hardin (2001) distinguishes between relations of trust based on beneficial acquaintance and based on love or close friendship, respectively. The latter are less problematic since they are based on strong

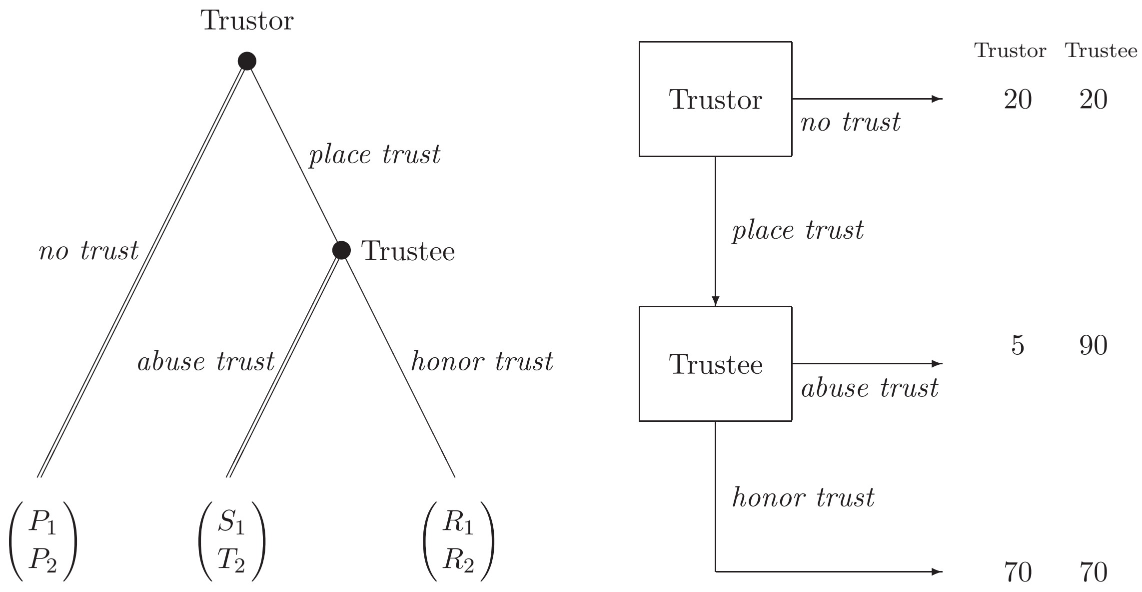


FIGURE 3.1 *Extensive form of the Trust Game ($R_1 > P_1 > S_1$ and $T_2 > R_2 > P_2$; double lines indicate equilibrium path of play) and an example as used in the experiment.*

In this chapter, we will try to gain more insight into the trustor's decision of placing trust. The basic scenario follows Coleman (1990: 97–98) who defines a *situation of trust* by four distinct elements. First, placing trust by the trustor allows an action of the trustee otherwise not possible. Second, placing trust puts resources in the hand of the trustee which he can use to his own or to the mutual benefit. Third, the trustor voluntarily puts resources at the trustee's disposal without any safeguard by the trustee. And, fourth, there exists a time–lag between placing trust and the action of the trustee.

Dasgupta (1988) and Kreps (1990a) provide a game theoretic representation of such a trust situation (see FIGURE 3.1). This *Trust Game* is sometimes referred to as a one–sided Prisoner's Dilemma Game (e.g. Rasmusen 1994: 129–131). It is a game of complete and perfect information. Actors in the Trust Game move sequentially. The trustor moves first and decides between placing trust or withholding trust. If the trustor refrains from placing trust, the game ends and both actors receive their punishment payoffs P_i $(i = 1, 2)$. If the trustor decides, for whatever reason, to place trust, the trustee must in turn decide between honoring trust or abusing trust. The game ends after the trustee's move. If the trustee abuses placed trust, he walks away with the temptation payoff T_2 while the trustor earns the sucker's payoff S_1. If the trustee reciprocates by honoring placed trust, both actors receive their reward payoffs R_i. The ordinal ranking of the payoffs ($T_2 > R_2 > P_2$) makes it favorable for the trustee to abuse trust whenever placed by the trustor. The trustor, on the other hand, would prefer an outcome where trust is placed and honored over trust not being placed since $R_1 > P_1 > S_1$. However, since the trustor knows that a rational and selfish

moral commitments. Since most trust relationships cannot bank upon such moral commitments of others, the focus should, according to Hardin, be on the former.

trustee abuses trust if placed, she is better off by withholding trust beforehand. The outcome of the Trust Game is the subgame–perfect Nash equilibrium formed by the strategy combination (withhold trust; abuse trust if placed).[4]

In the remainder of this chapter, we focus on the explanation of the behavior of the trustor, namely, on her *trustfulness* in a setting of *finitely repeated games* without exit options. The trustor's decision on whether or not to place trust in a current game is then based on different aspects of temporal embeddedness: the trustee's observed behavior in past games, the trustor's anticipation of the trustee's behavior in possible future games, and the trustor's anticipation of the trustee's behavior in the current game. When talking about *trust*, we thus refer to the trustor's decision to 'place trust', not to the entire relationship, as it is often used in commonplace. Relationships of reciprocal trust—or in our simple scenario of placed and honored trust—will be referred to as *trustworthiness* (for a conceptual discussion of the term 'trust', see Hardin 2001; Snijders 1996).

This chapter is structured as follows. The next two sections will discuss the temporally embedded situation of the two actors and present a social capital interpretation of trust. Several hypotheses are put forward which will be tested using experimental data. The fourth section presents and discusses this empirical evidence. We conclude with a discussion of the results, highlight some connections to game theory and reflect on some restrictions of the chosen approach.

3.2 Trust and Dyadic Social Capital

In general, situations of trust do not simply occur once but repeatedly over time. Consequently, the actors share a common past and future. In addition to the effect of the current game's incentive structure, such a situation gives rise to two additional types of effects which may influence the trustor's behavior (e.g. Buskens 1999; Raub 1997). A common past allows the trustor to obtain information about characteristics of the trustee, such as for instance his trustworthiness. *Learning* thus offers the trustor the possibility to update her assessment about the trustworthiness of the trustee. A common future, on the other hand, allows the trustor to influence the trustee's current behavior through sanctions (if anticipated by the trustee) in future games. This is referred to as the *control effect*.

The focus of this chapter is on repeated interactions of one trustor and one trustee. The trustor's behavioral guideline is that she takes into account past outcomes and future opportunities while keeping in mind the incentive structure of the current game. The trustee, of course, has to take decisions as well. He too bases his moves on certain characteristics of the game and the behavior of the trustor (e.g. payoffs of the current game, past outcomes, credible threats of the trustor concerning future sanctions) to

[4]As long as the ordinal ranking of the payoffs is maintained, the punishment and reward payoffs of the trustor and the trustee need not be equal. Throughout this chapter, including the experimental setup, we use symmetric payoffs, $P_1 = P_2$ and $R_1 = R_2$, and subscripts are thus dropped.

decide between short–term benefits (‘golden opportunities’) from abusing trust and future flows of gains from honoring trust.

With these cornerstones in place, we need a suitable concept to ‘summarize’ the essential effects of such a temporally embedded two actor situation. We argue that such a concept is *social capital* (e.g. Coleman 1988, 1990). Coleman (1990: 302) defines social capital as those resources from interpersonal relations which make it possible to achieve certain ends. The notion of social capital, however, not only refers to social networks (on social capital and networks, see for instance Burt 1992, 2000; Flap 1988, 1999; Flap and Boxman 1999; Flap and De Graaf 1986; Lin 2001). Coleman (1988, 1990), Luhmann (1979) and Putnam (1995) argue that trust between two persons likewise accumulates as a sort of social capital. We will thus interpret ‘resources’ as a trustworthy partner in a given dyad and ‘certain ends’ as mutual cooperation, in the sense of placed and honored trust, in this relation. Hence, we focus on *dyadic social capital* embodied in the relation between the trustor and the trustee and neglect social capital that is embodied in relations with third parties.

Social capital is often seen as the sum of shared experiences in a number of relations (e.g. Coleman 1990: 300ff.). This, however, would point to the fact that social capital “is accumulated history” (Sandefur and Laumann 1998: 482) which forms a potential *stock of social capital*. If such a social capital stock is a function of the past, modeling such a stock variable could be related to basic models on expectation formation in macro economic theory (e.g. Neumann 1996). These ‘reaction models’ assume that actors extrapolate past experiences into the future or adapt their expectations about future outcomes by the difference between expected and observed current outcomes (adaptive expectations). See also Becker (1996) and Becker and Murphy (2000) for the use of social capital in economics. The basic idea of a capital stock and its formalizations in sociology (e.g. Lahno 1995a, 1995b; Snijders 1998) likewise assume that the formation of expectations is based on past observations or outcomes only.

This is not the conceptualization of a stock we will use. In the following, we rather advocate the idea that the *social capital stock is a function of the temporally embedded situation* of the trustor and the trustee as a whole, including not only the past through learning, but also the future through control. The possibility of adding considerations about the future into the concept of social capital seems to be supported by Sandefur and Laumann (1998: 483) who stress that social capital allows for control and social influence. From our point of view, social capital is therefore the trustor’s *belief* with respect to the behavior of the trustee. Updating the capital stock is therefore *belief learning*. Such an interpretation indicates that a trustor enters a relationship with some amount of social capital, reflecting experiences from other relations that she now generalizes to the current setting (e.g. Coleman 1990). Effectively, social capital denotes the trustor’s expectation that the trustee will use his resources, at least to some extent, to pursue both his and the trustor’s interest.

The following section develops an approach to trust as dyadic social capital. The general idea is that the past outcomes and the expected future form the trustor’s capital stock. In deciding whether or not to place trust, this capital stock will be put into relation to the peril the current game poses, namely, the consequences of misplaced trust. That is, the more severe the consequences of misplaced trust in the current game, the more social capital is needed to induce the trustor to place trust.

3.3 Social Capital and Temporal Embeddedness

3.3.1 Isolated Encounters: Effects of Risk and Temptation

Isolated encounters are the most simple two–party trust situations. In the absence of temporal embeddedness, the only characteristics describing and distinguishing different trust situations are the payoffs of the game to be played. If the trustor and the trustee behave in accordance with 'utility is own money', any trustee would abuse placed trust since $T > R$. In anticipation of such behavior, any trustor would withhold trust since $P > S$. However, it applies to both inequalities that, in some games, the difference between the respective payoffs can be small(er) while in other games the difference can be large(r). This means that if, for instance, P is close to S, a trustor loses little when placed trust is eventually abused. The peril to place trust is thus not necessarily the same across different games. Both the trustor and the trustee should take into account such possible differences in payoff differences when making their moves. The behavior of the trustor in a Trust Game then depends on what she can win or lose, but also on what she knows about what the trustee can win or lose. The behavior of the trustee, on the other hand, will by assumption follow from a model of guilt. These considerations lead to an approach which predicts behavior by payoff differences and ratios thereof (e.g. Coleman 1990: ch. 5; Rapoport and Chammah 1964: ch. 1; Rapoport, Guyer, and Gordon 1976).

We introduce two payoff ratios which we use to measure the peril a trustor faces when having to choose whether or not to place trust in the current game (see also Snijders 1996: ch. 2). The first ratio $\frac{P-S}{R-S}$ is denoted as RISK. It can be derived from a *Subjective Expected Utility* (SEU; Savage 1954) consideration. If the trustor withholds trust, her payoff is P. However, if she places trust, her expected payoff is $pR + (1 - p)S$, where p denotes the trustor's subjective probability that the outcome of the game provides her with payoff R (i.e. p is the probability that a trustee honors trust). Rearranging the inequality $pR + (1 - p)S > P$ shows that a trustor should place trust if $p > \frac{P-S}{R-S} =:$ RISK. It is the ratio of the potential loss of the trustor due to unjustified trust $(P - S)$ and her gain due to justified trust $(R - S)$. RISK corresponds to Coleman's (1990: 97–102) assertion about when a trustor should place trust. Taking this analysis literally, a trustor should place trust if and only if she assesses the probability of facing a trustworthy trustee to be larger than RISK. We propose that a trustor's inclination to place trust in a Trust Game increases as RISK becomes smaller and, by the same token, decreases as RISK becomes larger.

We represent the trustee's inclination to abuse trust by the payoff ratio $\frac{T-R}{T-S} =:$ TEMPTATION. TEMPTATION can be derived, for instance, from a *model of guilt* (e.g. Snijders 1996: ch. 2). Assume that the trustee's utility from abusing trust is not the temptation payoff T but $T - \gamma(T - S)$. Abusing trust triggers a bad consciousness and reduces the trustee's utility by (a fraction of) the difference of his and the trustor's payoff. How large the reduction in utility is depends on the size of the guilt–parameter $\gamma \in [0, 1]$. In such a guilt model, the trustee honors placed trust whenever $R > T - \gamma(T - S)$ and hence if the guilt–parameter $\gamma > \frac{T-R}{T-S}$. A trustee with a small guilt–parameter will abuse trust while a trustee with a large guilt–parameter will not. As with RISK, we propose that the trustor's inclination to place trust decreases in TEMPTATION. Snijders (1996) presents experimental evidence that the trustor's trustfulness decreases in RISK and TEMPTATION in one–shot Trust Games.

We sketched our general idea of social capital as being the stock of past outcomes and expected future encounters. How the stock variable is precisely influenced by the past and the future will be addressed in more detail in the following subsections. Before we go into the details, it is useful to list the basic premises on how a trustor decides between placing or withholding trust. The trustor comes to a decision regarding the placement of trust by comparing the stakes of the current game to the stock of dyadic social capital 'available' to her. She places trust whenever the stock outweighs the peril of the current situation as indicated by the stakes of the current game.

- The *stakes of the current game* are represented by the payoff ratios RISK and TEMPTATION.
- The stock of *social capital* emerges as a function of the mutual past and future of the trustor and the trustee (see below for a discussion of this stock).[5]
- The larger RISK and the larger TEMPTATION of the current game, the more social capital is needed to induce trust (i.e. the larger the trustor's capital stock needs to be).

Consequently, a small amount of social capital means that games of relatively low stakes already pose a problematic situation in which trust will not be placed, while a larger endowment of social capital can lead to a placement of trust even in high stakes situations.

Having addressed the relation between the stakes of the current game and the capital stock, we now present two conceptualizations of the stock variable.

A Simple Stock: The trustor sees as 'ingredients' to her stock the outcomes of the past games (i.e. was placed trust abused or honored), the length of the past (i.e.

[5] We do not exclude that trust likewise depends on a natural inclination or tendency of the trustor to trust a (yet unknown) trustee based on past experiences gathered in similar situations (e.g. Coleman 1990). Neither do we exclude that variations due to personal characteristics affect the probability of placing trust. We do take these features into account in the empirical analysis.

the number of games played before with the trustee) and the length of the future (i.e. the number of games expected to be played).

A Complex Stock: A more complex conceptualization of the stock additionally takes into account the incentive structure of the past and future games as represented by RISK and TEMPTATION. The past is a function of the length and outcomes of the past games as well as the stakes of these games. Furthermore, the future is no longer interpreted as simply the number of games to be played but also depends on the stakes of the future games.

Both versions of the capital stock are based on the assumption that outcomes of past interactions lead to an increase or decrease of the capital stock. Similarly, futures that differ in terms of length and eventually stakes should have different impacts on the stock. Below, experimental data are used to test whether the more elaborated conceptualization of the stock better explains the behavior of the trustor. Obviously, we prefer a more parsimonious theory that explains the data as well as a more complicated theory. Hence, we will prefer the complex conceptualization of the social capital stock only if it explains the data better than the simple version of the stock.

Independent of the conceptualization of the stock, it should hold that the capital stock must 'cover' the risk of getting placed trust abused. Therefore, we should always observe less trust in games of high stakes.

Hypothesis 3.1 [Simple and Complex Stock] *The probability that the trustor places trust in the current game decreases in the stakes of the current game, as represented by the payoff ratios* RISK *and* TEMPTATION.

3.3.2 Repeated Encounters: Effects of the Past

The effect of a positive past is frequently discussed in the literature. Positive experience leads to a positive adjustment of the beliefs about the trustee's trustworthiness. Granovetter (1985) argues that, in general, there is no better information about a partner's reputation than information "from one's own past dealings with that person" (p. 490). While ignored in standard economic assumptions on, e.g. forward–looking maximizers and sunk cost, individuals do rely on their knowledge from past relations of individual transactions (p. 491). Schelling's (1960: 134–135) argument that trustworthiness often evolves simply by the continuity of a relation seems to be based on the implicit assumption that past outcomes were more or less successful, while relations with unsuccessful outcomes were terminated. A similar point on the effect of the past is favored by Hardin (2001) and the authors in Lazaric and Lorenz' (1998a) book on trust and economic learning. In general, these authors promote that trustworthiness can only emerge through a process of learning via direct interactions.

Since we employ repeated games without possibilities to communicate, information about the trustworthiness of the trustee must stem from the observable outcomes of the past games. Situations in which trust was placed and subsequently honored will, by assumption, increase the stock while abused trust will decrease the capital stock.

Hypothesis 3.2.1 [**Simple and Complex Stock**] *The probability of placing trust in the current game increases with positive past experience, that is, placed and honored trust.*

Hypothesis 3.2.2 [**Simple and Complex Stock**] *The probability of placing trust in the current game decreases with negative past experience, that is, placed but abused trust.*

Two points related to these hypotheses need to be addressed briefly. First, deviant behavior of the trustee of course induces the stock of social capital to shrink and the smaller the stock, the less willing the trustor will be to place trust. Consequently, placing trust should then only be observed in games of relatively low stakes while in games of higher stakes, the trustor will withhold trust as long as the stock is not yet large enough (again).

Second, the effect of additional information diminishes, especially after a longer stream of information of the same kind. For example, the more trust is honored by the trustee, the less information each additional cooperative move provides to the trustor since her assessment about the probability of the trustee being trustworthy is already considerably large and not much influenced by yet another cooperative move. However, such cooperative moves are necessary to maintain trust. On the other hand, a non–cooperative move by the trustee after a long series of honored trust may be of little effect on the stock precisely for the same reason. The trustor's assessment about the trustee being trustworthy will not be affected too much. If a seemingly trustworthy trustee surprisingly abuses trust, one may not believe that he indeed is 'the bad guy'. Rather one believes that he had a 'bad day' and forgets the incident. If he honors placed trust, one will likewise not be surprised since this is what constantly happened in the past. Such moves by the trustee bear little new information. Schelling (1960) argues that successful long–term relations form a focal point from which it is hard to deviate. If he is right, then an occasional defection should not immediately drive the relationship away from mutual trust. Not though, however, if defection starts to happen repeatedly. Social capital would decline and one's assessment about the trustworthiness of the trustee would be more severely adjusted downwards. As a consequence, one would no longer trust him.[6]

Earlier in the paper we argued that the trustee's decision whether or not to honor placed trust depends on the payoff ratio TEMPTATION. This consideration will be taken into account regarding the influence of the past on the *complex capital stock*. We argue that the trustor adjusts her stock based on good or bad past outcomes (*Hypothesis 3.2.1* and *3.2.2*). In the more complex definition of the stock, however, she will also take into account the TEMPTATION of the past game.

We argue that the stock increases when the trustee honors trust, and the more so the higher the TEMPTATION. Similarly, the stock decreases when placed trust is

[6]This is a long run effect which we may not be able to observe in the experiments.

abused, but this decrease should be smaller with higher Temptation. Assume trust is placed. The trustor knows that the trustee's inducement to abuse trust clearly rises in Temptation. Therefore, she values honored trust more the higher Temptation, that is, the larger the potential short–term gain the trustee foregoes by honoring trust. By the same token, the stock decreases if trust is abused by the trustee. The higher Temptation, the less an abuse of trust comes as a surprise. Hence, the stock should shrink less. However, if trust is abused in a situation of low Temptation, that is, where little is to gain from abusing trust, it shows the selfish nature of the trustee and the stock should decrease more. The incentive structure of the past games, therefore, contains additional information beyond the information of the outcome itself.

Hypothesis 3.3.1 [Complex Stock] *Assume Hypothesis 3.2.1 is true. Then, the higher the* Temptation *of a past game, the more the capital stock increases with positive past experience in this game, and the larger the increase in the probability of placing trust in the current game.*

Hypothesis 3.3.2 [Complex Stock] *Assume Hypothesis 3.2.2 is true. Then, the higher the* Temptation *of a past game, the less the capital stock decreases with negative past experience in this game, and the less the decrease in the probability of placing trust in the current game.*

Information from past outcomes has another interesting property that we expect to find in the experimental data. Kahneman (1994) argues that experimental as well as real life evidence shows that peoples' evaluation of the past is biased towards recent outcomes. That is, people tend to put more weight on the more recent outcomes in assessing the overall experience from the past. To take this fact into account, we argue that past outcomes are discounted.[7] Discounting could also be interpreted as the trustor's recall of past outcomes. Not only do people 'forget' past outcomes, but memory could work selectively and, for instance, recall bad outcomes more easily than good ones. This would even make it necessary to define two different discount parameters for different outcomes (i.e. abused and honored trust, respectively). Kahneman (1994) also reports a similar effect: people seem to remember extreme past outcomes more than average ones.

Hypothesis 3.4 [Simple and Complex Stock] *Assume Hypotheses 3.2.1 and 3.2.2 are true and the probability of placing trust in the current game increases with positive past experience and decreases in bad past experience. The further back in the past these outcomes, the less their impact on the probability of placing trust in the current game.*

The assumption of discounting the outcomes forming the stock leads to three straightforward and important conclusions. First, a trustor will eventually 'forget' defective

[7]There may be other reasons to discount information from more distant outcomes. Older information can become less relevant than newer one if the object one is learning about changes over time.

moves by the trustee (i.e. she is not resentful). Second, even good outcomes will become less important and mutual trust must be maintained by the trustee by further honoring placed trust. Third, social capital deteriorates over time if no interactions between the trustor and the trustee take place. That is, the longer the breaks in between two interactions, the less likely it will be that the trustor places trust in a next encounter.

While the first two points should show up in the experimental data (e.g. smaller or statistically not significant coefficients), the latter will not (directly) be observable in the experiments since we did not consider breaks of different length in between the games in our experiment. We can, however, address a related consideration about 'not playing the game'. At the beginning of any new relation, the capital stock is per definition simply a function of the future. Assume now that the trustor withholds trust in the first game. Since the trustee is forced to refrain from any action, the trustor cannot learn anything from such a game. Hence, her social capital with this trustee will only change because the (length of the) future becomes different. By the same token, this holds true for any following and consecutive game(s) in which she withholds trust.

Hypothesis 3.5 [**Simple and Complex Stock**] *The probability of placing trust in a situation without a past (i.e. the current game is the first game to be played) must be equal to the probability of placing trust in a current game with a past of withheld trust only.*

Of course this hypothesis does not hold true for situations in which trust is withheld only after trust has been placed at least once beforehand. In such situations, social capital is no longer a function of the future only. To test *Hypothesis 3.5*, we furthermore have to make sure that the future is the same for both types of situations.

3.3.3 Repeated Encounters: Effects of the Future

Although Schelling argues in favor of mutual trust as a focal point from which it is hard to deviate, he further argues that even a successful past seems not to be enough to maintain trust (Schelling 1960: 45). He claims that many agreements are only enforceable because of the recognition of future opportunities which will be eliminated if trustworthiness is not maintained. Put differently, abusing trust can lead the trustor to withhold trust (for one or more transactions) in the future which, in turn, should be anticipated by the trustee. This *mechanism of control* is often called reciprocity (Blau 1964; Emerson 1981; Gouldner 1960; Homans 1961), conditional cooperation (Taylor 1987), or tit–for–tat by analogy with one of the strategies supplied to Axelrod's (1984) computer tournament. A common future provides a trustor with opportunities to retaliate and punish deviant behavior by the trustee. Likewise, the future also offers the possibility to reward cooperative behavior by the trustee through positive sanctions. If placed trust was honored by the trustee in preceding games, the trustor can continue to place trust in future games. Matthews, Kordonski, and Shimoff (1983) report evidence from experiments in which possibilities of bilateral punishment helped to maintain cooperative behavior even in the most tempting (large payoffs) situations. Oliver (1984) presents experimental results which show the significant impact of punishment on increasing cooperation. The 'successful' combination of reciprocity and punishment has long been put forward (based on animal experiments) in biology and psychology as 'reinforcement learning' (e.g. Walker 1987: ch. 7): rewards should be repeated while responses that bring about punishments should be stopped.

Under the assumption that the trustor's threat to withhold trust in the future is seen as credible by the trustee (control effect), the future enters the capital stock as the collection of possibilities for the trustor to retaliate deviant behavior by the trustee. The trustee should anticipate the likelihood of such punishment by the trustor and see his future gains at risk. Consequently, the more possibilities exist to punish abused trust, the more risk can be taken by the trustor in the current game.

Hypothesis 3.6 **[Simple and Complex Stock]** *The probability of placing trust in the current game increases in the length of the mutual future of a dyadic relation.*

According to *Hypotheses 3.2.1* and *3.2.2*, the past adds or subtracts from the stock through cooperative and defective moves of the trustee. Since his future moves are not known, the future cannot influence the stock via outcomes. However, the longer the future the more possibilities for the trustor to retaliate possible deviant behavior by the trustee and the larger the dyadic social capital stock.

The argument of *Hypothesis 3.6* uses the number of future games. The future, however, contains more information than its length only. Following the definition of the *complex stock*, we are now taking into account the incentive structure of the future games as well. Since a common future provides the trustor with opportunities to sanction the trustee's deviant behavior in past and current games, our argument to include payoffs again runs along the lines of TEMPTATION. For the trustee, punishment is something to be avoided, more so the more severe the sanctions will be. We argue that the higher the TEMPTATION of the expected future games, the stronger the trustor's punishment in cases she withholds trust in the respective future games. A trustor who can credibly 'threaten' the trustee with severe punishment should then be more inclined to take higher risks in the current game.

Hypothesis 3.7 **[Complex Stock]** *Assume Hypothesis 3.6 is true and the probability of placing trust in the current game increases in the length of the future. This probability is even higher the higher the* TEMPTATION *of the future games.*

Consequently, the capital stock of a trustor with a future of games of mainly high TEMPTATION will be larger than the stock of a trustor with games of mainly low TEMPTATION. Hence, everything else being equal, the former trustor will be more likely to place trust than the latter. TABLE 3.1 summarizes our hypotheses.

At this point, let us briefly relate our considerations about past and future effects to game theory (e.g. Fudenberg and Tirole 1993; Rasmusen 1994). In a finitely repeated

TABLE 3.1 *Hypotheses with respect to the probability of placing trust (I).*

Independent variable	Probability to place trust: Predicted[a]	Hypothesis
Current		
Stakes in current game	−	*3.1*
Past		
Honored trust	+	*3.2.1*
Honored trust × stakes	+	*3.3.1*
Abused Trust	−	*3.2.2*
Abused trust × stakes	−	*3.3.2*
Having a past × discounting honored trust	+	*3.4*
Having a past × discounting abused trust	−	*3.4*
No past v. trust withheld	0	*3.5*
Future		
Length of future	+	*3.6*
Length of future × stakes	+	*3.7*

Note: Stakes refer to RISK and TEMPTATION as defined in Subsection 3.3.1.
[a]− : the probability of placing trust decreases. + : the probability of placing trust increases. 0 : the probability of placing trust is not affected.

Trust Game with complete information, the unique equilibrium behavior is that trust will never be placed and would never be honored if it would be placed (viz. Luce and Raiffa 1957: 97–102). However, cooperation (trust being placed and honored) until near the last period can occur as equilibrium behavior if certain characteristics of the game or of the trustee are unobservable for the trustor ('incomplete information').[8] In an *infinitely repeated* Trust Game, conditional cooperation is possible as long as 'the shadow of the future' is long enough so that conditional cooperation of the trustor is credible (see also Friedman 1986: 88–89).

3.4 Empirical Evidence

To test our hypotheses, we use data from a laboratory experiment in which subjects played a series of repeated Trust Games using the 'strategy method'. We first introduce the experimental setup in more detail. Subsequently, we discuss the construction of the variables and the statistical model that we use. Finally, we present the results.

[8]How 'near' to the last period cooperation should be uphold depends on the parameters of the game such as its payoffs (Kreps et al. 1982).

3.4.1 Experimental Design and Subjects

Background: The experiment was conducted at the Ludwig–Maximilians University Munich, Germany, in June 2000. The experiment on the repeated Trust Game was part of a larger experimental session consisting of seven different parts (for more details, see 'Design' below).

Subjects: Students of an introductory sociology course were informed repeatedly for a period of about four weeks prior to the experiment about the possibility to participate in an experiment on decision making. Students could sign up for one of the eight experimental sessions scheduled to take place during a one week period. Each session was planned to include 20 students. In announcing the experiment, students were promised a participation fee of DM 10 (German Marks, about € 5.10 or US$ 5 at the time of the experiment). Moreover, they were informed that they would have the possibility to earn more than DM 100 in about 30 minutes time. One week prior to the experiment, a total of 144 students had signed up for one of the eight sessions. Eventually, 83 students showed up for the experiment. Subjects were mostly freshmen. On average, they were about 25 years old (s.d. 6 years). One third of the participants were women, 24% had had economics as part of their high school graduation package and 59% had heard about game theory.

On arrival at their session, participating students were briefly instructed orally by the experimenter about the tasks to perform—to make decisions with consequences depending on the decisions of an unknown other. The subjects were not informed about our interest in the effects of temporal embeddedness on trust. In addition, they were told that two subjects in the session would earn more than the participation fee, namely, up to a maximum of DM 150, depending on their choices in one of the parts of the experiment. All necessary instructions for the experiment were printed on the first three pages of the experiment booklet (see Appendix A). Additional instructions regarding the tasks in the parts of the experiment were outlined in short introductions at the beginning of each of those parts. In addition, all game parts contained graphical representations of the given situation. Care was taken to avoid references to 'trust', 'risk', or related concepts.

Design: The experiment was a paper and pencil experiment divided into five different game parts. The experiment booklet likewise contained two questionnaires, a set of items on trust and related concepts (Yamagishi and Yamagishi 1994) and 23 items measuring the reliability of the subjects answers ('test attitude scale', Wilde 1970),[9] and several questions on sociodemographic variables. As described in the previous section, our interest is in the behavior of the trustor. Because we paid in accordance with actual choices at the end of the experiment, we randomly chose one subject to

[9]The test attitude scale was designed to measure a subject's self–defensive versus his or her self-critique attitude in answering questions regarding opinions on certain statements or conceptions. The scale can also be used to judge the reliability of the subject's answers on a questionnaire of interest (i.e. the trust related questions in our case).

play as the trustee.[10] The experiment booklet for the trustee subjects was an appropriately adapted version of the booklet for the trustor. Throughout the experiment, the instructions in the booklet constantly reminded the subjects that they should imagine that the 'other person' they played against is actually an unknown other person in the room. It was not known to the subjects in which role (i.e. trustor or trustee) the other persons in the room would play.

Subjects first played two one–shot Trust Games, followed by a once repeated Trust Game (i.e. two games in a row). Following the first two game parts, they filled in the 'test attitude scale'–items. Subsequently they played a twice repeated Trust Game (i.e. three games in a row). Finally, they received the item–battery on trust and related concepts. Together with the sociodemographic variables, these trust–items were used to measure the subjects' basic tendency to trust (trustfulness, see note 5). Subjects finished the experiment by playing a one–shot Continuous Trust Game that will not be discussed or analyzed in this chapter. The experiment was scheduled to last 50 minutes. On average, subjects finished in about 37 minutes. Due to time problems, four subjects did not play the Continuous Trust Game.

The first two one–shot Trust Games only differed in terms of their payoffs, such that we had a game of low and high stakes, respectively. The first set of payoffs was as follows: $S = 10$, $P = 20$, $R = 40$, and $T = 140$, which implies that TEMPTATION $= \frac{T-R}{T-S} = 0.77$ and RISK $= \frac{P-S}{R-S} = 0.33$. We will refer to this set of payoffs as *high stakes*. The second set of payoffs was $S = 5$, $P = 20$, $R = 70$, and $T = 90$, which implies TEMPTATION $= 0.24$ and RISK $= 0.23$. We will refer to this set of payoffs as *low stakes*. We also used these low and high stakes payoffs in the once and twice repeated Trust Game. In each session we used different combinations of the low and high stake games. The combination of games of different length and stakes allows us to vary the parameters necessary to test our hypotheses. Since we have (repeated) games of length one, two, and three, we are able to test the subjects' behavior in situations where they were embedded in various temporal settings: no past, a past of one game played, and a past of two games played. The same holds true regarding the subjects' future: no future, a future of one game, and a future of two games. The use of two different sets of payoffs (low and high stakes) in combination with the two one–shot games, the once repeated game, and the twice repeated game leads to $2 \times 2 \times 4 \times 8 = 128$ payoff combinations, of which we used 8 (see TABLE 3.2). These payoff combinations were chosen so that the resulting set of experimental conditions permitted a test of all of our hypotheses. Stakes as well as length of play in all game parts of the experiment were common knowledge to the subjects.

Strategy Method: To collect information about the subjects' behavior in the repeated games, we used the so–called *strategy method*. This method was first proposed by Selten (1967). Rather than observing the subject's behavior in only those nodes he

[10]The random selection of the trustee as well as the trustor, and the respective part of the experiment for the additional payment of these two participants was made prior to the begin of each session. Of course, this choice was not revealed to the participants until the end of the experiment.

TABLE 3.2 *Experimental design and stakes in the Trust Games (TG).*

Session	First TG	Second TG	Once–repeated TG	Twice–repeated TG
1	low	high	low, low	low, low, low
2	high	low	low, low	low, low, high
3	low	high	low, low	high, low, low
4	high	low	low, high	low, low, high
5	low	high	low, high	high, low, high
6	high	low	high, low	high, low, high
7	low	high	high, low	high, low, low
8	low	high	high, high	low, low, low

Note: 'low' and 'high' refer to the respective TEMPTATION of the game, as defined by $\frac{T-R}{T-S}$. Payoffs for the low stakes game are $S = 5$, $P = 20$, $R = 70$, and $T = 90$ (TEMPTATION=0.24); the high stakes games have payoffs $S = 10$, $P = 20$, $R = 40$, and $T = 140$ (TEMPTATION=0.77).

or she reaches, the strategy method asks for a decision of the subject in every node of the game (see below for an example). Roth (1995: 322–323) discusses some pros (e.g. observing subjects' entire strategies) and cons (e.g. subjects are forced to think about behavior in each node which could lead to different outcomes than had the game been played in an ordinary manner) of the strategy method. The method surely adds some complexity. Subjects must clearly understand that their moves are not independent of the partner's moves (i.e. the simulated trustee in our experiment). Boone and Macy (1999) observed that already in ordinary laboratory experiments, subjects sometimes have problems understanding the fact that their decision is not independent of their partner's decision. Since the strategy method so much relies on the understanding of the interdependent situation, extensive written information and graphical representations of the games to be played were provided in the experiment booklet. In addition, the instructions of each of these part of the experiment contained a comprehension check.[11] Personal communication with a sample of subjects after the experiment indicated that none of them had experienced problems with the experimental design.
Example: We now illustrate the strategy method using the once repeated Trust Game as an example. FIGURE 3.2 shows a condensed version of the graphical representation of 4 pages of the experiment booklet. For each decision the subject had to take, we reserved one page in the experiment booklet. After answering the question, subjects were routed to the page (number) of the respective next question. The graphical representation was such that the decision node the subject was in was clearly marked by

[11] A possible path (i.e. moves of the trustor and the trustee) through the game(s) of the given part of the experiment was sketched. Subjects were asked to calculate and write down the payoffs of the trustor and the trustee for each game on the path. They could verify their answers on the following page. If their answers were right, subjects were asked to start with the respective part of the experiment. Otherwise, they were urged to re–read the instructions and contact the experimenter for further instructions.

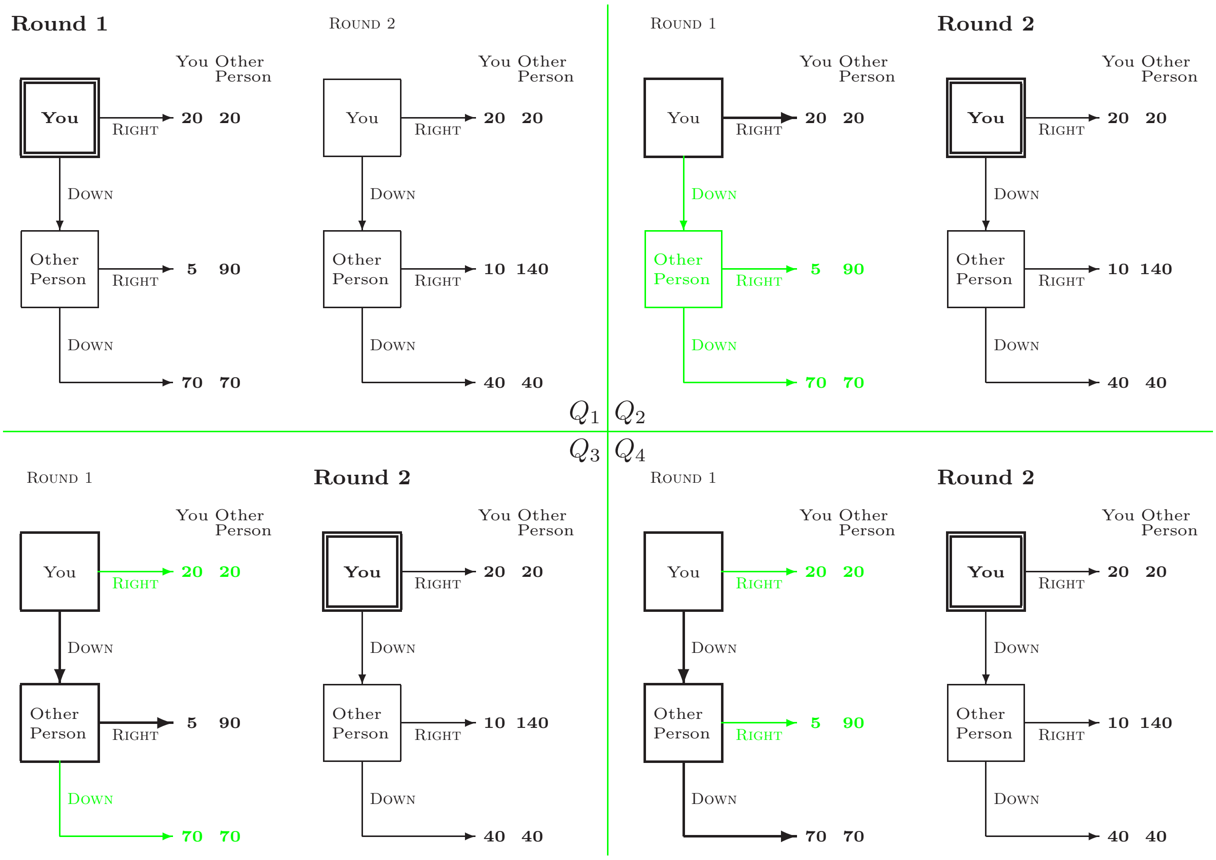


Figure 3.2 *The strategy method applied to the once repeated Trust Game.*

a double–edged box (see Figure 3.2). Paths not used in the (past) game were printed in shaded gray. The graphical representation belonging to question 1 is depicted in quadrant Q_1 of Figure 3.2. The subject (in the role of the trustor) was asked to decide between moving 'Right' or 'Down'. Assume first that the subject chose 'Right' (i.e. no trust). He was then routed to question 2 which was accompanied by the graphical representation in quadrant Q_2 of Figure 3.2. The subject additionally received the information that he chose 'Right' in round 1 and the 'Other Person' therefore had had no chance to move. Again the subject was asked whether he would choose 'Right' or 'Down' in round 2. The once repeated Trust Game ended after the subject's choice. He was routed to the first battery of questions to be answered.

Assume now that the subject did not choose 'Right' in round 1 but chose to play 'Down' (i.e. place trust). Under these circumstances, he was directly routed to question 3 graphically represented by the games in quadrant Q_3 of Figure 3.2. The subject received the information that he chose 'Down' in round 1 and that the 'Other Person' therefore had to choose between 'Down' and 'Right' as well. He was then told that he should first assume that the 'Other Person' had chosen 'Right' in round 1 (i.e. the other person abused trust). Given this situation, the subject was asked to choose between 'Right' and 'Down' in the second round. After answering this question, the subject was routed to question 4 which was accompanied by the graphical representation in quadrant Q_4 of Figure 3.2. It was carefully explained to the subject that because he chose ‘Down’ in round 1, the ‘Other Person’ could also choose between ‘Down’ and ‘Right’ and that in the previous question we assumed that this ‘Other Person’ decided to play ‘Right’ in round 1. We now asked the subject to assume that the ‘Other Person’ did not choose ‘Right’ in round 1 but ‘Down’. The subject was then again asked to choose between ‘Right’ and ‘Down’ in round 2, given this new situation. After making this decision, the game ended and the subject was routed to the next part of the experiment.

3.4.2 Variables

To test our hypotheses, we need variables that represent the stakes of the current game, the outcomes and the stakes of the past games, and the stakes and the length of the mutual future. Seven independent variables contain the necessary information of the experimental situations. What is to be explained is whether or not trust is placed in the current game. The dependent binary variable `TRUST` is 1 if trust is placed and 0 otherwise. With regard to the payoffs of the current game, we argued that the behavior of the trustor should depend on Risk and Temptation. Since only two types of payoffs are used in the experiment, Risk and Temptation do not vary independently of each other. For reasons of collinearity, one of them has to be omitted in the statistical analysis. We chose to include Temptation in the analysis.

`CURRENT STAKES` defines whether the current game (i.e. the game in which the decision whether or not to place trust is made) is a low stakes or a high stakes game, that is, whether the Temptation is low or high. `CURRENT STAKES` is a dummy variable where 0 = ‘low stakes’ and 1 = ‘high stakes’.

Past variables:

`PAST STAKES` is a nominal variable representing the stakes of the past game(s): 0 = ‘no past’, 1 = ‘one low stakes game’, 2 = ‘one high stakes game’, 3 = ‘1^{st} game low stakes; 2^{nd} game low stakes’, and 4 = ‘1^{st} game high stakes; 2^{nd} game low stakes’.

`PAST OUTCOMES` is a nominal variable representing the subject’s behavior in the past games. `PAST OUTCOMES` contains the information about the chosen path but not about the stakes of the past games. The labels describe the path through the game(s) before reaching the current game at t: 0 = ‘no past’, 1 = ‘no game at $t-2$; withheld trust at $t-1$’, 2 = ‘no game at $t-2$; placed and abused trust at $t-1$’, 3 = ‘no game at $t-2$; placed and honored trust at $t-1$’, 4 =‘withheld trust at $t-2$; withheld trust at $t-1$’, 5 = ‘withheld trust at $t-2$; placed and abused trust at $t-1$’, and so on till, 12 = ‘placed and honored trust at $t-2$; placed and honored trust at $t-1$’.

`PAST COMPLETE` combines `PAST STAKES` and `PAST OUTCOMES` and contains the information about the twelve possible paths in combination with the stakes of past games. Again, the nominal variable `PAST COMPLETE` is 0 if the current game is the first to be played (‘no past’). Including ‘no past’, `PAST COMPLETE` has 25 values.

Future variables:

`FUTURE DUMMY` defines a dummy variable which is 0 if there was no future after the current game (i.e. it is the last game) and 1 if either a future of one or two games laid ahead.

`FUTURE LENGTH` makes a difference between a short (i.e. one game) and a long (i.e. two games) future. The ordinal variable `FUTURE LENGTH` takes the values 0 = 'no future', 1 = 'short future', and 2 = 'long future'.

`FUTURE STAKES` additionally takes into account the stakes of the future games. We have again a nominal variable with values 0 = 'no future', 1 = 'short future: low stakes', 2 = 'short future: high stakes', 3 = 'long future: 1^{st} game low stakes, 2^{nd} game low stakes', and 4 = 'long future: 1^{st} game low stakes, 2^{nd} game high stakes'.

By construction of the variables, it follows that `PAST STAKES` and `PAST OUTCOMES` are 'nested' in `PAST COMPLETE`, `FUTURE DUMMY` is 'nested' in `FUTURE LENGTH` as well as `FUTURE STAKES`, and `FUTURE LENGTH` is 'nested' in `FUTURE STAKES`.

3.4.3 Statistical Model

To test the hypotheses presented in Section 3.2 by using these experimental data we need an appropriate statistical model. A subject's decision to either place or withhold trust is the outcome of a discrete choice. In the one–shot Trust Game, one would normally use an ordinary logit or probit regression model for the probability that trust is placed in terms of (a subset of) the independent variables described above. However, this is not entirely appropriate for our problem since the observations are repeated within subjects. To deal with this form of clustering, we consider an extension of the logistic regression model that incorporates (random or fixed) subject effects.

All subjects in the experiment played two one–shot games, a once repeated, and a twice repeated game. This adds up to a total of 19 possible decision nodes: one node in each of the one–shot Trust Game, 4 in the once repeated Trust Game, and 13 in the twice repeated Trust Game. If we additionally take into account the fact that subjects played games of different stakes (see Table 3.2), the number of different nodes increases to 70[12] out of which every subject reaches at maximum 19, representing the at most 19 decisions the subject has to make.[13] These 70 nodes are the experimental conditions and they can be defined in terms of the current stakes, the type of past,

[12]For the analysis it makes a difference whether subjects played, for instance, a once repeated game with (low, low) stakes or the same game with (low, high) stakes. Table 3.2 shows that two one–shot Trust Games with low and high stakes, respectively, were played. They were followed by a once repeated Trust Game with either (low, low), (low, high), (high, low), or (high, high) stakes. The twice repeated Trust Game, finally, had (low, low, low), (low, low, high), (high, low, low), or (high, low, high) stakes. Therefore, we have $2 + 4 \times 4 + 13 \times 4 = 70$ possible decision nodes in our experiment.

[13]A subject withholding trust throughout the experiment only needed to take 9 decisions. All subjects, therefore, had to made between 9 and 19 decisions.

and the type of future. For instance, one node would be the condition that ‘trust has been placed and abused in a high stakes Trust Game, while the decision now is to place or withhold trust in a low stakes Trust Game, knowing that there are no more Trust Games to be played afterwards’.

The starting point in formulating the statistical model is the simple Rasch model of item response theory. The Rasch model describes subject i’s binary response at item j via the difference between i’s personal parameter θ_i and a parameter β_j representing the ‘item-difficulty’ of that item,

$$\pi_{ij} \;=\; \Pr(y_{ij} = 1 \,|\, \theta_i, \beta_j) \;=\; \frac{e^{\,\theta_i - \beta_j}}{1 + e^{\,\theta_i - \beta_j}}\,. \tag{3.1}$$

Standard applications of the Rasch model involve estimating individual capabilities. Here we model the choices of subjects in a similar way. Subject i’s decision whether or not to trust is assumed to depend on the difference between i’s personal parameter θ_i, representing a person’s ‘general trustfulness’, and a parameter β_j representing the condition under which the person made his or her decision. Let $\boldsymbol{y}_i = (y_{i1}, \dots y_{i70})$ be subject i’s choices in each of the 70 conditions, taking the values 1 (placed trust) or 0 (withheld trust). For each subject, minimally $51 = 70-19$ values are missing by design. The simple Rasch model with the 70 condition parameters β_j is a *saturated model* for the differences between the conditions and therefore not suitable to test our hypotheses. It will, however, be useful as our baseline model with which all other models (i.e. models using combinations of past–future variables discussed in the preceding subsection) can be compared.

Our hypotheses are formulated as assertions about differences in trusting behavior between certain classes of nodes or conditions. To test these hypotheses, the conditions can be represented by the temporal embeddedness variables and the stakes in the current game. Thus, we model the item difficulty parameter β_j in terms of the temporal embeddedness variables and the stakes of the current game in condition j. Such a parsimonious model can be fitted to the data and compared to the saturated model in order to test whether we have accounted for the differences between the conditions via these variables. In psychometrics[14] this specialized version of the Rasch model is known as the *Linear Logistic Test Model* (LLTM, see Fischer 1997: 226–227) and it decomposes the item parameter β_j as

$$\beta_j \;=\; \sum_{\ell=1}^{m} \gamma_\ell z_{j\ell}\;, \tag{3.2}$$

where the $z_{j\ell}$’s are characteristics of the conditions in terms of temporal embeddedness and the stakes of the current game, and γ_ℓ the parameters to be estimated. For example, the decomposition of the item difficulty parameter into

$$\beta_j \;=\; \gamma_1\, \texttt{CURRENT STAKES}_j + \gamma_2\, \texttt{PAST OUTCOMES}_j + \gamma_3\, \texttt{FUTURE LENGTH}_j$$

[14] Biostatisticians and epidemiologists refer to this model as ‘conditional logistic regression for matched case-control groups’. Econometricians refer to it as fixed effects logistic regression (Chamberlain 1980).

reflects that a condition affects behavior via additive effects of the current stakes, past outcomes, and the length of the future, but not, for instance, of the stakes of the past and the future.

We now turn to the subject parameters θ_i. The first possibility is to use a *fixed–effects* specification, and estimate the model using the conditional maximum likelihood method; the standard maximum likelihood estimator would be inconsistent for the condition parameters β_j and γ_ℓ (see Fischer 1997; Greene 1993: 655–657; Stata manual, Vol. 1, 2001: 214–215). With this unstructured ('saturated') specification of subject effects, it is not possible to analyze the differences between the subjects. Effects of subject–level predictor variables such as age, sex, or education are not identified. By the same token, subjects who always withheld or placed trust do not show any within–subject variation, and hence carry no information regarding differences between the conditions. For the same reason, conditions in which all subjects either placed or withheld trust are likewise dropped from the estimation.

Alternatively, we may use a *random–effects* specification for the person parameters, and estimate the model using marginal maximum likelihood. As with the condition parameter β_j, a 'structure' is imposed on the person parameters in terms of characteristics of the subjects and a random component leading to a *latent regression model* for the person parameter

$$\theta_i = \sum_{k=1}^{K} \phi_k x_{ik} + \epsilon_i \,, \tag{3.3}$$

with $\epsilon_i \sim N(0, \sigma^2)$ and ϕ_k the parameters to be estimated. If we dispose of suitable variables $\mathbf{x}_i$ to describe individual differences, the random–effects estimator is often more efficient than the fixed–effects estimator, and we obtain inferences not only about how temporal embeddedness and stakes affect the decision to place or withhold trust, but also which kind of subjects are more likely to trust in general.

3.4.4 Results

This section offers statistical tests of our hypotheses. Results are reported as a number of *findings*, each of which is followed by an elaboration how we arrived at it. As a starting point and baseline, we fitted the simple Rasch model to our data, that is, the model in which both conditions and subjects are represented by specific parameters (fixed–effects). The saturated model is estimated on 673 observations (69 subjects) with an average of 9.8 observations per subject.[15] From the 83 subjects showing up for the experiment, eight are dropped since they played as a trustee. From the remaining 75 subjects, two (four) subjects were dropped from the analysis because they placed (withheld) trust in all nodes reached. 16 conditions were not taken into account since either no subject reached it (four conditions) or all subjects reaching the respective condition either placed or withheld trust (seven and five conditions, respectively).

---

[15]The saturated model has 51 estimable parameters and has a log likelihood of −213.2636. This number has no interpretation per se, but is used to compute likelihood ratio tests reported below.

The theory section proposes that the probability of placing trust depends on the trustor's dyadic social capital. Dyadic social capital itself is a function of temporal embeddedness that varies between the conditions. We operationalized temporal embeddedness in terms of three past variables and three future variables, varying in complexity. Which of these variables best describes the 'true' shadow of the past and the future of the experiments can only be determined by *comparing* the fit of models that differ in the past–future variables. As we were not able to model well the differences between the subjects using the characteristics of the subjects that we measured (a subject's sex, age, education, exposure to game theory), we will only report results for the 'fixed–effects' specifications. As there are nine combinations of the past–future variables, we therefore fitted nine conditional fixed–effects logit models, all including `CURRENT STAKES` as well (see TABLE 3.3). Comparison of these 9 models with each other and with the saturated model reveals a first important finding.

Finding 3.1 *Neither past stakes nor future stakes are necessary in explaining the trustor's probability of placing trust in the current game.*

This finding is based on a series of comparisons of models in TABLE 3.3. To find the best fitting LLTM, we perform likelihood–ratio tests between the nested models. For the comparison of non–nested models we applied Akaike's Information Criterion (AIC) and Schwarz' Bayesian Information Criterion (SC/BIC) (see Greene 1993; Raftery 1995). Our starting point is the saturated model. We compare the saturated model with the one with the most complex specifications of the past and the future. As is seen in TABLE 3.3, the model with additive effects of past and future fits the data as well as the saturated model (i.e. the model using `PAST COMPLETE` and `FUTURE STAKES`). Thus additive effects of stakes, past, and future suffice to describe the data. In particular, we do not 'need' interactions between the past and future variables, or between either the past or future and the current stakes. Next, we consider whether we can simplify the specifications of the past and future. We 'move down' within nested variables: from `PAST COMPLETE` to `PAST OUTCOMES`, and from `FUTURE STAKES` to `FUTURE LENGTH` and finally to `FUTURE DUMMY`. Comparing the likelihood–ratios, we found that the model using the variables `CURRENT STAKES`, `PAST COMPLETE`, and `FUTURE STAKES` fits the data just as well as the saturated model. In a first step, we now replace `PAST COMPLETE` by `PAST OUTCOMES`. This simplifies the model without loss of explanatory power. Simplifying from `FUTURE STAKES` to `FUTURE LENGTH` again leads to a simpler model with the same explanatory power. However, any further simplification in terms of using `FUTURE DUMMY` instead of `FUTURE LENGTH` decreases the fit of the model significantly. In a similar way we seek to simplify the specification of the past. Since `PAST STAKES` is not nested in `PAST OUTCOMES`, we cannot improve any further on the model containing `CURRENT STAKES, PAST OUTCOMES` and `FUTURE LENGTH`. This, therefore is the most parsimonious model which fits the data at least as good as the saturated model, however, using as little 'information' as possible.

TABLE 3.3 *Model selection based on likelihood–ratio tests: best model uses* CURRENT STAKES, PAST OUTCOMES *and* FUTURE LENGTH *($N = 673$).*

PAST[a]	FUTURE[a]	log–likelihood	df	LR test against saturated model LR χ^2	AIC	BIC
saturated model		−213.2636	53		532.53	771.65
STAKES	DUMMY	−268.3579	6	110.19 (0.0000)	548.72	575.79
OUTCOMES	DUMMY	−231.6306	13	36.73 (0.6181)	469.28	547.91
COMPLETE	DUMMY	−229.5676	23	32.61 (0.3398)	505.14	608.91
STAKES	LENGTH	−265.7337	7	104.94 (0.0000)	545.47	577.05
OUTCOMES	LENGTH	−229.2601	14	31.99 (0.7793)	486.52	549.68
COMPLETE	LENGTH	−227.0624	24	27.60 (0.5395)	502.12	610.41
STAKES	STAKES	−263.6334	9	100.74 (0.0000)	545.27	585.87
OUTCOMES	STAKES	−227.3500	16	28.17 (0.8513)	486.70	558.89
COMPLETE	STAKES	−225.1114	26	23.70 (0.6471)	502.22	619.53

Note: AIC=Akaike's Information Criterion and BIC=Bayesian Information Criterion. CURRENT STAKES is contained in all models.

[a] PAST STAKES is a factor with 5 levels, PAST OUTCOMES a factor with 13 levels, and PAST COMPLETE a factor with 25 levels. FUTURE DUMMY is a dummy, FUTURE LENGTH is a factor with 3 levels, and FUTURE STAKES a factor with 5 levels.

The model selection process, therefore, reveals that stakes of games played in the past and stakes of future games seem not to help us explain the subjects' behavior in the current game. The past affects behavior only through outcomes while the future affects behavior simply by its length. Since estimating any of the above models without CURRENT STAKES significantly decreases the model fit, we can conclude that the only payoffs that do influence the subjects' probability of placing trust are those of the current game. Consequently, *Hypotheses 3.3.1* and *3.3.2* as well as *Hypothesis 3.7* are not supported. Empirically, the simple conceptualization of the social capital stock seems to capture the effects of dyadic social capital on trust.

According to the simple conceptualization of the social capital stock, past stakes are irrelevant for the emergence or decline of dyadic social capital. If dyadic social capital is indeed merely a function of (past) outcomes but not of the stakes involved, we could conclude that small but regular favors are sufficient for creating and maintaining dyadic social capital. Favors, gifts and cooperative behavior involving high costs to the donor are then not necessary for creating and maintaining social capital. Commonplace wisdom manifested in the German saying that 'small gifts maintain the friendship' would then be at the core of social capital. The creation and maintenance of social capital occurs in a tit–for–tat like manner—cooperative behavior is repaid by cooperative behavior (reciprocity). Anthropologists, however, have observed that, in archaic societies, reciprocity evokes obligations between the parties which are sometimes repaid in ever increasing magnitude (e.g. Mauss 1954). Our experimental finding would suggest that these gifts and favors exchanged between parties indeed unnecessarily increase over time. In such archaic societies, however, the situation may be more complex since the obligations imposed by the norm of reciprocity may vary with the status, the resources, or the motives of the donor (Gouldner 1960), and the spirit of the gift itself (Mauss 1954). Moreover, some obligations to repay prescribe that the gift be returned with interest (Mauss 1994: 40–41) which may to some degree explain the increase in gifts or favors exchanged. In our Western society, on the other hand, gift giving is no longer a means to produce and maintain the basic social structures of society but a more personal and subjective matter to create or maintain social capital. In this light, and given our experimental findings, the increase in gifts and favors exchanged seems unnecessary.

Table 3.4 provides some details of the selected model as Model 1. `PAST OUTCOMES` was, of course, entered in the regression as a nominal variable, i.e. expanded into indicators. The same holds for the variable `FUTURE LENGTH`. That is, each value of the variables expands into a dummy variable equal to 1 if the respective node was reached and 0 otherwise.[16] For convenience of notation, we summarize past outcomes within brackets: '(outcome at $t-2$), (outcome at $t-1$)' where C_1 refers to placed trust and C_2 denotes honored trust. By the same token, D_1 refers to withheld trust and D_2 to trust that was placed but abused.

Finding 3.2 *The probability of placing trust in the current game decreases in the stakes of the current game.*

Model 1 in Table 3.4 reveals a statistically significant and negative effect of `CURRENT STAKES`. Consequently, the probability of the trustor to place trust is smaller if the current game is a high stakes game. *Hypothesis 3.1* is therefore supported.

Finding 3.3 *The probability of placing trust increases with placed and honored trust in the past but decreases in placed and abused trust in the past.*

In accordance with the simple conceptualization of social capital, the respective coefficients of Model 1 in Table 3.4 show that a purely positive past (i.e. (C_1, C_2) and (C_1, C_2), (C_1, C_2), respectively) significantly increases the likelihood of placing trust. Also, we find a positive effect of a (D_1), (C_1, C_2) past. Maybe somewhat surprising, the coefficient of (D_1), (C_1, C_2) is even slightly larger than the one of (C_1, C_2), (C_1, C_2). Based on a Wald test we can reject that the coefficients of these three 'good past'-dummies are different ($\chi^2(2) = 2.45$, $p = 0.2941$). The probability of placing trust rises in a good past and *Hypothesis 3.2.1* finds support.

Abused trust in the past significantly decreases the probability that a subject places trust in the current game (see the coefficients of (C_1, D_2) as well as (C_1, D_2), (C_1, D_2),

[16] Since no subject reached the third (and last) game via path (C_1,C_2), (D_1), no dummy for this possibility is entered in the regression.

TABLE 3.4 *Conditional fixed–effects logit model for the probability of placing trust.*

	Model 1		Model 2	
Regressor	Coefficient	p–value	Coefficient	p–value
CURRENT STAKES	−1.63	0.000	−1.62	0.000
PAST OUTCOMES				
(D_1)	0.45	0.309	0.45	0.314
(C_1, D_2)	−1.09	0.001	−1.11	0.000
(C_1, C_2)	0.90	0.016	0.89	0.016
(D_1), (D_1)	2.10	0.149	2.12	0.145
(D_1), (C_1, D_2)	0.90	0.194	0.90	0.193
(D_1), (C_1, C_2)	1.87	0.013	1.87	0.013
(C_1, D_2), (D_1)	0.15	0.793	0.16	0.786
(C_1, D_2), (C_1, D_2)	−1.64	0.001	−1.63	0.001
(C_1, D_2), (C_1, C_2)	0.22	0.659	0.22	0.654
(C_1, C_2), (C_1, D_2)	−1.09	0.010	−1.08	0.010
(C_1, C_2), (C_1, C_2)	1.64	0.002	1.64	0.002
Future				
SHORT FUTURE	0.73	0.005		
LONG FUTURE	1.57	0.000		
FUTURE LENGTH			0.77	0.000
	$LL = -229.2601$		$LL = -229.2813$	
	LR $\chi^2(14) = 160.97\,(0.000)$		LR $\chi^2(13) = 160.93\,(0.000)$	

Note: Reference category is a low stakes game with no past and no future. Both models are estimated on $N = 673$ observations and 69 subjects.

and (C_1, C_2), (C_1, D_2) of Model 1 in TABLE 3.4). As would have been expected, the impact of abusing trust twice is the largest. We conclude that *Hypothesis 3.2.2* also receives support. The probability of placing trust indeed increases after a positive past but decreases after a negative past. Somewhat surprising, however, is the positive, but not significant effect of a (D_1), (C_1, D_2) past. We will discuss this point in more detail below.

Finding 3.4 *The probability of placing trust in the current game increases in the length of the future.*

TABLE 3.4 shows that a positive and significant effect on placing trust can likewise be ascribed to the length of the future. Moreover, the longer the future, the higher the likelihood that trust is placed since the coefficient of `SHORT FUTURE` is smaller than the coefficient of `LONG FUTURE`. This finding is in line with *Hypothesis 3.6*. We can use a Wald test to check whether the coefficient of a short future is exactly half of the coefficient of the long future ($\chi^2(1) = 0.04$, $p = 0.8369$). Since this null hypothesis cannot be rejected, we may conclude that the effect of a having future is proportional

TABLE 3.5 *Conditional fixed–effects logit model for the probability of placing trust without an interaction between the near and distant past outcome.*

Regressor	Coefficient	p–value
CURRENT STAKES	−1.61	0.000
PAST OUTCOMES:		
(D_1) in $t-1$	0.53	0.179
(C_1, D_2) in $t-1$	−1.14	0.000
(C_1, C_2) in $t-1$	0.89	0.003
(D_1) in $t-2$	1.65	0.002
(C_1, D_2) in $t-2$	−0.52	0.130
(C_1, C_2) in $t-2$	0.32	0.372
FUTURE LENGTH	0.77	0.000
	$LL = -230.6722$	
	LR $\chi^2(8) = 158.15\,(0.000)$	

Note: Reference category is a low stakes game with no past and no future. The model is estimated on $N = 673$ observations and 69 subjects.

to the length (i.e. number of transactions) of that future. This leads to Model 2 in TABLE 3.4. Model 2 is even more parsimonious since it saves another degree of freedom without sacrificing the explanatory power of Model 1.[17] The coefficients for the effects of the past are almost identical between Models 1 and 2 in TABLE 3.4.

Both models in TABLE 3.4 can shed light on which past outcomes influence the behavior of the subjects. In addition, they support the alleged effect of the shadow of the future. However, these models cannot be used to test the hypothesis that the outcome of the more distant time $t-2$ is less important that the outcome at the closer time $t-1$. To test this hypothesis, we will seek to impose additional structure on the past. The fairly complicated specification of past in Models 1 and 2 allow for interactions between outcomes at times $t-2$ and $t-1$. Is such an interaction supported by the data? To study this, observe that at each previous time point four possible outcomes could occur, (1) there was no past, (2) trust was withheld, (3) trust was placed but abused, and (4) trust was placed and honored. We recode the PAST OUTCOMES variable into eight dummies, four each for the $t-1$ and $t-2$ past according to the above distinction. This further reduces the complexity of the conceptualization of the past. TABLE 3.5 displays the parameter estimates of the model without the interaction between the nearer and farther past. By a likelihood–ratio test we can conclude that the interaction is indeed not significant ($\chi^2(5) = 2.78$, $p = 0.7336$).

Comparing TABLES 3.4 and 3.5, we find comparable effects of CURRENT STAKES and FUTURE LENGTH. As could have been expected from the coefficients in Model 2, the effect of abused trust in the previously played game is negative while the effect of honored

[17]A likelihood–ratio test for the linearity hypothesis of the future yields a similar results as the Wald test ($\chi^2(1) = 0.04$, $p = 0.8368$).

trust at $t-1$ is positive. Both effects are highly significant and substantial. In addition, TABLE 3.5 clearly reveals that the same outcomes at $t-2$ do not significantly affect the probability of placing trust.[18] The effects are, however, in the right direction (viz. a negative effect of abused trust and a positive effect of honored trust). We take these non–significant coefficients of the more distant past as evidence that *Hypothesis 3.4* cannot be rejected. If subject's perspective of the past is biased towards recent outcomes due to discounting, the results would support *Hypothesis 3.4*.

Finding 3.5 *Past outcomes are discounted such that the trustor's behavior is mainly explained by the more recent $t-1$ outcomes while more distant $t-2$ outcomes are of less impact on the trustor's behavior.*

The effects of (C_1, D_2), (D_1) and (C_1, D_2), (C_1, C_2) in TABLE 3.4 are in line with the argument of discounting as well. These coefficients show that abused trust can be 'forgiven'. The coefficients of the two regressors are positive but not statistically significant. Unfortunately, the outcome (C_1, C_2), (D_1) has not been observed in the experiments. This would have been a situation in which trust was not maintained in the previous game. According to our discounting hypothesis, such a coefficient should clearly be smaller than (C_1, C_2), (C_1, C_2).

Unfortunately, the model in TABLE 3.5 does not facilitate a formal test of the hypothesis that more distant outcomes are less important than nearer outcomes, nor does it provide an estimate of the rate at which the past is discounted. For these purposes, we propose a final specification of the past. Here the effect of the past is modelled as the sum of the effect of the 'outcome at time $t-1$' and the discounted effect of the 'outcome at time $t-2$'.

$$\text{Past} = \omega(\text{outcome}_{t-1}) + \rho\,\omega(\text{outcome}_{t-2})$$

where ω is defined over the 4 possible outcomes with $\omega(\text{Null}) = 0$ and ρ is a discount parameter $0 < \rho < 1$.[19] The constraint that the past effects at time $t-1$ are proportional to those at time $t-2$ is a nonlinear one. Hence, it cannot be estimated as a LLTM. The 'alternating maximum likelihood' algorithm that is widely used to estimate bilinear models can be applied to estimate the bilinear version of the logistic regression model. TABLE 3.6 provides results for this model. The effects of the current

[18]Note that the effect of '(D_1) in $t-2$' is highly significant and positive. This coefficient may be 'connected' to the surprising and positive effect of a (D_1), (C_1, D_2) past we found in TABLE 3.4. We further look into these effects below.

[19]It is of course tempting to speculate that this model can be generalized to a past of more than three periods s:

$$\text{Past}_t = \sum_s \rho^s \omega(\text{outcome}_{t-s})$$

This specification can also be written as a *recursion* describing how the stock at time t depends on the stock at time $t-1$ and the outcome at time t (e.g. Lahno 1995a, 1995b):

$$\text{Stock}_t = \omega(\text{outcome}_t) + \rho\,\text{Stock}_{t-1}\,.$$

TABLE 3.6 *Non–linear conditional fixed–effects logit model for the probability of placing trust with discounting of past outcomes.*

Regressor	Coefficient	Standard Error
CURRENT STAKES	−1.60	0.2303
PAST:		
(D_1) past	0.91	0.3437
(C_1, D_2) past	−1.02	0.2044
(C_1, C_2) past	0.91	0.2253
FUTURE LENGTH	0.76	0.1602
Discount Parameter ρ	0.64	0.1984
	$LL = -234.0450$	
	LR $\chi^2(8) = 151.4050\,(0.000)$	

Note: Reference category is a low stakes game with no past and no future. The model is estimated on $N = 673$ observations and 69 subjects.

stakes as well as the future are comparable with the previous models in TABLE 3.4 and TABLE 3.5. These effects can be interpreted in the same straightforward manner as in the previous models. The effects of the three PAST variables, however, now have a somewhat more complex meaning. Take, for instance, the positive effect of the (C_1, C_2) past. It tells us that honoring placed trust in the past increases the probability of placing trust in the current game, irrespective of the time at which this outcome occurred. However, the further back in the past, the less the 'impact' of a positive past on the behavior in the current game since past outcomes need to be discounted by the parameter ρ. Honored trust at $t-2$ would therefore have an effect on the probability of placing trust in the current game of $0.91 \times 0.64 = 0.58$. The other two effects of PAST in TABLE 3.6 can be interpreted in the same way.

TABLE 3.6, however, not only allows to estimate the discount parameter and lends support to *Hypothesis 3.4*. Each coefficient of this parsimonious model corresponds to one of the other hypotheses we already tested using previous models.[20] The coefficient of CURRENT STAKES supports *Hypothesis 3.1*, the PAST coefficients of (C_1, C_2) and (C_1, D_2) support *Hypothesis 3.2.1* and *3.2.2*, respectively. Finally, *Hypothesis 3.6* on the length of the future is supported by the FUTURE LENGTH coefficient.

The only hypothesis left to test is *Hypothesis 3.5*. We argued that withholding trust and not playing at all (i.e. having no past) should not have a different effect on the probability of placing trust in a current game.[21] Remember that if placed trust

[20]This does not hold true for those hypotheses which proposed an interaction between the past or the future with its respective stakes. They were rejected by the model selection procedure which led to Model 1 in TABLE 3.4.

[21]I am well aware that in testing *Hypothesis 3.5*, we commit a sin against one of statistics' general foundations: never set out to prove anything. We turn the standard logic of hypotheses testing upside down by trying to prove our null hypothesis of no differences between 'no past' and 'withheld trust'.

TABLE 3.7 *A comparison of withholding trust and not having played.*

Regressor 1	Regressor 2	Null Hypothesis[a]	χ^2	df	p–value
(D_1)	no play	zero	0.61	1	0.4336
(D_1), (D_1)	no play	zero	2.12	1	0.1453
(D_1), (C_1, D_2)	(C_1, D_2)	equal	7.69	1	0.0055
(D_1), (C_1, C_2)	(C_1, C_2)	equal	1.45	1	0.2288
simultaneous			10.41	0.0341	

Note: Coefficients are taken from Model 2, TABLE 3.4.

[a] *zero*: the coefficient is 0; *equal*: the coefficient of regressor 1 is equal to the coefficient of regressor 2.

is honored or abused, this experience contributes information about who the trustee is—put simply, whether or not he is trustworthy. If the trustor withholds trust, she foregoes the possibility to learn something about the trustee's behavior. Therefore, it should not matter for the trustor's capital stock whether there was no past at all or whether she withheld trust in a past game.[22] The social capital stock remains the same in both situations and the trustor should not change (i.e. cannot update) her assessment about the trustee's trustworthiness. The respective '(D_1) past' coefficient in TABLE 3.6 is significantly unequal to zero so we should reject this hypothesis. A past of withheld trust *increases* the probability of placing trust in a current game. There are two reasons why we nevertheless should investigate the assessment of *Hypothesis 3.5* in some more detail. First, the effect of '(D_1)' is based on different D_1–past outcomes (viz. TABLE 3.4). It seems wise to look at them separately as well. Second, TABLE 3.5 revealed a counterintuitive effect of a 'D_1–past at $t-2$'. Its probable effect on the general parameter in TABLE 3.6 should also be addressed.

To test *Hypothesis 3.5* we turn our attention to Model 2 in TABLE 3.4 and test whether the coefficients representing subjects starting off the (repeated) game by a D_1–move are equal to the respective 'no past' counterparts. We provide Wald tests for these four comparisons in TABLE 3.7. In addition, we simultaneously test the four hypotheses (last row in TABLE 3.7). Such a simultaneous test is actually 'equivalent' to testing whether or not the (D_1)–parameter in TABLE 3.6 is zero. That we have to reject the null hypothesis that there is no difference between 'not having played' and 'no past' does not surprise if one remembers the significant (D_1)–coefficient in TABLE 3.6.

Let us now consider the separate hypotheses. The first row in TABLE 3.7 reports the comparison of the one–shot game (no past and no future) with a once repeated game in which trust was withheld in the first game. The coefficient is not statistically different from zero. Further, we compare a past of twice withheld trust (i.e. (D_1), (D_1)) with the one–shot Trust Game (second row in TABLE 3.7). Again, we cannot reject

[22]Since the capital stock varies with the future, it is of course of importance to keep the future constant when comparing 'no past' and 'withheld trust' situations.

the null hypothesis. The comparisons of a one–shot game with, respectively, a once and twice repeated game of withheld trust thus support *Hypothesis 3.5*.

Finally, we compare a past where trust was not withheld entirely with its respective 'no past' game. Results from these two comparisons, however, are twofold. In case of a positive past (i.e. (D_1), (C_1, C_2) versus (C_1, C_2), see row four in TABLE 3.7), the null hypothesis cannot be rejected. A negative past (i.e. (D_1), (C_1, D_2) versus (C_1, D_2), see row three in TABLE 3.7), however, leads to a strong rejection of the null hypothesis that 'no past' equals 'withholding trust'. It is precisely this difference that drives the result of the simultaneous test in the last row of TABLE 3.7.

It might be that the latter result is driven by the same underlying process as the statistically significant and positive effect of '(D_1) at $t-2$' of TABLE 3.5. According to this coefficient, independent of the outcome at $t-1$, not placing trust at $t-2$ increases the probability of placing trust in the current game relative to a situation in which no decision had to be taken at time $t-2$. How can we understand this? We select those subjects who withheld trust at $t-2$ ($n = 33$) and look at their behavior in the current and the $t-1$ game. We see that three subjects withheld trust not only at $t-2$ but also at $t-1$. In the current game, one of them placed trust while the other two withheld trust. Their behavior makes sense. Furthermore, there are 16 subjects who withheld trust at $t-2$ and had placed trust honored at $t-1$. Out of these 16 subjects, 12 placed trust in the last game, too. Such behavior is consistent with their good experience at $t-1$. The other four subjects withheld trust in the last game and thus behaved rationally from a game theoretic point of view (viz. end game effect). The remaining 14 subjects withheld trust at $t-2$ but experienced that placed trust was abused at $t-1$. Surprisingly, nine out of these 14 subjects again placed trust in the last game. Why did they place trust in the last game? For one thing, their trust had been abused in the previous game. Moreover, they could not learn from the first game since trust was withheld. Subjects in this situation should have less reason to assume that they were facing a trustworthy trustee. In addition, they had to make a decision knowing that no more future games would be played. One would expect that trust is not placed in such a last game. It is of little help to provide ad hoc explanations for their behavior. We must settle with the conclusion that these nine subjects are clearly responsible for the positive and significant effect of the 'D_1 at $t-2$' coefficient in TABLE 3.5 and the statistically significant difference between the '(D_1), (C_1, D_2)' and '(C_1, D_2)' coefficient of TABLE 3.4.[23]

Based on this elaborated comparison of the 'withheld trust' versus 'no past' situations, we put forward the final finding of this chapter.

[23]We can run the regression model of TABLE 3.5 without these nine subjects and test whether the coefficients of this model are systematically different from the coefficients reported in TABLE 3.5. Using Hausman's specification test, we find no systematic difference in the parameter estimates between this 'reduced' (i.e. only 60 subjects) model and the model in TABLE 3.5. The significant coefficient of 'D_1 at $t-2$', however, now vanishes.

Finding 3.6 *The data offer strong but not entirely conclusive evidence that there is no difference in the trustor's behavior between situations of 'no past' and situations of a past of withheld trust.*

To conclude the empirical section, let us in short address a point raised in connection with *Hypothesis 3.4*. It has been claimed that social capital should decrease more, the longer the breaks in between two interactions of the trustor and the trustee. In other words, the shorter the time in between games, the smaller the discount rate. Consider the following situation. Trustor A and trustee B play two Trust Games with one time period between games. Another trustor C plays two Trust Games with trustee D with two time periods in between. Both trustors place trust in the first game and both trustees react the same. Given that past outcomes are discounted, trustor A's social capital is larger (smaller) than trustor C's if the trustee honored (abused) placed trust. Consequently, the probability for trustor A to place trust in the second game should be higher (smaller) than the respective probability of trustor C.

The experimental design did not allow for different amounts of time between two games. In other words, we studied the behavior of trustor A and trustee B but cannot say anything about behavior of trustor C and trustee D. An approximation of the situation of trustor C and trustee D may nevertheless be possible with the data at hand. Assume that withholding trust in the second of three games 'transforms' the situation into one in which two Trust Games are played, however, in between which two time periods elapse. We can then compare the effects of a (C_1, D_2) past and a (C_1, D_2), (D_1) past on placing trust, taking the (D_1)–move as a proxy for an additional time period elapsing in between two games. A Wald test shows that the null hypothesis of equal coefficients can indeed be rejected at a 5% significance level ($\chi^2(1) = 4.29$, $p = 0.0383$). A quick glance on the respective coefficients of Model 2 in Table 3.4 shows that the probability that trustor A places trust is indeed smaller than the one of trustor C.[24]

3.5 Conclusion

We proposed a social capital approach to the study of trustworthiness in repeated interactions between two actors. This idea is based on Coleman's (1988, 1990) consideration that trustworthiness in dyadic relations should be seen as social capital. We conceptualized dyadic social capital as a capital stock used by the trustor to determine her behavior. The trustor's capital stock is not only a function of her past experience with the trustee, but also a function of the expected future interactions. Since this capital stock is a 'summary' of the past outcomes and expectations about the future, it can be used by the trustor to decide whether or not to place trust in a current game. More precisely, we postulated that trust is placed whenever the capital stock is large

[24]Unfortunately, the '(C_1, C_2), (D_1)' path was not observed in the experiment and can therefore not be compared with (C_1, C_2).

TABLE 3.8 *Hypotheses with respect to the probability of placing trust (II).*

	Probability to place trust		
Independent variable	Predicted[a]	Hypothesis	Results
Current			
Stakes in current game	–	*3.1*	supported
Past			
Honored trust	+	*3.2.1*	supported
Honored trust × stakes	+	*3.3.1*	not supported
Abused trust	–	*3.2.2*	supported
Abused trust × stakes	–	*3.3.2*	not supported
Having a past × discounting honored trust	+	*3.4*	supported
Having a past × discounting abused trust	–	*3.4*	supported
No past v. trust withheld	0	*3.5*	supported
Future			
Length of future	+	*6*	supported
Length of future × stakes	+	*3.7*	not supported

Note: Stakes refer to RISK and TEMPTATION as defined in Subsection 3.3.1.
[a] – : the probability of placing trust decreases. + : the probability of placing trust increases. 0 : the probability of placing trust is not affected.

enough to cover the peril of the current game, as indicated by the payoff ratios RISK and TEMPTATION.

TABLE 3.8 summarizes the results of all tests of our hypotheses. The experimental data reveal three important findings. First, as in the analysis of one–shot Trust Games (e.g. Snijders 1996; Snijders and Keren 2001), the stakes of the game to be played are an important factor in the trustor's decision whether or not to place trust. The higher these stakes, the smaller the probability of placing trust.

Second, effects of the past can be interpreted as reinforcement learning: abusing trust in the past decreases the probability of placing trust in a current game while placed and honored trust in the past increases this probability. Moreover, we also found evidence that past outcomes are discounted. In addition, the data support our hypothesis that a longer future increases the probability of placing trust in a current game.

Third, besides past outcomes and the length of the future, neither the stakes of the past games nor the stakes of the future games affect the trustor's behavior in the current game. A simple conceptualization of the capital stock seems to capture the effects of dyadic social capital on trust. Placing or withholding trust seems to be a simple function of whether trust was honored or abused and of the length of the future.

Chapter 4

Effects of Temporal Embeddedness in Buyer–Supplier Relations*

> "In Contracts, the right passeth, not onely where the words are of the time Present, or Past; but also where they are of the Future: because all Contract is mutuall translation, or change of Right; and therefore he that promiseth onely, because he hath already received the benefit for which he promiseth, is to be understood as if he intended the Right should passe: for unlesse he had been content to have his words so understood, the other would not have performed his part first. And for that cause, in buying, and selling, and other acts of Contract, a Promise is equivalent to a Covenant; and therefore obligatory."
>
> — Thomas Hobbes (1651: 95)

4.1 Introduction

The procurement management of firms is a good example of how real–world actors deal with issues of trust. This can be seen most easily by acknowledging the fact that many economic transactions involve a time delay between the promise to deliver and the actual delivery. Contingencies may arise in the time in between, and there are many other possible reasons why what was promised may eventually not materialize. Being willing to deal with this uncertainty, handing over the control of the situation to another party, is an act of trust. For some transactions, uncertainty is high and a lot of trust is asked for. This arises, for instance, when the transaction is of a large volume and would therefore have a substantial impact on a firm's profit if things went wrong, or when the quality of a delivered good is known beforehand to be hard to assess. Our assumption is that firms can actively try to decrease the amount of trust necessary by investing time and effort in *ex ante management* of a transaction. Examples of ways to do this are investing time and effort in the search for the right business partner, investing time and effort in the writing of a detailed contract, or investing time and effort in knowledge about the technical aspects of a product (so

*This chapter is co–authored with Chris Snijders.

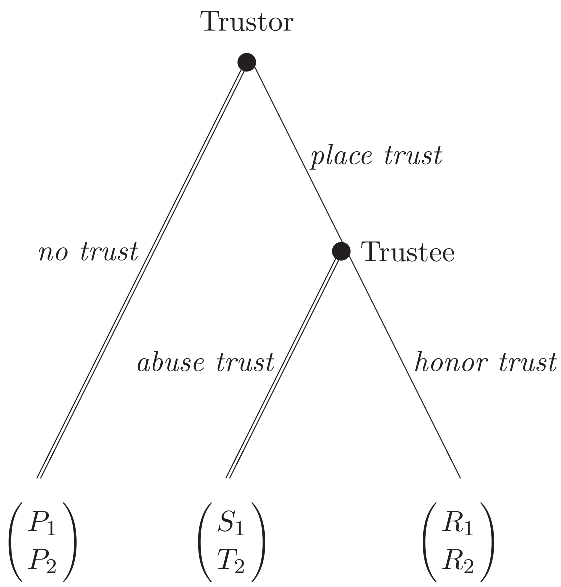


FIGURE 4.1 *Extensive form of the Trust Game ($R_1 > P_1 > S_1$ and $T_2 > R_2 > P_2$; double lines indicate equilibrium path of play).*

that the quality of the product delivered can be adequately measured). Implementing some kind of ex ante management therefore diminishes the trust that is necessary to complete the transaction. Hence, ceteris paribus, the extent to which firms invest in ex ante management, can be seen as a measure of (lack of) trust (cf. Batenburg, Raub, and Snijders 2000).

The correspondence of firms' ex ante management with trust provides a real life setting in which to test the hypotheses on the dynamics of trust relations from the previous chapter. For convenience, and to ensure that the present chapter is self-contained as much as possible, we briefly reiterate the setup of the previous chapter.

To study trust in the sociological laboratory, we used the Trust Game (Dasgupta 1988; Kreps 1990a). In this sequential two–player game depicted in FIGURE 4.1, the trustor chooses first between 'trust' or 'no trust'. The game ends if the trustor chooses not to trust. If she chooses to trust, the trustee can choose between 'honor' or 'abuse'. Abusing trust leads to a larger payoff for the trustee than honoring trust, but to the smallest payoff for the trustor. This makes that the trustor must deal with the idea that the trustee has a real incentive to abuse trust if it is given. Hypotheses were developed with regard to behavior in repeated Trust Games, and tested on the basis of experimental results of subjects playing such Trust Games at most three times in a row (see Chapter 3).

In the present chapter, we model a purchase transaction between a buyer and a supplier as a (variant of a) Trust Game (see FIGURE 4.2). The buyer has to choose to trust the supplier to deliver the product in due time and according to specifications, and the supplier has a subsequent incentive to 'shirk'. The supplier could try to get away with, for instance, late delivery, inferior quality of the product, or a lack of

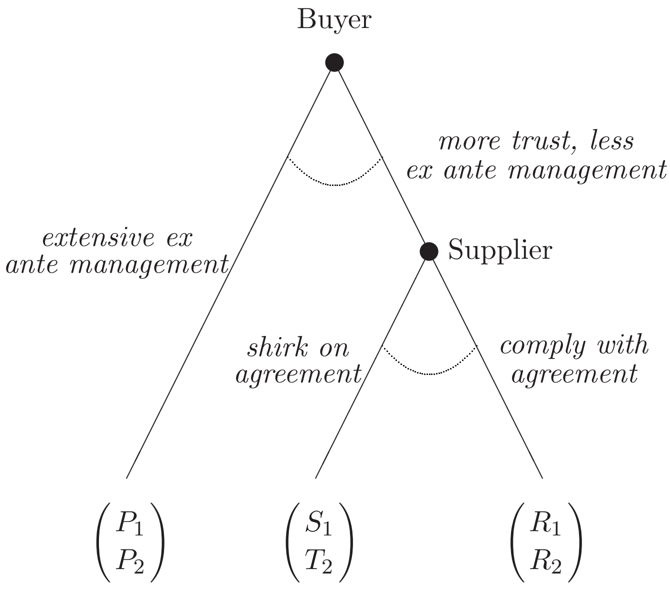


FIGURE 4.2 *Purchase transaction between buyer and supplier as an extensive form Trust Game ($R_1 > P_1 > S_1$ and $T_2 > R_2 > P_2$). Curved dotted lines indicate that the buyer's and the supplier's decision is a continuous one where the solid lines indicate their respective extreme decisions.*

after–sales assistance. Whereas we feel the Trust Game captures the essentials of a purchasing transaction quite well, the fit between real life purchasing and the abstract Trust Game is not seamless. Whereas trust is an all–or–nothing decision in the Trust Game, it is actually an interval measure in the case of buyer and supplier, because, as mentioned above, we consider the degree of ex ante management as the measure of the trust of the buyer in the supplier. Moreover, the trustee in the Trust Game is the only player who has a possibility to abuse trust. For most purchasing transactions, it is quite imaginable that not only the supplier (the trustee) but also the buyer (the trustor) has at least some incentives to abuse trust. For instance, a buyer might change or even cancel his order, or delay payment.

Purchase transactions and other inter–firm relations are being studied extensively in (transaction cost) economics (Coase 1937; Williamson 1985, 1996), economic sociology (Granovetter 1985; Smelser and Swedberg 1994), and the sociology of law (Macaulay 1963). In fact, to a large extent our analysis of purchasing transactions is closely related to these literatures (see Batenburg, Raub, and Snijders 2000 for a detailed discussion). While this chapter can be conceived as a contribution to the literature on buyer–supplier relations, we choose to focus primarily on the opportunity to test the previous chapter's theory. That is, the main aim of the present chapter is to see whether the hypotheses from the previous chapter on trust in dynamic settings can withstand a test using 'real life' data on purchase transactions between a buyer and a supplier. In a way, one could say that we leave the sociological laboratory, setting out to find out how our theory on trust as a social capital stock compares with reality outside the laboratory.

4.2 Trust over Time: Theory and Experimental Evidence

There is ample evidence that actors' behavior in a Trust Game depends on the payoff structure of the game to be played (e.g. Buskens 1999; Coleman 1990; Snijders 1996). In a setting of repeated games, however, theoretical arguments suggest that there are also effects of the past (see arguments by, for instance, Granovetter 1985; Hardin 2001; Lazaric and Lorenz 1998a; Schelling 1960) and the future (Axelrod 1984; Blau 1964; Gouldner 1960; Homans 1961; Schelling 1960; Taylor 1987) on actors' behavior in the current Trust Game. While focusing on the behavior of the trustor in a repeated Trust Game, the previous chapter proposed that learning and control effects contribute to the formation of a *dyadic social capital stock*. The core of this idea is that behavior of the trustor is guided by a comparison of the available social capital stock within the relation with the stakes involved in the current game ("Can our relationship withstand a trust problem of this magnitude?"). Technically, the 'stakes involved' refer to the probability that placed trust is abused given the payoffs of the game as represented by RISK and TEMPTATION. These ratios are defined as RISK $:= \frac{P_1 - S_1}{R_1 - S_1}$ and TEMPTATION $:= \frac{T_2 - R_2}{T_2 - S_1}$ and are further referred to as the *stakes of the game*.[1]

Based on this capital stock model, the previous chapter deduces several hypotheses on the effects of the temporal embeddedness of actors on the trustor's behavior. We recap the hypotheses and their underlying arguments in brief.

The behavioral premise that the capital stock must outweigh the stakes of the current game straightforwardly leads to the first hypothesis: *The probability of placing trust in the current game decreases in the stakes of the current game* (H3.1). A common past allows the actors to rely on their knowledge from earlier transactions (learning effects). The trustor uses the past to update her beliefs about the trustee's trustworthiness, thereby adapting the social capital stock. From this, the next two hypotheses follow directly: *The probability of placing trust in the current game increases with positive past experiences* (H3.2.1), while *the probability of placing trust in the current game decreases with negative past experiences* (H3.2.2). *Hypotheses 3.2.1* and *3.2.2* state the main effects of sharing a past with your interaction partner. A good past increases trust and a bad past decreases it. For our buyer and supplier this means that a good past implies, for instance, that time and money can be saved on the writing of detailed contracts since the increased trust is supposed to cover the loss of detail in the contract. In this scenario, a bad past would logically lead to an increased effort in ex ante management of the next transaction: the buyer has been fooled once, but will take care to prevent this from happening a next time.[2] Both hypotheses do not take

[1] On how to derive them from a given Trust Game, see for instance Snijders (1996) or page 56 in the previous chapter.

[2] As shown below, our data reveal that most firms follow a different strategy. When not satisfied with the behavior of the supplier, firms do not increase their management effort but instead find another supplier.

into account the payoff structure of the past games. Each of these past games once was a 'current game' and back then, stakes did matter for the trustor's decision. How come they should not matter in retrospect? When considering more complex versions of a social capital stock, not only past outcomes but also past stakes should be taken into account when updating the stock. *Hypothesis 3.3.1* can then be formulated as an interaction effect as follows. Suppose that a trustor has placed trust and saw his trust honored by the trustee; trustor and trustee subsequently interact in a Trust Game again. *Then, the higher the* TEMPTATION *of a past game, the more the capital stock increases with positive past experience in this game, and the larger the increase in the probability of placing trust in the current game* (H3.3.1). It also works the other way around. Suppose a trustor has had his trust abused by the trustee; trustor and trustee subsequently interact in a Trust Game again. *Then, the higher the* TEMPTATION *of a past game, the less the capital stock decreases with negative past experience in this game, and the less the decrease in the probability of placing trust in the current game* (H3.3.2). Based on the idea that peoples' evaluation about the past is biased towards recent outcomes (e.g. Kahneman 1994), we argued in the previous chapter that *the further back in the past these outcomes, the less their impact on the probability of placing trust in the current game* (H3.4). In other words, past outcomes are discounted and an abuse of trust can be forgiven by the trustor as time progresses; the downside is that positive outcomes are 'forgotten' as well. For completeness—unlike the others, this hypothesis will not be tested in the present chapter—we add that the capital stock model argues that *the probability of placing trust in a situation without a past (i.e. the current game is the first game to be played) must be equal to the probability of placing trust in a current game with a past of withheld trust only* (H3.5).

Axelrod's (1984) claim that a mutual future can establish trust is at the heart of *Hypothesis 3.6*. A longer expected future or, more precisely, a larger probability of future business dealings, makes opportunistic behavior less attractive because the punishment that will follow on uncooperative behavior in the current Trust Game will last longer (Matthews, Kordonski, and Shimoff 1983; Oliver 1984). Stated this way, *reciprocity* ensures that likely future interactions have a positive effect on trust. Taking this argument loosely, this implies that *the probability of placing trust in the current game increases in the length of the mutual future of a dyadic relation* (H3.6). A more subtle effect of future interactions is laid out in *Hypothesis 3.7*. As in *Hypotheses 3.3.1* and *3.3.2*, we claim that not only the likelihood of future interactions, but also the *kind* of future interactions matter. The argument runs as follows. When future Trust Games involve relatively large temptations to behave opportunistically, the trustee can be more certain that opportunistic behavior in the current transaction will have a detrimental impact on future business deals. This is because the trustee can anticipate that the trustor needs a relatively large capital stock in these future interactions and is therefore less likely to be able to afford to lose some of its capital stock now. In other words: when the trustee knows that future dealings will require a large amount of trust from the trustor, it makes sense for the trustee to invest in a good relationship.

The trustor knows this and can rest assured that the trustee will not abuse trust in the current transaction.[3] This leads to the following hypothesis. Suppose trustor and trustee face each other in a Trust Game, knowing that n other Trust Games will follow after the current one. *Assume Hypothesis 3.6 is true and the probability of placing trust in the current game increases in the length of the future. This probability is even higher the higher the* TEMPTATION *of the future games* (H3.7).

4.3 The Link with Trust in Buyer–Supplier Relations

We now 'translate' the hypotheses for interactions in (repeated) Trust Games in Chapter 3 to business dealings between a buyer and a supplier. For each hypothesis we add whether it was supported or rejected using the experimental data from the previous chapter. A summary is shown in TABLE 4.1.

Hypothesis 4.1 *A buyer's ex ante management increases with increasing volume and increasing opportunism potential (supported in the experimental data).*

As when Trust Games were concerned, this is the most basic of all our hypotheses. It states that buyer and supplier's behavior are affected by what is at stake in the current transaction.

Hypothesis 4.2.1 *A buyer's ex ante management decreases when buyer and supplier have done business to their mutual satisfaction before (supported in the experimental data).*

Hypothesis 4.2.2 *A buyer's ex ante management increases when buyer and supplier have had bad experiences with each other in the past (supported in the experimental data).*

Hypotheses 4.2.1 and *4.2.2* state the main effects of sharing a past with your interaction partner. A good past increases trust and a bad past decreases it. For our buyer and supplier this means that a good past implies, for instance, that time and money can be saved on the writing of detailed contracts since the increased trust is supposed to cover the loss of detail in the contract. In this scenario, a bad past would logically lead to an increased effort in ex ante management of the next transaction: the buyer has been burnt once, but will take care to prevent this from happening a next time.

[3] Note that this line of reasoning differs from Axelrod's (1984) reasoning on cooperation in the indefinitely repeated Prisoner's Dilemma. There, cooperation depends on the threshold ratio $\frac{T-R}{T-P}$, where the T stems from the focal transaction and *not* from the future transactions. This implies that for a given shadow of the future, the threshold for cooperation depends on T in the focal transaction only (or, if you prefer, on $T-R$ in the focal transaction only), suggesting that the probability of cooperation does not depend on future temptation (cf. Raub and Weesie 1993).

Hypothesis 4.3.1 *Suppose a buyer and supplier have done business to their mutual satisfaction before. Then, ex ante management in the current transaction is less when the opportunism potential in the previous transaction was larger (not supported in the experimental data).*

Hypothesis 4.3.2 *Suppose a buyer and supplier have done business before that was not to their liking. Then, ex ante management in the current transaction is less when the opportunism potential in the previous transaction was larger (not supported in the experimental data).*

Hypotheses 4.3.1 and *4.3.2* differentiate between the kinds of past dealings buyer and supplier may have had. *Hypothesis 4.3.1* suggests that when a buyer has done business with a supplier and was satisfied with the supplier's performance, this satisfaction will be larger when the temptation for the supplier to behave opportunistically was larger. In other words: the buyer is extra pleased in the situation where he knew he was at a huge risk of being treated opportunistically, but finds out the supplier was able to withstand the temptation. *Hypothesis 4.3.2* uses a similar argument. When a supplier has abused trust in the past, this does not reflect badly on the supplier when his temptation to abuse was large ("anyone would have caved in for such a temptation"). Therefore, the capital stock will not decrease as much as compared to the case where the supplier abused trust in a situation with only a small temptation to abuse ("he caved in even for small temptations").

Hypothesis 4.4 *Suppose the ex ante management in the current transaction is indeed influenced by experiences between buyer and supplier in the past. Then, the longer ago these experiences, the smaller the impact on ex ante management (supported in the experimental data).*

Hypothesis 4.4 is our equivalent of the idea that memories fade. Nice and not so nice behavior of the partner get forgotten over time; moreover, the longer ago the previous behavior was, the less it is taken as indicative for future behavior.

Hypothesis 4.5 *Ex ante management is independent of whether firms have been potential (but not actual) suppliers to a buyer (supported in the experimental data).*

The straightforward idea is that as long as you have not interacted with your business partner (and have no information about his behavior in encounters with others), you cannot learn about this partner. This hypothesis is not tested in this chapter.

Hypothesis 4.6 *Ex ante management decreases with increasing probability of future business dealings (supported in the experimental data).*

Although the generalization of Axelrod's (1984) reasoning on the effect of reciprocity in buyer–supplier relationships may seem logical, there are compelling arguments to argue against this effect in the case that buyer and supplier have not done business with each other before. An important difference between the abstract iteration of Trust Games in the sociological laboratory and the real world repeated encounters between buyers and suppliers is that in repeated Trust Games, nothing tangible that was done or paid out in a previous game is of use in the next one.[4] Typically, buyers and suppliers who face each other for the first time have an extra incentive to make more ex ante investments. This *investment effect* occurs because (part of) the investments can be carried over to following business deals. The most clear–cut example that comes to mind is that parts of contracts can be used again, so that investment in writing one is relatively attractive the more likely it is that it can be of use in future interactions. As shown in Batenburg, Raub, and Snijders (2000), this may be the reason why an increasing expectation of future business dealings decreases the investment in ex ante management, but only when buyer and supplier have done business before. By that time, the set–up investments have already been carried through, which ensures that the effect of the 'shadow of the future' becomes more important.

Hypothesis 4.7 *Suppose buyer and supplier deal with each other in a focal transaction, both anticipating that they will deal with each other again in the future. The ex ante management in the focal transaction decreases with increasing opportunism potential in the anticipated future transactions (not supported in the experimental data).*

A subtle effect of future interactions is laid out in *Hypothesis 4.7*. As argued above, we claim that not only the likelihood of future interactions, but also the kind of future interactions matter. The argument runs as follows. When future business deals involve relatively large temptations to behave opportunistically for the supplier, the supplier can be more certain that opportunistic behavior in the current transaction will have a detrimental impact on future business deals. This is because the supplier needs a relatively large capital stock in these future interactions and is therefore less likely to be able to afford to lose some of its capital stock now. In other words: when the supplier knows that future dealings will require a large amount of trust from the buyer, it makes sense for the supplier to invest in a good relationship with the buyer. The buyer knows this and can rest assured that the supplier will not abuse trust in the current transaction. On the other hand, arguments following from the investment effect elucidated in connection with *Hypothesis 4.6* might also be applicable here. If the buyer expects future business with large opportunism potential, it would eventually pay for the buyer to invest more into current ex ante management and carry the investment over to future deals.

4.4 Data and Variables

The data at hand ('The External Management of Automation', MAT) result from a survey on the purchase of IT–products (software and hardware) by Dutch SMEs (small

[4] We write "nothing tangible" instead of just "nothing" because one could, of course, rightfully argue that what can be of use in next rounds is, for instance, knowledge about the behavior of the partner.

Table 4.1 *Hypotheses with respect to the buyer's investment in managing the current transaction (I).*

Independent variable	Investment in management Predicted[a]	Hypothesis	Experimental results[b]
Focal transaction			
Volume and opportunism potential	+	*4.1*	supported
Past transactions			
Successful past transactions	–	*4.2.1*	supported
Unsuccessful past transactions	+	*4.2.2*	supported
Successful past transactions × past volume and opportunism potential	–	*4.3.1*	not supported
Unsuccessful past transactions × past volume and opportunism potential	–	*4.3.2*	not supported
Having a past × discounting past experience	+	*4.4*	supported
Potential vs. actual supplier	0	*4.5*	supported
Future transactions			
Future transactions	–	*4.6*	supported
Future transaction × future volume and opportunism potential	–	*4.7*	not supported

[a] – : the investment in management decreases. + : the investment in management increases. 0 : the investment in management is not affected.
[b] Experimental results refer to the analyses in Chapter 3.

and medium sized enterprises; between 5 and 200 employees) that was collected in two waves, namely in 1995 (MAT95) and in 1998 (MAT98). See Batenburg (1997a, 1997b) for a brief introduction to MAT95, Batenburg and Van de Rijt (1999) for an overview on MAT98, and Buskens and Batenburg (2000) for information on the combined data set of MAT95 and MAT98. In the following, we provide a brief introduction to the data, based on the descriptions mentioned above. For a more elaborated discussion of the data see Rooks (2002: ch. 3). The data analyzed in this chapter are based on the combined MAT95 and MAT98 data set.

4.4.1 IT–Transactions of SMEs: MAT95 and MAT98

The Dutch IT–market at the time of the data collection was a typical seller's market: the substantial risk associated with the purchase of IT–products (e.g. the complexity of the products, monitoring problems of the buyer, opportunistic behavior of suppliers, high switching costs for the buyer once a product has been bought, and relatively little expertise of the SMEs) can lead to problematic transactions for the buyer (Raub, Batenburg, and Snijders 2000; Rooks 2002: ch. 3). We therefore choose to see these

IT–transactions as one–sided trust problems with the supplier in the role of the trustee and the SME in the role of the trustor.[5]

Sampling Frame and Procedure

The sampling frame was a business–to–business database of Dutch SMEs.[6] The database contains about 80% of all Dutch SMEs and is representative for the population of Dutch small and medium sized enterprises (see Batenburg 1997a).

To be able to collect information on different IT–transactions (i.e. software as well as hardware, complex as well as less complex transactions), the sampling procedure in 1995 was subject to three stratification criteria:

1. The type of IT–products in use by the firm: either standard IT–products, or standard as well as complex (tailor made) IT–products.
2. The number of IT–specialists employed by the firm: no specialist, only part–time specialists, or one or more full–time specialists.
3. The network embeddedness of the buyer, that is, the extensiveness of the network of interfirm relations within the sector (whether or not firms communicate with each other about other firms in the sector, about other suppliers, or about mutual business contacts) as judged by 21 independent business experts. This lead to sectors being classified as having either weak, medium, or strong interfirm relations.

Following this sampling procedure results in a sampling design with $2\times3\times3 = 18$ cells. To sample the respondents, a randomization procedure was used until at least 30 cases per cell were collected. This results in a minimum of $18 \times 30 = 540$ questionnaires to be distributed. Due to good response rates half way into the field work, it was decided to oversample transactions on complex IT–products.

Questionnaire and Field Work in 1995

In 1995, the field work began with a short computer assisted telephone interview (CATI). By this, some key information of the sampled firms was collected and cooperation was asked for an interview on an IT–purchase. The key information of those firms not willing to take part in the survey was used for a non–response analysis. For the firms that wanted to participate, an IT–transaction from within the last 10 years was randomly selected. In most cases, an interviewer personally brought the questionnaire to the respondent and returned with the filled out questionnaire as soon as the respondent had completed it. Some respondents preferred to have the questionnaire

[5] That SMEs have little chance to avoid such risky transactions is supported by their lack of the possibility of an in–house production of the necessary IT–products (i.e, make or buy decision). Only in 4% of all transactions could the buyer have easily or very easily produced the purchased product in–house.

[6] The database is developed and owned by Directview, a Dutch firm specialized in IT marketing.

mailed. Most respondents were IT–managers. At the end of the questionnaire it was asked whether or not the respondent was willing to fill in a second questionnaire on another IT–purchase. If yes, an additional IT–transaction was chosen by selecting the most innovative and complex transaction from within the last 10 years. In this case, the questionnaire was, without exception, sent by mail. This leads to an additional batch of questionnaires on complex IT–products.

The questionnaire used in 1995 contained several segments. First, it covered the main characteristics of the product and the transaction, followed by questions regarding the search for a suitable product and supplier, and the relation with the supplier previous to the focal transaction. Next, the questionnaire covered the agreement used and the terms of trade regarding the transaction. Finally, there were questions regarding the period after the transaction had been completed and some questions on characteristics of the buyer firm and the respondent (see also Rooks 2002: ch. 3).

In total, the 1995 sample consisted of 1788 SMEs of which 1325 were suited for the survey. From those 1325, 892 were willing to take part in the survey (673 were visited and 219 received the questionnaire by mail). In total, 778 questionnaires were returned (response rate = 59%). From the 892 respondents, 327 were willing to fill out an additional questionnaire regarding a second IT–transaction, of which 183 questionnaires were returned. Including a pretest of 10 transactions, this leads to a data set with information from 788 firms on 971 transactions. A non–response analysis showed no significant differences between the response and non–response group on the key variables gathered during the CATI–interview.

Questionnaire and Field Work in 1998

In 1998, a second survey with an almost identical questionnaire as in 1995 was conducted. All 1995–respondent firms were again contacted by use of CATI. From the 778 respondent firms in 1995, 664 firms were reached. If possible, the same respondent as in 1995 was contacted. If this was not possible, a respondent within the firm was searched who was well informed on the 1995 transaction(s). In this telephone interview, additional questions regarding the 1995 transaction(s) and cooperation for a new survey was asked. This procedure resulted in 560 informants who were willing to answer additional questions on their 1995 transaction(s). From those 560 respondents, 463 were willing to fill in the 1998 questionnaire on a new IT–transaction. If possible, a transaction with the same supplier as in 1995 was selected. Otherwise, a transaction with a new supplier was chosen. In both cases a transaction as complex as possible was selected. The questionnaires were sent by mail and 281 were returned (response rate = 42%). Again, a non–response analysis (also with the response and non–response groups of 1995) showed no significant differences between key variables of the participating and non–participating firms.

In total, combining the field work in 1995 and 1998, we have information from 788 firms on a total of 1252 IT–transactions. The number of questionnaires per firm ranges from one to three. More precisely (see Table 4.2), there are 404 firms that only

TABLE 4.2 *Number of buyers and transactions per buyer.*

Questionnaire	Number of buyers	Transactions per buyer	Total number of transacions
First in 1995	404	1	404
First and second in 1995	103	2	206
First in 1995 and 1998	201	2	402
First and second in 1995 and, 1998	80	3	240
Total	788		1252

answered the first 1995 questionnaire (404 transactions). There are 103 firms that filled out two questionnaires in 1995 (206 transactions). There are 201 firms that filled out one questionnaire in 1995 and one in 1998 (402 transactions). Finally, there are 80 firms that filled out all three questionnaires, two in 1995 and one in 1998 (240 transactions).

4.4.2 Transactions and Buyer–Supplier Dyads

The MAT–data on IT–transactions resemble the experimental data of the previous chapter in the sense that the data contain single transactions between a buyer and supplier ('one–shot games') as well as repeated transactions between the same buyer and supplier (once and twice 'repeated games').

The combined MAT95 and MAT98 data set contains information on 1252 transactions (or cases). TABLE 4.2 reveals 404 buyers with one transaction only, $103 + 201 = 304$ buyers with two transactions (i.e. $206 + 402 = 608$ cases) and 80 buyers with three transactions (i.e. 240 cases). The 404 respondents with one questionnaire filled out form the basis of the 'one–shot transaction' cases. For the remaining 384 firms, we thus have information on more than one transaction (namely on two or three). In the following, we rearrange the data such that we end up with dyads of buyer–supplier couples which have done repeated transactions. Buyers with more than one transaction, but each transaction with a different supplier, will be added to the 404 buyers with only one transaction.

The combined MAT–data disclose that out of the 304 buyers with two transactions (second and third row in TABLE 4.2), there are 197 buyers with two transactions with two different suppliers ($146/2 = 73$ with two transactions in 1995 and $248/2 = 124$ buyers with one transaction in 1995 and one in 1998), and $30 + 77 = 107$ buyers with two transactions with the same supplier. The latter, therefore, form 107 'real dyads' in the sense of our Trust Game experiments with two transactions. The right–hand side of TABLE 4.3 (row three and five, respectively) shows that 30 buyers reported on those two transactions in MAT95 while the other 77 buyers informed on one transaction in MAT95 and the other in MAT98. Those 197 buyers that reported information on two IT–transactions with different suppliers can be treated as 394 additional one–shot buyer–supplier relations. The right–hand side of TABLE 4.3 (row two and four,

TABLE 4.3 *Number of transactions within a buyer–supplier couple (dyad).*

Questionnaire	Total number of buyers	Total number of transactions	Repeated transactions in dyad: First TA	Second TA	Third TA
One in 1995	404	404	404		
Two in 1995	103	206	146		
			30	30	
One in 1995, one in 1998	201	402	248		
			77	77	
Two in 1995, one in 1998	80	240	121		
			34	34	
			17	17	17
Total	788	1252	1077	158	17

respectively) shows that 146 transactions (from 73 buyers) were reported in MAT95 while 248 transactions (from 124 buyers) were reported first in MAT95 and second in MAT98. This increases the number of 'one–shot transactions' to $404 + 394 = 798$.

This leaves us with 80 buyers with three transactions (i.e. 240 cases; see row four in TABLE 4.2). Out of these 80 buyers, 29 buyers reported on three transactions with three different suppliers. We consider them as $29 \times 3 = 87$ 'one–shot transactions'. Further, there are 34 buyers that reported on two transactions with the same supplier (row seven on the right–hand side of TABLE 4.3) and a third transaction with another supplier. The latter 34 transactions are added to the 'one–shot transactions' ($87 + 34 = 121$, row six on the right–hand side of TABLE 4.3), which increases the one–shot games to a total of $404 + 394 + 87 + 34 = 919$. Consequently, we also increase the buyer–supplier dyads with two transactions to a total of $107 + 34 = 141$. Finally, there are 17 buyers in the data set with three transactions with the same supplier. These form 17 dyads with three transactions (row eight on the right–hand side of TABLE 4.3).

In sum, we now have 919 transactions that serve as the equivalent of our one–shot Trust Games, 141 buyer–supplier relations with two transactions serving as the equivalent of a once repeated game, and 17 buyer–supplier relations with three transactions serving as the equivalent to a twice repeated game. TABLE 4.3 summarizes how TABLE 4.2 'translates' into transactions within buyer–supplier dyads.

Though the parallel with the setup of the experimental data is straightforward, the differences are just as apparent. In the experimental data, subjects formed a fixed dyad as trustor and trustee; they had no choice but to interact repeatedly. This is of course not the case in the MAT–data. Although perhaps at some cost, buyers can choose the supplier they prefer and can also switch to another supplier. The MAT–data therefore show us behavior under incentive problems when one is free to choose a partner, whereas this was clearly not the case in the experiments of the previous chapter.[7]

This construction of buyer–supplier couples immediately points out a problem with how we treat the data. We suppose that the buyer had had no other contacts to any other third parties. We neglect the fact that the buyer might have gotten information on possible suppliers from other buyers or suppliers and, therefore, had been given the chance to process information on the supplier (learning) even before they first interacted. More important, we also disregard the circumstance that the buyer may have done business with the actual or another supplier prior to the first transaction, between the transactions, or after the last transaction we have in the data. Unfortunately, the MAT–data do not provide information on any transactions not reported in the respective questionnaire.

4.4.3 Operationalization of Variables

As mentioned before, we consider the buyer's level of investment in safeguarding a transaction the equivalent of (lack of) trust. Our dependent variable in the buyer–supplier analyses is therefore not binary as in Trust Games, but on an interval scale.

On average, buyers invested 4 mendays in arranging and negotiating the contract with the supplier. About 13% of all buyers invested 5 or more mendays in the ex ante management of the transaction. The financial volume of a transaction was on average NLG 117,000 (Dutch Guilders, about € 53,100) and 25% of all transactions had a financial volume larger than NLG 75,000. The buyer, on average, was a firm with about 80 employees and roughly 10% of all buyers had 150 or more employees.

These figures are not very different for those buyers that reported to have done business with the current supplier in the past. These buyers have done business with the current supplier for, on average, 6.5 years, and indicated that business with this supplier has occurred regularly. They invested a little bit more than 3 mendays in arranging and negotiating the contract with the supplier. About 9% of these buyers invested 5 or more mendays in the ex ante management of the transaction. The financial volume of a transaction was about NLG 134,000 and 25% of all these transactions had a financial volume larger than NLG 75,000. At the time of the transaction, there were on average 80 people employed in the buyer's firm and 10% of the buyer firms had 160 or more employees.

[7] In fact, for those dyads in the MAT–data which represent repeated transactions, it seems likely that these are all relations with a mainly positive past. The MAT–data support this idea. On a satisfaction scale ranging between 1 and 10, the supplier on average scored a 7.21 (7.33 for the suppliers in dyads of repeated transactions), and about 89.4% of all suppliers received a grade of 6 or higher (for the repeated transactions, 91.2% scored 6 or higher). Roughly, grades of 6 and higher indicate 'satisfactory' or 'good' performance; a grade below 6 indicates 'bad' performance.

INVESTMENT IN MANAGEMENT

Our dependent variable is the principal component of questions regarding the number of person–days involved in contracting, whether or not the used contract was tailor–made or standard, the number of departments involved in contracting, whether or not external legal help was used, the number of product related technical specifications arranged in the contract, the number of contract issues discussed, and the number of contract issues finally agreed on orally or written (cf. Batenburg, Raub, and Snijders 2000).[8,9]

We next need to identify suitable variables that can represent what in the experiments fulfilled the role of the stakes of the current transaction as well as the temporal embeddedness of the buyer and the supplier.

VOLUME AND OPPORTUNISM POTENTIAL

We argued that an IT–transaction can be seen as a one–sided Trust Game with the buyer in the role of the trustor. The behavior of the trustor in a Trust Game was said to be dependent on RISK $:= \frac{P_1-S_1}{R_1-S_1}$ and TEMPTATION $:= \frac{T_2-R_2}{T_2-S_1}$ while the trustee's behavior was assumed to be dependent on TEMPTATION only (previous chapter, Section 3.3.1; Snijders 1996). We choose to represent these stakes through two variables: the volume of the focal transaction and its opportunism potential.[10] The volume of the transaction serves as an overall measure for the buyer's loss if trust is abused (i.e. $P_1 - S_1$ versus $R_1 - S_1$) and also as a measure for the supplier's incentive to abuse trust after it is given (i.e. $T_2 - R_2$). The supplier's opportunism potential is derived as a weighted average of the damage potential and the buyer's monitoring problems (as well as the financial volume of the transaction). The damage potential covers 'how bad it would be' if it turns out that the delivered product cannot be used productively and directly reflects the S_1–value. It is measured using questions on the importance of the product for the automation and profitability of the firm, the importance that the product was delivered on time, the long–time perspectives regarding support, suitability and compatibility of the product, and the damage in terms of time and money (new product, new data entry, training personnel, idle production) inflicted on the firm

[8] See Appendix B: Q4.11, Q4.12, Q4.13, Q4.14a, Q4.14b, Q4.16, Q4.17, and Q4.18

[9] In a similar analysis using only MAT95, Buskens (1999: ch. 5) operationalizes 'lack of trust' by using the number of safeguards specified in the buyer's contract (i.e. the number of orally agreed or written contract issues). He intentionally neglects the buyer's effort in writing up the contract since he argues that "[...] predictions of the game–theoretic model are mainly based on the incentives of the supplier to abuse trust" and that therefore the focus should be on "the type and number of safeguards [rather] than on the effort invested by the buyer." (p. 131). However, we argue that the buyer's assessment about the supplier's incentive to abuse trust is not only represented by the completeness of the contract but also by the buyer's time invested in writing up the contract. The buyer's (lack of) trust in the supplier is therefore not only a function of how complete a contract is but also how carefully it is planned and executed.

[10] The financial volume and the opportunism potential are also referred to as the 'problem potential of a transaction'.

if the product had failed.[11] Monitoring problems, on the other hand, summarize the buyer's ability to judge the quality of the product, to screen the market for alternative suppliers or products, and to assess the complexity of the product as a whole, thereby reflecting the difference $T_2 - R_2$ and thus the supplier's incentive to abuse trust if placed.[12]

Temporal Embeddedness

We choose two different approaches to the operationalization of the temporal embeddedness of the transaction between the buyer and the supplier. The first approach will make use of the information on the past and future as provided by the respondent in the questionnaire on the focal transactions (Method A). The second approach makes use of the fact that we can identify true buyer–supplier dyads in the combined data of MAT95 and MAT98 (Method B). We first introduce these two approaches in more detail.

Method A: Method A is actually the standard way of using cross–sectional questionnaires. To operationalize the temporal embeddedness of the buyer and the supplier, we simply use the information on the past[13] and future[14] as collected in each of the 1252 questionnaires on the focal transactions.

Method B: Since we can identify buyer–supplier dyads in the data, we can make use of this in constructing the temporal embeddedness of the buyer and the supplier. Method B actually postulates that the different dyads in the data (single transactions, once repeated transactions with the same buyer, and twice repeated transactions with the same buyer) can be seen as the direct equivalent of the experimental setup. To construct the dyad's past and future, we now apply the following procedure. First note that any transaction is either the first, second, or third in a stream of transactions with the same supplier: we have no buyer–supplier dyads in the data with more than three joint transactions. For a focal transaction that is a first transaction, we boldly assume that buyer and supplier have no past. If a transaction is the second or third in a stream of transactions, the past variables are constructed using the information provided in the questionnaire(s) on the previous transaction(s), and we disregard the information about the past in the focal questionnaire. To define the future we follow a similar procedure. For transactions that are last in a stream of two or three we assume that buyer and supplier have no mutual future. For other transactions the variables describing the future are constructed by using the information provided in the

[11] See Appendix B: Q1.6a, Q1.6b, Q1.10, Q1.12a, Q1.12b, Q1.12c, Q1.13a, Q1.13b, Q1.13c, and Q1.13d.

[12] See Appendix B: Q1.2, Q1.7, Q1.8, Q1.15, Q2.8, Q2.9, Q2.10, Q6.3ab, and Q6.3bb.

[13] See Appendix B: Q3.7, Q3.8, and Q3.9.

[14] See Appendix B: Q3.13.

questionnaire(s) on the future transaction(s), thereby disregarding the information on the possibility of future interactions in the questionnaire on the focal transaction.

In short, in Method B all 919 buyer–supplier couples with only one transaction in MAT95 and MAT98 are, regardless of the information in the questionnaire, treated as 'one–shot transactions' with no mutual past and no mutual future. And, all 'x-times repeated' transactions are treated as a set of $x + 1$ interactions between a buyer and a supplier who had not done business before, and who knew they would do business exactly $x + 1$ times.

Both methods have their pros and cons (see below). As far as possible, we test each hypothesis using both methods, so that support or rejection of each hypothesis is backed up by different ways of looking at the data. We now provide an overview of the variables used in the analyses and for each variable we provide the numbering of questions as used in Appendix B.

Temporal Embeddedness using Method A

Method A is easiest, both to implement and understand. When operationalizing the variables, we stick to the information in the questionnaire as provided by the respondent.

Past: We choose the simplest operationalization here: a binary variable covering whether or not buyer and supplier had done business before (question Q3.7). In the analysis of those dyads with more than one transaction, we also include the frequency of former transactions with the supplier (Q3.8). Note that we have information on past transactions even for those buyers with only one transaction in MAT95 or MAT98.

Past Stakes: The stakes of past transactions are not directly measured in the questionnaire. We approach the stakes of the past transactions by using the information on how extensive past business with the supplier was (Q3.9).

Future: The future is operationalized by the buyer's expectation, prior to the current transaction, about possible future business with the same supplier (Q3.13).

Future Stakes: Only the MAT98 questionnaire contained information on the expected volume of future transactions. We therefore have information on the future stakes in 281 cases only. Exclusion of missing cases in the regression analyses decreases the number of cases to an unacceptable level which prohibits the use of Method A in testing hypotheses regarding future stakes.

Temporal Embeddedness using Method B

For Method B we operationalize the variables by extracting the necessary information from the repeated transactions between the buyer and the supplier. For instance, data about future business for a focal transaction are taken from later transactions in the database between the same buyer and supplier. This implies, as mentioned before, that one–shot transactions in MAT95 and MAT98 are treated as having no temporal embeddedness (even though information as provided in the questionnaire on this focal transaction might indicate, for instance, that there have been prior business dealings between buyer and supplier).

Past: Again, we introduce a dummy–variable to test for the effect of having a shared past of business dealings versus not having done business before. In addition, we operationalize length of the past by using a count variable representing the length of the past, constructed as the number of past transactions between the supplier and the buyer (either 0, 1 or 2). Further, we include a variable that gives the years elapsed since the first transaction between the buyer and the supplier.

Past Stakes: We can construct precise variables for the stakes of the past transactions by using the volume and the opportunism potential of the previous transactions (for the variables used, see description above). To make this more clear: suppose we consider a transaction that is the last one of a set of three between buyer and supplier. For 'past stakes' we then use the financial volume and opportunism potential as indicated in the questionnaires on the two previous transactions.

Future: As with the past, we construct a count variable for the length of the future by counting the number of future transactions between buyer and supplier in the data (either 0, 1 or 2).

Future Stakes: As with past stakes, we use the information on the volume and the opportunism potential as provided in the questionnaires on the transactions at $t+1$ and, if applicable, $t+2$ to indicate the future stakes of the current transaction in time period t.

Using Method B, we create—though admittedly in a somewhat artificial way—a data set on buyer–supplier dyads 'playing Trust-Games' either once, twice, or three times.

Pros and Cons of Method A and Method B

Method A offers a simple and straightforward approach to operationalize variables regarding the temporal embeddedness. We make use of the summarized information on the whole past and expected future (Q3.13) with a given supplier. Consequently, we obtain information on the whole past and the future for all respondent firms, even for those with only one questionnaire filled out. However, this approach does ask for inferences of the respondent about the past or future *some time ago* ('what kind of past did you have a year ago' or 'how likely was it that you would do business again before the transaction actually started'), which are difficult for respondents to make. Clearly, this is a downside of Method A. Method B does not have this flaw—it has others, no doubt—because it measures actual information about past and future dealings directly,

Table 4.4 *Overview of variables and descriptive statistics.*

Variable name	N	Mean	s.d.	Min	Max
Dependent variable					
Investment in management[a]	1252	0	1	−2.08	3.00
Focal transaction					
Volume[b]	1236	−0.80	1.21	−2.08	5.77
Opportunism potential[a]	1252	−0.04	1.37	−3.18	3.82
Past transactions					
Past (dummy)[d]	1252	0.53	0.50	0	1
Volume[b]	1252	−0.04	1.37	−3.18	3.82
Frequency[c]	1252	1.76	1.86	0	5
Length[e]	1252	0.15	0.40	0	2
Year of first transaction[f]	1227	5.57	2.97	0	20
Volume at $t-1$[c]	1252	0.34	0.99	0	5
Volume at $t-2$[c]	1252	0.03	0.32	0	4
Opportunism potential at $t-1$[a]	1252	0.01	0.49	−3.18	2.98
Opportunism potential at $t-2$[a]	1252	0.00	0.14	−1.79	2.83
Future transactions					
Expected future business[c]	1228	2.89	1.32	1	5
Future (dummy)[d]	1252	0.14	0.35	0	1
Length[e]	1252	0.15	0.40	0	2
Control variables					
Standardization in contracting[c]	1179	2.57	1.15	1	5
Availability own legal expertise[d]	1252	0.19	0.39	0	1
Size buyer[b]	1213	3.69	1.03	0	8.70
Size supplier[b]	1228	3.36	1.41	1	5
Network activities[a]	1220	−0.05	1.76	−3.82	5.86

Note: N = Number of cases; s.d. = standard deviation; Min = minimum value; Max = maximum value.

[a] Standardized factor score.

[b] Natural log of financial volume in 100,000 guilders; natural log of number of employees, respectively.

[c] Five point scale.

[d] Dummy, 1=yes.

[e] Number of transactions.

[f] Years since first transaction with the current supplier.

using the available information in the questionnaires prior and past the focal transaction between the same buyer and supplier. By making use of these variables, we also secure compliance with the setup of the experiments in the previous chapter: the data format is then such that we have one–shot, once repeated and twice repeated interactions as in the experimental data from the previous chapter. This immediately points out the drawbacks of Method B. We artificially put a 'time frame' of at maximum three transactions on a buyer–supplier relationship that may in fact involve many more transactions. It is possible that buyer and supplier have done business either before the transaction that is the first in the data, between the transactions that we have in the data, or after the last of their transactions that we have in the data.

Control Variables

We control for several transaction and firm specific characteristics available in the data, most of them related to possible differences between buyer–supplier dyads with regard to the efficiency with which ex ante management can be implemented. First, the size of the respondent buyer firm and the supplier (Q3.2 and Q6.1b). Second, the availability of legal expertise for the buyer (Q6.3ac, Q6.3bc, and Q6.3cc). Additionally, we control for the buyer firm's standardization in its procedures around contracting (Q4.20) and the buyer's network activities (Q6.4, Q6.5, Q6.6, and Q6.7). We refer the reader to Batenburg, Raub, and Snijders (2000) for a more thorough discussion of these control variables.

Table 4.4 presents an overview of the variables. The scale of most of the variables is rather meaningless since they are either a weighted average of several questions (factor score) or measured on a five point scale (see also Batenburg, Raub, and Snijders 2000).

4.5 Results

As much as possible, we test all hypotheses using both Method A and Method B. Moreover, we separately run all analyses on two sets of cases. First, we analyze all buyer–supplier dyads in MAT95 and MAT98 (Table 4.5). Second, we restrict the analyses to buyer–supplier dyads with a history of shared business dealings (Table 4.6). This leaves us with 650 cases in the Method A analysis. In the analyses using Method B (Table 4.6), we also exclude the transactions that are the first in a stream of dealings between buyer and supplier, but where the data reveal that buyer and supplier had done business in the past before. Loosely speaking, these are the cases where the setup in the data does not match the experimental setup of the previous chapter. When we exclude these transactions (where there was a history of business dealing between buyer and supplier, but information regarding this history is not available in the data), we use Method B in a 'clean' way: only those cases are included where information about past transactions is taken from previous questionnaires on business dealings between buyer and supplier. The downside is that this leaves us with only 161 cases in the analysis.

The MAT–data surveyed 788 firms with a maximum of three transactions per firm. We thus have a data set that is comparable to the experimental data of the previous chapter. The experimental subjects made repeated decisions on whether or not to place trust while the buyers in the MAT–data made repeated decisions on their investment in ex ante management. The dependent variable in the analyses of this chapter is thus no longer a binary variable (place or withhold trust) but an ordinal variable measuring the buyer's investment in his ex ante management of the focal transaction. This forbids to use a logistic model such as the fixed–effects logistic regression of the previous chapter (see Section 3.4.3). However, we can use a similar method for interval dependent variables, namely, a *cross–sectional time–series regression model* (Stata's `xtreg`). Besides the necessary variables on temporal embeddedness and the 'stakes

TABLE 4.5 *Maximum likelihood random–effects linear regression coefficients on the buyer's investment in management.*

	Method A		Method B	
Regressor	Coefficient	p–value	Coefficient	p–value
Focal transaction				
Volume	0.129	0.000	0.127	0.000
Opportunism potential	0.267	0.000	0.272	0.000
Past transactions				
Past (dummy)	−0.178	0.091	−0.096	0.032
Volume	0.017	0.626		
Future transactions				
Expected future business	−0.009	0.605		
Future (dummy)			0.184	0.317
Length[a]			−0.202	0.215
Control variables				
Standardization in contracting	0.065	0.001	0.064	0.001
Availability own legal expertise	0.118	0.039	0.115	0.045
Size buyer	−0.006	0.803	−0.006	0.814
Size supplier	0.060	0.001	0.059	0.001
Network activities	0.035	0.008	0.035	0.008
Constant	−0.074	0.566	−0.114	0.348
Variance on dyad level (σ_v)	0.4997 (0.000)		0.5051 (0.000)	
Variance on transaction level (σ_ϵ)	0.5527 (0.000)		0.5503 (0.000)	
Fraction of variance due to σ_v (ρ)	0.4497		0.4573	
	$LL = -1271.4335$		$LL = -1281.5655$	
	LR $\chi^2(11) = 511.82\,(0.000)$		LR $\chi^2(10) = 509.43\,(0.000)$	

Note: Method A is based on $N = 1147$ observations from 989 buyers with on average 1.2 transaction. Method B is based on $N = 1154$ observations from 996 buyers with on average 1.2 transactions.

[a] The length of the future is an ordinal variable counting the number of future transactions between the buyer and the supplier.

of the current game', the MAT–data also contain information on characteristics of buyer firms. Given that we have more and adequate measurements of buyer firm characteristics, we can now use a *random–effects* estimator (see Section 3.4.3 for more information). We thus obtain inferences not only about how temporal embeddedness and financial volume and opportunism potential affect the buyer's ex ante management, but also which buyer firms are more likely to trust the supplier in general (see section CONTROL VARIABLES).

In the following, we discuss the results of the statistical tests. First we present the influence of the current stakes on investment in management. We then outline the effects of the past as well as the future[15] and finally discuss the effects of the control variables.

[15] Note that TABLE 4.6 reports past and future variables not used in the analyses of TABLE 4.5. The analyses in TABLE 4.6 refer to those buyer–supplier dyads with a past only. This allows, especially for Method B, to construct additional variables such as, for instance, the length of the mutual past and future. For more information, see the discussion on TEMPORAL EMBEDDEDNESS above.

Table 4.6 *Maximum likelihood random–effects linear regression coefficients on the buyer's investment in management for buyer–supplier dyads with a past.*

	Method A		Method B	
Regressor	Coefficient	p–value	Coefficient	p–value
Focal transaction				
Volume	0.140	0.001	0.053	0.556
Opportunism potential	0.263	0.000	0.341	0.000
Past transactions				
Frequency	−0.079	0.010		
Length			0.008	0.988
Year of first transaction			−0.001	0.969
Volume	0.046	0.228		
Past × ...				
... Volume at $t-1$			0.029	0.687
... Volume at $t-2$			−0.028	0.882
... Opportunism potential at $t-1$			0.006	0.943
... Opportunism potential at $t-2$			0.013	0.948
Future transactions				
Expected future business	−0.016	0.573		
Future (dummy)			−0.099	0.612
Control variables				
Standardization in contracting	0.058	0.019	0.117	0.037
Availability own legal expertise	0.083	0.248	0.129	0.403
Size buyer	0.002	0.957	−0.111	0.116
Size supplier	0.046	0.036	0.011	0.805
Network activities	0.016	0.349	−0.004	0.907
Constant	0.011	0.947	0.186	0.790
Variance on dyad level (σ_υ)	0.4791 (0.000)		0	
Variance on transaction level (σ_ϵ)	0.5421 (0.000)		0.7051 (0.000)	
Fraction of variance due to σ_υ (ρ)	0.4385		0	
	$LL = -701.5884$		$LL = -172.1850$	
	LR $\chi^2(10) = 285.96$ (0.000)		LR $\chi^2(14) = 70.87$ (0.000)	

Note: Method A is based on $N = 650$ observations from 558 buyers with on average 1.2 transactions. Method B is based on just $N = 161$ observations from 147 buyers with on average 1.1 transactions because it also excludes the cases that are the first in a stream of transactions between buyer and supplier in the data.

Effects of the Stakes of the Focal Transaction

Table 4.5 and Table 4.6 both reveal a statistically significant and positive effect of the financial volume as well as of the opportunism potential on ex ante management. Moreover, the effects of the financial volume and the opportunism potential are similar across all analyses. *Hypothesis 4.1* is therefore clearly supported, as it was in the previous chapter and in earlier analyses on the MAT–data using different operationalizations (e.g. Batenburg, Raub, and Snijders 2000; Buskens 1999). Trust in the supplier decreases as the current stakes become larger. Or, perhaps more aptly put, the need to increase trust between buyer and supplier through the use of ex ante management increases when the stakes of the current transaction are larger. This result is also supported by Blumberg (1997) who focuses on the explanation of ex ante management in R&D alliances.

MAIN EFFECTS OF PAST TRANSACTIONS

Hypothesis 4.2.1 argues that a positive past decreases the buyer's investment in ex ante management. In the data, 'having a past' is equal to 'having a good past' with the supplier (in general, 'not having a good past' causes buyers to use a different supplier for future transactions). The dummy whether buyer and supplier have done business dealings before shows the predicted negative effect on the buyer's investment in management, both under Method A ($p = 0.091$) and under Method B ($p = 0.032$). These findings support *Hypothesis 4.2.1.* The data suggest, as mentioned before, that most buyers are satisfied with their suppliers. Consequently, the effect of a negative past on ex ante management (*Hypothesis 4.2.2*) cannot be tested.

Now we turn our attention to whether it matters—given that buyer and supplier have done business before—what the stakes were in the previous transactions. TABLE 4.6 discloses that Method B does not lend support to the predicted negative effect of past stakes on current investment in management (*Hypothesis 4.3.1*). We tried various other ways to test the effect of past stakes, none of which showed any evidence that the kind of past matters.[16] We must therefore reject *Hypothesis 4.3.1.* Since buyers are generally satisfied with their past, the interaction effect between a negative past and the respective past stakes on ex ante management cannot be tested with the MAT–data (*Hypothesis 4.3.2*).

EFFECTS OF FUTURE TRANSACTIONS

Next, we focus on the effect of (the probability of) future transactions. None of our analyses lends support to the idea that (the probability of) future transactions are related to the investment in ex ante management. This means that neither evidence in favor of the investment effect nor the reciprocity effect can be found in the data. Moreover, regardless of the cases we use, the size of the coefficient is small in all analyses except in Method B of TABLE 4.5, where it is somewhat larger but still far from significant. This finding is in line with results in Batenburg, Raub, and Snijders (2000), Blumberg (1997) as well as Buskens (1999) to the extent that they also do not find a main effect of future business when all MAT–cases are used. Nevertheless, the result is at odds with Batenburg, Raub, and Snijders to the extent that they do find an effect of (the probability of) future business dealings for cases with a shared past, which supports the existence of a reciprocity effect.

On grounds of the insignificant future coefficients in all models, we have to reject *Hypothesis 4.6.* Our results are at odds with the experimental findings in the previous

[16]For instance, we also tried to test the effect on a set of cases being a hybrid of Method A and Method B. We additionally included the one–shot transactions where the questionnaire informed us that business with the same dealer has been done in the past. For lack of information on the precise past volume, we coded past volume of such 'first transactions' as zero. Also we tried to impute past volume of these additional cases as a function of the information we have. However, none of these extra analyses showed an effect of past stakes. Additionally, we tested past volume using Method A. However, for lack of suitable variables in the MAT–data, we could only use the extensiveness of past business with the supplier (see Appendix B: Q3.9) as a proxy for past volume. TABLE 4.5 and TABLE 4.6 show that neither analyses support *Hypothesis 4.3.1.*

chapter, where the length of the future significantly increased the probability of placing trust in the current game. Remember, however, that in the laboratory only the reciprocity effect characterized the influence of the future on the probability of placing trust.

Using Method B, we can test the interaction between having a future and its respective stakes (*Hypothesis 4.7*). We argued theoretically that larger future stakes should go together with less ex ante management in the focal transaction. The MAT–data, however, give no indication whether or not the stakes of possible future transactions were known at time of the current transaction. In the focal questionnaire, it was only asked to what extent the buyer would expect to continue business with the supplier before the purchase of current product (see Appendix B: Q3.13). Since this question cannot be used for the construction of the future stakes, we boldly decided to construct the financial volume and the opportunism potential of future transactions according to the construction of the past stakes. That is, we used the information from the second and, if available, the third transaction to define the future volume and opportunism potential in the current transaction. Unlike the experimental data where subjects knew the stakes of the future games, assuming the buyer's knowledge of future stakes during or even prior to the current transaction is a daring move of us. We nevertheless run the analysis (not reported) but found no effect of the future stakes on current investment in management. However, since neither the reciprocity effect nor the investment effect is clearly visible in the data, it comes as no surprise that neither the financial volume nor the opportunism potential of future transactions affect the effect of the shared future. Since the construction of the variables measuring the future stakes can be considered a 'wild guess', we are forced to conclude that a test of *Hypothesis 4.7* using these variables would be inadequate.

Unfortunately, the structure of the MAT–data does not allow to test *Hypothesis 4.4* on the discounting of past experience and *Hypothesis 4.5* on the effect of current suppliers being potential but not actual suppliers in the past.

Effects of Firm–Specific Control Variables

We did not formulate any hypotheses regarding the control variables. They were simply added to account for possible heterogeneity of the respondent firms (see discussion above). Adding these control variables indeed improves the fit of the models (Method A: all cases Wald $\chi^2(5) = 342.46, p = 0.000$ and past only Wald $\chi^2(5) = 199.36, p = 0.000$; Method B: all cases Wald $\chi^2(5) = 448.24, p = 0.000$ and past only Wald $\chi^2(5) = 65.95, p = 0.000$) as compared to the models without control variables. Remember that in the previous chapter, individual characteristics did not improve the estimation of the probability to place trust in the experiments discussed in the previous chapter.

It is worthwhile to first mention that the effects of all control variables are of about the same size independent of the method used. Only the effects of standardization in contracting and the size of the buyer in Method B in Table 4.6 are somewhat larger in magnitude. The control variables reveal several interesting effects. First, we find that standardization in contracting and the availability of own legal expertise *increases* the investment in management. Apparently, when one has 'cheap' or 'easy' access to necessary contracting resources, this goes with an increased investment in ex ante planning. This is in line with the hypothesis and the empirical results in Batenburg, Raub, and Snijders (2000). Second, we find that investments in ex ante management increase when larger suppliers are concerned. One could argue that this is in line with the idea that larger firms make for larger internal communication costs (Macaulay 1963) and therefore tend to incur higher ex ante investments, but this is completely post–hoc. Finally, there is some evidence that the buyer's network activities have a significant and positive effect on his investment in contracting if the focus is on all cases (see TABLE 4.5), but not so much when only buyer–supplier dyads with a shared history of business dealings are concerned. It seems as if the availability or the amount of third–party information leads to more carefulness in managing transactions. The statistical significance of this effect vanishes when we focus on those dyads with a past only (see TABLE 4.6). It looks as if information from the network becomes less important as soon as own experience with the supplier is available.[17] TABLE 4.7 shows an overview of the tests of our hypotheses.

4.6 Conclusion and Discussion

We analyzed the extent to which buyers of IT–products invest in ex ante management of their transactions, as a means to test the previous chapter's hypotheses on trust over time in a real life setting. We assume that a transaction between a buyer and a supplier of IT can be seen as a (repeated) Trust Game and that the size of the investment in ex ante management of the current transaction can be interpreted as a measure for trust in the supplier. In the previous chapter we analyzed how the behavior of people in Trust Games depends on previous and future Trust Games to be played, together with characteristics of the focal Trust Games. Likewise, we analyze here how a shared history and shared future between buyer and supplier, together with characteristics of a transaction, influence the decision on the investment in ex ante management.

The analysis is based on 1252 transactions between buyers and their suppliers in the Dutch IT–business. The data contain several buyers with two or even three transactions with the same supplier. The presence of such dyads in the data is precisely what allows for a comparison with theoretical and experimental evidence stemming from repeated Trust Games. TABLE 4.8 shows an overview of the results from the survey test and the experimental test conducted in the previous chapter.

[17]Theoretically, one would expect that a higher network embeddedness leads to more trust since information on malevolent behavior by the trustee (i.e. the supplier) is spread in the network, damaging his reputation and preventing him from doing future business with any of the trustors in the network (for a discussion of this puzzling effect, see Buskens 1999)

TABLE 4.7 *Coefficients from the analyses on the buyer's investment in managing the current transaction.*

	Investment in management		Results[b]			
			All cases		Past only	
Independent variable	Predicted[a]	Hypothesis	A	B	A	B
Focal transaction						
Financial volume	+	*4.1*	$\oplus$	$\oplus$	$\oplus$	+
Opportunism potential	+	*4.1*	$\oplus$	$\oplus$	$\oplus$	$\oplus$
Past transactions						
Successful past (dummy)	−	*4.2.1*	$\ominus$	$\ominus$		
Frequency of past transactions	−	*4.2.1*			$\ominus$	
Length of past	−	*4.2.1*				+
Year of first transaction	−	*4.2.1*				−
Successful past × ...						
... past volume	−	*4.3.1*	+		+	
... volume $t-1$, $t-2$	−	*4.3.1*				+, −
... opportunism potential $t-1$, $t-2$	−	*4.3.1*				+, +
Future transactions						
Expected future	−	*4.6*	−		−	
Future (dummy)	−	*4.6*		+		−
Future length	−	*4.6*		−		

[a] − : the investment in management decreases. + : the investment in management increases.
[b] Reported are the signs (+ or −) of the respective coefficients. $\oplus$ and $\ominus$ refer to positive and negative coefficients, respectively, if significant on at least $p = 0.1$. 'A' and 'B' refer to the coefficients of Method A and Method B, respectively.

Our results can be summarized as follows. First of all, investment in ex ante management increases with what is at stake, both in abstract Trust Games and in buyer–supplier relations. This is in fact a robust finding in all our experimental and field studies (Batenburg, Raub, and Snijders 2000; Buskens 1999; Blumberg 1997). More important for our purposes, investment in ex ante management decreases when buyer and supplier have done business together before, both in Trust Games and in buyer–supplier relations. In fact, when we focus on buyer–supplier dyads with a shared past, we see that it is not the length of the past but the frequency of transactions that decreases the buyer's effort in contracting. However, we do not find any evidence for the importance of the *kind* of shared history between actors (as long as it is a successful one). Both in the Trust Games and in buyer–supplier relations we see that when actors have had a successful interaction before, it does not matter to what extent this interaction was a transaction that necessitated a high or a low degree of trust. In other words, if a previous interaction was extremely 'difficult' (in the sense that there was a high incentive for the abuse of trust) but nevertheless completed successfully, then ex ante management will indeed be lower in following interactions. But, ex ante

Table 4.8 *Hypotheses with respect to the buyer's investment in managing the current transaction (II).*

	Investment in management		Results	
Independent Variable	Predicted[a]	Hypothesis	Experiment[b]	Survey
Focal transaction				
Volume and opportunism potential	+	*4.1*	supported	supported
Past transactions				
Successful past transactions	−	*4.2.1*	supported	supported
Unsuccessful past transactions	+	*4.2.2*	supported	not tested
Successful past transactions × past volume and opportunism potential	−	*4.3.1*	not supported	not supported
Unsuccessful past transactions × past volume and opportunism potential	−	*4.3.2*	not supported	not tested
Having a past × discounting past experience	+	*4.4*	supported	not tested
Potential vs. actual supplier	0	*4.5*	supported	not tested
Future transactions				
Future transactions	−	*4.6*	supported	not supported
Future transaction × future volume and opportunism potential	−	*4.7*	not supported	not tested

[a] − : the investment in management decreases. + : the investment in management increases. 0 : the investment in management is not affected.
[b] Experimental results refer to the analyses in Chapter 3.

management will be just as much lower in following interactions in the case where the previous transaction did not have such a high problem potential. To take this point home: it matters that actors had a successful transaction before; given that, it does not matter much what the characteristics of that transaction were. With respect to the shadow of the future the findings are less conclusive. We focused on two effects. The first one is an investment effect: the longer the shared future, the longer the period over which one can depreciate the investments, and hence the more attractive it is to invest in ex ante management. The experimental data show evidence in favor of the other effect, that of reciprocity along the lines of Axelrod (1984): the longer the shared future, the more trust and therefore the less ex ante management is necessary because mutually cooperative behavior is safeguarded by possible future sanctioning. This reciprocity effect was also found in Batenburg, Raub, and Snijders (2000) after controlling for the investment effect. However, in our setup we hardly find evidence supporting either the reciprocity or the investment effect. In any case, as with the effect of the shared past, there does not seem to be any support for the assertion that the *kind* of future matters much. By and large, we can conclude—as we did in Chapter 3—that we find some support for what was there called the 'simple stock' model: what is at stake influences trust, and so does having a positive past. We do not find support for the ‘complex stock’ model: information about the *kind* of past or the *kind* of future cannot be shown to have an effect on trust.

One can come up with evident points of criticism regarding the analysis based on the experimental data of the previous chapter, as well as regarding this chapter’s field data. To emphasize the latter (see the previous chapter for a discussion on the experimental setup): to ensure compatibility with repeated Trust Games, we assumed that the buyer alone decides on the amount of ex ante management invested in securing a transaction. This may not be a realistic approach for all IT–transactions, let alone for transactions on other markets.[18] The analysis would improve if one would be able to take into account more explicitly that both buyer and supplier influence that amount of ex ante management (Raub and Snijders 2001). Second, as outlined above in Section 4.4.3, both ways of analyzing the data (Methods A and B) are mere approximations of the kind of data one would want to have. Method B measures the temporal embeddedness of the buyer and the supplier on the basis of actual data on transactions between a buyer and supplier, but neglects that transactions can have taken place before, in between, or after the transactions we have represented in the data. Method A does not suffer from this, but only allows to approximate the temporal embeddedness of buyer and supplier on the basis of information on previous transactions elicited through retrospective subjective estimates. In short, by using any of these two methods we pay a certain price. The strong point of our analyses lies in the fact that under different circumstances (experimental data, and field data under different ways of analysis), we find results that are remarkably similar with respect to the stakes involved and the effect of a shared past.

[18]Though, as pointed out by Rooks (2002), at least for the Dutch IT–market during the period of research, this assumption may not be that far–fetched.

Chapter 5

Who Gets How Much in Which Relation? A Non–Cooperative Bargaining Approach to Exchange Networks*

> "Justice of Actions, is by Writers divided in *Commutative*, and *Distributive*: [...] And Distributive, in the distribution of equall benefit, to men of equall merit. As if it were Injustice to sell dearer than we buy; or to give more to a man than he merits."
> — Thomas Hobbes (1651: 105)

5.1 Introduction

Potential employers and employees, lawyers and clients, or different business firms often determine their contracts in dyadic negotiations. Exchange theories reflect this pattern—they explain individual profits as the result of dyadic bargaining on the distribution of at least one perfectly divisible resource (e.g. cake, dollar). Following Cook et al. (1983), sociologists usually consider a situation in which a given network structure limits matches between bargaining partners, propose formal models to predict outcomes of negotiated exchange, and test those predictions in laboratory experiments. However, there are not just different theories, but also many controversies—a selection of important contributions and discussions includes Bienenstock and Bonacich (1992, 1993, 1997), Bonacich and Bienenstock (1995), Bonacich (1998, 1999), Bonacich and Friedkin (1998), Burke (1997), Friedkin (1992, 1993, 1995), Lovaglia et al. (1995), Markovsky et al. (1993, 1997), Markovsky, Willer, and Patton (1988, 1990), Skvoretz and Fararo (1992), Skvoretz and Lovaglia (1995), Skvoretz and Willer (1991, 1993),

*This chapter is an adapted version of Braun and Gautschi (2000, 2001).

Thye, Lovaglia, and Markovsky (1997), Yamagishi and Cook (1990), Yamagishi, Gillmore, and Cook (1988) as well as Yamaguchi (1996, 1997, 2000).[1]

Despite their differences, sociological exchange theories (see, for overviews, several contributions in Willer 1999) have common features. First, practically all theories neglect interindividual heterogeneity (in terms of, e.g. age, education, gender, or wealth) in favor of the effects the given network structure has for exchange outcomes. More precisely, they explain how the structural positions in the bargaining network affect the exchange patterns between adjacent actors and their bilateral splits of cakes of given sizes. Power inequalities due to different structural positions manifest themselves in the negotiated distributions of exchange profits and, at least partly, in the actual trading patterns between connected actors.

Second, according to Bonacich and Friedkin (1998), only few theoretical approaches can account for variations in the value of relationships (i.e. heterogeneity in terms of the size of the cake to be partitioned). And, experimental research has focused almost exclusively on those relations which concern the split of an identical surplus. Since such 'equally valued' relations are rare in everyday life, Bonacich and Friedkin point out correctly that a sufficiently realistic approach to the study of exchange networks should refer to a scenario in which bilateral relations may vary with respect to the amount of surplus to be divided.

Third, the application of sociological exchange models often is not straightforward. Partly due to additional experimental evidence, older models have been successively adjusted or revised such that ad hoc assumptions and difficult prediction procedures characterize their updated versions. And, profit predictions from new approaches require either specific computer simulations (Burke 1997) or applications of an iterative computational algorithm in which a key parameter has to be set by the researcher (Yamaguchi 1996).

Fourth, few theories (e.g. Willer and Skvoretz 1999; Yamaguchi 1996) allow for different types of network connections. Most theories for exchange networks exclusively deal with a scenario in which an actual exchange in one relation tends to prevent transfers in others. Put differently, they refer to exchange networks with substitutable relations only. This focus is narrow because, in accordance with Cook et al. (1983), Cook and Emerson (1978), as well as Emerson (1972), one can distinguish between positively and negatively connected exchange relations—a positive connection exists if a resource transfer in one relation promotes transfers in others (e.g. communication networks), whereas a negative connection exists if a resource transfer in one relation tends to preclude transfers in others (e.g. dating networks).[2] Yamagishi, Gillmore, and Cook (1988) justly emphasize that real networks often are mixtures of both types. It

---

[1]Special issues of journals (cf. *Social Networks* 14, No. 3–4, 1992 and, at least partly, *Rationality and Society* 9, No. 1–2, 1997) contain additional articles and controversies.

[2]Willer and Skvoretz (1999) present an alternative categorization of network connections which, as they emphasize, differs from the positive/negative classification. Their distinction embraces five types of connections. It draws on the ranking and values of several node–specific parameters which are, at

makes sense that researchers, while designing experiments, decide to look at positively and negatively connected relations separately. Theory formation, however, was closely associated with experiments on negatively connected systems only. As a consequence, there is just one exchange model which explicitly deals with negatively and positively connected relations (Yamaguchi 1996).[3]

Fifth, the experimental evidence indicates that negatively connected networks with structurally distinct positions vary in terms of exchange outcomes.[4] Sociologists have suggested several explanations for this variation (e.g. Bienenstock and Bonacich 1992; Cook and Yamagishi 1992; Friedkin 1992, 1993; Lovaglia et al. 1995; Markovsky, Willer, and Patton 1988; Markovsky et al. 1993). According to Skvoretz and Willer (1993), these contributions have in common that they identify, more or less explicitly, the exclusionary potential associated with a particular network position as the crucial explanatory factor. And, Markovsky et al.'s (1993) distinction of weak and strong power networks has become a popular classification in this context (cf. Bonacich 1999; Lovaglia et al. 1995; Skvoretz and Willer 1993; Thye, Lovaglia, and Markovsky 1997).

Another common feature of sociological exchange theories is the assumption that negotiation partners pursue their self–interests. Although social psychological considerations sometimes are taken into account as well (e.g. Lovaglia et al. 1995, Skvoretz and Willer 1993), most exchange theorists prefer a rational actor perspective. Apart from a few exceptions (e.g. Yamaguchi 1996), however, they usually do not specify an optimization problem. It is thus often not clear where network features enter the choice calculus and how they influence decision–making. And, since interactive choices characterize negotiations, Bienenstock and Bonacich (1992, 1997) justly emphasize the relevance of game–theoretic ideas for the analysis of exchange networks. However, only a few contributions explicitly refer to the game–theoretic bargaining literature (cf. for a recent review, Muthoo 1999) in this context.

---

least partly, unique to their theoretical approach. Like other authors (e.g. Bonacich 1999; Yamaguchi 1996), we therefore adopt the standard terminology of negative or positive connections.

[3]Modifying and extending Coleman's (1973, 1990) competitive equilibrium approach, Yamaguchi (1996) equates negative (positive) connections with closely substitutable (complementary) exchange relations and introduces a flexible continuous parameter for substitutability/complementarity (viz. the elasticity of substitution). Yamaguchi's theory thus embraces situations in which exchange in one relation tends to prevent or promote transfers in others. Its application requires, however, an ad hoc specification of the elasticity of substitution. And, it is limited to the analysis of either substitutable or complementary relations in a given network. Combining basic ideas of his original model with additional assumptions, Yamaguchi (2000) presents a theoretical analysis of structures characterized by the simultaneous presence of both substitutability and complementarity among the multiple exchange relations of an actor.

[4]We use the term *structurally distinct*, and thus also *structurally equivalent*, in a slight abuse of its true meaning. We use 'structurally equivalent' to refer to actors that occupy indistinguishable structural locations in the network. According to Wasserman and Faust (1994: 468–473), this definition actually refers to *automorphic equivalence*. We will denote 'structural network positions' using capital letters A,B,C,D,E. Actors on structurally equivalent positions will be distinguished by numeric subscripts, for instance, A_1 and A_2.

Moreover, just a subset of current exchange theories (e.g. Bonacich and Bienenstock 1995; Friedkin 1995; Simpson and Willer 1999) systematically predicts why and when specific links to bargaining partners lead to actual deals if, as is often assumed in experiments, actors may complete at most one exchange per round. In such an extreme case of negatively connected networks, actors may have several bargaining partners, but face an exogenously given restriction with respect to the acceptable number of exchange partners (e.g. monogamy rule). That is, they may have to select their actual exchange partners from a larger set of potential exchange partners. An appropriate theory for negatively connected relations thus has to predict the exchange patterns between adjacent actors. In particular, it should answer why and when individual choices of actual exchange partners create a relational structure which deviates from the given network of potential exchange partners. Such answers are particularly desirable if one follows Cook and Whitmeyer's (1992) suggestion and advocates a theoretical approach that uses exchange theory to explain network structure.

The literature additionally reveals that competing theories differ with respect to their performance in predicting experimental profit splits. Systematic comparisons of established approaches suggest that among the best fitting theories are usually those based on 'resistance' (e.g. Heckathorn 1983; Willer 1981) concepts. According to Skvoretz and Willer (1993), a relatively successful model for negatively connected networks with uniformly valued relations rests on the idea that two bargaining partners agree on the outcome to which they are equally resistant (Exchange Resistance). In another test series for such networks, Lovaglia et al. (1995) found that the best predictions result from combining the resistance logic with a measure for the chance that a network position can be excluded from exchange. Specifically, they modified the Graph–theoretic Power Index (GPI, see Markovsky, Willer, and Patton 1988; Markovsky et al. 1993) for negotiation resistance and positional degree in the network of bargaining partners. Although it exclusively focuses on negatively connected relations, the application of this GPI–Resistance Degree (hereafter, GPI–RD) model is cumbersome, however.

In sum, this discussion indicates properties a sufficiently general theory of exchange networks should have. First of all, such a model should be parsimonious and simple to apply. At the same time, it should offer unique point predictions for negotiation outcomes which closely fit the available experimental evidence. It also should allow for unequally valued relations (i.e. variations in terms of the surplus to be partitioned). Moreover, an appropriate theory should systematically predict whether and, if so, how actual exchange structures deviate from the exogenously given bargaining network when the focus is on settings with negative connections. However, it should not be limited to the analysis of negatively connected settings. Finally, if the theory is based on the rationality postulate, it should clarify how structural features (i.e. the network) affect the actors' decision–making and how their interactive choices determine the negotiation outcomes. And, since interactive choices characterize negotiations, Bienenstock and Bonacich (1992, 1997) justly emphasize the relevance of game–theoretic ideas for the analysis of exchange networks. Such a strategic approach should explain, however, how selfish bargaining partners arrive at mutual cooperation.[5] As a consequence, the focus should be on a non–cooperative bargaining scenario.[6]

We present and apply such a model in the remainder of this chapter. In contrast to previous approaches, this model generates point predictions for negotiation outcomes in negatively connected, positively connected, and mixed networks. Its predictions result from an equilibrium analysis in the sense of non–cooperative game theory. And, contrary to other work on exchange networks, the approach also uniquely predicts whether and, if so, how the actual exchange structures deviate from the exogenously given bargaining network.[7]

The presentation and application of the model requires several steps. The next section describes Rubinstein's (1982) Alternating Offers Game and Binmore's (1985, 1987, 1998) refinement. This reflects that our approach draws heavily on their ideas about non–cooperative bargaining between two rational egoists. The third section specifies how the network embeddedness in exogenously given bargaining networks affects the behavior of negotiation partners and the exchange outcomes. It combines the non–cooperative bargaining logic with the sociological idea that structural positions matter for negotiation and exchange.

Following the presentation of the model in the next section, we substantiate its empirical relevance via three applications. In doing so, we follow the practice of exchange theorists (e.g. Burke 1997; Friedkin 1995; Yamaguchi 1996) and rely on published experimental results for empirical validation. More precisely, we compare predictions from the new model with available experimental findings and relevant predictions from other approaches.

After a description of typical features of experiments on exchange structures and our data selection criteria, the first application refers to experimental results on bilateral profit splits in negatively connected networks (Lovaglia et al. 1995; Skvoretz and Fararo 1992; Skvoretz and Willer 1993). It shows that, while predictions of the new model are straightforward to obtain, they are at least as consistent with the evidence as predictions of the currently best fitting theories.

---

[5]Bienenstock and Bonacich (1992, 1997) conceptualize a network of potential exchange partners as a cooperative game with transferable utility such that binding agreements and commitments are, by definition, always feasible before distribution problems are solved. No matter which solution concept from cooperative game theory (e.g. core, kernel) is applied, this approach thus does not explain how cooperation between rational egoists emerges in the first place. This argument loses its force if the focus is on a non–cooperative scenario in which, by definition, binding agreements and commitments are not feasible at the outset (e.g. Rasmusen 1994).

[6]There is an additional argument in favor of the non–cooperative approach: if one follows Nash (1950, 1951, 1953) in regarding non–cooperative games as more fundamental than cooperative games, the analysis of a non–cooperative bargaining game may suggest an appropriate cooperative solution concept for the negotiation problem at hand.

[7]Also, the model can be extended to the case in which interindividual variation (i.e. heterogeneity not attributable to structure) matters as well. This point will be addressed in footnotes.

The second application refers to experimental findings on power distributions in negatively and positively connected settings as reported by Cook et al. (1983), Markovsky et al. (1997), Markovsky, Willer, and Patton (1988), Skvoretz and Fararo (1992), and Yamagishi, Gillmore, and Cook (1988). It demonstrates that the new model fits better than Yamaguchi's (1996) theory, the only other model for positively and negatively connected exchange relations.

The third application refers to experiments on deviations between bargaining and exchange structures in negatively connected settings (Markovsky, Willer, and Patton 1988; Simpson and Willer 1999; Skvoretz and Willer 1993). It identifies misinterpretations of experimental results in the literature and shows that the new model correctly predicts the empirical observations.

Despite its generality, the new approach does not address all aspects of network exchange, however. And, like any other model, it rests on several strong premises. We therefore conclude with a brief discussion of desirable extensions.

5.2 Dyadic Bargaining Situation

Consider an exogenously given network with m mutual ties between a finite number of rational egoists ($i, j, k = 1, 2, \ldots, n$). Assume that each actor knows the symmetric relations between all network members.[8] These relations limit the matches of potential partners for bilateral negotiations and exchanges. To specify the bargaining situation for each separate pair of actors, we rely on Rubinstein's (1982) and Binmore's (1985, 1987, 1998) game–theoretic work on non–cooperative bargaining. Specifically, we assume that the actors i and j bargain over the partition of a surplus of given value $v_{ij} = v_{ji}$. When x_{ij} represents i's negotiated share of the value v_{ij}, it holds that $0 \leq x_{ij} \leq v_{ij}$ and $x_{ij} + x_{ji} = v_{ij}$.[9] Put differently, x_{ij} denotes i's negotiated exchange profit in the relation with j.

5.2.1 The Alternating Offers Game

Each bargaining session refers, by postulate, to the bilateral distribution of a fixed quantity (i.e. surplus) of a perfectly divisible resource (e.g. money, pie). Exchange appears here, in accordance with sociological approaches (e.g. Skvoretz and Willer 1993), as an agreement of two rational actors on the division of a given surplus. A player i's profit from a dyadic exchange relation can also be expressed as his fraction of the surplus, p_{ij}, obtained. Note that p_{ij} coincides with x_{ij} if and only if $v_{ij} = 1$.

[8] This assumption of complete information about the overall shape of the network structure may be replaced by a weaker postulate. As will become clear below, it suffices to assume that each actor is informed about his 'immediate vicinity' in the network under study.

[9] If the actors perpetually disagree, they do not get a proportion of the surplus. That is, the payoff associated with disagreement is 0 for both network partners. Put differently, there is no outside option which would pay more than zero.

To further describe the situation, we follow Rubinstein (1982) and assume that, once two selfish actors i and j are matched, they play a non–cooperative game in which they alternate in proposing how to divide the pie with one time period elapsing between each offer. This Alternating Offers Game refers to a bargaining situation with complete information in the sense of common knowledge (i.e. everybody knows the payoffs and rules of the game and everybody knows that everybody knows, and so on). It explicitly models the bargaining procedure as a dyadic negotiation with an indeterminate time horizon. That is, i offers a piece of cake to player j which j can either accept or reject; if j rejects i's offer, j makes a proposal which i can either accept or reject, and so on. Note that, despite the open time horizon, the alternating offers game is not an infinitely repeated game. Rather, it ends as soon as one player accepts an offer (e.g. Rasmusen 1994).[10]

As a consequence, only the accepted offer is important for the players' payoffs (but not earlier proposals). Specifically, each actor's utility linearly increases with the share obtained and decreases with the time elapsed until an agreement. The latter assumption deviates from sociological exchange approaches. It reflects that time is modelled explicitly—each bargaining partner is assumed to be impatient such that getting a specific piece of cake in the future is less beneficial than now. More precisely, i's utility of receiving the profit share p_{ij} in the match with j during the t–th bargaining period is

$$u_{ij}\left(p_{ij}, t\right) \,=\, \delta_i^t\, p_{ij} \quad \text{for all } \; i \neq j, \tag{5.1}$$

where t denotes discrete time units ($t = 0, 1, 2, ...$) and the discount factor δ_i, $0 < \delta_i < 1$ measures i's impatience: the smaller δ_i gets, the less patient or more impatient i becomes.

The Rubinstein model therefore refers to bargaining situations in which delays in agreement are costly for negotiation partners (e.g. losses due to a strike's duration in the case of wage bargaining between unions and employer organizations). It apparently does not fit situations in which bargaining occurs quickly such that the amount of time between offers and the delay of agreement are negligible. The latter features characterize laboratory experiments on network exchange (e.g. Bienenstock and Bonacich 1993; Lovaglia et al. 1995; Skvoretz and Willer 1993). As will become clear below, however, the Alternating Offers Game can be adjusted to the problem of dyadic exchange in exogenously given negotiation networks. To introduce this adjustment, some implications of the original bargaining game are important.

A fundamental result of the game–theoretic analysis concerns the bargaining strategy played by rational actors with complete information. Rubinstein (1982) showed that, everything else being constant, only 'stationary strategies' matter for the bargaining solution in the alternating offers game with open time horizon. Actor i plays

[10] A finite–horizon version of this game was proposed by Ståhl (1972).

a stationary strategy if he makes the same proposals when it is his turn to make offers and if he rejects proposals unless they give him at least his reservation level.[11]

Starting from this insight, Rubinstein proved that the alternating offers bargaining game with indeterminate time horizon has a unique subgame–perfect Nash equilibrium.[12] Accordingly, the surplus is to be divided in the first bargaining period, where the actors' profit shares solely depend on their discount factors. This Pareto–efficient equilibrium for the match between i and j rests on the postulate that always one time period elapses between proposals. As a consequence, the equilibrium payoffs depend on the sequence of offers.[13] From actor i's perspective, the negotiation with j in the Rubinstein game thus yields the equilibrium profit share

$$p_{ij} = \begin{cases} z_{ij} & \text{if } i \text{ makes the first offer} \\ \delta_i \, z_{ij} & \text{if } j \text{ makes the first offer} \end{cases}, \quad \text{where} \quad z_{ij} := \frac{1-\delta_j}{1-\delta_i\delta_j}\,. \tag{5.2}$$

A closer inspection of these equilibrium payoffs reveals that even if both players have identical preferences (i.e. if they have a common discount factor $\delta_i = \delta_j$), they will not receive equal benefits—the first mover gets a larger piece of cake even if the players are equally patient. Since in many bargaining situations the order of play will be arbitrary, the first mover advantage appears as a serious shortcoming of Rubinstein's game. This artificial asymmetry of payoffs seemingly prevents, together with the afore–mentioned observation that the alternating offers game apparently does not apply to situations in which bargaining occurs quickly, that Rubinstein's alternating offers game is relevant for the analysis of network exchange. Fortunately, Binmore has provided a plausible solution for these problems.

5.2.2 The Limiting Equilibrium Solution

Binmore (e.g. 1985, 1987) starts from the intuition that, if the reduction in utility from deferring agreement by one period becomes smaller, so that delay becomes less important, the first mover advantage should vanish. To clarify this idea, we follow Gravelle and Rees (1992) and imagine that the original bargaining period (say, a day) is split into ℓ shorter periods (say, minutes). Let δ_i continue to be i's daily discount factor and let $\delta_{i\ell}$ denote i's discount factor for a period of $(1/\ell)$–th of a day. If player i is to have the same attitude to pie in one day's time, it must be true that $\delta_{i\ell} =$

[11] More precisely, a rational egoist i offers $1 - z_{ij}$ to his bargaining partner j and accepts any offer which does not fall short of his reservation value $\delta_i \, z_{ij}$, where z_{ij} denotes the maximum fraction of the pie i can obtain in the match with j (cf. eq. (5.2)).

[12] A Nash equilibrium refers to a stable situation in which sufficiently experienced actors play their best reply strategies (Nash 1951). Since the alternating offers scenario is an extensive form game with many Nash equilibria (viz. any agreement), Rubinstein additionally requires subgame perfectness. A strategy profile is a subgame–perfect Nash equilibrium if it is a Nash equilibrium for the entire game and every subgame (Selten 1965). This standard refinement of the Nash equilibrium concept rules out strategy profiles based upon incredible threats.

[13] If actor i makes the first offer, the equilibrium payoff vector is $(z_{ij}, 1-z_{ij})$; if not, it is $(1-z_{ji}, z_{ji})$.

$(\delta_i)^{(1/\ell)}$. Inserting $\delta_{i\ell}$ into the above payoff functions (eq. (5.2)) gives subgame–perfect equilibrium outcomes as functions of the length of period $(1/\ell)$. Taking the limit as $\ell \to \infty$ (and applying L'Hopital's rule), player i's equilibrium profit share in the match with j tends to (see, e.g. Binmore 1998; Osborne and Rubinstein 1990)

$$p_{ij} = \frac{\ln \delta_j}{\ln \delta_i + \ln \delta_j} \quad \text{for } i \neq j\,, \tag{5.3}$$

where ln denotes the natural logarithm. This limiting equilibrium solution ensures that neither player enjoys a first mover advantage. And, it predicts an equal division of the surplus if the players have identical preferences (i.e. a common one–day discount factor $\delta_i = \delta_j$).

Notice that the limiting equilibrium of the alternating offers game results if the amount of time that elapses between proposals is allowed to recede to zero. This scenario corresponds with the typical experimental setting of network exchange in which the amount of time between offers is negligible and the delay of agreement imposes practically no costs. Hence, as long as the focus is on its limiting equilibrium solution, we can adopt the alternating offers game as a strategic model for bargaining sessions between two selfish network partners.

In doing so, it has to be recognized that the discount factors δ_i and δ_j refer to a much longer time horizon than the bargaining periods actually considered. From the perspective of the typical network exchange setting, δ_i and δ_j thus apply to a counterfactual scenario—they measure the actors' impatience if the sequence of offers and counteroffers would delay agreements for days (instead of, say, minutes) and therefore would impose non–negligible costs to the partners.[14] According to eq. (5.3), however, the one–day discount factors δ_i and δ_j affect negotiated exchanges even when the focus is on situations with extremely short bargaining periods (e.g. seconds)—δ_i and δ_j remain important parameters for the bargaining solution if there are no costly delays of agreement. Binmore (1985: 273) therefore suggests a definition of i's bargaining power b_i in terms of i's impatience at long delays in agreement:[15]

$$b_i := -\frac{1}{\ln \delta_i} \quad \text{for all } i\,. \tag{5.4}$$

Accordingly, there is always a positive relationship between i's bargaining power b_i and his one–day discount factor δ_i. For expository convenience, we therefore will often refer to the concept of bargaining power only. Combining eqs. (5.3) and (5.4) gives a

[14]This indicates, among other things, that the discount factors δ_i and δ_j fundamentally differ from resistance concepts—neither Skvoretz and Willer's (1993) 'Exchange Resistance' nor Lovaglia et al.'s (1995) 'Negotiation Resistance' have something to do with time preferences in alternating offers bargaining. Put differently, resistance concepts do not measure utility losses associated with long delays of agreement.

[15]Formally, b_i is the reciprocal of the discount rate ρ_i corresponding to the discount factor δ_i. It thus holds $\rho_i = -\ln \delta_i$ or $\delta_i = e^{-\rho_i}$, where e denotes the base of natural logarithm. Impatient actors have high discount rates, but low discount factors.

version of the limiting equilibrium solution for the Alternating Offers Game in terms of the actors' bargaining powers when they negotiate over the partition of a given cake of size $v_{ij} = v_{ji}$:

$$x_{ij} = \left(\frac{b_i}{b_i + b_j} \right) v_{ij} = p_{ij}\, v_{ij} \quad \text{for } i \neq j, \tag{5.5}$$

where $p_{ij} := b_i/(b_i + b_j)$ defines i's relative bargaining power in the relation with j.[16] And, since $p_{ji} = 1 - p_{ij}$ holds by definition, i's partner j can receive $x_{ji} = (1 - p_{ij})\, v_{ij} = p_{ji} v_{ji}$.

Accordingly, the optimal partition of the given surplus depends critically on the combination of b_i and b_j. Put differently, the bargaining power of just one partner is irrelevant for the negotiation outcome—it is the relative bargaining power (i.e. p_{ij} or $p_{ji} = 1 - p_{ij}$) which matters for the profit split. Notice, however, that p_{ij} does not coincide with x_{ij} when $v_{ij} \neq 1$. Since $x_{ij} = p_{ij}\, v_{ij}$ and $x_{ji} = (1 - p_{ij})\, v_{ij}$, a comparison of actor i's profit share with that of his bargaining partner j yields the following chain of equivalent conclusions:

$$x_{ij} \gtreqless x_{ji} \iff x_{ij} \gtreqless \frac{1}{2} v_{ij} \iff p_{ij} \gtreqless \frac{1}{2} \iff b_i \gtreqless b_j \quad \text{for } i \neq j. \tag{5.6}$$

Put verbally, a symmetric distribution of bargaining powers ($b_i = b_j$ or $p_{ij} = \frac{1}{2} = p_{ji}$) always yields an equal split of the pie ($x_{ij} = \frac{1}{2}\, v_{ij} = x_{ji}$). There will be an unequal profit division, however, when the power of the two negotiation partners differs. Specifically, i's exchange profit x_{ij} dominates j's exchange profit x_{ji} such that i gets more than half of the pie if and only if p_{ij} exceeds p_{ji}. Because of $p_{ij} + p_{ji} = 1$, the latter is satisfied if and only if i's relative bargaining power in the relation with j exceeds $\frac{1}{2}$. And, this is equivalent to the condition that i's absolute bargaining power b_i exceeds j's absolute bargaining power b_j.

Concrete predictions for such cases require, of course, information about the bargaining powers which, by definition, depend on the one–day discount factors. Before we can predict outcomes of negotiated exchanges via the limiting equilibrium solution, we thus need to determine these exogenous parameters of the alternating offers bargaining game. It is precisely here where the basic idea of theories for exchange networks (see, for overviews, Lovaglia et al. 1995 as well as Skvoretz and Willer 1993) comes in—from a sociological perspective, the actors' bargaining powers result from their structural positions in the given negotiation network.

---

[16] As Binmore (1985, 1998) observes, the limiting equilibrium outcome (eq. (5.5)) solves the optimization problem max $x_{ij}^{b_i}\, (v_{ij} - x_{ij})^{\,b_j}$. The latter is a weighted version of Nash's (1950) bargaining solution from cooperative game theory. The subgame–perfect Nash equilibrium for the non–cooperative Alternating Offers Game therefore suggests a corresponding cooperative solution concept if the amount of time between proposals vanishes (see Braun and Gautschi 2001).

5.3 Network Structure and Dyadic Bargaining

Bargaining relations reflect, as assumed, the existence of m mutual ties between n rational egoists. Taking the negotiation network as a starting point for the determination of the actors' bargaining powers, we focus on a situation in which each pair of connected actors separately plays an Alternating Offers Game with negligible amounts of time between proposals. That is, we combine Emerson's (1971, 1982) idea that network embeddedness matters for negotiation and exchange (see Chapter 1) with the limiting equilibrium solution of the Alternating Offers Game for each exogenously given relation between two rational egoists. As will become clear, this approach gives point predictions for negotiation outcomes in negatively connected, positively connected, and mixed networks. And, it helps to uniquely predict whether and, if so, how the actual exchange structures deviate from the exogenously given bargaining network.

To reach these goals, we have to introduce additional assumptions and concepts. Before we go into details, it is useful to list our basic premises:

1. Depending on their structural positions in the exogenous bargaining network and the values of their relations (i.e. the size of the cakes to be partitioned), actors differ in terms of their 'network control' (i.e. the extent to which an actor controls the relations to him by his relations to others).

2. Depending on an actor's assessment of the relation he has with a potential exchange partner (i.e. the type of connection he perceives), his network control positively or negatively affects his individual bargaining power.

3. Depending on the actors' individual bargaining powers in the bilateral match under consideration, their profit shares result via the limiting equilibrium solution of the alternating offers game.

4. Depending on the ranking of negotiated profit shares, actors with negatively connected relations select their actual exchange partners from the relevant set of bargaining partners.

In accordance with this chain of postulates, we now successively present the assumptions and implications of our approach. Starting from a given negotiation structure, the actors' network control is to be determined first.

5.3.1 Bargaining Structure and Network Control

Let the $n \times n$ matrix $\mathbf{V}$ with main diagonal elements $v_{ii} = 0$ for all i and off–diagonal elements $v_{ij} \geq 0$ for all $i \neq j$ represent the exogenously given network of m valued bargaining relations between the n actors. While the relation between the bargaining partners i and j is always symmetric, an actor's relations with distinct partners may differ with respect to the values at stake—the corresponding off–diagonal elements $v_{ij} = v_{ji}$ express whether i and j are bargaining partners and, if so, how large the cake is they can divide. Formally, it holds $v_{ij} = v_{ji} > 0$ in the presence of a bargaining relation between $i \neq j$, but $v_{ij} = v_{ji} = 0$ in its absence.[17]

Even if matrices of valued adjacencies differ, they may represent the same relational structure. A standardization thus is reasonable. Let $\mathbf{R}$ be the $n \times n$ matrix of standardized actor relations such that $r_{ij} := v_{ij} / \sum_{k=1}^{n} v_{kj} \geq 0$ for all i, j, and $\sum_{k=1}^{n} r_{kj} = 1$ for all j. That is, $\mathbf{R}$ is the column–stochastic matrix derived from the valued graph. Its off–diagonal element r_{ij} measures i's fraction of the systemwide valued relations to j. In other words, r_{ij} represents i's degree of 'control' over the valued relations to j in the system. Specifically, it holds $0 \leq r_{ij} \leq 1$, where $r_{ij} = 0$ indicates that i has no control over j (i.e. absence of a tie between i and j) and $r_{ij} = 1$ reflects that i has complete control over j (i.e. i is j's only bargaining partner).

The i–th row of the matrix $\mathbf{R}$ informs about i's control over each other actor in the system, whereas the i–th column of $\mathbf{R}$ informs about each other's control over i. Adding up the relevant pairwise elements of $\mathbf{R}$ defines the 'network control' of actor i:

$$c_i := \sum_{k=1}^{n} r_{ik} \, r_{ki} \quad \text{for all } i. \tag{5.7a}$$

Put verbally, c_i is the degree to which i controls the valued relations to him by his valued relations to others. For example, $c_i = \frac{3}{4}$ means that actor i controls, via his valued relations to others, 3/4–th of their valued relations to him. The control fraction c_i thus may be interpreted as i's 'structural autonomy' as well.[18] And, given information about i's valued relations and those of his partners, its calculation is straightforward (see, for illustrations, FIGURE 5.1 below). As a consequence, we do not have to assume that every actor has complete information about the overall shape of the network structure. For the determination of the control distribution in the system, it suffices to postulate that everyone has complete information about his own valued relations and those of his network partners.

Equally Valued Relations

Nearly all laboratory experiments focus on networks with equally valued relations. That is, each bilateral bargaining session concerns the division of an identical surplus $v_{ij} = v_{ji} = v$ for all $i \neq j$. In such cases, the $n \times n$ matrix $\mathbf{V}$ reduces to the $n \times n$ adjacency matrix $\mathbf{A}$ with main diagonal elements $a_{ii} = 0$ for all i and off–diagonal elements $a_{ij} \in \{0, 1\}$ for all $i \neq j$ representing the exogenously given and symmetric

[17] In the basic scenario, each pair of connected actors bargains over the partition of just one cake with a specific size. In a slightly more complicated case, there may be more than just one reciprocated tie between each pair of connected actors in the network (e.g. each dyad can divide two pies per round of negotiated exchanges). If so, the sum of the relevant surpluses determines the off–diagonal elements of matrix $\mathbf{V}$.

[18] The definition of network control (cf. eq. (5.7a)) shows that the c_i's are the positive main diagonal elements of the $n \times n$ matrix $\mathbf{C} := \mathbf{RR}$. That is, $c_i = c_{ii} > 0$ for all i. Like $\mathbf{R}$, $\mathbf{C}$ is a column–stochastic matrix. That is, $0 < c_i \leq 1$ and $1 - c_i = \sum_{k \neq i} c_{ki} \geq 0$ for all i. Since the upper bound of c_i is 1, its complement $1 - c_i$ measures i's 'structural dependence' (i.e. the degree to which the other system members affect, via their valued relations to one another, the valued relations to actor i).

bargaining relations. More precisely, a_{ij} is a binary measure for the absence or presence of a mutual tie between the actors i and j (i.e. $a_{ij} = a_{ji}$ is coded as 0 or 1 for all $i \neq j$).

Let again **R** be the $n \times n$ matrix of standardized actor relations such that $r_{ij} := a_{ij}/\sum_{k=1}^{n} a_{kj} \geq 0$ for all i, j, and $\sum_{k=1}^{n} r_{kj} = 1$ for all j. In the case of equally valued relations, **R** is the column–stochastic matrix derived from the adjacencies with the same properties as described above.

Then, the calculation of c_i is even more straightforward. A closer look at eq. (5.7a) then shows that c_i may be alternatively expressed as the mean of the i–th row in the matrix **R**. In other words, i's network control is the mean of i's control over the systemwide relations to his partners. Formally,

$$c_i = \frac{1}{n_i} \sum_{k=1}^{n} r_{ik} = \frac{1}{n_i} \sum_{k=1}^{n} \frac{a_{ik}}{m_k} \quad \text{for all } i, \tag{5.7b}$$

where n_i denotes the number of i's negotiation partners and m_i the number of i's ties in the network structure under consideration.[19] Since the number of positive elements in the i–th row in **R** is always n_i, the relevant control distribution can be practically read off from the standardized actor relations (see, for illustrations, Figure 5.1 below).

The number of i's bargaining partners, n_i, coincides with the number of positive elements in the i–th row of the adjacency matrix **A**, too. As indicated by the far right–hand side of eq. (5.7b), information about i's network relations and the number of his partners' ties therefore suffices for the computation of i's network control as well. Put differently, the distribution of network control can be determined when each actor knows his connections and the connections of his partners.

The concept of network control thus requires only weak assumptions about the structural information of network members. This becomes more obvious if c_i is expressed in still another way. When S_i denotes the set of the n_i bargaining partners of actor i, it is possible to rewrite eq. (5.7b) as follows:

$$c_i = \frac{1}{n_i} \sum_{k \in S_i} \frac{1}{n_k} = \frac{1}{(n_i / \sum_{k \in S_i} (1/n_k))} \quad \text{for all } i. \tag{5.7c}$$

Stated differently, i's network control c_i reflects how many negotiation partners actor i and his partners have. Hence, information about the number of i's bargaining partners and the numbers of their partners allows the calculation of i's network control.[20]

---

[19] In the basic scenario of an adjacency matrix, the number of each actor's negotiation partners n_i is equal to the number of his ties m_i (i.e. $n_i = m_i$ for all i). If, however, more than just one cake per round is to be divided by each pair of connected network members, every actor has more ties than partners (i.e. $n_i < m_i$ for all i). For instance, if i has $n_i = 3$ bargaining partners where two cakes per tie can be divided, we have that $m_i = 6$. We can alternatively write $r_{ij} := a_{ij}/\sum_{k=1}^{n} a_{kj} = a_{ij}/m_j \geq 0$ for all i, j.

[20] For the determination of the distribution of network control, we thus do not have to assume that every actor has complete information about the overall shape of the network structure. It suffices to postulate that either everyone has complete information about his own relations and his partners' ties or that everyone knows the number of his bargaining partners and the numbers of their partners.

A closer inspection of the far right–hand side of eq. (5.7c) shows, moreover, that the concept of network control is compatible with a rational actor perspective—c_i is simply the reciprocal of the estimate a rational actor i will have for the mean number of partners of his partners in a given network.[21] For example, $c_i = \frac{3}{4}$ expresses that the average number of bargaining partners of i's partners is $\frac{4}{3} = 1.333$. Clearly, i's network control c_i decreases if the mean number of partners of i's partners increases. And, c_i rises if the mean number of bargaining partners of i's partners falls. In the limit case, the mean number of bargaining partners of i's partners equals 1 (i.e. actor i is their only negotiation partner) such that i has full network control $c_i = 1$. So, when the assumption is made that i's behavior reflects his network control or structural autonomy, it is postulated, in effect, that he takes account of the mean number of partners of his partners.

Whether or not we focus on bargaining structures with equally valued relations, the weak informational requirements reflect that i's network control captures the structure only two steps from i (but neglects structural effects which are three or more steps away). The degrees of network control thus may be seen as parsimonious indicators for the actors' structural positions. They are, by postulate, essential determinants of the individual bargaining powers. That is, actor i's network control c_i affects b_i, i's level of individual bargaining power. The direction of the relationship between network control c_i and bargaining power b_i depends, however, on i's categorization of the respective network relation.[22]

5.3.2 Relational Assessments and Bargaining Power

As described, most work in the field exclusively focuses on situations in which each agreement tends to prevent other resource transfers (due to, e.g. exchange restrictions like the monogamy rule). The respective theories for such negatively connected exchange relations rest on the (usually implicit) assumption that everyone knows the relevant restrictions and behaves accordingly. From a more general perspective, the type of network connections is, in effect, an exogenously given component in these theories.

Our approach deals not only with negatively connected relations, but with positively connected and mixed networks as well. For that purpose, we need a more explicit assumption about network connections than previous theories. Our postulate reflects that the people decide about the type of their connections in reality—we assume that every system actor who has two or more ties in the given negotiation network classifies

[21]Feld (1991) explains why friends always seem to have more friends than oneself. For the scenario in which this 'class size paradox' is fully understood, he also derives the appropriate estimate for the mean number of friends of an individual's friends. The latter corresponds with the denominator of the far right–hand side of eq. (5.7c).

[22]Irrespective of i's categorization of his network relations, i's relative bargaining power in match with j (eq. (5.5)) is affected by c_i as well as c_j. From i's point of view, the calculation of his *relative* bargaining power thus requires the knowledge of the network structure three steps away from i.

each of his relations as either a negative or a positive connection, while each actor with just one bargaining relation adopts the classification of his only partner.[23] To keep things as simple as possible, however, the actors' assessments of relations are exogenous components in our model.[24]

It is clear that the individual classifications of potential exchange relations may coincide or deviate. A match between i and j thus will be either one of the pure types (i.e. a pure negatively connected or a pure positively connected relation) or a mixture of them. And, the whole system may be composed out of subnetworks which embrace either negatively or positively connected relations, but overlap somewhere. A mixed exchange network thus may easily result. According to Yamagishi, Gillmore, and Cook (1988), a mixture of positive and negative connections often characterizes real exchange systems (e.g. exchange in the Kula Ring as described by Malinowski 1922).

Starting from this insight, we now can specify how network control affects individual bargaining power. For that purpose, imagine first a system with negatively connected relations. Such a network has two essential features: depending on the values involved, i's bargaining relations are more or less substitutable and compete with the relations of i's partners to others ('friends of friends are enemies').[25] Both features suggest that i's absolute bargaining power rises with his network control—by definition, more control, that is, more structural autonomy means that i depends less on his current negotiation partner for exchange and that i's bargaining partners tend to have fewer and/or less valued relations. For the case of negatively connected relations, it thus can be assumed that i's bargaining power b_i rises with i's network control c_i.[26]

It is reasonable to postulate just the contrary for the opposite scenario of positively connected relations. In a setting with positive connections, concluded exchanges pro-

[23]Notice that this assumption does not exclude that the individual classifications of network connections simply reflects systemwide incentives. Our postulate therefore should not create problems in experimental work as long as test persons (have learned to) systematically react to incentives. Suppose, in accordance with the usual design of network exchange experiments (e.g. Skvoretz and Willer 1991), that subjects' monetary compensation for participation explicitly depends on their bargaining success. We expect that these incentives ensure, at least after several training rounds, relational assessments in the sense of the experimenters.

[24]Following the usual practice in the field, we thus predict exchange profits and structures only. This clearly simplifies the model—apart from the actors' structural positions (e.g. just one potential exchange relation or several bargaining relations), the assessments of relations may reflect the network type (e.g. dating or communication network), the number and type of goods to be divided (e.g. substitutes or complements), and the existence of systemwide restrictions (e.g. laws, norms, rules).

[25]Negatively connected relations usually refer to a situation in which bargaining sessions with different partners concern the distribution of just one homogeneous good. Note that, if there are at least two distinct resources, negative connections are also possible—negotiation partners then have to be willing to substitute one good for the other on a one–to–one basis (perfect substitutes).

[26]In a network of equally valued relations, actor i's network control c_i increases if the mean number of partners of i's partners decreases (cf. eq. (5.7c)). Assuming, in general, a positive relationship between bargaining power and network control thus reflects a basic property of a negatively connected system: friends of friends are enemies.

mote transfers with others.[27] Relations with different partners are complementary for the resource flow through the system. If an actor's structural autonomy is lower, the others more affect, via their links to one another, the relations to him. And, the exchange partners of an individual's partners appear as intermediaries ('friends of friends are friends'). These features of positively connected relations suggest that i's bargaining power increases if i's network control decreases. If an actor's structural autonomy is lower, the others more affect, via their valued links to one another, the valued relations to him. Put differently, if his network control is smaller, an individual has, by definition, a higher structural dependence. This creates additional opportunities in a positively connected setting. Due to the others' transactions, a less autonomous actor may serve as a broker. And, he may crucially affect the resource flow through the system. For the case of positively connected relations, it thus can be assumed that i's bargaining power b_i rises if i's network control c_i falls.[28]

To formalize these ideas, we take into account that b_i depends positively on the one–day discount factor δ_i (see, eq. (5.4)). It thus suffices to relate δ_i and c_i in a way which corresponds with the above reasoning. Moreover, it holds $0 < c_i \leq 1$, but $0 < \delta_i < 1$. Actor i's discount factor thus has to be a weighted version of i's network control. It is therefore assumed that

$$\delta_i = \begin{cases} wc_i & \text{if } i \text{ classifies a relation as negatively connected} \\ 1 - wc_i & \text{if } i \text{ classifies a relation as positively connected} \end{cases} \tag{5.8}$$

where we use the shorthand

$$w := \frac{m+n}{1+m+n} . \tag{5.9}$$

The latter is a network–specific fraction which rises with the number of mutual ties in the network, m, and the number of network members, n. In eq. (5.9), the weight w scales the degrees of network control or structural autonomy such that δ_i and b_i, respectively, are always positive numbers.[29]

---

[27] Positively connected relations may be characterized by the property that exchanges with distinct partners increase individual benefits. Such a situation exists, for example, if the focus is on pairwise distributions of different goods which are always consumed together in fixed proportions (perfect complements)—since every actor only cares about the combination of those goods, he will successively engage in a series of dyadic bargaining sessions to obtain as much as he can from each good. Note, however, that positive connections do not necessarily involve distinct goods: exchanges between, say, a professional athlete and his agent as well as complementary relations between the agent and organizers of athletic events may involve money only.

[28] As mentioned, c_i increases if the mean number of partners of i's partners decreases. The postulate of a negative relationship between bargaining power and network control therefore reflects a basic property of a positively connected system, regardless of the surplus values: friends of friends are friends.

[29] There are three reasons for eq. (5.9), the specific definition of w. First, since network control is measured on a ratio scale (see, eqs. (5.7a), (5.7b), and (5.7c)), the only admissible transformations for it are those which just change the unit of scale. The weight w clearly fulfills this requirement. Second, any bargaining network may be characterized by the number of mutual ties, m, and the number of system actors, n. It thus is reasonable to define the scaling factor w in terms of these system

Starting from the distribution of network control, each actor's bargaining power results from combining eq. (5.4) with eqs. (5.8) and (5.9) for either relational classification he may have. In accordance with our basic ideas, actor i's bargaining power b_i increases with i's network control c_i when he sees a potential exchange relation as a negative connection. If he classifies such a relation as a positively connected one, however, his bargaining power b_i decreases with his network control c_i.[30]

These assumptions ensure that the network structure determines the actors' bargaining powers. In accordance with other theories for exchange networks, we thus focus on the effects structure has for bargaining results.[31] This becomes clear when we specify the implications of our postulates for the negotiation outcomes in the different types of relations.[32]

5.3.3 Negotiation Outcomes and Relational Types

The actors' bargaining powers depend, by postulate, on their network embeddedness and relational classification. As a consequence, the model implications for the distributions of relative bargaining power and surplus in any given match reflects these determinants. The individual profit from a specific match results, as mentioned, via the limiting equilibrium solution of the alternating offers game.[33] Substituting eqs.

---

parameters. Third, weighting should preserve the essential role of network control in our approach. The weight w is a systemwide constant which, at most, moderately changes the original values of network control—because of $m \geq 1$ and $n \geq 2$, it holds that $0.75 \leq w < 1$.

[30]Stated differently, i's bargaining power b_i will be larger if he sees a relation as a positive connection and has a higher structural dependence $1 - c_i$.

[31]While the study of structure justly is at the center of theories for exchange networks, experimental results from other disciplines show that divisional bargaining does not vary with structure alone. Eckel and Grossman (1998) find, for example, that men and women bargain differently—while men apparently accept lower offers from women than from men, women seem to accept lower offers from women. In the future, it thus may be important to assess the relative strength of individual and structural effects on exchange outcomes or to ask similar questions (e.g. which individual effects reinforce which structural effects, which structural effects 'survive' the introduction of individual effects?).

[32]A more sophisticated approach would postulate that discount factors depend on both individual and structural variables. For example, if i classifies a relation as a negatively connected one, the assumption $\delta_i = \gamma w c_i + (1 - \gamma) d_i$ could be made. Here, d_i with $0 < d_i < 1$ would represent i's 'baseline' patience which, in turn, may be understood as a function of i's characteristics. In accordance with Becker and Mulligan's (1997) conclusions from an endogenization of time preference, for example, it can be expected that younger, less educated, and poorer persons will be less patient. And, the parameter γ with $0 \leq \gamma \leq 1$ would measure the degree to which structural positions determine the bargaining partners' discount factors. From this perspective, the above model focuses, due to eq. (5.8), on the extreme case $\gamma = 1$. While restrictive, the latter scenario ensures compatibility with the available experimental evidence on exchange networks. Such results usually rest on experimental designs that eliminate effects of interindividual differences.

[33]Stable transaction patterns thus are not an assumption, but a conclusion of our approach. This is important because, as Bonacich and Bienenstock (1995) point out, theories for exchange networks often just assume the emergence of stable exchange patterns.

(5.4) and (5.8) into eq. (5.5), we can distinguish four relational types each of which allows specific conclusions about the effects structure has on relative bargaining powers and negotiated exchange profits. For clarification and later reference, it is useful to consider the scenario of a pure negatively connected relation in greater detail.

Negative–Negative Connection: If the relation between actors i and j is, from their perspective, a pure negatively connected one, then i's relative bargaining power over j is

$$p_{ij} = \frac{b_i}{b_i + b_j} = \frac{(-1/\ln(wc_i))}{(-1/\ln(wc_i)) + (-1/\ln(wc_j))} = \frac{\ln(wc_j)}{\ln(wc_i) + \ln(wc_j)} . \quad (5.10)$$

As a consequence, i's negotiated exchange profit in the match with j results from

$$x_{ij} = p_{ij}\, v_{ij} = \left(\frac{\ln(wc_j)}{\ln(wc_i) + \ln(wc_j)} \right) v_{ij} \quad \text{for } i \neq j. \quad (5.11)$$

Hence, i's relative bargaining power and exchange profit in a negatively connected relation with j rise, everything else being constant, when either i's network control c_i increases or j's network control c_j decreases.[34]

Combining eqs. (5.4), (5.8) and (5.5) for the other configurations of relational classifications in an analogous way, we obtain three other predictions for the relative bargaining powers and the negotiated exchange profits.

Negative–Positive Connection: If i classifies the relation to j as negatively connected and j categorizes the relation as positively connected, then actor i's profit in the match with j is

$$x_{ij} = p_{ij}\, v_{ij} = \left(\frac{\ln(1 - wc_j)}{\ln(wc_i) + \ln(1 - wc_j)} \right) v_{ij} \quad \text{for } i \neq j. \quad (5.12)$$

For such a mixed relational orientation, actor i's relative bargaining power and negotiated profit increase if, everything else being constant, either c_i or c_j rises.

Positive–Negative Connection: If i classifies the relation to j as positively connected and j categorizes their relation as negatively connected, then actor i's profit in the match with j is

$$x_{ij} = p_{ij}\, v_{ij} = \left(\frac{\ln(wc_j)}{\ln(1 - wc_i) + \ln(wc_j)} \right) v_{ij} \quad \text{for } i \neq j. \quad (5.13)$$

In this mixed relation i's relative bargaining power and profit share increase, everything else being constant, when either c_i or c_j falls.

---

[34] These conclusions reflect that $\partial p_{ij}/\partial c_i > 0$, $\partial p_{ij}/\partial c_j < 0$, $\partial x_{ij}/\partial c_i > 0$ and $\partial x_{ij}/\partial c_j < 0$. The signs of these partial derivatives reflect the reaction of i's relative bargaining power p_{ij} and i's profit share x_{ij} in the match with j when exogenous structural changes affect either i's network control c_i or j's network control c_j, but preserve the valued relation between i and j.

Positive–Positive Connection: If the relation between actors i and j is, from their perspective, a pure positively connected one, then actor i's profit in the match with j results from

$$x_{ij} = p_{ij}\, v_{ij} = \left(\frac{\ln(1 - wc_j)}{\ln(1 - wc_i) + \ln(1 - wc_j)} \right) v_{ij} \quad \text{for } i \neq j. \tag{5.14}$$

Consequently, if their relation is a pure positive connection, actor i's relative bargaining power and negotiated profit increase, everything else being constant, if either i's network control c_i falls or j's network control c_j rises.

Since $w := (m + n)/(1 + m + n)$ and $c_i := \sum_k r_{ik}\, r_{ki}$ hold by definition and r_{ij} measures i's fraction of the systemwide valued relations to j, we thus may uniquely predict the distributions of relative bargaining power and negotiated profit for any given combination of relational orientations in each exogenously given bargaining network with m valued ties between n actors. A comparison of the four model conclusions shows, moreover, that the negotiated profits associated with the distinct relational types just differ in terms of p_{ij}, i's relative bargaining power in the match with j. And, the calculation of relative bargaining powers and negotiated profits requires, in principle, just a pocket calculator.

However, some of the negotiated profits may not be realized—negatively connected relations may reflect the existence of exogenous restrictions (e.g. one–exchange rule). In such settings (and mixed exchange networks as well), at least one system member may select his actual exchange partners from a larger set of potential exchange partners. That is, depending on the negotiation outcomes, he may never complete transfers with one or more of his negotiation partners. The analysis of such exclusionary behavior yields testable conclusions about coincidences or deviations between bargaining and exchange structures.[35]

5.3.4 Bargaining and Exchange Structures

Suppose that a bargaining network persists over several rounds of negotiated exchanges and that at least one system member i classifies his valued relations as negative connections. By assumption, i knows his negotiation outcomes with two different partners j and k, but can complete (for whatever reason) just one exchange per connection and round.[36]

[35] The term "exclusionary behavior" refers to a situation in which at least one actor never exchanges with one or more of his bargaining partners (consistent or permanent exclusion). It does not refer to a situation in which actors alternately exclude each of their potential partners in different rounds of negotiated exchanges (temporary exclusion).

[36] This simple scenario captures the essential features of all situations in which network partners may be consistently excluded. If the number of pies to be divided is smaller than the number of negotiation partners, potential exchange partners may be successively excluded via pairwise comparisons of the negotiation results.

We are interested in conclusions about situations in which i never completes transactions with his bargaining partner j, but always exchanges with j's competitor k. If i is a selfish profit–maximizing actor with substitutable exchange relations, he will decide about j's permanent exclusion on the basis of the exchange profits he could realize. It thus is assumed that i chooses his actual exchange partners via a systematic comparison of the negotiation results he can realize with each potential partner.[37]

Theorem 5.1 *If i is as selfish profit–maximizing actor with substitutable exchange relations, i will consistently exclude j if*

(a) there is an alternative bargaining partner k which guarantees, from i's perspective, a more favorable deal than j, and

(b) i's profit from the match with k in the case of exclusionary behavior towards j exceeds the profit i could obtain in the relation with k if he would at least occasionally exchange with his negotiation partner j as well.

Proof: The proof of this theorem can be found in Appendix C. □

Notice that Theorem 5.1 is fairly general—it is not restricted to pure negatively connected relations and a specific calculation of negotiated profits.[38] In fact, the result holds true for mixed and pure negatively connected networks and any model of negotiated exchanges which generates, based upon the postulate of individual rationality, point predictions for profit divisions.

Having specified such a model and the respective network type, we therefore can predict differences between the given bargaining network and the actual exchange structures by examining the whole system on its basis. According to Theorem 5.1, the exchange structures will differ from the negotiation network if and only if (see Appendix C)

$$x^*_{ik} > x_{ik} > x_{ij} \quad \text{for at least one } i \text{ with negative connections to } j \neq k, \qquad (5.15a)$$

where x^*_{ik} denotes i's negotiated profit in the match with k if i always exchanges with actor k (i.e. x^*_{ik} refers to the subnetwork which results from removing the link between i and j in the initial structure). Otherwise, the exchange network will coincide with the given bargaining network.

In combination with our exchange model, the inequality (5.15a) may be used to predict, as will be illustrated below, whether and, if so, how actual exchange structures

[37] Of course, there are other ways to deal with exclusionary behavior. One could, for example, follow Binmore (1985) and analyze a non–cooperative bargaining game with 'outside options' (i.e. exchange opportunities with others) for at least one partner. Osborne and Rubinstein (1990) discuss some basic results associated with this very complicated scenario.

[38] These features reflect that the formal derivation of this conclusion (see, Appendix C) does not depend on the modeling of bargaining and exchange. In addition, its derivation does not require assumptions about non–trivial probability values (e.g. probability of inclusion).

deviate from given bargaining relations.[39] Before the condition is routinely applied, however, its most important implications have to be discussed.

First, from the perspective of our approach, inequality (5.15a) is equivalent to a condition in terms of relative bargaining powers and cake sizes.[40] Accordingly, deviations between exchange and negotiation structures may occur because of different cake sizes and/or different relative bargaining powers in i's relations with j and k. Suppose, for purpose of clarification, that i's relative bargaining power in the match with j is much smaller than that in the match with k. Then, actor i may refrain from exchanging with j if the cakes to be divided with j and k have a similar size. If, however, i and j bargain over the partition of a large cake, whereas i and k negotiate over the division of a small surplus, it may be favorable for i to keep his relation with j—even if i does not get much of the large pie he divides with j, this share may be absolutely larger than that he could obtain in the relation with k.

Second, the inequality (5.15a) reflects that deviations between negotiation and exchange structures emerge because restricting attention to a smaller set of partners is more beneficial for at least one individual i than concluding exchanges with all initial partners. The permanent exclusion of a partner thus means that i can improve his bargaining situation for the remaining matches without suffering losses. In other words, actor i can change, without incurring costs to himself, the distribution of network control to his advantage. And, if i selects a proper subset of his initial partners for transactions, he affects the network control of his partners and their partners (cf. eq. (5.7a)). The permanent exclusion of at least one negotiation partner from exchange therefore also influences the profit divisions in others' relations.

Third, the original condition for differences between exchange and negotiation structures may be split into the necessary part $x_{ik} > x_{ij}$ and the sufficient part $x^*_{ik} > x_{ik}$ (see, Appendix C). The former identifies the subset of substitutable initial relations with a potential for exclusion. The latter selects those relations from this subset which, in fact, will be characterized by non–exchange.[41] Analyses of concrete network struc-

[39] From the perspective of our approach, the inequality (5.15a) is equivalent to a condition in terms of bargaining powers. Accordingly, the exchange structures will differ from the given negotiation network if and only if $b^*_i/b^*_k > b_i/b_k > b_i/b_j$ for at least one i with negative connections to $j \neq k$. Here, b^*_i and b^*_k denote the bargaining powers of i and k in the subnetwork which results if the tie between i and j in the initial structure is removed.

[40] This reformulation of inequality (5.15a) reflects eq. (5.5). Accordingly, the exchange relations will differ from the given negotiation network if and only if $p^*_{ik}\, v_{ik} > p_{ik}\, v_{ik} > p_{ij}\, v_{ij}$ for at least one i with negative connections to $j \neq k$. Here, p^*_{ik} denotes i's relative bargaining power in the relation with k if the tie between i and j has been removed from the initial structure.

[41] In larger negotiation structures, several of actor i's ties may fulfill the necessary part of inequality (5.15a). If so, actor i will successively compare all relevant bargaining partners with respect to the feasible payoffs. To realize the most favorable situation, he will apply inequality (5.15a) 'bottom–up', starting with the least favorable deal and moving up in the ranking of negotiation outcomes as long as the sufficient part of (5.15a) is not met.

tures illustrate, as will become clear below, the straightforward application of (the necessary and sufficient part of) the condition for exclusionary behavior.

Equally Valued Relations

In the case of equally valued relations $v_{ij} = v_{ji} = v$ for all $i \neq j$, inequality (5.15a) reduces to an expression in relative bargaining powers

$$p^*_{ik} > p_{ik} > p_{ij} \text{ for at least one } i \text{ with negative connections to } j \neq k, \qquad (5.15b)$$

where p^*_{ik} refers to i's relative bargaining power in relation with k if the link between i and j has been removed. Analogue to inequality (5.15a), (5.15b) can be split into the necessary part $p_{ik} > p_{ij}$ and the sufficient part $p^*_{ik} > p_{ik}$. In the simple (and experimentally relevant) scenario of equally valued relations, exclusionary behavior may occur, if at all, in specific networks only. To avoid superfluous work, essential properties of the relevant networks therefore should be explicitly specified. And, since exclusion usually is discussed in the context of pure negatively connected systems, such a characterization is particularly desirable for networks with negative connections and equally valued relations only.

Theorem 5.2 *From the perspective of our exchange approach, the relevant negatively connected networks with equally valued relations for the application of inequality (5.15b) are those in which*

(i) at least one actor i has three or more negotiation partners, and

(ii) these partners have two or more structurally distinct positions, one of which may be equivalent to i's.

Proof: The proof of this theorem can be found in Appendix C. □

That is, when a negatively connected network does not fulfill (i) and/or (ii), the condition for deviations between negotiation and exchange networks (i.e. inequality (5.15b)) is violated for sure. If, however, (i) and (ii) are met, there is at least one actor i who may avoid exchanges with at least one of his original negotiation partners—analyzing the negatively connected network via inequality (5.15b) then will be informative.[42]

In the following, analyses of concrete network structures illustrate, among other things, the straightforward application of (the necessary and sufficient part of) the condition for exclusionary behavior. To compare such theoretical predictions with empirical observations, we first need to describe and select relevant experimental studies.

[42]Exclusionary behavior reduces, in effect, one's own subnetwork to a smaller and more beneficial exchange structure. From the perspective of our approach, it is clear that only specific network positions provide opportunities for a profitable selection of exchange partners. Important features of such structural positions are described by (i) and (ii).

5.4 Experiments and Data Selection

There are various experimental results with regard to profit distributions in exchange networks (e.g. Bienenstock and Bonacich 1993; Lovaglia et al. 1995; Skvoretz and Fararo 1992; Skvoretz and Willer 1993; Yamagishi, Gillmore, and Cook 1988). There are also different experimental results with regard to deviations between negotiation and exchange relations (e.g. Markovsky, Willer, and Patton 1988; Simpson and Willer 1999). Since many experiments have been conducted, several exchange theorists (e.g. Burke 1997; Friedkin 1995; Yamaguchi 1996) substantiate the empirical relevance of new theoretical ideas by referring to an appropriate subset of the published findings. We also adopt this strategy of empirical validation. To specify the relevant data selection criteria, it is reasonable to briefly describe features of experimental studies on exchange networks.

5.4.1 Experimental Work

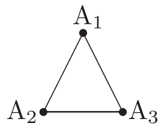
Starting point of an exchange network experiment is a specific bargaining structure which limits matches between potential exchange partners. Experimenters are interested in the effects the network embeddedness has on profit splits and/or exchange patterns between adjacent nodes. The presentation of experimental findings therefore refers to the types of structural positions (A,B,C,D,E) the different actors have. Individuals located at structurally equivalent positions are normally distinguished by numeric subscripts (e.g. A_1, A_2).

FIGURE 5.1 depicts negotiation structures most of which are popular in laboratory experiments. Nearly all of those experiments focus on networks with equally valued relations (i.e. each bilateral bargaining session concerns the division of an identical surplus). To prepare the application of our model to such structures, FIGURE 5.1 also informs about the relevant column–stochastic matrices $\mathbf{R}$ with elements r_{ij} and the associated control vectors $\mathbf{c}$ with entries $c_i := \sum_k r_{ik} r_{ki}$ for all i.

The design of experiments on exchange networks has common features (see, e.g. Skvoretz and Willer 1991). All experiments consist of several rounds of negotiation and exchange, while the relational structure is kept constant. Bargaining sessions involve adjacent network positions only, where usually a surplus of identical size (normally 24 ‘profit points’) is to be split in any bilateral match. After a series of offers and counteroffers, negotiations stop when an agreement is reached. Partly due to a computerized setting, proposals can be made within seconds and bargaining sessions do not last long (viz. agreement in less than a few minutes). Most experiments concern negatively connected networks. The experimental design often restricts the number of exchanges per connection and round—empirical research usually refers to those negatively connected settings in which every actor is subject to the same exchange restriction (e.g. one–exchange rule).

Subjects are usually undergraduates who participate for pay (monetary compensation according to bargaining success). They receive general information about the

TRIANGLE



	A1	A2	A3
A1	0	**.5**	**.5**
A2	**.5**	0	**.5**
A3	**.5**	**.5**	0

A1	.5000
A2	.5000
A3	.5000

3–LINE

	A1	A2	B
A1	0	0	**.5**
A2	0	0	**.5**
B	**1**	**1**	0

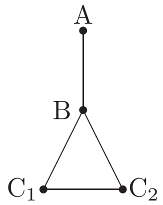
A1	.5000
A2	.5000
B	.9999

4–LINE

	A1	A2	B1	B2
A1	0	0	**.5**	0
A2	0	0	0	**.5**
B1	**1**	0	0	**.5**
B2	0	**1**	**.5**	0

A1	.5000
A2	.5000
B1	.7500
B2	.7500

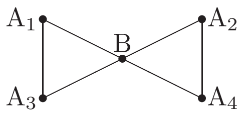
STEM



	A	B	C1	C2
A	0	**.33**	0	0
B	**1**	0	**.5**	**.5**
C1	0	**.33**	0	**.5**
C2	0	**.33**	**.5**	0

A	.3333
B	.6667
C1	.4167
C2	.4167

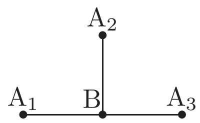
KITE



	A1	A2	A3	A4	B
A1	0	0	**.5**	0	**.25**
A2	0	0	0	**.5**	**.25**
A3	**.5**	0	0	0	**.25**
A4	0	**.5**	0	0	**.25**
B	**.5**	**.5**	**.5**	**.5**	0

A1	.3750
A2	.3750
A3	.3750
A4	.3750
B	.5000

3–BRANCH



	A1	A2	A3	B
A1	0	0	0	**.33**
A2	0	0	0	**.33**
A3	0	0	0	**.33**
B	**1**	**1**	**1**	0

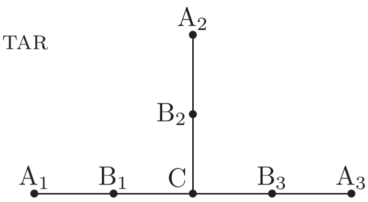
A1	.3333
A2	.3333
A3	.3333
B	.9999

5–LINE

	A1	A2	B1	B2	C
A1	0	0	**.5**	0	0
A2	0	0	0	**.5**	0
B1	**1**	0	0	0	**.5**
B2	0	**1**	0	0	**.5**
C	0	0	**.5**	**.5**	0

A1	.5000
A2	.5000
B1	.7500
B2	.7500
C	.5000

31–STAR



	A1	A2	A3	B1	B2	B3	C
A1	0	0	0	**.5**	0	0	0
A2	0	0	0	0	**.5**	0	0
A3	0	0	0	0	0	**.5**	0
B1	**1**	0	0	0	0	0	**.33**
B2	0	**1**	0	0	0	0	**.33**
B3	0	0	**1**	0	0	0	**.33**
C	0	0	0	**.5**	**.5**	**.5**	0

A1	.5000
A2	.5000
A3	.5000
B1	.6667
B2	.6667
B3	.6667
C	.5000

FIGURE 5.1 *Network structures, relational matrices* **R**, *and control vectors* **c**.

purpose and the number of rounds of the experiment. Training and practice rounds serve to ensure their understanding of the bargaining situation. And, there is no misdirection. Subjects normally possess complete information about the bargaining rules, the earnings of their partners, the shape of the negotiation structure, and their own positions within the network.[43]

To eliminate possible effects of interindividual heterogeneity, researchers apply different procedures, however. Some authors (e.g. Bonacich and Friedkin 1998) randomly assign subjects to fixed network positions and prevent face to face encounters (i.e. subjects sit in separate rooms and interact through computer terminals). Others (e.g. Lovaglia et al. 1995) ensure an impersonal experimental setting as well, but rotate subjects through all network positions during the experiment. In addition, both procedures should help inhibit, partly in combination with other aspects of the experimental design (e.g. lack of information on other players' earnings), possible effects of individual preferences for equality in earnings over the entire experiment. It is not clear, however, whether these precautions are fully effective or, if not, which procedure is more appropriate.[44]

Nevertheless, experimental results are often interpreted as effects the given network structure has on exchange patterns and/or profit divisions between adjacent positions. As a matter of fact, however, they are averages of what happened over a limited number of negotiated exchanges. Experimental findings on deviations or coincidences of negotiation and exchange structures are reported as observed proportions of exchanges between specific positions over a series of rounds. And, results on exchange profits are routinely represented by the means of profit points the advantaged positions in given matches could realize over several rounds of the experiment. Occasionally, they are reported as systemwide fractions of power associated with structural positions (i.e. in our terminology, the relative bargaining powers of position holders are standardized over all matches such that they add up to unity over all network members). Before we can compare predictions from our theoretical model with such experimental observations, it is necessary to say a few words about our choice of data.

5.4.2 Selection of Experiments

Several aspects characterize our choice of empirical data. First, although our model can handle exchange networks with heterogeneously valued relations, we focus on those experiments which concern identically valued relations. The reason is simple: Bonacichand Friedkin (1998) present the only experimental study which systematically varies, for each of four network structures, the size of the cake to be partitioned in specific bilateral relations. They do not report precise numerical findings, however.

[43]There exist some experimental studies in which the subjects did not know the overall shape of network structures, their own network positions, or the other players' earnings (e.g. Bonacich and Friedkin 1998; Thye, Lovaglia, and Markovsky 1997).

[44]This conclusion reflects statements of experimenters who apply the different procedures. To rationalize specific empirical results, Lovaglia et al. (1995) refer to the potential problem of subjects' preferences for equity in experimental research. Bonacich and Friedkin (1998) also mention that their results may have been affected by such preferences.

We therefore cannot systematically compare their data with predictions from our model and other theories.[45]

Second, we restrict our attention to those experimental studies in which subjects possess complete information about the bargaining rules, the earnings of their partners, the shape of the negotiation structure, and their own positions within the network. This decision simply reflects the assumptions of our theory—actors have to know at least their own valued relations and those of their partners as well as the numbers of mutual ties and actors in the network.

Third, apart from one exception, we compare theoretical predictions with experimental results on negatively connected structures. This reflects, as noted by Yamaguchi (2000), that empirical research on mixed exchange networks is absent and that Yamagishi, Gillmore, and Cook (1988) is the only experimental study of a positively connected setting. In addition to empirical work on negative connections (e.g. Cook et al. 1983; Markovsky, Willer, and Patton 1988), we therefore refer to the latter in our comparison with Yamaguchi's (1996) theory, the only other model for both negatively and positively connected relations.

Fourth, because of the relatively large number of experiments on negatively connected structures, we select a subset of those studies. This subset is characterized by a similar design—our focus will be on experiments in which (I) each position holder faces the same exchange restriction (e.g. one–exchange rule), and (II) subjects are rotated through all network positions. The selection criterion (I) reflects that the above version of our model does not refer to networks in which actors at distinct structural positions differ with respect to the number of exchanges they can complete per period.[46] Rather, it is concerned with networks in which each actor is subject to the same restriction.

The selection criterion (II) reflects a methodological point. As mentioned, experimenters try to inhibit possible effects of subjects' desire for equality in earnings via two distinct procedures (random assignment to fixed positions versus rotation of sub-

[45] Bonacich and Friedkin (1998) present diagrams to contrast their findings with predictions from Bienenstock and Bonacich's (1992, 1993) game–theoretic core model, Cook and Yamagishi's (1992) Equi–Dependence approach, and Friedkin's (1993, 1995) Expected Value theory. A rough comparison with their figures shows that our model fits their data worse than the latter, but better than the Equi–Dependence approach. And, the model predictions for the relative bargaining powers always fall into the range predictions of the core.

[46] This feature does not characterize a slightly generalized version of our model which is based on the introduction of an experimentally adjustable parameter. Let $g_i \geq 1$ represent the exogenously given number of actor i's admissible exchange relations per period. Using the additional parameter g_i, we can extend our assumption on individual bargaining power b_i. In the case of negatively connected relations, eq. (5.4) may be replaced by $b_i = -1/(g_i \ln \delta_i)$. Starting from this more general postulate, we obtain, after substituting into eq. (5.5) and some algebra, $x_{ij} = p_{ij} v_{ij} = g_j \ln(wc_j) v_{ij} / (g_i \ln(wc_i) + g_j \ln(wc_j))$. This equation for i's negotiated exchange profit in a negatively connected relation with j reduces to eq. (5.11) whenever $g_i = g_j$ for $i \neq j$. Additionally, it allows profit predictions for those negatively connected networks in which actors at distinct structural positions differ with respect to the number of exchanges they can complete per period. A systematic comparison with empirical findings for such networks (Lovaglia et al. 1995; Skvoretz and Willer 1993) shows that the relevant profit predictions correspond at least as well with the experimental data as those of the competing resistance theories.

jects through all structural positions)—the more appropriate procedure is not known yet, however. To avoid biases as far as possible, we therefore fix the procedure and mainly refer to such studies on negatively connected structures in which subjects were rotated through networks positions.[47] While this selection criterion may seem restrictive, there are enough published data for a systematic comparison of predictions from our Non–Cooperative Bargaining (hereafter, NCB) model and other theories.

5.5 Applications

5.5.1 Profit Splits in Negatively Connected Networks

Using experimental results obtained by Skvoretz and Willer (1993) as reference point, TABLE 5.1 compares NCB–predictions for dyadic exchange outcomes in popular networks (4–LINE, STEM, KITE, 3–BRANCH, see FIGURE 5.1) with those of alternative approaches. The latter are Lovaglia et al.'s (1995) GPI–RD, Yamaguchi's (1996) power model (Y), Skvoretz and Willer's (1993) Exchange Resistance theory (ER), Friedkin's (1986, 1992) Expected Value theory (EV), Cook and Yamagishi's (1992) Equi–Dependence theory (ED), and Burke's (1997) Identity Simulation model (IS).[48] The observations listed in TABLE 5.1 refer to estimated means of profit points of structural position B, where 24 profit points per match had to be divided.

The associated NCB–predictions follow, for the given size of the cake of 24, from inserting the actors' degrees of structural autonomy (cf. FIGURE 5.1) and the network–specific weight into eq. (5.11), the equation for dyadic profit shares in a pure negatively connected relation. To illustrate the procedure, let us take a look at the 4–LINE structure and compute the relevant profit points. FIGURE 5.1 informs about the structural autonomy levels or degrees of network control: $c_{A_1} = 0.5 = c_{A_2}$ and $c_{B_1} = 0.75 = c_{B_2}$. It also gives the number of mutual ties ($m = 3$) and the number of system actors ($n = 4$) such that $w = (m + n)/(1 + m + n) = 0.875$. Substituting into eq. (5.11) yields, when numeric subscripts are dropped for notational convenience, the profit point prediction for B in the match with A:

$$x_{BA} = p_{BA} \times 24 = \left(\frac{\ln(0.875 \times 0.5)}{\ln(0.875 \times 0.75) + \ln(0.875 \times 0.5)} \right) \times 24 = 15.9$$

such that $x_{AB} = (1 - p_{BA}) \times 24 = 24 - 15.9 = 8.1$ gives the profit points for A in the relation with B. In addition to such NCB–predictions, TABLE 5.1 contains profit point predictions from other models. To compare the empirical performance of the theories, we use two 'goodness–of–fit' measures, viz. the 'Absolute Deviation'

[47] The only exception is Cook et al.'s (1983) study. In their experimental setting, subjects were neither rotated nor informed about their partners' earnings. Following Yamaguchi (1996) and Markovsky et al. (1997), we include the Cook et al. data in our comparison with Yamaguchi's theory.

[48] We selected those approaches because, in contrast to other published theories (e.g. Bienenstock and Bonacich 1992), they offer profit point predictions for all networks under study.

TABLE 5.1 *Observed and predicted dyadic profits for negatively connected network structures.*

		Theoretical Predictions[a]							
Network	Match	NCB	GPI–RD	Y	ER	EV	ED	IS	Observed[b]
4–LINE	B:A	15.9	14.5†	13.3†	16.0	21.1	16.0	13.8†	14.1 (0.40)
STEM	B:A	16.8†	15.6†	14.4†	18.3	22.0	18.0	15.6†	15.3 (0.82)
	B:C	15.7†	13.7†	13.2†	15.2†	19.5†	14.4†	12.9†	16.5 (2.64)
KITE	B:A	13.9†	13.7†	12.8†	12.5†	12.0	12.0	12.1	14.1 (0.77)
3–BRANCH	B:A	21.7†	23.0[c]	18.0	21.2†	22.0†	24.0	21.9†	21.6 (0.49)
AD[d]		0.88	1.06	1.98	1.64	3.84	2.24	1.30	
MD[d]		0.50	0.64	1.04	0.83	2.07	1.01	0.83	

Notes: Experiments took place under an overall one–exchange rule, where 24 profit points per match had to be divided. Predictions and observations are for profit points of structural position B. Profit points for positions A and C in the match with B result from (24 – profit of B), respectively.

NCB = Non–Cooperative Bargaining model; GPI–RD = Graph–theoretic Power Index with Resistance and Degree (Lovaglia et al. 1995); Y = Yamaguchi's (1996) power model, predictions based on the parameter choice $s = 8$ (elasticity of substitution); ER = Exchange Resistance theory (Skvoretz and Willer 1993); EV = Expected Value theory (Friedkin 1986, 1992); ED = Equi–Dependence theory (Cook and Yamagishi 1992); IS = Identity Simulation model (Burke 1997).

[a] Profit predictions, except for GPI–RD, Y, and IS, are taken from Skvoretz and Willer (1993, Table 2); predictions for GPI–RD, Y, and IS are taken from Lovaglia et al. (1995, Table 1), Yamaguchi (1996, Table 3), and Burke (1997, Table 1), respectively.

[b] Experimental results with standard errors in brackets, as reported by Skvoretz and Willer (1993, Table 2).

[c] Own calculation, based on Lovaglia et al. (1995).

[d] AD = Absolute Deviation (the sum of absolute distances between observed and predicted profit points relative to number of comparisons); MD = Mean Deviation (the Euclidean distance between observed and predicted profit points relative to number of comparisons).

† Daggers indicate that predicted values fall within the 95% confidence interval of the observed values. Put differently, these predictions fit the observations at the $p < 0.05$ significance level (two-tailed tests).

(AD) and the 'Mean Deviation' (MD). Whereas AD is the average absolute difference between observed and predicted profit points over all matches under consideration, MD is the Euclidean Distance between observed and predicted profit points relative to the number of comparisons.[49]

According to the MD and AD values in TABLE 5.1, all theories make acceptable profit point predictions. Nevertheless, they differ in terms of predictive accuracy. In particular, Friedkin's (1986, 1992) Expected Value theory (EV) seems, at first sight, to be an outlier. A closer look reveals, however, that EV predicts poorly with respect to just two matches (viz. the 4–LINE and the B:A relation in the STEM network). These relations also turn out to be problematic for Skvoretz and Willer's (1993) Exchange

[49]While the calculation of AD is straightforward, an example for the computation of MD may be helpful. To obtain the value of MD for the NCB model (see, TABLE 5.1), we determine the sum of squared deviations between observations and predictions: $(14.1-15.9)^2+(15.3-16.8)^2+(16.5-15.7)^2+(14.1-13.9)^2+(21.6-21.7)^2=6.18$. Taking the square root and dividing the result by the number of comparisons (i.e. $\sqrt{6.18}/5$) gives the reported MD of 0.50.

Resistance theory (ER) as well as Cook and Yamagishi's (1992) Equi–Dependence theory (ED). While Burke's (1997) Identity Simulation model (IS), Lovaglia et al.'s (1995) GPI–RD, and Yamaguchi's (1996) power model (Y) offer much better predictions for those links, they do not closely fit the empirical results for others.[50] And, even though the NCB model has the lowest AD and MD, it is also not completely in line with the experimental evidence. This can be seen from the more standard test of fit reported in TABLE 5.1 as well. Following Skvoretz and Willer (1993) and Yamaguchi (1996), it is asked there whether predictions fall within the 95% confidence intervals of the empirically observed means. According to this test, the NCB–prediction for the 4-LINE structure does not fit the experimental data. The other NCB–predictions correspond with the evidence, however.

This is important for a specific reason. The empirical observations in TABLE 5.1 indicate that negatively connected networks with equally valued relations and structurally distinct positions may considerably differ in terms of profit divisions. Compare, for example, the mean profit points obtained by position B in matches with actors at position A in the 3–BRANCH and KITE networks: B gets, on average, 21.6 profit points (i.e. 90% of the available resources) in the 3–BRANCH network, but 14.1 profit points (i.e. 58% of the available resources) in the KITE structure. Because of the experimental design, such variations of profit splits can be attributed to structural effects.[51] According to the NCB model, they reflect that, apart from unequal numbers of actors and ties, the structures under consideration are associated with different distributions of network control—while the KITE network has more nodes and links than the 3–BRANCH structure, the distribution of network control is more unequal in the latter than in the former. As a consequence, the profit differentiation in the 3–BRANCH network will be more extreme than in the KITE structure.[52] For a more elaborated discussion on

[50] All theories fail to predict that the structural position B in the STEM network apparently gets a larger profit in the match with C than in the relation with A. It should be noted, however, that the mean of profit points associated with the B:C relation in TABLE 5.1 is, in comparison to the empirical results for all other matches, based on a low number of exchanges and associated with a large standard error. Skvoretz and Lovaglia (1995) discuss, among other things, the relatively rare occurrence of B:C exchanges in the STEM structure.

[51] As Skvoretz and Willer (1993) point out, most established exchange theories identify, more or less explicitly, the 'exclusionary potential' associated with a particular network position as the crucial explanatory factor for the profit differentiation in negatively connected networks with equally valued relations and structurally distinct positions.

[52] The NCB model predicts a maximum differentiation of profits in specific networks only. For negatively connected structures with equally valued relations between distinct positions, eq. (5.11) implies that the profit division approaches maximum differentiation if and only if there is a sufficiently large number of network members and mutual ties and the advantaged actor has full control over his relations (monopoly position). An example is the negatively connected 100–BRANCH structure in which 100 actors solely depend on actor i for the split of an identical value v. This straightforward extension of the 3–BRANCH logic from FIGURE 5.1 is characterized by $m = 100$, $n = 101$, $c_i = 1$, and $c_j = 0.01$ for all $j \neq i$ such that $x_{ij} = p_{ij}\, v = 0.999\, v$ determines i's profit share in the negatively connected relation with j. Notice that, while the predicted profit division in small networks (i.e. structures with few actors and/or ties) deviates from maximum differentiation, the condition $c_i = 1$

the variation of negotiation results among negatively connected structures, see Braun and Gautschi (2000). They address the question why exchange outcomes in different negatively connected networks vary as much as they do. Notice that, if the empirically observed means of profit points are taken as measuring rods, the NCB model closely predicts the profit splits in both networks.[53]

And, the acceptable fit of the NCB model is not limited to these empirical findings. This becomes clear if one follows Burke (1997) and combines the data of Skvoretz and Willer (1993) with those of Lovaglia et al. (1995) as well as Skvoretz and Fararo (1992).[54] Specifically, we calculate the means of all the experimental results for each of the four network structures: 4–Line 13.7; Stem B:A 15.2, B:C 15.9; Kite 13.2; 3–Branch 20.7. A systematic comparison of these means with the theoretical predictions listed in Table 5.1 shows that the NCB model produces the smallest deviations (AD = 1.14; MD = 0.60). Again, the NCB predictions are relatively close to the average observations for the networks under consideration. It is to be asked whether this conclusion may be extended to other networks and types of connections.

5.5.2 Power Distributions in Negatively and Positively Connected Structures

Negotiated distributions of exchange profits reflect power inequalities due to different structural positions. Usually, the power of one position over another is measured by the proportion of value obtained by the former position in the match with the latter. This procedure is compatible with the NCB approach: it holds $p_{ij} = x_{ij}/v_{ij}$ for $i \neq j$ because of eq. (5.5). Additionally, the NCB model provides, for each of four relational configurations, a simple formula according to which i's relative bargaining power in the relation with j depends on structural features (cf. the different expressions for p_{ij} in eqs. (5.11) – (5.14)). Actor i's systemwide power share p_i thus may be obtained via

always ensures a relatively extreme profit differentiation in favor of i if the focus is on negatively connected structures.

[53]It is sometimes argued that structures like the 3–Branch are characterized by a process of competitive bidding between those position holders who are threatened by exclusion. Accordingly, there will be steadily increasing differences in profits between exchanging actors in consecutive rounds of experimental bargaining and exchange—since theoretical predictions usually are understood as equilibrium outcomes, means of profit points may be inappropriate indicators for comparison. If this argument is valid, however, it is surprising that experimenters report means of profit points only. Clearly, the median would be a more appropriate measure for the typical experimental outcome whenever the distribution of observed profit splits is highly skewed or if errors are not evenly distributed about the mean.

[54]The combination of these data neglects, however, that the studies vary, despite similar designs, in terms of the experimental results reported. For example, while Lovaglia et al. (1995) report simple arithmetic means of profit splits in network relations from the last rounds of experiments only, Skvoretz and Willer (1993) estimate means of profit points via a constrained regression technique from all experimental observations.

TABLE 5.2 *Observed and predicted power of positions for selected networks.*

Network	Network Positions C	B	A	WD
5–LINE, NEGATIVELY CONNECTED				
Predicted (NCB)	.1649	.3351	.0825	.0043[a]
Observed (Cook et al. 1983)	.2097	.3059	.0892	—
Observed (Cook et al. 1983, corrected)[b]	.1436	.3303	.0978	—
Observed (Markovsky et al. 1997)	.1159	.3464	.0957	—
Mean of Observations	.1564	.3275	.0942	—
Predicted (Yamaguchi 1996)				
parameter choice $s = \infty$	.1667	.2500	.1667	.0301[a]
parameter choice $s = 8$	.1764	.2500	.1618	.0294[a]
31–STAR, NEGATIVELY CONNECTED				
Predicted (NCB)	.1923	.2051	.0641	.0188[a]
Observed (Skvoretz and Fararo 1992)[c]	.1021	.2583	.0410	—
Observed (Markovsky et al. 1988, 1997)	.0635	.2355	.0767	—
Mean of Observations	.0828	.2469	.0589	—
Predicted (Yamaguchi 1996)				
parameter choice $s = \infty$	.1250	.1667	.1250	.0264[a]
parameter choice $s = 8$	.1383	.1667	.1205	.0262[a]
5–LINE, POSITIVELY CONNECTED				
Predicted (NCB)	.3264	.1736	.1632	.0071
Observed (Yamagishi et al. 1988)[d]	.3133	.1931	.1503	—
Predicted (Yamaguchi 1996)				
parameter choice $s = -\infty$	.5000	.2500	.0000	.0588
parameter choice $s = -8$	.4961	.2500	.0020	.0579

Notes: WD = Weighted Deviation (the average weighted Euclidean Distance between observed and predicted systemwide profit shares, where the weights refer to the number of indicated positions in the network). NCB = Non–Cooperative Bargaining model. Applying Yamaguchi's theory requires a specification of the elasticity of substitution (s).

[a] The calculation of WD refers to the mean of observations.

[b] Corrections by Markowsky et al. (1997, Table 1).

[c] The original values have been re–scaled to sum to 1.0.

[d] Following Yamaguchi (1996) and Markowsky et al. (1997), the result from the first trial block (see Table 1 in Yamagishi, Gillmore, and Cook 1988, p. 843) was considered as an extreme outlier and therefore excluded.

adding up i's relative bargaining power for all matches and standardizing the result by m, the number of mutual ties in the respective network. Formally,

$$p_i = \frac{1}{m} \sum_{k=1}^{n} p_{ik} \quad \text{for all } i \tag{5.16}$$

such that $\sum_{k=1}^{n} p_k = 1$. TABLE 5.2 reports relevant predictions of the NCB model for the negatively connected networks 5–LINE and 31–STAR as well as for the positively connected 5–LINE structure (see again FIGURE 5.1 for graphs, relational matrices, and control vectors).

Using experimental observations for each network (Cook et al. 1983; Markovsky, Willer, and Patton 1988; Markovsky et al. 1997; Yamagishi, Gillmore, and Cook 1988) as reference points, TABLE 5.2 contrasts NCB–predictions with those from

Yamaguchi's (1996) power model (i.e. the other theoretical approach dealing with both negatively and positively connected relations). For the comparison of the theories' predictive accuracy, we use the average weighted Euclidean distance between observed and predicted systemwide profit levels associated with distinct positions in any particular network as the 'goodness–of–fit' measure. This 'Weighted Deviation' (WD) takes into account that specific network positions may exist several times in the structure under consideration. To demonstrate its computation, let us take a look at the NCB–predictions for the negatively connected 5–LINE. As displayed in TABLE 5.2, NCB assigns the systemwide power levels $p_C = 0.1649$, $p_B = 0.3351$, and $p_A = 0.0825$ to the positions C, B, and A, respectively.[55] The latter are compared with the average observations (TABLE 5.2) for this structure. Since the 5–LINE consists of just one position C, but two positions A and B, the sum of squared deviations is $(0.1564 - 0.1649)^2 + 2 \times (0.3275 - 0.3351)^2 + 2 \times (0.0942 - 0.0825)^2 = 0.000462$. Taking the square root and dividing the result by the number of comparisons (i.e. $\sqrt{0.000462}/5$) yields the reported WD of 0.0043.

Together with the other WD values from the last column of TABLE 5.2, this result indicates that the NCB–predictions are closer to the experimental observations than predictions from Yamaguchi's model for different parameter choices. In particular, the NCB model performs well in regard to the only published experiment on a positively connected exchange network (Yamagishi, Gillmore, and Cook 1988).[56] The latter focused on the flow through exchange of two different goods in a 5–LINE structure, with one flowing from the left–hand side and the other flowing from the right–hand side of the network. Positively connected relations were ensured by incentives—the extent to which an actor acquired both goods determined his payoff. Following Yamaguchi (1996) and Markovsky et al. (1997), TABLE 5.2 reports the experimental observations in standardized form. The NCB predictions for the power levels of the three positions

[55]The calculation of the systemwide power levels may be illustrated as well. Take, for example, position B in the negatively connected 5–LINE network. Because of the structural features and the degrees of network control displayed in FIGURE 5.1, position B has the relative bargaining power $p_{BA} = 0.6701 = p_{BC}$ in matches with positions A and C, respectively. Positions in the 5–LINE structure are connected by four mutual ties ($m = 4$). Evaluating eq. (5.16), B's systemwide proportion of power is then $p_B = (1/4) \times (0.6701 + 0.6701) = 0.3351$.

[56]Willer and Skvoretz (1999) report experimental results for specific 'inclusive' exchange networks. The latter are related, but not identical to positively connected settings. Loosely speaking, inclusive exchange networks are structures in which each actor has as much admissible and beneficial relations as he has exogenously given network ties. Because of the complementarity aspect associated with such a configuration, this specific class of networks may be analyzed via a slightly generalized version of eq. (5.14). Specifically, if $g_i \geq 1$ represents the experimentally adjustable number of exchange relations per period in which actor i can benefit, we can extend our assumption on individual bargaining power b_i for the case of a positively connected relation: $b_i = -g_i/\ln\delta_i$ for all $i \neq j$. Starting from this extended version of eq. (5.4) (and taking into account eqs. (5.8) and (5.9)), we obtain, after substituting into eq. (5.5) and some algebra, $x_{ij} = p_{ij}v_{ij} = g_i\ln(1 - wc_j)v_{ij}/(g_i\ln(1 - wc_j) + g_j\ln(1 - wc_i))$. This equation for profit predictions in inclusive structures reduces to eq. (5.14) if $g_i = g_j$ for $i \neq j$. And, its application to Willer and Skvoretz's inclusive networks gives predictions which closely fit their empirical findings.

result from combining the network features ($m = 4$, $n = 5$) and the degrees of network control (cf. FIGURE 5.1) in eq. (5.8) and inserting the results into eq. (5.14). They correspond closely with the empirical evidence.

In comparison to its predictive performance for 5–LINE structures, the NCB model gives a relatively inaccurate prediction for the power distribution in the 31–STAR network. TABLE 5.2 lists the NCB predictions and informs about the relevant observations of Skvoretz and Fararo (1992) and Markovsky et al. (1988, 1997). While the NCB predictions for the power levels of the positions A and B are reasonably close to the empirically observed means, the prediction for C's power is too large. But, as far as we know the literature, neither theory does better with respect to point predictions for all positions in this particular structure.

In sum, the NCB model predicts the effects of network embeddedness for exchange profits at least as well as the best fitting published theories. It also answers whether a given bargaining structure gives rise to the same exchange network and, if not, which position will be responsible for such a deviation and what exchange structures will result. To clarify, we now focus on such structural questions.

5.5.3 Deviations between Bargaining and Exchange Structures

As mentioned, exclusionary behavior exists if at least one network member always refuses to exchange with a bargaining partner. Such non–exchange may be expected if there are, like in most experiments, exchange restrictions (e.g. one–exchange rule). That is, negatively connected settings may be associated with deviations between the given bargaining network and the resulting exchange structures. As was already described in detail, there is a condition for exclusionary behavior which may be split into a necessary and a sufficient part (see, Appendix C). Accordingly, exclusionary behavior towards a specific network partner requires that (a) exchange with an alternative partner guarantees a higher profit (necessity), and (b) the profit from the match with the alternative partner in the case of exclusionary behavior exceeds the profit from the relation with the alternative partner in the case of non–exclusionary behavior (sufficiency).

The inequality (5.15a) concisely expresses these requirements. From the perspective of the NCB model, its application is straightforward when the focus is on the empirically relevant case of negatively connected structures with uniformly valued relations. This simplicity has two reasons. First, if each matched pair of actors divides an identical cake (i.e. $v_{ij} = v_{ji} = v$ for all $i \neq j$), inequality (5.15a) reduces to the simpler condition in terms of relative bargaining powers, that is, inequality (5.15b): $p^*_{ik} > p_{ik} > p_{ij}$ for at least one i with negative connections to $j \neq k$. Second, it can be shown (see Appendix C) that this condition can only be fulfilled in those negatively connected networks with uniformly valued relations in which (*i*) at least one actor i has three or more negotiation partners, and (*ii*) i's partners have two or more structurally distinct positions, one of which may be equivalent to i's (cf. Theorem 5.2). That is, the NCB model immediately predicts that a given negatively connected structure with equally valued relations will be robust (i.e. bargaining and exchange ties will coincide) if all actors have at most two negotiation partners or only structurally equivalent partners.

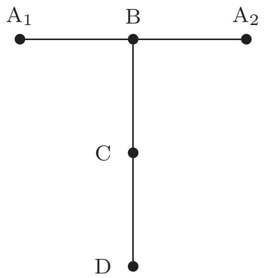
To illustrate the application of this result, take a look at the structures with equally valued relations displayed in FIGURE 5.1 and assume negative connections only. There is just one network for which we cannot immediately predict a coincidence of bargaining and exchange structures—apart from the STEM network, all structures in FIGURE 5.1 are robust because (i) and/or (ii) are not fulfilled. We thus have to apply the inequality (5.15b) just to the STEM structure. Such a closer analysis reveals that the STEM network is also a robust structure in the sense of a coincidence of bargaining and exchange network—no position holder will consistently exclude a specific bargaining partner.

For the subset of networks listed in TABLE 5.1, these structural predictions are in line with the available evidence on the number of exchanges between positions (cf. Lovaglia et al. 1995; Skvoretz and Willer 1993). Several authors (e.g. Bienenstock and Bonacich 1993; Markovsky, Willer, and Patton 1988; Simpson and Willer 1999; Skvoretz and Lovaglia 1995) report and discuss 'breaks' of other experimental networks, however. To illustrate such deviations and the application of the NCB model for their identification, we now analyze a specific network (T–SHAPE) with negative connections and uniformly valued relations in detail and then contrast model predictions for different structures with experimental findings.

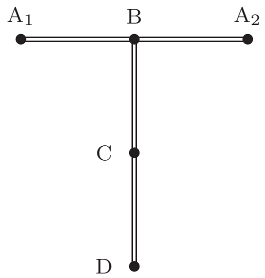
5.5.3.1 Identifying Structural Deviations

FIGURE 5.2 depicts the single–link and double–link T–SHAPE negotiation structures each of which refers to equally valued relations. The former (latter) represents a situation in which connected position holders engage in one (two) bargaining session(s) per round. While the value matrices $\mathbf{V}_S$ and $\mathbf{V}_D$ in FIGURE 5.2 capture these differences, they actually represent the same relational structure. As displayed in FIGURE 5.2 as well, there is a common matrix of standardized actor relations $\mathbf{R}$ and, consequently, just one vector of network control $\mathbf{c}$. Because of the different number of mutual ties ($m_S = 4$, $m_D = 8$), however, there are different weights for the calculation of profit shares (viz. $w_S = 0.900$, $w_D = 0.929$). The concrete profit predictions for both negotiation networks thus differ.

The profit rankings for the analysis of exclusionary behavior are identical in both structures, however. We therefore can postulate the existence of a one–exchange rule and examine just the potential decay of the single-link T–SHAPE structure into at least two separate exchange subnetworks. Its inspection shows that the requirements (i) and (ii) are met—it is worthwhile to ask for exclusionary behavior on the basis of inequality (5.15b). It also shows that the positions A and D cannot induce a structural deviation between negotiation and exchange patterns—both have just one potential partner so



Single–Link T–SHAPE



Double–Link T–SHAPE

$\mathbf{V}_S$	A1	A2	B	C	D
A1	0	0	**24**	0	0
A2	0	0	**24**	0	0
B	**24**	**24**	0	**24**	0
C	0	0	**24**	0	**24**
D	0	0	0	**24**	0

$\mathbf{V}_D$	A1	A2	B	C	D
A1	0	0	**48**	0	0
A2	0	0	**48**	0	0
B	**48**	**48**	0	**48**	0
C	0	0	**48**	0	**48**
D	0	0	0	**48**	0

$\mathbf{R}$	A1	A2	B	C	D
A1	0	0	**.33**	0	0
A2	0	0	**.33**	0	0
B	**1**	**1**	0	**.5**	0
C	0	0	**.33**	0	**1**
D	0	0	0	**.5**	0

$\mathbf{c}$	
A1	.3333
A2	.3333
B	.8333
C	.6667
D	.5000

FIGURE 5.2 *Single–link and double–link* T–SHAPE *bargaining structures with corresponding value matrices* $\mathbf{V}_S$ *and* $\mathbf{V}_D$*, common relational matrix* $\mathbf{R}$*, and associated control vector* $\mathbf{c}$*.*

that the inequality (5.15b) does not apply. The latter is to be used, however, for analyzing B's and C's selection of actual exchange partners.

To identify a potential deviation between negotiation and exchange structures via inequality (5.15b), we determine the actors' relative bargaining powers by inserting the weight and the degrees of network control (cf. FIGURE 5.2) into eq. (5.10). In the structure under consideration, position B would get a more profitable deal with actors at position A than with the actor at position C because $p_{\mathrm{BA}} = 0.8071 > p_{\mathrm{BC}} = 0.6397$. And, due to $p_{\mathrm{CD}} = 0.6099 > p_{\mathrm{CB}} = 0.3603$, C would realize a higher profit share with D than with B. That is, the necessary condition for exclusionary behavior is fulfilled for both B and C.

We further have to examine whether the sufficient condition for exclusionary behavior is satisfied for at least one position as well. First, let us discuss the decision problem of C. If C exclusively exchanges with D, they will end up in a DYAD structure. Given C's exclusion of B, $p^*_{\mathrm{CD}} = 0.5$ thus represents C's relative bargaining power in the relation with D.[57] But if C keeps the link to B, C's profit level from the relation to D will be larger (because $p_{\mathrm{CD}} = 0.6099 > p^*_{\mathrm{CD}} = 0.5$). Hence, C's exclusion of B would

[57] Since both actors in a DYAD structure have a network control of 1, their relative bargaining powers are 0.5 for all relational types.

not pay and, consequently, C will not exclude B (i.e. the sufficient condition for C's exclusionary behavior towards B is not met).

The same line of reasoning now is applied to position B. If the latter always refrains from exchange with C, a 3–Line network will result—by excluding C, B thus can increase his relative bargaining power in the match with position A from $p_{\mathrm{BA}} = 0.8071$ to $p^*_{\mathrm{BA}} = 0.8276$. Since the necessary and the sufficient condition for B's exclusionary behavior are fulfilled, B's choice of exchange partners induces a deviation of bargaining and exchange structures. That is, the NCB model predicts that the single–link T–Shape network is not robust, but decays into two robust exchange structures (viz. the 3–Line A–B–A and the Dyad C–D). And, because of the invariance of profit rankings, this conclusion holds true for the double–link T–Shape as well.

Both predictions correspond closely with empirically observed exchange frequencies and profit point distributions. Markovsky, Willer, and Patton (1988, p. 227) report that in 100 negotiation rounds within a single–link T–Shape network only three exchanges occurred between B and C. Moreover, for a cake of 24 profit points, B realized, on average, 19.14 profit points in matches with A (NCB predicts 19.86) and C received 11.88 profit points in matches with D (NCB predicts 12.0). Skvoretz and Willer (1993, pp. 809–810) present similar results for a double–link T–Shape bargaining network. In their experiments, only four exchanges (out of a total of 307) involved both B and C. Here, the relevant mean profit points between B and A were 20.67 profit points in favor of B (NCB predicts 20.66) and 12.86 profit points for C in the relation with D (NCB predicts 12.0). Due to the corroborant evidence, the NCB model seems to be a useful tool for structural predictions which may be applied to other networks as well.

5.5.3.2 Structural Predictions and Observations

Figure 5.3 depicts negatively connected networks with equally valued relations for which the NCB approach predicts a deviation between bargaining and exchange structures in the presence of a systemwide one–exchange rule. Here, dotted lines indicate non–exchange between positions. Apart from H–Shape, the necessary part of the condition for exclusionary behavior is fulfilled for different positions in all structures. The sufficient part of the condition is met only for position B, however. That is, according to the NCB model, B's choice of partners is responsible for the deviation between bargaining and exchange structures.

In most of the displayed cases, the NCB model predicts that B dissolves the initial structure to end up in the profitable center of a 3–Line exchange network. There are only two exceptions: in the Stack 4–Line structure, B increases his payoffs by excluding C, but cannot realize an extremely beneficial situation for himself; in the VBox–XBox network, B further improves his profitable situation without splitting the initial network into separate exchange networks.

All negotiation structures are simple enough to become robust after one tie is removed. Of course, larger structures may first decay into non–robust subnetworks. With

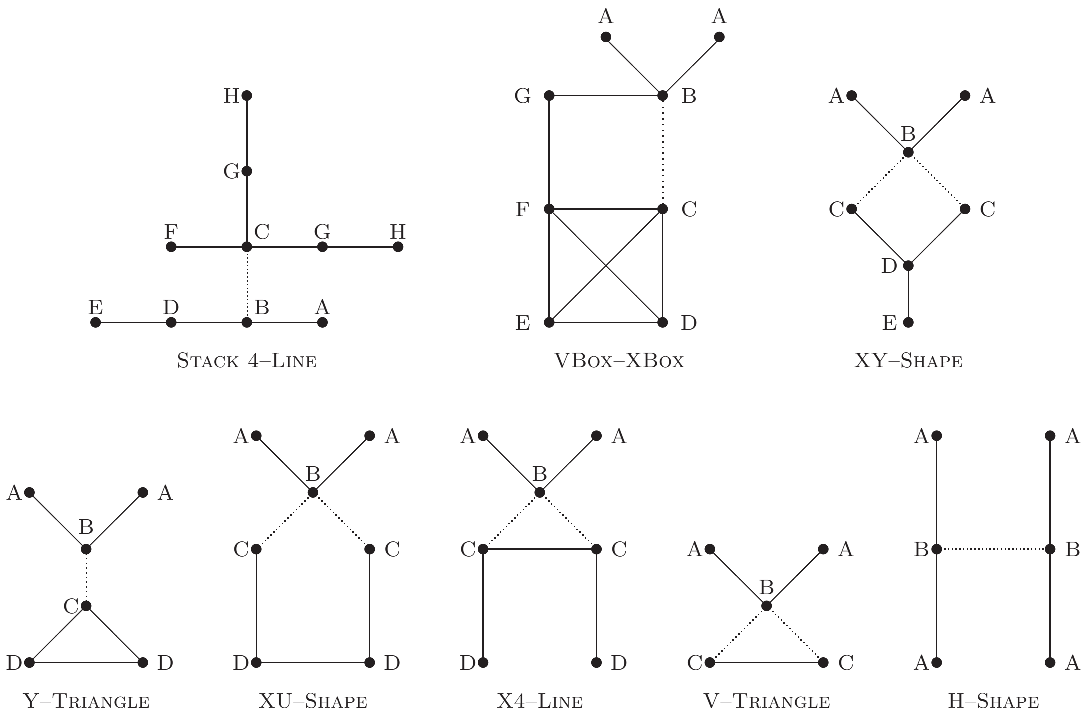


Figure 5.3 *Non–robust network structures with equally valued and negatively connected relations under a one–exchange rule (dotted lines indicate non–exchange).*

respect to large networks, it therefore may be necessary to iteratively apply the logic of the NCB model until robust exchange structures emerge.[58]

Before larger structures are examined, however, one should ask for the empirical relevance of the NCB–predictions in Figure 5.3. While there is no evidence for the three structures in the upper part of Figure 5.3, Simpson and Willer (1999) report relevant experimental results for the other five networks. Apart from one exception to be addressed shortly, Table 5.3 reproduces these findings—it contains data from Simpson and Willer's experiments for the first four networks displayed in the lower part of Figure 5.3. Specifically, Table 5.3 reports the relative frequencies or observed proportions of exchange (P) in B:C ties of those structures as well as the number of experimental rounds (N).

In addition, it informs about Π, the expected proportion of exchange when the subjects at positions B and C make independent random draws from their subsets of potential partners. Since Π is due to chance, it can serve as the baseline against which to assess whether the observed proportion of exchange (i.e. the relative frequency P) significantly differs from the NCB prediction. The test idea is simple: if a standard z–test for equality of proportions suggests that P is significantly greater than Π, then the NCB prediction (no exchanges in the B:C ties) can be rejected. The last column in Table 5.3 shows the test results. For each of the networks under consideration, the

[58]For larger networks the straightforward computations thus may become tedious. There is, however, a Microsoft Visual Basic program for NCB calculations which runs under any version of Microsoft Windows. The program is available from the author on request.

TABLE 5.3 *Proportions of exchange in B:C ties in specific negatively connected networks.*

	Experimental Data[a]		Random Choice	
	Rounds	Observed Proportion	Proportion[b]	
Network	N	P	Π	z–test[c]
V–TRIANGLE	450	.053	.125	−4.618*
X4–LINE	270	.031	.083	−3.110*
XU–SHAPE	310	.046	.125	−4.250*
Y–TRIANGLE	360	.094	.111	−1.027†

Notes: The Non–Cooperative Bargaining (NCB) model predicts the absence of exchanges in all of these ties. Simpson and Willer's (1999) Optimal Seek (OS) method predicts no exchanges in the B:C ties of the V–TRIANGLE and the X4–LINE networks, but exchanges between B and C in the XU–SHAPE and the Y–TRIANGLE structures.
[a] The experimental data (i.e. the number of rounds, N, and the relative frequency or observed proportion of exchange, P) are taken from Simpson and Willer (1999, p. 278 as well as Table 10.1 and Table 10.2).
[b] The proportion Π is calculated by assuming that each actor randomly chooses his exchange partner from available network partners (see also Markowsky et al. 1993; Skvoretz and Lovaglia 1995). More precisely, Π represents the expected proportion of exchange when the subjects at positions B and C make independent random draws from their subsets of potential partners.
[c] The expected proportion of exchange between randomly choosing actors, Π, serves as the baseline against which to assess whether the observed proportion of exchange (i.e. the relative frequency P) in a specific network is compatible with the prediction of no exchange. The test statistic is $z = (P - \Pi)/\sqrt{(\Pi\,(1 - \Pi)/N)}$. If P is significantly greater than Π, the prediction of no exchange in the B:C tie can be rejected.
* The observed relative frequency of exchange is significantly lower than the proportion of exchange which results if trading partners randomly choose one another ($p < .001$, one–tailed test).
† At the .05 level, the observed relative frequency of exchange does not significantly differ from the proportion of exchange which results if trading partners randomly choose one another ($p = .152$, one–tailed test).

NCB prediction cannot be falsified—the observed proportion P is never significantly greater than Π, the expected proportion of exchange if the actors would randomly select partners.

Interestingly, Simpson and Willer's (1999, Table 10.1 and Table 10.2) data analysis does not yield the same conclusions for all four networks under study. Contrary to our findings, Simpson and Willer report that the B:C ties in the XU–SHAPE and the Y–TRIANGLE networks are robust. To rationalize the different conclusions, it is necessary to briefly describe their data analysis. Simpson and Willer start from the null hypothesis of no exchange (i.e. $H_0 : E(P) = 0$, where E denotes the expectation operator with respect to the proportions of exchange). They conduct adjusted z–tests for proportions to falsify the assertion that the expected value falls within the confidence limits of the observed.[59] However, this procedure neglects important statistical facts—since z–tests for proportions are based on an approximation of a binomial distribution by a normal distribution, they only can be used if specific conditions are met:

[59]Simpson and Willer (1999, p. 284, note 10) face the problem that the test statistic of the standard z–test is not defined if the expected proportion vanishes (i.e. $z = (P - E(P))/\sqrt{(E(P)(1 - E(P))/N)}$ does not exist for $E(P) = 0$). To avoid division by zero, they treat the expectation as the observation and vice versa. Because of this adjustment, they argue that their test refers to the assertion that the expected falls within the confidence limits of the observed.

$NP \geq 5$, $N(1-P) \geq 5$, and $E(P)(1-E(P))N \geq 10$ (cf. Blalock 1972; Riedwyl 1992). Simpson and Willer's test procedure always violates at least one of these conditions. Therefore, their test results do not allow any valid conclusion about the robustness of ties.

The latter statement holds true for Simpson and Willer's analysis of the B:B relation in the H–Shape structure as well. In this case, however, our test method cannot be applied either. Because $N = 405$ and $P = 0.007$ are the empirical data for the B:B tie in this network (cf. Simpson and Willer 1999, Table 10.1), we have $NP = 2.835 \approx 3 < 5$ such that a z–test for the equality of the proportions P and Π is inappropriate. To check the fit between the NCB prediction and the empirical observation of B:B exchanges in the H–Shape structure, the theoretical probability for three or more exchanges has to be computed. Such a calculation gives a value of 0.008.[60] Since the latter corresponds closely to the empirically observed proportion of $P = 0.007$, the NCB prediction (no exchange in the B:B tie of the H–Shape network) cannot be rejected.

In sum, the data analyses suggest that the NCB predictions for the five networks displayed in the lower part of Figure 5.3 correspond with the empirical observations. This is important for an additional reason. In their article, Simpson and Willer (1999) suggest the so–called 'Optimal Seek' (OS) method for the prediction of deviations between bargaining and exchange networks. Their application of the OS method suggests, in accordance with the NCB predictions, that exchanges will be absent in the B:B tie of the H–Shape structure as well as the B:C ties of the V–Triangle and the X4–Line networks. In contrast to the NCB model, however, OS predicts exchanges between B and C in the XU–Shape and the Y–Triangle structures. That is, Simpson and Willer's theory implies that the XU–Shape and the Y–Triangle networks are robust structures. As was demonstrated above, however, appropriate statistical tests indicate the contrary. The empirical evidence therefore supports the NCB model over Simpson and Willer's OS approach.

In addition, the NCB model is comparatively simple to apply. And, as shown by the non–robust networks in the upper part of Figure 5.3, it encourages further experimental work. Like any other model, however, the NCB theory rests on several strong premises. It thus is reasonable to conclude with a brief discussion of the pros and cons of the model.

[60]From a statistical point of view, the NCB prediction suggests that exchanges between actors at position B are extremely rare events. The number of exchanges thus may be interpreted as a random variable Y with discrete values $y = 0, 1, 2, ...$ which approximately follows a Poisson distribution $f(Y = y) = (\lambda^y/y!)\, e^{-\lambda}$ with rate $\lambda = \pi N$, where e is the base of natural logarithm and π denotes the small probability of exchange. Suppose, in accordance with the NCB prediction, that $\pi = 0.001$. Then, $\lambda = 0.405$ and $f(Y \geq 3) = 0.008$.

5.6 Conclusion and Discussion

In this chapter, we have presented an alternative approach to the study of exchange structures. In accordance with other theories, the latter reflects the idea that rational actors take advantage of their structural positions in negotiations. Contrary to other sociological models, however, we have combined Rubinstein's (1982) game–theoretic work on non–cooperative bargaining and its refinement (Binmore 1985, 1987, 1998) with the assumption that both relational features and network positions affect exchange outcomes. More precisely, negotiated exchange was conceptualized as a one–shot Alternating Offers Game in which the actors' levels of structural autonomy or degrees of network control determine their bargaining powers. And, since we have differentiated with Cook et al. (1983) between negatively and positively connected relations, the approach refers to situations in which one exchange relation prevents and/or promotes another.

The resulting model uniquely predicts, via an equilibrium analysis in the sense of non–cooperative game theory, power and profit distributions for all relational configurations when, in addition to the surpluses to be divided, the actors' network ties and the types of their relations are fixed. It identifies, moreover, coincidences or deviations between exchange and negotiation structures. That is, the robustness or decay of a structure results as the product of the interdependent choices of position holders in a network. In accordance with Cook and Whitmeyer's (1992) suggestion, the model thus takes a modest first step in using exchange theory to explain network structure.

At the same time, the model has desirable properties. First, it is as parsimonious as established ones, but simpler to apply. Second, the approach precisely specifies how structural aspects influence choices. Third, it links sociological exchange ideas to the game–theoretic bargaining literature and, by doing so, takes into account that interactive choices characterize negotiations. Fourth, its predictions correspond with optimizing behavior of bargaining partners such that agreements between selfish rational actors are not assumed, but explained. Fifth, it allows for variations in the value of relationships (i.e. heterogeneity in terms of the size of the cake to be partitioned). Sixth, while broader in explanatory scope, the model predicts experimental findings with respect to the effects of network embeddedness for exchange outcomes at least as well as the best fitting established theories.

It has to be emphasized, however, that the approach rests on strong premises. The latter ensure, on the one hand, a close correspondence between the theoretical model and the artificial conditions in laboratory research on exchange networks. They prevent, on the other hand, that the current model captures all relevant aspects of real negotiation and exchange systems. It thus is reasonable to discuss briefly important restrictions and desirable modifications of the approach.

A first shortcoming of the model is its focus on the effects of network positions and relational assessments for exchange outcomes. More specifically, position holders are assumed to be homogenous up to their relational categorization. A similar postulate characterizes most sociological work on exchange networks. However, interindividual heterogeneity with respect to, say, age or gender may matter for negotiation results as well (e.g. Eckel and Grossman 1998). A more realistic model version therefore would start from the idea that bargaining powers depend on both individual and structural variables.

A second limitation of the approach is its behavioral postulate. According to the model, each actor is a selfish profit maximizer. This postulate fits the typical design of experiments on exchange networks. However, it excludes fairness considerations and/or effects of relative standing in terms of profits. According to experimental research from different disciplines, such additional motivations seem to play an important role in specific settings (see, for a review, Rabin 1998). The recent progress in behavioral game theory (e.g. Bolton and Ockenfels 2000; Camerer 1997) may be a helpful starting point for their incorporation into the analysis of exchange networks.

A third shortcoming of the model is its neglect of potential side–effects of ongoing exchanges in dyads. While this feature is typical to many theories of network exchange and their laboratory tests, it is surely not a realistic one. According to Lawler and Yoon's (1996) experiments, for instance, people who engage in repeated exchanges may form attachments and make commitments. Such an increase in relational cohesion can stabilize established network relations and, possibly, change long run outcomes. These aspects of durable exchange relations deserve attention because of their importance in everyday life. Their analysis requires, from a theoretical perspective, a dynamic modeling approach. For example, one could follow Muthoo (1999) who studies, among other things, an infinitely repeated bargaining game and its differences to the static bargaining scenario.

A final shortcoming of the model is its neglect of network extensions. While it predicts whether or not a given structure reduces to one or more subnetworks, it ignores the potential growth of networks. For the analysis of positively connected relations, however, the issue of network extension may be as important as exclusion is in the context of negatively connected relations. And, as emphasized by Yamagishi, Gillmore, and Cook (1988), real exchange systems often are mixtures of positive and negative connections. Therefore, an appropriate theory of exchange networks has to address the issue of structural growth. To deal with network extensions, the model may be generalized such that actual system members at distinct positions have, for example, random contacts with potential bargaining partners and decide, depending on their relational assessments and profit expectations, whether they accept or reject the newcomers. Such a model generalization is a worthwhile task for future work.

Chapter 6

Summary and Conclusions

In this book, we examine how a surplus is provided and divided between two actors. A surplus is an 'additional something' which, if produced, can be divided between two actors in some way with the division leaving both actors better off than if they had not produced and divided the surplus. The book addresses the provision and division of a surplus based on two scenarios. In Chapter 1, we introduce these scenarios. The first part of this book, Chapters 2 through 4, analyzes the provision and division of a surplus in the context of *delayed exchange*. This scenario can be seen as a *trust problem* between two actors (e.g. Coleman 1990). The main focus here is on whether trust will be placed, and whether the provision of valuable resources for the production of a surplus will occur. More specifically, we study *effects of temporal embeddedness on the provision of a surplus*. Section 6.1 stipulates the reasons for incorporating temporal embeddedness into the analysis of trust and delayed exchange. This section also discusses the theoretical and empirical results of Chapters 2 to 4 and makes some suggestions for further research.

The second part of the book, Chapter 5, focuses on the division of a given surplus rather than on the provision. The scenario involves two actors who determine their shares of the surplus by bargaining where a *delay in agreement* on the division of the surplus is costly for them. The bargaining scenario corresponds with the *Alternating Offers Game* introduced by Rubinstein (1982). The question is who gets how much of the surplus if its division is negotiated. The costs for the actors caused by a delay in agreement determine the division. Though the costs are exogenous in Rubinstein's game, we try to derive them endogenously. We argue that the costs can be determined if the focus is not on a single dyad but on a *negotiation network* in which dyadic negotiations on the division of a surplus are possible. An actor's costs of delaying an agreement are a function of the actor's position in such a negotiation network. Section 6.2 sheds some light on the theoretical background of bargaining and network embeddedness. We summarize the experimental tests on network bargaining with delayed agreement and conclude with suggestions for further research.

6.1 Delayed Exchange: The Provision of a Surplus

Consider the basic scenario of delayed exchange with a trustor who has valuable resources to invest in the production of a surplus by the trustee. A trustee, once he receives the resources, can either use them for his own benefit or for the mutual benefit of the trustor and trustee. In other words, he can divide the surplus produced by the trustor's investment of resources in two ways: either keep the whole surplus or divide it in such a way that both actors benefit. If the trustor has no right of co–determination in the division of the surplus, she must anticipate how the trustee will divide the surplus when deciding on whether to make an investment. The provision and division of a surplus can then be seen as a trust problem in the sense of Coleman (1990: 97–99). In this book, we conceptualize this type of trust problem as a Trust Game (Dasgupta 1988; Kreps 1990a).

In our analysis of the basic scenario of delayed exchange, we make three crucial assumptions. First, an actor's utility is determined by the actor's share of the cake and an actor prefers a larger share of the cake to a smaller one. Second, actors behave according to game theoretic rationality or at least follow some other principle of incentive guided behavior.[1] Third, an encounter between two actors is isolated in the sense that it is not connected to any previous or future exchanges or interactions between the trustor and trustee, nor to any previous or future interactions involving other actors. Hence, we consider the basic scenario of a one–shot Trust Game. The prediction then is that trust is not placed, and if it were placed, it would be abused. To see this, assume the trustee has received the trustor's resources. Given our first assumption, he would rather use them to his own benefit than to the mutual benefit of himself and the trustor. Given our second assumption, the trustor can anticipate the trustee's behavior. She will thus not provide the resources, and, consequently, the surplus will not be provided. Both actors are then worse off than if the resources are provided (trust is placed) and the surplus is divided in such a way that both actors benefit (trust is honored).

Empirical research and everyday observation show that the Pareto improvement with trust placed and honored, that is, a surplus is provided and divided so that both actors benefit, sometimes occurs. In the first part of this book, we investigate how this can be explained. Several approaches are possible. One could, for instance, relax the assumption about 'utility is own share of the cake' and incorporate fairness, relative payoffs or other 'soft' aspects into the actors' utility functions (see, for example, Bolton and Ockenfels 1998, 2000; Camerer 1997; Kelley and Thibaut 1978; Nelson 2001; Rabin 1993, 1998; Snijders 1996). Theorists such as Emerson (1972, 1981) argue that a more sociological approach would focus on variations in social conditions rather than utility. In this book, we choose this more sociological approach and rather than primarily

[1] Game theoretic rationality is basically conceptualized as Nash equilibrium behavior (Nash 1951). This means roughly that each actor maximizes utility, given the other actor's behavior. See, for instance, Rasmusen (1994: 22–28) for technical details.

challenge assumptions about utility, we investigate the role of social conditions. We therefore relax the third assumption but adhere as far as possible to the first and the second assumption. We consider situations where two actors repeatedly meet over time and exchange between them is embedded in a temporal sense (Granovetter 1985; Raub and Weesie 2000b). 'Temporal embeddedness' thus refers to a sequence of exchanges between the two actors. In terms of the abstract Trust Game, this means the one–shot game is repeated a finite or infinite number of times. Consequently, the general research question of the first part of the book is: *Given that the provision and division of a surplus is conceptualized as a situation of trust between two actors, which aspects of temporal embeddedness affect the placement of trust and how (i.e. positively or negatively) do these aspects affect the placement of trust?*

Trust: History Effects

Chapter 1 presents our reasons to focus on temporal embeddedness in explaining the provision of a surplus. One element of temporal embeddedness, namely 'the shadow of the future', has been relatively well studied in the literature (e.g. Axelrod 1984; Maynard Smith 1982; Taylor 1987; Trivers 1971). An actor's current behavior has been shown to be dependent on the anticipated effects of this behavior in future interactions. **Chapter 2** takes this result as given. We then examine whether the 'shadow of the past' additionally influences an actor's current behavior. These past effects have been less well studied in the literature. However, if we advocate temporal embeddedness as a possible 'solution' to the provision and division of a surplus, we need to know whether and how the past affects an actor's current behavior. To begin with, we choose to study the effects of the past on decision making in social dilemma situations. A social dilemma is a situation where individual and collective interests clash and rational and selfish actors end up in an inefficient (Pareto inferior) equilibrium. The Trust Game and the Prisoner's Dilemma Game are examples of dilemma situations. The focus of Chapter 2 is on *the comparison of two situations* that are identical in all but one aspect: one situation has a past, or history and the other does not. In all the other aspects, that is, alternatives and payoffs, the situations are the same. Moreover, at least one of the situations is a social dilemma.

We present a simple theory based on the theory of games with complete information that predicts no difference in behavior between the 'no history' and the 'with history' situation. The argument is straightforward. If actors behave in accordance with the behavioral assumptions presented above, the history does not contain any information to induce actors to make different decisions than in the 'no history' situation. Their behavior should consequently be the same in the 'no history' as in the 'with history' situation.

More formally, we first introduce a situation referred to as Game A and then present a second, more complicated situation referred to as Game B. This latter game resembles the situation captured by Game A but is characterized by preceding actions of the actors of Game A or other actors. These preceding actions, called history, extend Game A in such a way that Game A is a subgame of the larger Game B.[2] This subgame is called Game A_B (read 'Game A in B'). Since our simple theory requires that Game A_B and Game A are identical, the 'preceding actions' in Game B only extend Game A in two ways. First, Game B may be a repeated version of Game A. Second, for at least one actor, the 'preceding actions' are not available in Game A. Consequently, Game B becomes a game where Game A_B is one of several possible subgames. Predictions that behavior will not differ in Game A and Game A_B are confronted with findings from three different experiments.

First, we compare a variant of a Dictator Game (Game A) with a Trust Game (Game B). In both games, there are two actors, ego and alter. In the variant of the Dictator Game, alter can split a surplus equally between himself and ego or keep the surplus himself. Ego is a dummy player who does not move in the Dictator Game. In the Trust Game, ego starts by choosing between withholding or placing trust. If ego withholds trust, the game ends with certain payoffs to both actors. However, if ego places trust, a subgame Game A_B starts that is identical to the variant of the Dictator Game. If alter behaves in accordance with the above assumptions, his behavior should be equal in Game A_B and Game A, even though it is ego's move in Game B that gives alter the chance to make a decision in Game A_B. However, experimental findings (Snijders and Keren 1997) contradict our theory. Experimental subjects playing as alter in Game B divide the surplus equally in 73.5% of the cases. In other words, they mainly honor placed trust. In Game A, however, only 35.5% of the alters choose an equal distribution of the surplus.

Second, we focus on ego's behavior in a Trust Game (Game A) compared to his behavior in a Hostage Trust Game (Game B). In the Hostage Trust Game, alter decides whether to post a hostage prior to ego's decision on whether to place trust. Posting a hostage makes it less favorable for alter to keep the surplus rather than split it equally.[3] This in turn makes placing trust safer for ego. For our study of history effects, we focus on situations where alter does not post the hostage. Ego and alter then play a subgame Game A_B which is identical to Game A, namely the Trust Game. Experimental evidence (Snijders 1996) shows that alter's decision not to post a hostage decreases ego's probability of placing place trust by 0.10 compared to her probability of placing trust in Game A, averaged over different game payoffs.

Third, we compare ego's behavior in a Prisoner's Dilemma Game (Game A) with his behavior in a Hostage Prisoner's Dilemma Game (Game B). In both games, ego and alter simultaneously decide between cooperation or defection. Prior to this decision,

---

[2]Loosely speaking, a subgame is the part of the entire game that follows upon the history played so far. For a formal definition of a subgame see, for example, Rasmusen (1994: 94).

[3]A hostage is of value to the trustee and by abusing trust if placed, the trustee loses the hostage, consequently decreasing the gain from keeping the whole surplus. See, for example, Schelling (1960) for a discussion of hostages in social interactions, Raub and Keren (1993) for a theoretical analysis and prior experimental evidence, and Raub and Weesie (2000a) and Weesie and Raub (1996) for a general discussion on hostages and incentive problems.

the latter game allows ego and alter to post a hostage, which later makes defection an unattractive move. If both ego and alter refrain from posting a hostage, they reach the subgame Game A_B identical to Game A, the Prisoner's Dilemma Game. Experimental evidence from Mlicki (1996) shows that defection is chosen more often in Game A_B than in Game A: namely about 91% versus 74%.

The analysis of these three different experiments shows that behavior in the 'no history' situation differs from behavior in the 'with history' situation. We conclude from the analysis in Chapter 2 that *history does indeed matter*. A theory on behavior that does not take the history of a game into account obviously leads to inadequate predictions about behavior. We suggest two main elaborations on the theory presented in Chapter 2. First, we address how to explain the experimental findings using more complex utility arguments that relax the assumption on 'utility is own share of the cake'. Since it is not our primary aim to challenge assumptions about utility and we would much rather explain variations in behavior by variations in social situations, we concentrate on the *more complex information structure* that evolves from the game's history. In other words, the history of Game B is assumed to disclose information to an actor in Game A_B. According to our simple theory however, this information should not be relevant. Experimental subjects, on the other hand, seem to behave as if they attach importance to such information and adapt their behavior accordingly.[4] Chapter 3 builds upon the finding that *history matters*. We present a model in which actors learn from past information and, additionally, take their expected future into account to determine their behavior in a current situation.

Trust and Dyadic Social Capital: Theory

Chapter 3 addresses the effects of the past and future on behavior alike. We explore which effects of temporal embeddedness facilitate or prevent the provision of a surplus. The provision of a surplus is once again conceptualized as a Trust Game. By repeating the Trust Game, we can systematically study which aspects of the past and future affect the provision of a surplus. Our hypotheses on the trustor's behavior are based on a simple idea. Whether a trustor invests in the production of a surplus depends on her *amount of social capital in relation to the trustee*. We assume that social capital accumulates through past interactions with the trustee (i.e. learning) but also reflects the mutual future of their relation (i.e. control). The trustor's behavior is then supposed to follow a simple *behavioral guideline*: the higher the stakes of the current game, the more social capital is needed to induce the trustor to provide the necessary resources for the production of a surplus. The stakes of the game give an indication of how much the trustor can win or lose by placing trust, and of what the trustee can win by abusing placed trust. We use the payoff ratios $\text{Risk} := \frac{P_1 - S_1}{R_1 - S_1}$

[4] An example may help illustrate this. Ego's decision in a Trust Game where trust has been placed resembles a Dictator Game (see first experiment described above). Since alter's move in the Trust Game leaves ego's alternatives and payoffs unchanged, ego's behavior should not differ from his behavior in a Dictator Game. Obviously, however, alter's decision to place trust induces ego to 'deviate' from his Dictator Game behavior.

and TEMPTATION := $\frac{T_2-R_2}{T_2-S_1}$ to measure the stakes of a game.[5] RISK can be derived from a Subjective Expected Utility (SEU) consideration (Snijders 1996: ch. 2; Coleman 1990: 97–102) and TEMPTATION follows from a model of guilt (Snijders 1996: ch. 2). Snijders presents experimental evidence from one–shot Trust Games that the trustor's probability of placing trust decreases in RISK and TEMPTATION.

The outcomes of past interactions, that is, the realized division of a surplus, provide information which allow for *learning*. Information from past experience can either be negative or positive. Information is negative if the trustee 'confiscated' the full surplus. Information is positive if the surplus is divided in such a way that the trustor and the trustee both benefit. We assume that the capital stock decreases in negative outcomes but increases in positive outcomes. How much the capital stock shrinks or grows is assumed to depend on the situation in the past, that is, how high or low the stakes of the past game were. Assume, for instance, that the trustee keeps the entire surplus and abuses trust. The larger the surplus, the larger the temptation for the trustee and the more the trustor can expect the trustee to keep all the surplus. An outcome where the trustee keeps a large surplus is thus less surprising to the trustor. Consequently, social capital is assumed to decline, but less than in a situation where the trustee keeps a small surplus. A trustee who keeps a small surplus is thus considered less trustworthy than a trustee who keeps a large surplus. In the same vein, dividing a large surplus contributes more to an increase in the trustor's social capital than dividing a small surplus. Moreover, we assume that a trustor considers recent outcomes on the division of a surplus more important than those experienced further back in the past: the trustor discounts past outcomes.

The second aspect of temporal embeddedness is the mutual future of the trustor and trustee. The future enables the trustor to *reciprocate* the trustee's current behavior. Successful dealings with the trustee can be rewarded by further providing the necessary resources for the production of more surpluses. Unsuccessful dealings, on the other hand, can be penalized by the failure to provide further resources for one or more future interactions. The trustee would then forfeit possible shares of future surpluses. The longer the mutual future, the larger the trustor's social capital since there will be more opportunities to reciprocate the trustee's behavior. This is called the *control effect* of the future. The effectiveness of this control effect depends on the size of the possible future surpluses. For example, eventually forfeiting a large surplus is a harsher penalization for the trustee than forfeiting a small surplus. So, the more large surpluses may materialize in the future, the larger the trustor's current capital stock.

The social capital theory developed in Chapter 3 proposes conditions for trust, given temporal embeddedness. A real–world application of how actors deal with issues of trust is illustrated in the procurement management of firms. **Chapter 4** therefore

[5]In the Trust Game, if the trustor withholds trust, she and the trustee earn P_1 and P_2 respectively. If placed trust is honored, the trustor and trustee earn R_1 and R_2, respectively. If placed trust is abused, the trustor receives S_1 and the trustee T_2. Payoffs are ordered $R_1 > P_1 > S_1$ and $T_2 > R_2 > P_2$.

addresses relations between buyers and suppliers. It is assumed that their transactions can be seen as a (variant of a) Trust Game. In general, economic transactions involve some kind of time delay between the promise to deliver and the actual delivery. This time delay leaves room for all kinds of things going wrong that are beyond the buyer's control (problem potential). For example, goods may be delivered late, broken, or not all. For certain transactions, this problem potential is high, for instance, if the transaction volume is large or the impact on the buyer's performance is enormous if delivery is late. A buyer who is willing to accept uncertainties hands over control to the supplier and thus commits an act of trust. Chapter 4 assumes that buyer firms seek to decrease the necessary trust by investing time and effort in *ex ante management* of the transaction. For example, they invest time and effort in negotiating contractual arrangements. More thorough transaction management indicates a lack of trust in the supplier. On the other hand, Macaulay (1963) notes that buyers and suppliers often reduce their contractual investments in favor of trust. If the buyer deals repeatedly with the same supplier (temporal embeddedness), the buyer's investment in ex ante management is not only determined by the problem potential of the current transaction, it is also shaped by their mutual past and expected future.[6] Chapter 4 suggests that the social capital theory on trust can also be used to predict the buyer's investment in ex ante management of the focal transaction. Chapter 4 consequently uses variants of the hypotheses from Chapter 3 for the specific case of buyer–supplier relations.

Trust and Dyadic Social Capital: Empirical Results

First and foremost, the hypotheses on the trustor's behavior in a temporal setting are tested using laboratory experiments (Chapter 3). The experimental subjects were mainly freshmen, following an introductory course in sociology at the Ludwig–Maximilans University in Munich, Germany (see Appendix A for the experiment booklet). The subjects were confronted with several repeated Trust Games and had to make decisions in the role of the trustor: they thus had to decide whether to place trust. The experiment recorded the subjects' decisions using the strategy method (see Roth 1995). At any given moment in the game, a subject was asked to make decisions, one for each history possible up to the current game. This method made it possible to collect information on a subject's behavior in different settings of temporal embeddedness.

Second, we want to know whether our hypotheses on trust in a temporal setting can withstand a test outside the sociological laboratory. Chapter 4 uses data on buyer–supplier relations in the Dutch IT–industry ('The External Management of

[6] The problem potential of a transaction can be linked to payoffs or payoff differences in the Trust Game. The problem potential of a transaction is operationalized via the financial volume of the transaction and the supplier's opportunism potential inherent to the transaction. The financial volume serves as an overall measure of the buyer's loss if trust is abused (i.e. $P_1 - S_1$ versus $R_1 - S_1$) and as a measure of the supplier's incentive to abuse trust after it is given (i.e. $T_2 - R_2$). The opportunism potential is derived as a weighted average of the damage potential and the buyer's monitoring problems problems (as well as the financial volume of the transaction). The damage potential covers 'how bad it would be' if it turns out that the delivered product cannot be used productively and directly reflects the S_1–value (see Section 4.4.3).

TABLE 6.1 *Hypotheses on trust in temporally embedded relations.*

Independent Variable Experiment		Survey		Results[a] Exp	Sur
Current stakes					
H3.1 $^{-}$	Stakes	*H4.1* $^{+}$	Volume and opportunism potential	s	s
Past					
H3.2.1 $^{+}$	Honored trust	*H4.2.1* $^{-}$	Successful past transactions	s	s
H3.2.2 $^{-}$	Abused trust	*H4.2.2* $^{+}$	Unsuccessful past transactions	s	nt
H3.3.1 $^{+}$	Honored trust × stakes	*H4.3.1* $^{-}$	Successful past transaction × past volume and opportunism potential	ns	ns
H3.3.2 $^{-}$	Abused trust × stakes	*H4.3.2* $^{-}$	Unsuccessful past transaction × past volume and opportunism potential	ns	nt
H3.4 $^{-}$	Having a past × discount past outcomes	*H4.4* $^{+}$	Having a past × discount past outcomes	s	nt
H3.5 ^{0}	No past vs. trust withheld	*H4.5* ^{0}	Potential vs. actual supplier	s	nt
Future					
H3.6 $^{+}$	Length of future	*H4.6* $^{-}$	Future transactions	s	ns
H3.7 $^{+}$	Length of future × stakes	*H4.7* $^{-}$	Future transactions × future volume and opportunism potential	ns	nt

Note: The first column gives the numbering of the hypotheses used in Chapter 3. The signs refer to the proposed effect on the probability of trust being placed: increase (+), decrease (−) and no effect (0). The third column gives the numbering of the hypotheses used in Chapter 4. The signs refer to the proposed effect on the buyer's ex ante investment in management.

[a] Experimental results (Exp) refer to the analyses in Chapter 3 and survey results (Sur) refer to analyses in Chapter 4. s : hypothesis is supported. ns : hypothesis is not supported. nt : hypothesis is not tested.

Automation', MAT, see Buskens and Batenburg 2000). As outlined above, the buyer's investment in ex ante management of the transaction can be seen as a proxy for the buyer's lack of trust in the supplier. The data contain enough information to calculate the amount of time and effort the buyer invested in ex ante management of a focal transaction.

TABLE 6.1 gives an overview of our hypotheses and the results of the experimental and survey tests. The last two columns of TABLE 6.1 show that the findings from our two tests are quite similar. The hypothesis on a negative effect of increased risk in a current transaction (*Hypothesis 3.1* and *Hypothesis 4.1*, respectively) is supported in both data sets. Experimentally, it holds true that an increase in the stakes of the current game (i.e. RISK and TEMPTATION) decreases the trustor's probability of placing trust. Likewise, the survey data show that an increase in the volume of the current transaction as well as in the opportunism potential of the current transaction increases the buyer's investment in ex ante management. A risky current transaction thus reduces the buyer's willingness to trust the supplier and induces more investments in ex ante management of the current transaction.

There is also evidence confirming the proposed learning effect of the past in the experimental as well as the survey data. The experimental data show that the trustor's probability of placing trust in the current game increases if there was a successful past where placed trust was honored by the trustee (*Hypothesis 3.2.1*). There is also evidence of a negative effect on the trustor's probability of placing trust after placed trust was abused in the past (*Hypothesis 3.2.2*). Unfortunately the survey data do not make it possible to distinguish between positive and negative past dealings. Buyers depicted in the data set as having several transactions with the same supplier generally report successful past transactions. Negative past experience can be assumed to have led buyers to look for new suppliers. The survey data thus only allow for a direct test of *Hypothesis 4.2.1* but not of *Hypothesis 4.2.2*. Having a good past with the current supplier decreases the buyer's investment in management and thus supports *Hypothesis 4.2.1*. For those buyer–supplier dyads with a past, we further examine whether the frequency of past transactions additionally decreases the buyer's investment in safeguarding a focal transaction. The data do indeed support this proposition. Buyers who reported to have a mutual past with the current supplier significantly decreased their investment in ex ante management the more frequent their mutual transactions were (see Table 4.6 in Chapter 4).

Due to the absence of the necessary variables in the survey data, two hypotheses on the effects of the past can only be tested on the experimental data. First, we suggest that past outcomes are discounted (*Hypothesis 3.4*). Evidence to support this hypothesis is found in the experimental data. Second, we assume that the trustor's probability of placing trust does, ceteris paribus, not depend on whether trust was withheld in a past game or whether this game was not played at all (*Hypothesis 3.5*). The experimental data offer some, though not conclusive evidence that there was indeed no difference between the trustor's behavior in situations of 'no past' and situations of a past of withheld trust. In terms of the buyer–supplier data, this means that as long as the buyer has not interacted with the supplier (and has no information about the supplier's behavior in encounters with others), he cannot learn anything about the supplier. The survey data do not make it possible to test this hypothesis. The MAT–data do not contain information on whether the supplier of the focal transaction was a potential supplier, be it not an actual one, in the previous 'time period'.

Both data sets, however, make it possible to test the hypothesis about the control effect of the future. *Hypothesis 3.6* suggests that the longer the expected future, the greater the likelihood of a trustor placing trust. The experimental data support this control effect. The buyer–supplier data do not demonstrate an effect of the future. However, this result may reflect a substantive difference between the experiment discussed in Chapter 3 and the buyer–supplier relations from Chapter 4. Experimentally, we are testing the 'true' control effect that works via the tit–for–tat mechanism of reciprocity. In the buyer's relation with the supplier, more opportunities for reciprocity should likewise lead to less investment in management. However, the buyer's procurement management may also depend on what we call the investment effect. Such an effect should lead to more investment the longer the expected future. The idea is that investment in the management of the focal transaction can be re–used in safeguarding future transactions. The control effect may have been counteracted by the investment effect in such a way that the MAT–data no longer demonstrate an effect of expected future transactions on the ex ante management of the focal transaction.

Lastly, there are two hypotheses that are not supported by either of the data sets. On the one hand, we propose an interaction between past outcomes and the stakes of the respective past game. On the other hand, we suggest an interaction between the length of the expected future and the stakes of the future games. The experimental data do not show any interaction between positive past outcomes and TEMPTATION (*Hypothesis 3.3.1*), nor between negative past outcomes and TEMPTATION (*Hypothesis 3.3.2*). Likewise, the past volume and opportunism potential do not show a significant interaction with successful past transactions (*Hypothesis 4.3.1*).[7] There is similarly nothing in the experimental data to support the proposed interaction between the length of the expected future and the TEMPTATION in these future games (*Hypothesis 3.7*). Since the field data do not contain the necessary information to construct future stakes, *Hypothesis 4.7* cannot be tested.

In sum, the empirical test of our hypotheses suggests that whether a trustor or buyer places trust is theoretically best described as a function of three distinct characteristics of temporal embeddedness: the stakes of the current game, the discounted past outcomes, and the length of the future. A social capital stock that incorporates these three characteristics is referred to as the *simple social capital stock* in Chapter 3. The so–called complex social capital stock also takes past and future stakes into account. Our data indicate that the simple social capital model suffices to account for the observed behavior in experimental Trust Games and IT–transactions.

TRUST: SOME SUGGESTIONS FOR FURTHER RESEARCH

Neither the experimental nor the empirical data provide evidence of the importance of past and future stakes. As we discuss at greater length in Chapter 4, the survey data do not contain the necessary variables to properly construct past and future stakes (i.e. past and future monetary volume and opportunism potential). The experimental data do however allow for a clear and straightforward conceptualization of past and future stakes via RISK and TEMPTATION. How come we do not find an effect of past and future stakes? Bear in mind that only two different payoff structures were considered in the experiments. In planning the experiments, there were good reasons to confine ourselves to only two payoff structures (see Section 3.4.1). Further research, however, should study the effect of monetary stakes on behavior in repeated Trust Games more thoroughly. This requires setting up experiments using repeated Trust Games with nu-

[7] Bear in mind that the MAT–data do not contain information on unsuccessful past dealings with the current supplier, so *Hypothesis 4.3.2* can not be tested.

merous different Risk and Temptation ratios. If we deal with more than just low and high stakes, however, the possible payoff combinations increase tremendously.[8] Even though one need not use every payoff combination, an experimental design with a larger variance in the games' stakes consequently requires a larger number of experimental subjects.

Second, in formulating the capital stock model, past outcomes are assumed to be discounted. As is noted above, an assumption like this has some plausible consequences, for instance, that an abuse of trust can be forgiven or that even a functional (trustful) relation needs certain efforts to be made to maintain its good state. As the analyses in Chapter 3 show, we find evidence that past outcomes are discounted.[9] It is now tempting to generalize this finding to more than two time periods and propose a social capital model that assumes *exponential discounting* of past outcomes. A model with exponential discounting is indeed just one option. We know that forward–looking rational choice models generally assume time–consistent behavior that implies exponential discounting (e.g. Becker 1996). On the other hand, experimental as well as survey evidence (e.g. Braun 1998: ch. 4; Fehr and Zych 1998; Loewenstein and Elster 1992) suggests that time–inconsistent behavior is sometimes observed. This kind of behavior is predicted by hyperbolic discounting models. Would these empirical findings therefore suggest that the past is also discounted hyperbolically instead of exponentially? The experimental results of Chapter 3 are based on at most two repetitions of the constituent Trust Game. A conclusive answer on how the past is discounted cannot be given based on the experimental data at hand. The question, however, is an interesting one and it would be worthwhile to repeat the study using Trust Games that are repeated several more times. Of course, 'several more times' is vague. It is hard to stipulate a precise number of repetitions of the constituent Trust Game that would allow for a satisfactory test of different hypotheses on discounting. A Trust Game repeated four times would probably be a good start, though of course repeating the game ten times would be much better. A useful design for simultaneously addressing the effects of the stakes of past and future games as well as the discounting issue would be a four times repeated Trust Games with more than only two different payoff structures.

One final comment pertains to the deduction of the hypotheses in Chapter 3 and Chapter 4. In both chapters the trustor's and buyer's behavior, respectively, was assumed to be dependent on a social capital stock. As the empirical findings show, this assumption is largely confirmed. However, we arrived at the hypotheses basically by using *verbal theory*. Consequently, the model in Chapter 3 does not allow for point predictions on the trustor's behavior, such as predictions on precisely how much the capital stock grows after a positive outcome in a game of high stakes represented by

[8]Remember that we already had $2 \times 2 \times 4 \times 8 = 128$ payoff combinations using low and high stakes in combination with two one–shot games, a once repeated game, and a twice repeated game.

[9]To distill a discount parameter, we propose modeling the effect of the past as the sum of the effect of the 'outcome at time $t-1$' and the discounted effect of the effect of the 'outcome at time $t-2$' (see page 77).

a Temptation of say, 0.85. A formal model might also make it possible to predict the relative effect of learning versus control. In other words, one might be able to say whether the past or the future is more important in a given situation. Finally, such a model should conceptualize the trustor's decision on whether to place trust as a consequence of maximizing behavior given the specific temporal embeddedness.

6.2 Delayed Agreement and Bargaining: The Division of a Surplus

The second part of this book, **Chapter 5**, addresses the division of a surplus as an Alternating Offers Game in which two actors negotiate their shares of the surplus (Rubinstein 1982). The game assumes that ego and alter successively make offers and counteroffers on the division of a surplus until they reach an agreement. Whenever ego makes an offer, alter can either accept it or reject it. If she rejects it, she can make a counteroffer one time period later. Ego can then accept or reject alter's offer and proceed, again one time period later, to make a new counteroffer and so on and so forth. However, delaying the agreement until later time periods is costly to the actors. These costs are captured by the actors' discount factors that give an indication of the actors' patience in delaying the agreement on the division of the surplus.

As in the first part of the book, certain assumptions are made on the behavior of ego and alter. First, the actors are assumed to not have to worry about the provision of the surplus. The surplus to be divided is available to them beforehand. Second, an actor's utility is once again assumed to be determined by the actor's share of the cake and an actor is assumed to prefer a larger share of the cake to a smaller one. Finally, actors are assumed to behave according to game theoretic rationality in the sense of Nash equilibrium behavior.

In line with these assumptions, Rubinstein (1982) shows that in equilibrium, the division of the surplus depends on the ratio of the actors' discount factors only. Put differently, in equilibrium the surplus is divided according to the actors' patience in delaying their agreement on the division of the surplus. An actor who has lower costs is more patient and can consequently wait longer for the surplus to be divided. He or she then gets the larger part of the surplus. We show in Chapter 5 how being patient translates into an actor's absolute *bargaining power*.

The actors' costs of delaying an agreement, and thus their absolute bargaining powers, are exogenous in Rubinstein's model. To make predictions on the division of the surplus and test them using experimental evidence, we need information on the actors' bargaining powers. We enrich the scenario of dyadic bargaining with a sociologically relevant extension from which the actors' bargaining powers can be inferred. In Chapter 5, we follow Emerson (1972, 1981) and assume that exchange between two actors should not be considered independently of their relations to other actors. Dyadic bargaining is affected by the actors' embeddedness in a larger network of other bargaining partners. Following Emerson, we propose that ego's and alter's network embeddedness can be seen as the source of their bargaining powers. From an actor's point of view, networks determine the alternatives for dyadic exchange. Consequently, networks define differences in opportunity structures between actors that are then reflected in the actors' bargaining powers. The main research question of the second part of the book is: *How does an actor's embeddedness in a network of negotiation partners determine his or her share of the surplus in relations with his or her partners?*

Being embedded in a network gives actors with more than one partner access to different bargaining partners with whom they can divide a surplus. As is explained in detail in Chapter 5, in some networks, actual exchange may be restricted to one relation in keeping with exogenously given rules (e.g. a monogamy rule). Thus, an additional research question is: *With whom of his or her connected partners does an actor divide a surplus if restrictions in a network prescribe that the actor can only divide one surplus at a time?*

Bargaining and Exchange: Structurally Embedded Relations

As a starting point in determining the actors' absolute bargaining power, we focus on an exogenously given network of actors. This network describes bargaining relations between connected actors. A position in a negotiation network can be 'described' by three distinct elements. It is only by simultaneously considering them that an actor's absolute bargaining power can be defined. First, each position in a network is characterized by a specific number of available negotiation partners. Relations provide opportunities for exchange. An actor with numerous negotiation relations is less dependent on one specific adjacent exchange partner. Second, each of the bargaining partners has their own bargaining partners with whom they can negotiate the division of surpluses as well. So it is also important how many partners one's own partners have. The fewer partners ego's partners have, the more they depend on ego for exchange. Lastly, it is important whether bargaining ties in a network are negatively or positively connected (relational assessment). Ties are negatively connected if, from one's own point of view, the partners are alternatives (Emerson 1972: 70–71). Exchange in one of these relations then prevents exchanges in the rest of the relations. Emerson argues that relations can be positively connected as well. This means that exchange in one relation tends to stimulate exchanges in other relations. In that case, partners are complementary for exchanges.

Once the network is known and its relations are classified as either negative or positive, we can determine each actor's absolute bargaining power. In a negatively connected network, ego's absolute bargaining power increases in the number of his relations but decreases in the number of his partners' relations. In a positively connected network, however, relations are complementary and an exchange in one relation promotes transfers in other relations. In a network of this kind, being dependent on others is an advantage. Ego's absolute bargaining power thus increases in the number of his partners.

However, the absolute bargaining power does not determine ego's share of the surplus in relation with alter. Ego's share of the surplus in relation with alter is dependent on his relative bargaining power with respect to alter, that is, ego's absolute bargaining power in relation to alter's absolute bargaining power (see eq. 5.5 in the previous chapter). Thus a high absolute bargaining power does not automatically secure a larger share of the pie. Actors with equal absolute bargaining powers (such as structurally equivalent actors) split the cake evenly.

Inserting the absolute bargaining powers into the equilibrium solution of the Alternation Offers Game yields point predictions for the division of a surplus in all the negotiation relations of the network under study. To test our theoretical predictions, we rely on published results from sociological network exchange theory (for an overview, see for instance Willer 1999). Moreover, we confront our predictions with those of prominent exchange theories.

Bargaining and Exchange: Experimental Results

We test our predictions on the division of a surplus using experimental results published in the literature (Burke 1997; Cook et al. 1983; Lovaglia et al. 1995; Markovsky, Willer, and Patton 1988; Markovsky et al. 1997; Skvoretz and Willer 1993; Yamagishi, Gillmore, and Cook 1988; Yamaguchi 1996). Laboratory experiments on network exchange normally follow a standardized protocol. Subjects are assigned to a specific position within a network structure determined by the experimenter. The network structure is common knowledge. After a thorough introduction and a test round, subjects can negotiate the division of a surplus of 24 points with each of their linked neighbors. In general, about 20 rounds of negotiation and exchange are played. With one exception (Yamagishi, Gillmore, and Cook 1988), experiments assume negative connections and generally restrict the number of exchanges per actor and round to one (i.e. one–exchange rule).

The empirical performance of our model is tested in three distinct steps. In each of these steps, we also compare our predictions to those of alternative approaches: Lovaglia et al.'s (1995) GPI–RD, Yamaguchi's (1996) power model, Skvoretz and Willer's (1993) Exchange Resistance theory, Friedkin's (1986, 1992) Expected Value theory, Cook and Yamagishi's (1992) Equi–Dependence theory, and Burke's (1997) Identity Simulation model. In a first step, we compare our profit point predictions on the division of a surplus in negatively connected networks to the respective empirical results. In a second step, we compare our predictions on profit splits in positively connected networks to the one experiment on positively connected networks reported in the literature. These two steps thus focus on the first research question of Chapter 5. Some negatively connected network structures with a one–exchange rule have been experimentally observed to 'break'. In other words, at least one actor was observed to permanently refrain from exchange with at least one bargaining partner. In a third step, we therefore focus on the second research question of Chapter 5. We investigate whether our model consistently predicts observed deviations between negotiation and exchange structures.

Predictions on the division of a surplus in negatively connected networks are compared to experimental findings on four network structures (see Section 5.5.1): 4–Line, Stem, Kite, and 3–Branch.[10] Observed splits of the surplus in these negatively connected networks match fairly well with the predictions of all the theoretical approaches we consider in our analysis. However, the theories differ in terms of predictive accuracy. Not all the theories perform equally well with regard to the same networks. Every theory has at least one network where predictions fit the data poorly. Our theory, for instance, is less on target in the 4–Line structure. From a surplus of 24 points, we predict that ego will get 15.9 points. Experiments, however, only report 14.1 points in favor of ego. With regard to the other three networks, our theory makes predictions that all fall within the 95% confidence interval of the experimentally observed values. To compare the overall empirical performance with regard to all four negatively connected networks, we use two distinct goodness–of–fit measures: the mean deviation and the absolute deviation. Comparing the theoretical approaches on these two measures shows that our predictions correspond closest to the experimentally observed results.

In a second step, we test our predictions on dividing a surplus in positively connected networks (see Section 5.5.2). Empirical results on positively connected networks are sparse and we have to rely on one single experiment on a 5–Line structure (Yamagishi, Gillmore, and Cook 1988). From the vast body of theories, only Yamaguchi's (1996) power model is capable of making profit point predictions for positively connected networks. Yamagishi et al. report their findings as an actor's systemwide profit share. This means they add up an actor's profits attained in all of his or her dyadic bargaining relations and standardize the result in such a way that the systemwide profit shares of all the actors add up to one. Profit shares for positions in the positively connected 5–Line A–B–C–B–A are reported to be: A = 0.1503, B = 0.1931, and C = 0.3133. For all positions, we predict profit splits that correspond closely with this empirical evidence: A = 0.1632, B = 0.1736, and C = 0.3264. These predictions are also better in line than predictions by Yamaguchi: A = 0.0020, B = 0.2500, and C = 0.4961.

In a third step, we test our predictions on deviations between bargaining and exchange structures (see Section 5.5.3). In some exchange experiments, the experimenter kept track of the number of times linked actors participated in exchange with each other. Based on these results, some authors (e.g. Bienenstock and Bonacich 1993; Markovsky, Willer, and Patton 1988; Simpson and Willer 1999; Skvoretz and Lovaglia 1995) report and discuss deviations between the negotiation and exchange structure. A prominent candidate for such 'breaks' is the T–Shape structure (see Figure 5.2 in Chapter 5), which connects two peripheral actors A_1 and A_2 to actor B who is further connected to actor C, who is also connected to peripheral actor D. Markovsky, Willer, and Patton (1988: 227), for instance, report a break of this structure. In 100

[10] See Figure 5.1 in the previous chapter for a graphical depiction of these networks.

negotiation rounds, only three exchanges occurred between B and C. The network thus degenerates into a 3–LINE A_1–B–A_2 and a DYAD C–D. Our model not only correctly predicts this decay, it also accurately predicts the division of the surpluses in the resulting subnetworks. Out of a surplus of 24 profit points, B was observed to realize, on average, 19.14 profit points in relation to the actors at position A.[11] We predict 19.86 profit points in favor of B. C was observed to receive 11.88 profit points in match with D. We predict an equal division of 12.0 profit points each. Simpson and Willer (1999) report alternative networks for which experimental evidence on the number of exchanges in all relations is available, namely, Y–TRIANGLE, XU–SHAPE, X4–LINE, V–TRIANGLE, and H–SHAPE (see FIGURE 5.3 in the previous chapter). Our model predicts breaks in all these negotiation networks. Statistical tests on Simpson and Willer's data support our predictions. For the XU–SHAPE and the Y–TRIANGLE network, however, Simpson and Willer do not draw the same conclusion. Their statistical tests suggest that these two networks do not break. As we argue in Chapter 5, their conclusion may be due to a statistical mistake.

Apart from profit point predictions on the division of a surplus and predictions on network breaks, our model allows for further theoretical conclusions. First, actors in a DYAD always split their surplus evenly since they are structurally equivalent and do not have any alternative exchange partners. Due to this lack of alternative exchange partners, the DYAD is a robust network (i.e. the exchange structure coincides with the given negotiation structure and does thus not break). The available experimental data support these predictions. Second, not all negatively connected networks can break. If we focus on those negatively connected networks where all the relations negotiate a surplus of equal value, the structure can only break if it meets with the following condition: (i) the network has to contain at least one actor with three or more exchange partners, and (ii) these exchange partners have to be located on at least two structurally different positions. Consequently, networks with equally valued surpluses, where all the actors have at most two exchange partners, are thus always robust. Networks of this kind are either ring or line structures. The experimental evidence supports this conjecture. Third, relations in positively connected networks are complementary and these structures are thus always robust. Since exchange in one relation stimulates exchange in other relations, rational and selfish actors never exclude any of their possible exchange partners. The lack of experiments on positively connected networks makes it impossible to either confirm or refute this prediction. Fourth, whether or not an actor initiates a break depends on the relative bargaining power in all his matches with his partners. If the surpluses to be divided are not equal in all the relations, the possibility of a break also depends on the size of the surpluses to be divided. In unequally valued relations, it may sometimes be more advantageous to obtain a smaller part of a large surplus than a larger part of a small surplus or parts of

[11] The B subjects realized almost identical profits with both A_1 and A_2 throughout the experiment. The cited 19.14 profit points are an average of profits realized in relation to A_1 and A_2.

one or more smaller surpluses. An actor might thus sometimes choose to be exploited in a large surplus relation. There is no empirical evidence to enable us to address this prediction. Fifth, there are negotiation networks where an actor has a latent monopoly position. Such an actor has full control over two or more of his bargaining partners—they have no other bargaining partner—but is also connected to at least one partner he does not completely control, since this partner has one or more alternative bargaining partners. An actor with this type of latent monopoly position always excludes at least one of his less dependent bargaining partners. In other words, he would like to establish a pure monopoly position. Experiments using network structures with latent monopoly positions do indeed demonstrate the predicted behavior of the 'monopoly actor'.

Bargaining and Exchange: Some Suggestions for Further Research

We saw that the theory developed in Chapter 5 performs better or at least as well as established sociological exchange theories. These conclusions are based, however, on experimental evidence from small and simple network structures. We offer some predictions for larger networks, especially in connection with our results on network breaks. We will only know whether these predictions on the division of a surplus are accurate after further experiments with more complex network structures have been conducted. In addition, new experiments could reveal whether the weak information assumption of our model ('locally complete information' instead of complete information on the network structure) is also sufficient for precise predictions in large structures.

Further, predictions of our model are based on the assumption that bargaining power is a function of structure alone. This assumption is in line with Emerson's (1972) basic idea and also fulfills the premises of network exchange theory. Experimental results suggest that our theory clearly illustrates the influence of structure on individual behavior. This is why it is tempting to add a further step, especially if one seeks to explain real world bargaining behavior. Eckel and Grossman (1998) note, for instance, that men and women behave differently when bargaining. They report that men tend to accept lower offers from women than from other men. Women, on the other hand, apparently accept lower offers from women. One the grounds of Becker and Mulligan's (1997) results on the endogenization of time preference, younger or less educated people, for example, can be expected to be less patient. It can thus be concluded that they make less advantageous deals when bargaining. Moreover, it does not seem too farfetched to assume that wealth effects can also play an important role in bargaining. Rich people, for instance, can be expected to bargain less forcefully if the cake is small, and poor people will gladly accept a tiny share of a larger pie. This suggests that bargaining power is also a function of effects other than structure alone. Future research on network exchange should therefore also address effects of actor characteristics.

Finally, if Emerson's (1972) suggestion that exchange theory explains exchange structure is taken seriously, new theories should be able to explain how and when actors initiate exchange relations with new partners outside the given negotiation structure ('network extension'). Cook and Whitmeyer (1992) also suggest that network exchange theory should put more emphasis on explaining changes in network structures rather than solely predicting profit splits. In dealing with positively connected network structures, it might be as important to focus on network extensions as it is to study network breaks in negatively connected networks. In positively connected networks, actors are complementary in exchange, and having more relations might increase one's own bargaining power. A more complete theory on network exchange should therefore 'allow' new actors to enter a given negotiation structure, and should define a behavioral concept on the individual level for the acceptance or rejection of these new negotiation partners. Whether actors in positively connected networks behave according to such assumptions can subsequently be tested in laboratory experiments.

Appendix A

The Repeated Trust Game Experiment

This appendix contains a concise translation of selected parts of the German experiment booklet used to collect the data discussed in Chapter 3. The original experiment booklet contained the following parts:

Introduction The basic game, a one–shot the Trust Game, was explained and further information on the experiment was given to enable the experimental subjects to perform the experiment without further assistance. This part is contained in this appendix.

Part 1 A decision as the trustor in a one–shot Trust Game had to be taken by the experimental subjects. Payoffs of this game were either high stakes or low stakes (cf. Chapter 3). This part is contained in this appendix.

Part 2 A second decision as the trustor in a one–shot Trust Game had to be taken by the experimental subjects. If Part 1 contained the low stakes game, this part contained the high stakes game and vice versa. Otherwise, Part 2 was identical to Part 1 and is thus omitted from this appendix.

Part 3 Decisions as the trustor in a once repeated Trust Game had to be taken by the experimental subjects. The strategy method (see Chapter 3, page 65) was used to route the subjects through the game and record their decisions. This part is contained in this appendix.

ABV Questions A battery of 23 questions from the *Amsterdamse Biografische Vragenlijst* (ABV, 'test attitude scale', Wilde 1970) was presented to the subjects. The ABV scale measures how reliable subject's answers on questions regarding opinions and statements are, for example, on the Trust Items presented to the subjects in a following part of the experiment. This part is omitted from this appendix.

Part 4 Decisions as the trustor in a twice repeated Trust Game had to be taken by the experimental subjects. The strategy method was again used to route the subjects through the game and record their decisions. Besides the fact that the Trust Game is repeated twice instead of once, Part 4 is identical to Part 3 and thus omitted from this appendix.

Trust Items An item–battery on trust and related concepts (Yamagishi and Yamagishi 1994) was presented to the subjects. This item–battery was developed by Yamagishi and Yamagishi to measure how trustful and trustworthy, respectively, subjects are. This part is omitted from this appendix.

Part 5 A decision as the trustor in a Continuous Trust Game had to be taken. This game is also called an Investment Game (e.g. Berg, Dickhaut, and McCabe 1995). This part was only added to the experiment booklet as a pre–test for further experiments. It is thus omitted from this appendix.

Demographics This part of the experiment booklet contained questions on the subjects' age, sex, secondary education, being a blood donor, and whether they have heard about game theory prior to the day of the experiment. This part is omitted from this appendix.

Payment Procedure The last part of the experiment booklet contained a detailed explanation on how two subjects were chosen to be paid in accordance with their decisions in one part of the experiment. This part is omitted from this appendix.

The parts of the experiment booklet omitted from this appendix were either not relevant for the analysis in Chapter 3 or follow the same logic as those parts contained in the appendix. Part 1 and Part 2 only differed in terms of the payoffs used. Part 3 and Part 4 both used a repeated version of the Trust Game; these parts thus differed in the number of repetitions of the constituent game. Part 3 and Part 4 were further different with respect to the payoff combinations used and we conducted the experiment using eight different versions of the experiment booklet that varied with regard to the sets of payoff combinations used in Part 3 and Part 4. The logic of the strategy method used in the once repeated Trust Game of Part 3 was adapted to the twice repeated game of Part 4.

The original experiment booklet in German as well as a concise English translation is available for download in Adobe PDF file format from http://www.fss.uu.nl/soc/iscore as ISCORE Paper No. 193: *An Experiment on Decision Making in Repeated Trust Games: Experimental Booklet.*

Introduction to the Experiment

This experiment can be done with only the material just given to you. We ask you not to speak during the session. Conferring with others is also not allowed during the experiment. If you have any questions reading the instructions or during the course of the session, please rise your hand. The experimenter will try to help you out. Please try not to disturb the others.

This experiment consists of different independent parts, as you will find out while reading the instruction text. You can go about the instruction text and these different parts of the experiment in your own pace. In the **main part of the experiment**, you will be asked to make choices that have to do with money. We want to emphasize that there is no "right" nor "wrong" choice. We are only interested in what *you* consider to be the most appropriate choice. **Your choices are important because it is possible that you will be paid in accordance with your choices in the main part of the experiment.** After everyone has completed the experiment, two persons will be chosen by lot. These two persons will be paid in accordance with their choices made in one of the experiments in the main part (hereafter called "situations"). Which situation is chosen is again determined by lot. Everyone of you has the same probability of being one of these two persons. The payment procedure will be explained in more detail in the instruction text at the end of the experiment. Independent of your performance, everyone will be paid a participation fee of DM 10.–[1]

Explanation of the Basic Game

We would now like to make you familiar with the game that forms the basis of the experiment. The game is played by two players, you ("Player 1") and an unknown other person ("Player 2"). Both players have to make choices in this game. **In all situations of this experiment, we will ask you some questions about your behavior as "Player 1".** The questions will, in general, be of the following kind:

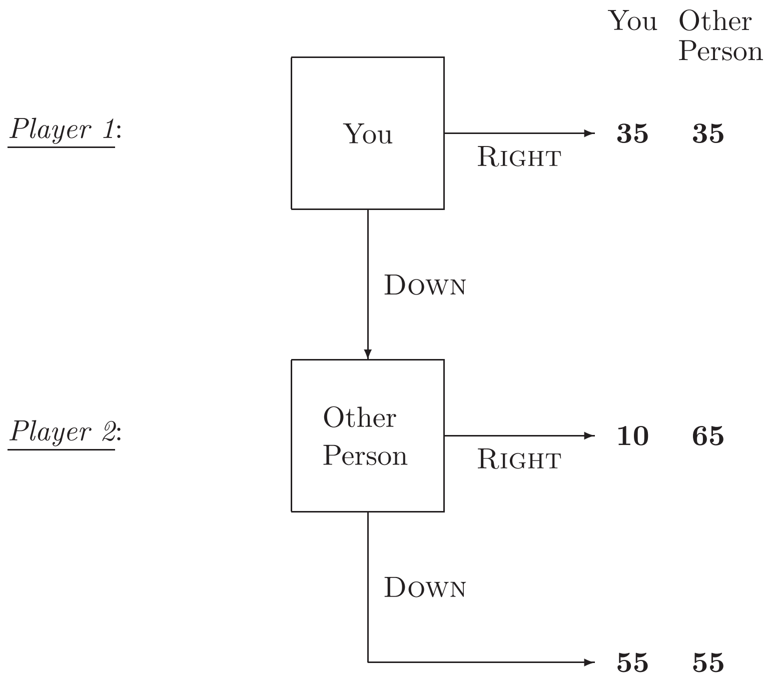
In this situation, I choose
☐ Right
☐ Down

You will only have to decide how you would choose as Player 1. Eventually, your decisions will depend on previous decisions of the other person ("Player 2"). We will thus sometimes ask you what you would do, given the other person has behaved in a certain manner.

In Figure 1, you see a graphical representation of the basic game used in this experiment. We explain the game on the basis of this example. The game starts with a decision by you ("Player 1"). You can choose between two directions: Right or Down. If you choose Right, the game ends. You and the other person both earn

[1] TG: At the day of the experiment, DM 10.– was about US$ 5.– or € 5

FIGURE 1 *The Basic Game of the Experiment*



DM 35.– as you can see in FIGURE 1. The other person's choice does not matter in this case. If you choose DOWN, however, the other person also gets the chance to choose between RIGHT and DOWN. In both cases, the game ends. If the other person chooses RIGHT, you earn DM 10.– and the other person earns DM 65.– If the other person chooses DOWN, both of you earn DM 55.-

As you can see in FIGURE 1, your and the other person's choice lead to monetary rewards. The value of the monetary rewards depend on the choices of you and the other person. The monetary rewards are located to the right of the graphical depiction of the game. Your rewards are in the column labelled "You" and the other person's rewards are in the column labelled "Other Person". Note that the monetary rewards used in the experiment will be different from the ones just used in the above example. Carefully study the picture and make sure you completely understand the game. If you have any questions regarding the basic game, please rise your hand. The experimenter will try to answer your questions.

To make sure that you understood the game, please answer the following question about FIGURE 1. Assume that you choose DOWN and that the other person chooses RIGHT. What are the rewards to you and the other person? Please write them down here:

You receive	DM __________
The other person receives	DM __________

Check your answer at the bottom of the following page. If your answer was correct, please continue reading the instructions. Otherwise, please study the basic game again. If you have any further questions, consult to the experimenter.

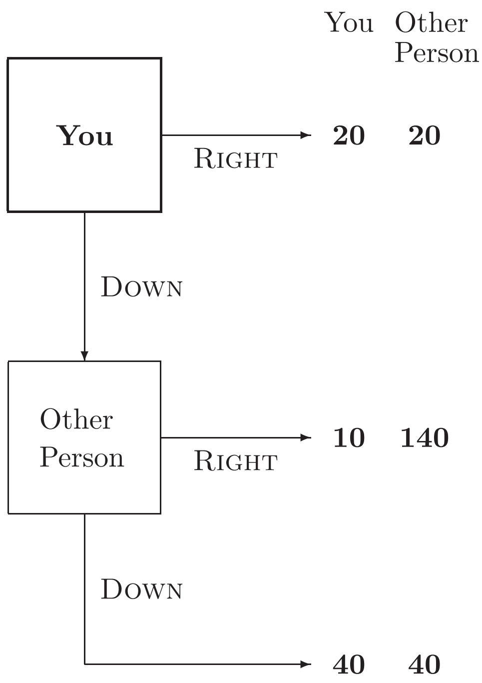
If you do not have any other questions, please start with the experiment on the next page. Again, there are neither right nor wrong decisions. We will only ask you to take your decisions as Player 1. Always make the choice you think to be the best one, given the respective situation. **For reasons of clearness, we will address Player 1 with "You" and Player 2 with "Other Person" throughout the rest of the experiment.** Again, it may be that you will be paid in accordance with your choices in one of the situations in this experiment.

Please start with the experiment on the next page.

Answer to Question: You receive DM 10.– and the other person receives DM 65.–

Part 1

You are in a situation described by the game depicted below. Please study the game carefully and take a decision: either Right or Down. If you take your decision, think of the other person as somebody in this room. However, you do not know which person. Indicate your choice by marking the respective box.



In the situation depicted above, I choose[2]

☐ Right
☐ Down

Please continue with the experiment on page ...

[2]TG: Depending on the subject's answer, the experiment booklet routed the subject to the next part or to the respective question in case of the repeated games in Part 3 and Part 4 of the experiment.

Part 3

We will confront you with yet another situation. Please take a thorough look at the figure "The Game in Part 3" on the opposite page. As you can see, it is the game we discussed in the introduction to this experiment and which you have already played in the first and second part of this experiment. **In Part 3, however, this game is repeated once.** You are playing **a total of two rounds.** You are playing the same other person in both rounds. Think of the other person as somebody in this room. However, you do not know which person. It is, however, not the same person as in the previous two parts. Please take a good look at both rounds of the game depicted on the opposite page. Take into account that the monetary rewards may differ between rounds.

You start in Round 1. Each question will be accompanied by a graphical representation. These representations will tell you in which round you are and which decisions have been taken by you and the other person. In each of these representations, a box with doubled edges will mark your position in the game. In addition, the branches of the game you and the other person did not choose are marked in gray. After each of your answers, we will tell you with which question you have to continue.

Take a look at the game on the opposite page and answer the following question. You can check yourself whether you understood the game. Assume that in Round 1, you choose RIGHT. The other person, therefore, cannot take a decision. In Round 2, you choose DOWN and the other person chooses RIGHT. What are the monetary rewards to you and the other person? Please write them down here:

	Round 1	Round 2
You receive	DM ________	DM ________
The other person receives	DM ________	DM ________

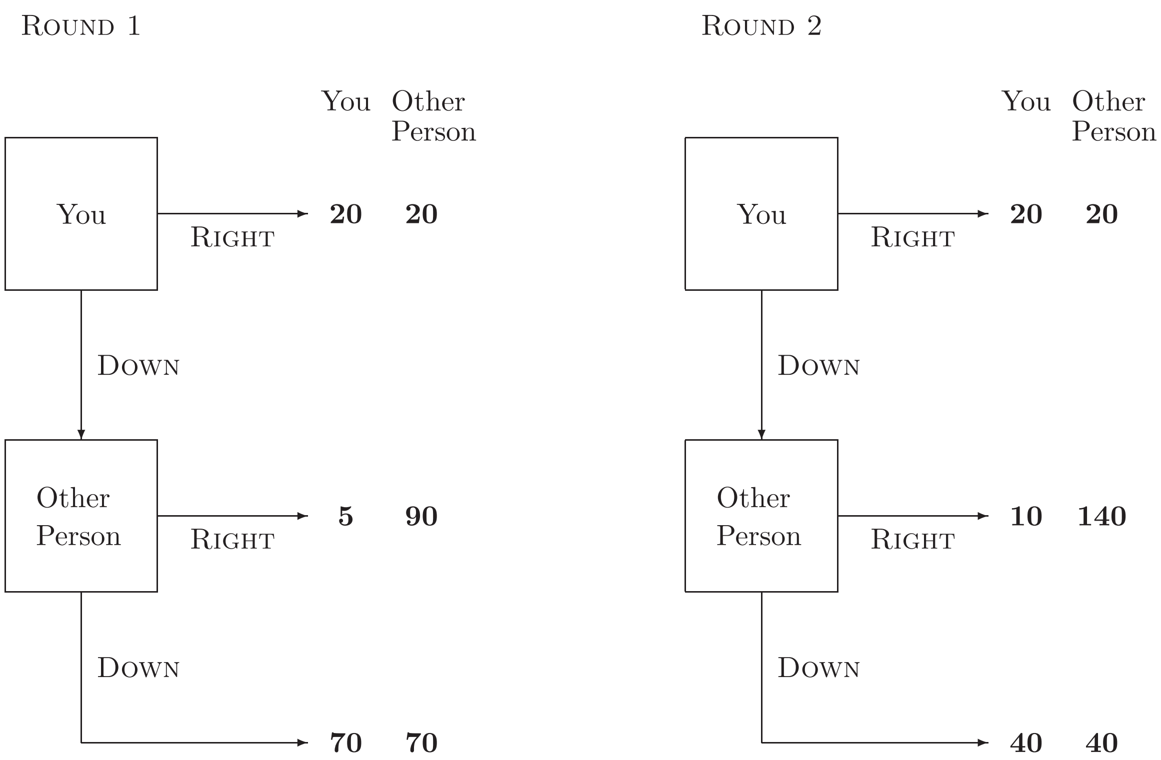
Check your answer at the bottom of this page.[3] If your answer was right, please continue with the third part of the experiment on page 179. Indicate your choice by marking the respective box. Otherwise please study the game again. If you have any further questions, consult to the experimenter.

Answer to Question Part 3:

You and the other person earn DM **20.–** each in Round 1. In Round 2, you earn DM **10.–** and the other person earns DM **140.–**

[3]TG: In the experiment booklet, the answer was on a blank page prior to Question 1 of Part 3. Throughout the experiment booklet, all questions, as well as the different parts, were separated by blank pages.

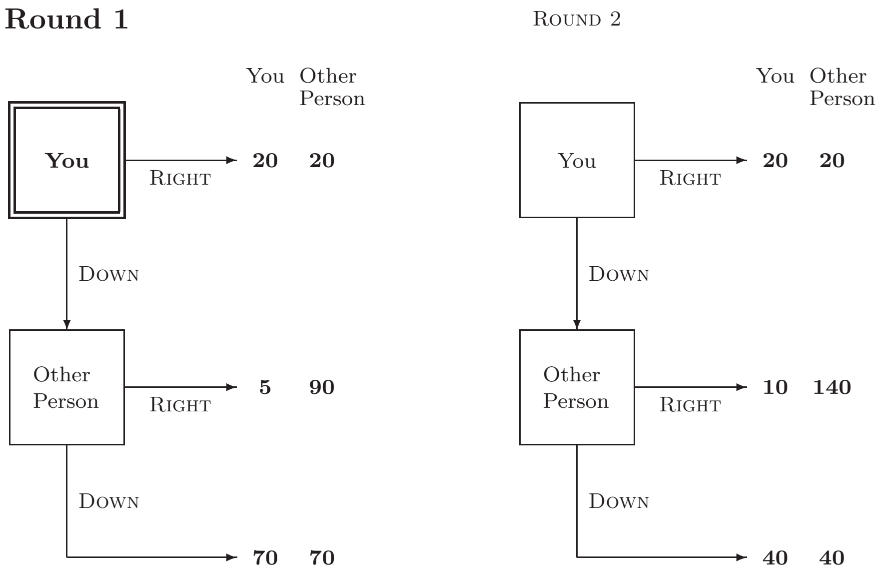
The Game in Part 3



Question 1

You start in Round 1 *and continue to* Round 2. *In both rounds, you play with the same other person. Always indicate your choice by marking the respective box.*

Remark: *Assume that all your decisions are immediately disclosed to the other person. This also holds true for all decisions by the other person, on which you will be informed immediately. You and the other person are therefore, at any time in the game, informed about the choices made in the course of the game and the respective rewards.*



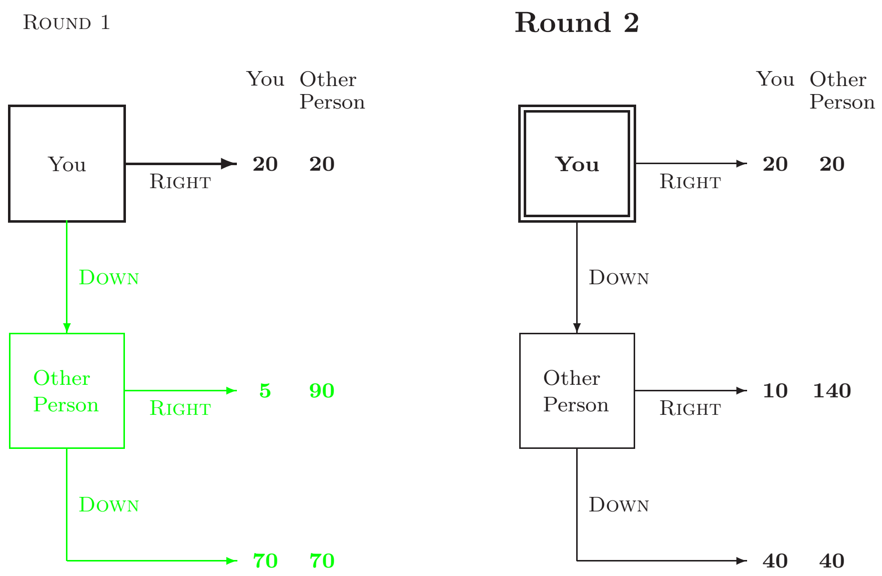
You are in Round 1. You start the game. Please take a decision. Either choose Right or Down.

In Round 1, I choose

☐ Right **(continue with question 2 on page 180)**

☐ Down **(continue with question 3 on page 181)**

Question 2



You are in Round 2. Due to your choice of Right in Round 1, the other person has no possibility to choose. What are you doing in Round 2? Either choose Right or Down.

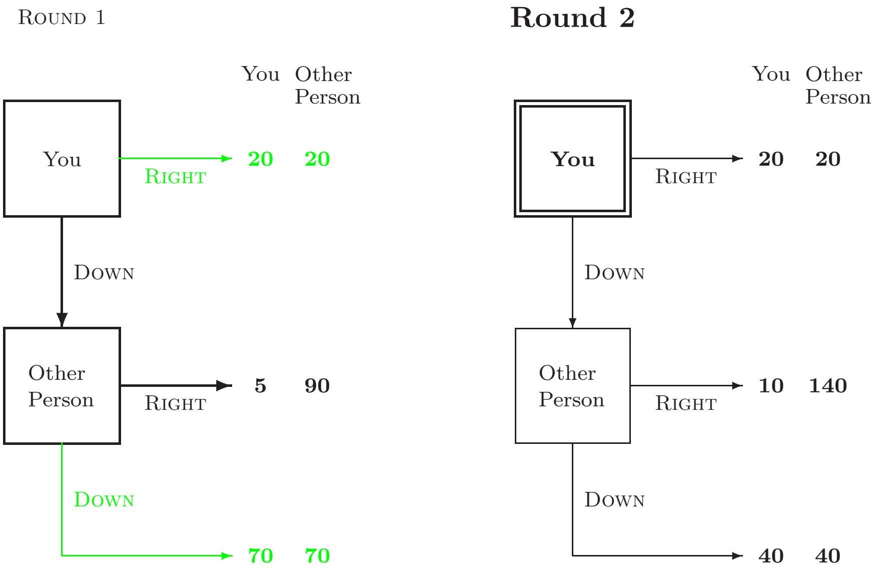
In Round 2, I choose

☐ Right

☐ Down

(This was the last question of Part 3. Continue on page ...)

Question 3



You are in Round 2. Due to your choice of Down in Round 1, the other person could choose between Right and Down in Round 1. Assume the other person has chosen Right. What are you doing in Round 2? Either choose Right or Down.

In Round 2, I choose

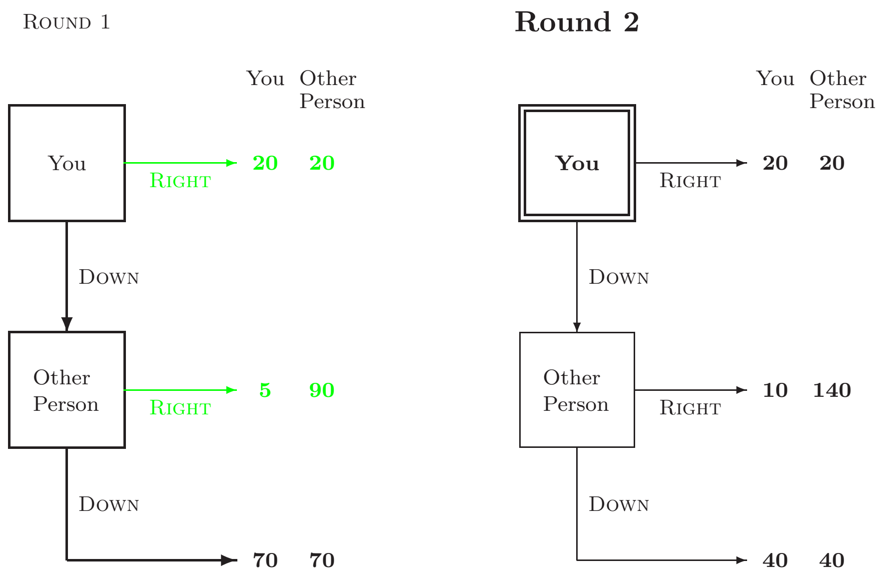
☐ Right
☐ Down

(continue with question 4 on page 182)

Question 4

Due to your choice of Down *in Round 1, the other person could also choose between* Right *and* Down *in Round 1. In the previous Question 3, you took a decision in Round 2 under the assumption that the other person has chosen* Right *in Round 1.*

Now assume that in Round 1, the other person, after your choice of Down*, has chosen* Down *as well instead of* Right*. Please again indicate your decision in Round 2 given these new circumstances.*



You are in Round 2. Due to your choice of Down in Round 1, the other person could choose between Right and Down in Round 1. *Assume the other person has not chosen* Right *but* Down *as well in Round 1.* What are you doing in Round 2? Either choose Right or Down.

In Round 2, I choose

☐ Right

☐ Down

(This was the last question of Part 3. Continue on page ...)

Appendix B

Information on the MAT Questionnaire

This appendix contains translations of questions from the Dutch survey on IT–transactions ('The External Management of Automation', MAT) as far as they are relevant for Chapter 4. For the most part, these translations are taken from Buskens (1999). Below, the questions are given followed by the response categories in parentheses. The questions are ordered as they were actually included in the questionnaire. The numbering of questions follows the codebook by Buskens and Batenburg (2000).

For further information on and analyses with the MAT–data, see for instance, Batenburg (1997b), Batenburg, Raub, and Snijders (2000), Blumberg (1997), Buskens (1999), Raub (1997), and several articles in Weesie and Raub (2000).

Section 1. The Product

Q1.2: Which of the following products/services were delivered at that time? (standard software, adjusted software, tailor-made software, sector-specific software, personal computers, workstation, network configuration, mini computer, mainframe, computer-controlled machines, side equipment, cabling, design, training, instruction, consultation, documentation, support)

Q1.4: How much was paid to the supplier, not including later supplements? (up to NLG 25,000, NLG 25,000–50,000, NLG 50,000–100,000, NLG 100,000–200,000, more than NLG 200,000)

Q1.6a: How important was this product for the automation of your firm? (unimportant, hardly important, moderately important, very important, of major importance)

Q1.6b: How important was this product for the profitability of your firm? (unimportant, hardly important, moderately important, very important, of major importance)

Q1.7: Was it easy or difficult for you and your employees to judge the quality of the product at the time of delivery? (very easy, easy, somewhat difficult, difficult, very difficult)

Q1.8: Compared to other firms in your sector, how much experience did your firm have with automation? (none, little, some, much, very much)

Q1.10: How important was it that the product delivery time was met? (unimportant, hardly important, moderately important, very important, of major importance)

Q1.12a: How important was the long–term suitability of this product? (unimportant, hardly important, moderately important, very important, of major importance)

Q1.12b: How important was long–term support by the supplier? (unimportant, hardly important, moderately important, very important, of major importance)

Q1.12c: How important was the long–term compatibility of this product with other hardware and software? (unimportant, hardly important, moderately important, very important, of major importance)

Q1.13a: What would have been the damage, in terms of money and time spent on purchasing a new product, if the product had failed to function and had had to be replaced? (very small, small, moderate, large, very large)

Q1.13b: What would have been the damage, in terms of money and time spent on training personnel, if the product had failed to function and had had to be replaced? (very small, small, moderate, large, very large)

Q1.13c: What would have been the damage, in terms of money and time spent on data entry, if the product had failed to function and had had to be replaced? (very small, small, moderate, large, very large)

Q1.13d: What would have been the damage, in terms of money and time wasted in idle production, if the product had failed to function and had had to be replaced? (very small, small, moderate, large, very large)

Q1.15: Did your firm have the possibility to make or adapt this product? (no, very difficult, difficult, easy, very easy)

Section 2. The Choice of Product and Supplier

Q2.8: Was it difficult for your firm to compare tenders? (very easy, easy, somewhat difficult, difficult, very difficult)

Q2.9: Was it difficult for your firm to compare the product with other products? (very easy, easy, somewhat difficult, difficult, very difficult)

Q2.10: Was it difficult for your firm to compare the price–quality relation of potential suppliers? (very easy, easy, somewhat difficult, difficult, very difficult)

Section 3. The Relation with the Supplier

Q3.2: How many employees were working at the supplier's firm at that time? (less than 5, 5–9, 10–19, 20–49, more than 50)

Q3.7: For how long has your firm done business with the supplier before the purchase of this product? (about ... years)

Q3.8: How frequently has your firm done business with the supplier before the purchase of this product? (once, occasionally, quite regular, regular, often)

Q3.9: How extensively has your firm done business with the supplier before the purchase of this product? (minimal, small, quite extensive, extensive, very extensive)

Q3.13: To what extent did you expect, before the purchase of this product, that your firm would continue business with this supplier? (no business, incidental business of limited size, some business of limited size, regular and/or extensive business, very regular and/or very extensive business)

Section 4. The Contract

Q4.11: Was the main contract a more or less standard contract or was it a more or less tailor made contract? (standard contract, more or less written by ...; tailor made contract, more or less written by ...)

Q4.12: Was the main contract a more or less adapted version of a contract that your firm used before? (no, yes)

Q4.13: For each of the following technical specifications of the product, can you indicate how detailed they were arranged in the contract? (hardly, to some extent, to a moderate extent, detailed, very detailed)

General issues (4): security, user friendliness, definition of system limits, definition of system functions.
Hardware issues (8): mainboard, memory (RAM and harddrives), speed (CPU and RAM), communication with other hardware, hardware environment, additional hardware (co–processor, videocard etc.), installation procedure, quality monitor (resolution).
Software issues (12): operating system, additional software, implementation, memory and harddrive usage, system analysis, system methodology, structure and type of data, modules and procedures, conversion old data, operation and performance monitoring, benchmark and system tests, programming language.

Q4.14a: For each of the following financial and legal clauses, can you indicate how much attention was payed to these issues during negotiation of the contract? (less, normal, much)

Issues (24): price determination, price level, price changes, terms of payment, sanctions on late payment, delivery time, liability supplier, force majeure, warranties supplier, quality (norms), intellectual property, protection against privacy, restrictions on product use, non–disclosure, insurance supplier, duration service, reservation spare–parts, duration maintenance, updating, arbitration, calculation R&D costs, joint management during transaction, technical specifications, termination.

Q4.14b: For each of the following financial and legal clauses, can you indicate how they were arranged? (only verbally, in a written document, not at all arranged)

Issues (24): see Q4.14a.

Q4.16: During negotiating and arranging the contract, did your firm use external legal help? (no, yes)

Q4.17: Which of yours and the supplier's departments or employees were involved in contracting? (our firm, the supplier)

Issues (7): CEO's, IT department or employees, financial department or employees, production, purchases, selling, legal department or employees.

Q4.18: In total, how much time have you and your colleagues invested in arranging and negotiating the contract with the supplier? (about . . . mendays)

Q4.20: Every firm has its standard procedures. Regarding the negotiations and agreements with this supplier concerning the product as a whole: To what extent could these be considered to be standard procedures? (hardly, to some extent, to a moderate extent, largely, completely)

SECTION 6. ABOUT YOUR FIRM AND YOURSELF

Q6.1b: How many full–time employees were working at your firm at the time of the purchase of this product? (number of full–time employees: . . .)

Q6.3ab: Does your firm have employees with expertise on automation? (no, yes)

Q6.3bb: Does your firm have an automation department? (no, yes)

Q6.3ac: Does your firm have employees with expertise on legal issues? (no, yes)

Q6.3bc: Does your firm have a legal department? (no, yes)

Q6.3cc: Does your firm have external advisors with expertise on legal issues? (no, yes)

Q6.4: Some firms are very active in their own industry and are involved in many activities such as congresses, exposition, meetings etc. Others are less active. How active would you consider your firm to be? (not active, little active, moderately active, active, very active)

Q6.5: To what degree is your firm an active member of different organizations within the industry? (not active, little active, moderately active, active, very active)

Q6.6: How often do you or other IT employees of your firm visit events in the field of IT? (never, incidentally, once in a while, regularly, very regularly)

Issues (3): Expositions, seminars or workshops, congresses.

Q6.7: How often do you or other IT employees of your firm visit clubs where business partners meet each other in an informal environment? (never, incidentally, once in a while, regularly, very regularly)

Issues (3): Lions and Rotary Club, local business club, social clubs.

Appendix C

Mathematical Details*

C.1 Condition for Structural Differences between Negotiation and Bargaining Structure

Proof of Theorem 5.1 (page 130). Starting from the assumptions in Chapter 5 and using the same notation, we show that inequality (5.15a) is the condition for differences between exchange and bargaining structures. Note that the proof of inequality (5.15b) is equivalent to the proof of (5.15a) since $x_{ij} = p_{ij}\, v_{ij}$. In this case, the focus is on equally valued relations, $v_{ij} = v_{ji} = v = 1$.

To derive inequality (5.15a), suppose that i is a selfish and rational actor with substitutable relations in a given negatively connected or mixed network. His overall welfare thus may be conceptualized as an additively separable function of the utility contributions from each match. And, if the latter are linear in the shares of the pie under consideration, i's maximization of total welfare just involves the comparison of potential profit levels. Actor i thus will decide about j's exclusion on the basis of the profits he could realize in actual exchanges. From this perspective,

$$x_{ik} > x_{ij} \tag{C.1}$$

is a necessary condition for i's consistent exclusion of j.

It is, however, not a sufficient condition. To see that, assume that actor i decides about his actual exchange partners by comparing the (expected) profit levels from j's non–exclusion and exclusion. When actor i exchanges with probability q_{ij} with j, his profit from non–exclusionary behavior is $q_{ij}\, x_{ij} + (1 - q_{ij})\, x_{ik}$. The latter is compared to x^*_{ik}, i's profit from the relation with k if i exclusively exchanges with actor k, that is, when i never exchanges with j such that the exchange network differs from the bargaining network. Clearly, i will exchange with both j and k if and only if his profit from non–exclusionary behavior does not fall short of his profit from exclusionary behavior:

$$q_{ij}\, x_{ij} + (1 - q_{ij})\, x_{ik} \geq x^*_{ik}\,. \tag{C.2}$$

*This appendix is adapted from Braun and Gautschi (2000).

Now, if $x_{ik} > x_{ij}$, rearranging yields the following condition for i's non–exclusionary behavior towards actor j:

$$q_{ij} \leq \frac{x^*_{ik} - x_{ik}}{x_{ij} - x_{ik}} =: q^*_{ij}, \tag{C.3}$$

where q^*_{ij} denotes the critical value for q_{ij}, the probability of i's exchange with j. As long as $q_{ij} \leq q^*_{ij}$ is met, actor i will complete transfers in different rounds with j and k although $x_{ik} > x_{ij}$ is fulfilled. Hence, the latter inequality does not suffice for i's consistent exclusion of j.

To obtain a sufficient condition, we take into account that $q_{ij} = 0$ holds by definition when i always avoids exchanges with j. And, since $q_{ij} \leq q^*_{ij}$ has to be satisfied in the case of non–exclusionary behavior, i's consistent exclusion of j requires then $q^*_{ij} < 0$. Due to the definition of q^*_{ij} and the necessary condition $x_{ik} > x_{ij}$, $q^*_{ij} < 0$ if and only if $x^*_{ik} > x_{ik}$. The latter inequality is the sufficient condition for i's consistent exclusion of j. Hence

$$x^*_{ik} > x_{ik} > x_{ij} \tag{C.4}$$

for at least one actor i with negative connections to distinct negotiation partners j and k is the general condition for differences between exchange and negotiation structures. This proofs that inequality (5.15a) represents the necessary and sufficient condition for a deviation between the negotiation and exchange structures. □

C.2 Network Characterization

Starting from the non–cooperative bargaining model for profit divisions as presented in Chapter 5, we now prove that the condition for structural differences (viz. inequality (5.15b)) can only be fulfilled for those negatively connected networks *with equally valued relations* in which (*i*) at least one actor i has three or more negotiation partners, and (*ii*) i's partners are located on two or more structurally distinct positions (one of which may be equivalent to i's).

Proof of Theorem 5.2 (page 132). The proof of (*ii*) is straightforward. It suffices to show that those negatively connected networks with equally valued relations are robust (i.e. bargaining and exchange relations will coincide) in which the actors' bargaining partners do not vary in terms of structural position. For that purpose, we start from the fact that $p_{ik} > p_{ij}$ is, in the presence of an alternative partner k, necessary for i's permanent exclusion of j. Since both p_{ik} and p_{ij} can be expressed in terms of network control (see, eq. (5.11)), substituting and rearranging yields

$$p_{ik} > p_{ij} \iff c_j > c_k\,. \tag{C.5}$$

That is, the necessary condition for i's consistent exclusion of j is fulfilled if and only if j's network control c_j exceeds k's network control c_k. Structurally equivalent partners always possess the same degree of network control (cf. eqs. (5.7a), (5.7b), (5.7c)). So, if network members have only structurally equivalent partners (like, e.g. the actor at the structural position B in the 3–Line A–B–A), they will never consistently exclude a specific actor and the respective negatively connected network will be robust. This completes the proof of (*ii*).

To prove (*i*), it suffices to show that those negatively connected structures with distinct positions are robust in which all actors have at most two negotiation partners. Such a bargaining structure is always a line, the smallest being the 3–Line (which, as we already know, is robust). We therefore have to prove that exclusionary behavior does not occur in a structure in which $n > 3$ actors are chained.

Three features of line or chain structures simplify this task. First, there is always a (fictitious or actual) center at which the left–hand side of the line may be distinguished from its right–hand side. The sequence and type of structural positions (A,B,C,D) on both sides of this center coincide, of course, and the position associated with the center depends on the number of actors. Second, each line has two endpoints—an actor located at the very end of the line (i.e. at position A) is, in contrast to all others, tied to just one partner. Such an actor will never consistently exclude his only partner. Third, if at all, exclusionary behavior will be displayed by those actors that are at most two positions away from the end of the line. The reason is simple: the neighbor of the actor residing at the end of the line has no incentives to consistently exclude his fully dependent partner and his other neighbor (if he would permanently exclude the latter, he would end up in a less profitable Dyad structure with the former).

We therefore can restrict the analysis to the part of the left–hand side of line structures that begins with position B. And, the focus can be on just three actors (j, i, k) who are consecutively chained (additional actors are indexed by h). Specifically, we assume that actor j is always at position B and that i, his neighbor at the right–hand side, always is adjacent to actor k. This scenario reflects that only the decision situation of actor i with respect to the consistent exclusion of j has to be examined. A closer look at short line structures is instructive:

4–Line: A–B–B–A or, written in terms of actors, h–j–i–k
The actors j and i are located at the same position, whereas j and k are not structurally equivalent. Combining eqs. (5.7b), (5.8), (5.9) and (5.10), we therefore find

$$p_{ik} > p_{ij} = 0.5 \iff c_j = c_i > c_k \,. \tag{C.6}$$

That is, the necessary condition for i's consistent exclusion of his neighbor j is fulfilled. However, the sufficient condition for the permanent exclusion of j is violated—i's exclusionary behavior towards j would put him into a Dyad structure with k (i.e. he would obtain just half of the pie in the match with k). Since he cannot improve his situation by exclusionary behavior, i maintains the relation to j.

5–LINE: A–B–C–B–A or, written in terms of actors, h_1–j–i–k–h_2
Here, the actors j and k are structurally equivalent. Because of eqs. (5.7b), (5.8), (5.9) and (5.10), we therefore have

$$p_{ik} = p_{ij} \iff c_j = c_k \,, \tag{C.7}$$

the necessary condition for exclusionary behavior is not met. So, i does not exclude j.

6–LINE: A–B–C–C–B–A or, written in terms of actors, h_1–j–i–k–h_2–h_3
The actors j and k are not structurally equivalent, but i and k are. Combining eqs. (5.7b), (5.8), (5.9) and (5.10), we therefore obtain

$$p_{ik} = 0.5 > p_{ij} \iff c_j > c_k = c_i \,. \tag{C.8}$$

That is, the necessary condition for i's consistent exclusion of j is fulfilled. However, if i would always exclude j, he would be located at the endpoint of a 4–LINE. And, since k would be then i's only exchange partner, this situation would be associated with a lower profit p^*_{ik} than the status quo (in which he gets $p_{ik} = 0.5$). This violation of the sufficient condition for exclusionary behavior ensures, of course, that i maintains his relations with j and k.

7–LINE: A–B–C–D–C–B–A or, written in terms of actors, h_1–j–i–k–h_2–h_3–h_4
The actors j, i, and k now reside at structurally distinct positions. Nevertheless, i and k control their relations to the same degree ($c_i = 0.5 = c_k$). This 'control equivalence' results because either actor has two partners each of which has the same number of relations as well. Because of eqs. (5.7b), (5.8), (5.9) and (5.10), we therefore have

$$p_{ik} = 0.5 > p_{ij} \iff c_j > c_k = c_i \,. \tag{C.9}$$

Again, the necessary condition for i's consistent exclusion of j is met. But if i would permanently exclude j, he would be located at the very end of a 5–LINE. And, since k would be then his only exchange partner, his profit share p^*_{ik} would be lower than $p_{ik} = 0.5$. As a consequence, i will not exclude j.

In sum, neither of these lines decays into separate subnetworks. To complete the proof, it remains to be shown that longer chains are also robust. Interestingly, the scenario of the 7–LINE is paradigmatic for each longer line structure. That is, in all chains with $n > 7$ actors, the pattern of the 7–LINE can be found: while the actors j, i, and k have structurally different positions, there is always a control equivalence of i and k. It therefore holds that $p_{ik} = 0.5 > p_{ij} \iff c_j > c_k = c_i$ (C.9) such that the necessary condition for i's consistent exclusion of j is met. If i would exclude j, however, he would be located at the very end of a line with $n - 2$ actors such that $p^*_{ik} < p_{ik} = 0.5$; this violation of the sufficient condition prevents, of course, that i excludes j. In combination with the above results, this conclusion completes the proof of (i). □

Samenvatting*

In dit boek bestuderen we de productie van een surplus door, en de verdeling van dit surplus tussen twee actoren. Een surplus is 'iets extra' dat, indien geproduceerd, op de een of andere manier tussen twee actoren verdeeld kan worden waarbij beide actoren beter af zijn dan wanneer ze het surplus niet geproduceerd en verdeeld hadden. Het boek stelt de productie en verdeling van een surplus volgens twee scenario's aan de orde. In hoofdstuk 1 introduceren we deze scenario's. In hoofdstuk 2 tot en met 4 worden de productie en verdeling van een surplus in de context van *sequentiele ruil* geanalyseerd. Dit scenario kan gezien worden als een *vertrouwensprobleem* tussen twee actoren (Coleman 1990). De vraag of vertrouwen al dan niet wordt geschonken en of daarom de productie van waardevolle hulpbronnen voor de productie van een surplus al dan niet tot stand komt, staat in deze hoofdstukken centraal. Preciezer gezegd: we bestuderen *de effecten van temporele inbedding op de productie van een surplus*.

Het tweede deel van het boek, hoofdstuk 5, richt zich op de verdeling van een gerealiseerd surplus in plaats van op haar productie. Het scenario betreft twee actoren die hun aandeel in het surplus middels onderhandelen bepalen, waarbij een *vertraging in het tot een overeenkomst komen* over de verdeling van het surplus kostbaar is voor de actoren. Dit onderhandelingsscenario komt overeen met het *Alternating Offers Game*, geïntroduceerd door Rubinstein (1982). De vraag waar wij een antwoord op proberen te vinden is wie hoeveel van het surplus krijgt als over de verdeling wordt onderhandeld. De kosten van de actoren die worden veroorzaakt door een vertraging in het tot een overeenstemming komen, bepalen de verdeling. Deze kosten zijn exogeen in Rubinstein's spel, terwijl wij ze juist endogeen proberen af te leiden. We stellen dat deze kosten kunnen worden bepaald wanneer we niet een enkel paar actoren maar een heel *onderhandelingsnetwerk* centraal stellen waarin paarsgewijze (dwz. dyadische) onderhandelingen over de verdeling van een surplus mogelijk zijn. De kosten van het uitstellen van een overeenkomst zijn voor een actor dan een functie van de positie van die actor in zo'n onderhandelingsnetwerk.

*I would like to thank Arount van de Rijt for translating the English version of this summary.

Sequentiele ruil: de productie van een surplus

Laten we het basisscenario van sequentiele ruil beschouwen waarin een *vertrouwengever* waardevolle hulpbronnen bezit die zij kan investeren in de productie van een surplus door de *vertrouwennemer*.[1] Een vertrouwennemer kan, zodra hij de hulpbronnen heeft ontvangen, deze in zijn eigen belang of in het gemeenschappelijke belang van zichzelf en de vertrouwengever gebruiken. Met andere woorden, hij kan het surplus dat geproduceerd is door de vertrouwengever op twee manieren verdelen: het hele surplus voor zichzelf houden of het zo verdelen dat beide actoren ervan profiteren. Wanneer de vertrouwengever geen medebeschikkingsrecht heeft bij de verdeling van het surplus, moet zij bij het besluiten tussen wel of niet investeren inschatten hoe de vertrouwennemer het surplus zal verdelen. De productie en de verdeling van een surplus kunnen zo gezien worden als een vertrouwensprobleem in de zin van Coleman (1990: 97–99). In dit boek conceptualiseren we zo'n vertrouwensprobleem als een *vertrouwensspel* ('Trust Game', Dasgupta 1988; Kreps 1990a).

In onze analyse van het basisscenario van sequentiele ruil maken we drie cruciale aannames. Ten eerste wordt het nut van een actor bepaald door zijn 'stuk van de taart' en geeft de actor de voorkeur aan een grote 'taartpunt' boven een kleinere. Ten tweede gedragen actoren zich volgens speltheoretische rationaliteit of volgen zij in ieder geval een of ander principe van doelgericht handelen. Ten derde nemen we aan dat een ontmoeting tussen twee actoren geïsoleerd plaats vindt. Dat wil zeggen, er waren geen eerdere of toekomstige interacties tussen de vertrouwengever en de vertrouwennemer, noch waren er eerdere of toekomstige interacties waarbij andere actoren betrokken waren. Onder deze drie aannames is de voorspelling dan dat vertrouwen niet wordt geschonken, en als dat toch zou gebeuren, dat het gegeven vertrouwen zou worden geschonden. Als de vertrouwennemer de hulpbronnen van de vertrouwengever heeft ontvangen zou hij ze liever voor zijn eigenbelang aanwenden dan voor het gemeenschappelijke belang. Dit volgt uit onze eerste aanname. Volgens onze tweede aanname kan de vertrouwengever op het gedrag van de vertrouwennemer anticiperen en zal zij dus geen hulpbronnen ter beschikking stellen. Daardoor zal het surplus niet tot stand komen. Beide actoren zijn dan echter slechter af in vergelijking tot de situatie waarbij de hulpbronnen wel ter beschikking worden gesteld (vertrouwen wordt geschonken) en het surplus zo verdeeld wordt dat beide actoren ervan profiteren (vertrouwen wordt gehonoreerd).

Empirisch onderzoek en alledaagse waarneming leren dat de Pareto–verbetering, waarbij vertrouwen wordt geschonken en gehonoreerd, dus het surplus tot stand komt en zo wordt verdeeld dat beide actoren profiteren, soms feitelijk voorkomt (en dus dat iets aan onze aannames onjuist is). In het eerste deel van het boek onderzoeken we hoe dit kan worden verklaard. Er zijn een aantal benaderingen mogelijk. Men kan bijvoor-

[1] De vertrouwengever is degene die vertrouwen kan geven. De vertrouwennemer kan het gegeven vertrouwen wel of niet kan honoreren. Ik gebruik vertrouwengever en vertrouwennemer als vertaling van de gebruikelijke Engelse termen 'trustor' en 'trustee'.

beeld de aanname dat nut gelijk staat aan de grootte van de 'taartpunt' wat verzachten en eerlijkheid, relatieve uitbetalingen of andere 'zachte' aspecten in de nutsfuncties van de actoren opnemen (zie bijvoorbeeld Bolton en Ockenfels 1998, 2000; Camerer 1997; Kelley en Thibaut 1978; Nelson 2001; Rabin 1993, 1998; Snijders 1996). Theoretici zoals Emerson (1972, 1981) stellen dat een meer sociologische benadering zich moet toespitsen op variatie in sociale condities in plaats van op variatie in nut. In dit boek kiezen we voor deze meer sociologische benadering: we betwisten niet zozeer de aannames over nut, maar onderzoeken de rol die sociale condities spelen. We verzachten zo de derde aanname terwijl we zo veel mogelijk vasthouden aan de eerste en de tweede. Hiervoor nemen we situaties onder de loep waarbij actoren elkaar herhaaldelijk ontmoeten en de ruil die plaatsvindt tussen hen *temporeel ingebed* is (Granovetter 1985; Raub en Weesie 2000). 'Temporele inbedding' verwijst dus naar een serie uitwisselingen tussen de twee actoren. In termen van het abstracte vertrouwensspel betekent dit dat het eenmalige spel een aantal keren wordt herhaald. Dientengevolge is de onderzoeksvraag van het eerste deel van het boek: *Gegeven dat de productie en verdeling van een surplus geconceptualiseerd kunnen worden als een situatie van vertrouwen tussen twee actoren, welke aspecten van temporele inbedding zijn dan van invloed op het schenken van vertrouwen en zijn deze aspecten positief of negatief van invloed op het schenken van vertrouwen?*

Vertrouwen: Effecten van het verleden

In **hoofdstuk 1** worden argumenten besproken waarom we ons richten op temporele inbedding bij het verklaren van het tot stand komen van een surplus. Een specifiek aspect van temporele inbedding, 'de schaduw van de toekomst', is tamelijk uitvoerig bestudeerd in de literatuur (zie bijvoorbeeld, Axelrod 1984; Maynard Smith 1982; Taylor 1987; Trivers 1971). Men heeft laten zien dat het gedrag van een actor afhankelijk is van het verwachte effect ervan op toekomstige interacties. In **hoofdstuk 2** wordt deze bevinding als gegeven beschouwd. De vraag is hier of naast de 'schaduw van de toekomst' ook de 'schaduw van het verleden' van invloed is op het gedrag van een actor. Zulke effecten van het verleden zijn minder uitvoerig bestudeerd in de literatuur. Echter, wanneer we temporele inbedding beschouwen als een mogelijke 'oplossing' voor de productie en verdeling van een surplus, moeten we ook weten of, en wanneer dat zo is, en hoe het verleden het huidige gedrag van een actor beïnvloedt. Om te beginnen kiezen we ervoor de effecten van het verleden op beslissingen te bestuderen in de context van sociale dilemma's. Een sociaal dilemma is een situatie waarin individuele en collectieve belangen conflicteren en rationele en egoïstische actoren uitkomen op een inefficiënt (Pareto suboptimaal) evenwicht. Het vertrouwensspel en het gevangenendilemma zijn voorbeelden van zulke dilemmasituaties. Hoofdstuk 2 richt zich op *het vergelijken van twee situaties* die op één aspect na identiek zijn: de ene situatie kent een verleden, de andere niet. In alle andere opzichten, zoals de alternatieven en uitbetalingen, zijn de situaties hetzelfde. Bovendien is minstens één van deze situaties een sociaal dilemma.

We presenteren een simpele theorie die gebaseerd is op speltheorie met complete informatie. Deze theorie voorspelt dat er geen verschil in gedrag zou moeten zijn tussen de situatie zonder verleden en de situatie met verleden. Het argument hiervoor is eenvoudig. Wanneer actoren zich gedragen in overeenstemming met onze gedragsaannames zoals hierboven beschreven, dan bevat het verleden geen informatie die actoren ertoe zou brengen hun beslissingen zoals genomen in de situatie zonder verleden te veranderen. Hun gedrag zou daarom hetzelfde moeten zijn in de situatie zonder verleden en die met verleden.

We zetten dit argument iets formeler op en introduceren eerst een situatie genoemd spel A. Dan presenteren we een tweede, gecompliceerdere situatie, spel B. Dit tweede spel lijkt op de situatie bij spel A, maar wordt gekarakteriseerd door voorafgaande acties door de actoren van spel A of andere actoren. Deze voorafgaande acties vormen een uitbreiding van spel A in de zin dat spel A een deelspel is van het grotere spel B.[2] Dit deelspel wordt spel A_B (lees: 'spel A in B') genoemd. Aangezien het voor onze simpele theorie nodig is dat spel A_B en spel A identiek zijn, vormen de 'voorafgaande acties' in spel B slechts op twee manieren een uitbreiding van spel A. Ten eerste kan spel B een herhaalde versie van spel A zijn en ten tweede staan de 'voorafgaande acties' niet ter beschikking aan minstens één actor in spel A. Op die manier wordt spel B een spel waarvan spel A_B een van de mogelijke deelspelen is. De voorspelling dat het gedrag in spel A en spel A_B niet zal verschillen wordt geconfronteerd met bevindingen van drie verschillende experimenten.

Eerst vergelijken we een variant van een dictatorspel (spel A) met een vertrouwensspel (spel B). In beide spelen zijn er twee actoren: ego en alter. In de variant van het dictatorspel kan alter een surplus evenredigf verdelen tussen zichzelf en ego of het gehele surplus voor zichzelf houden. Ego is een 'dummy' speler die in het dictatorspel geen handeling verricht. In het vertrouwensspel begint ego met de keuze tussen het schenken dan wel onthouden van vertrouwen. Als ego geen vertrouwen schenkt, eindigt het spel met een zekere uitbetaling voor beide actoren. Echter, wanneer ego wel vertrouwen schenkt, begint deelspel spel A_B dat identiek is aan het dictatorspel. Wanneer alter zich gedraagt volgens onze bovengemaakte aannames, zou zijn gedrag in spel A_B en spel A gelijk moeten zijn, ook al is het ego's handeling in spel B die alter de kans geeft een beslissing in spel A_B te maken. Experimentele bevindingen (Snijders en Keren 1997) spreken onze theorie echter tegen. Zij die in experimenten zoals in spel B de rol van alter vervullen verdelen het surplus in 73.5% van de gevallen gelijk. Zij honoreren dus in het algemeen vertrouwen. In spel A, echter, kiest slechts 35.5% van alle alters voor een gelijke verdeling van het surplus.

Op soortgelijke wijze richten we onze aandacht op ego's gedrag in een vertrouwensspel (spel A) in vergelijking met zijn gedrag in een vertrouwensspel met onderpand (spel B). In het laatstgenoemde spel besluit alter om al dan niet een onder-

[2] Simpel gesteld is een deelspel dat stuk van een geheel spel dat volgt op de speelgeschiedenis tot dat moment. Zie voor een formele definitie van een deelspel bijvoorbeeld Rasmusen (1994: 94).

pand in te zetten voordat ego de beslissing maakt wel of niet vertrouwen te schenken. Het inzetten van het onderpand maakt het minder gunstig voor alter om het surplus te houden in plaats van het evenredig te verdelen.[3] Dit maakt vervolgens het schenken van vertrouwen veiliger voor ego. In onze studie over effecten van het verleden, richten we ons op situaties waar alter geen onderpand inzet. Ego en alter spelen dan een deelspel spel A_B dat identiek is aan spel A, het vertrouwensspel. Experimentele evidentie (Snijders 1996) wijst uit dat een beslissing van alter om geen onderpand in te zetten de kans dat ego vertrouwen schenkt, gemiddeld genomen over verschillende speluitbetalingen, met 0.10 vermindert vergeleken met de kans dat zij dat doet in spel A.

Ten slotte vergelijken we ego's gedrag in een gevangenendilemma zonder onderpand (spel A) met zijn gedrag in een gevangenendilemma met onderpand (spel B). In beide spelen kiezen ego en alter tegelijkertijd tussen cooperatie en defectie. In het laatste spel mogen ego en alter voor deze beslissingen een onderpand inzetten dat defectie een onaantrekkelijke zet maakt. Als zowel ego en alter het nalaten een onderpand in te zetten bereiken ze het deelspel spel A_B dat identiek is aan spel A, het gevangenendilemma zonder onderpand. Experimentele evidentie van Mlicki (1996) toont aan dat defectie vaker in spel A_B wordt gekozen dan in spel A: ongeveer 91% versus 74%.

Uit de analyse van deze drie verschillende experimenten blijkt dat het gedrag in de situatie zonder verleden verschilt van het gedrag in de situatie met verleden. We concluderen uit de analyse in hoofdstuk 2 dat *verleden er inderdaad toe doet.* Een theorie over gedrag die geen rekening houdt met het verleden van een spel leidt duidelijk tot inadequate gedragsvoorspellingen. We stellen daarom twee hoofduitbreidingen voor van de theorie zoals beschreven in hoofdstuk 2. Ten eerste wordt besproken hoe de experimentele bevindingen verklaard zouden kunnen worden wanneer we gebruik maken van complexere nutsargumenten die de stricte aanname dat nut gelijk is aan de grootte van de 'taartpunt' verzachten. Aangezien we in eerste instantie liever niet nutsaannames problematiseren, maar verschillen in gedrag willen verklaren met verschillen in sociale situaties, concentreren we ons verder op de *complexere informatiestructuur* die voortkomt uit het verleden van het spel. Met andere woorden, we nemen aan dat het verleden van spel B een actor in spel A_B informatie verschaft. Zulke informatie zou echter niet relevant moeten zijn volgens onze simpele theorie. Proefpersonen gedragen zich echter zo dat het lijkt alsof ze belang hechten aan zulke informatie en hun gedrag daaraan aanpassen.[4] In hoofdstuk 3 bouwen we verder op de bevinding dat *verleden*

[3] Een onderpand is de vertrouwennemer wat waard en bij het schenden van vertrouwen verliest hij het onderpand wat de waarde van het houden van het gehele surplus vermindert. Zie bijvoorbeeld Schelling (1960) voor een bespreking van onderpanden in sociale interacties, Raub en Keren (1993) voor een theoretische analyse en eerdere empirisch evidentie, en Weesie en Raub (1996) voor een algemene verhandeling over onderpanden.

[4] Een voorbeeld ter illustratie: Ego's beslissing in een vertrouwensspel waar vertrouwen is geschonken lijkt op een dictatorspel (zie het eerste hierboven beschreven experiment). Aangezien alter's zet in het vertrouwensspel ego's alternatieven en uitbetalingen onveranderd laat, zou ego's gedrag niet mogen verschillen van zijn gedrag in een dictatorspel. Alter's beslissing brengt echter ego er toe van zijn gedrag in het dictatorspel af te wijken.

ertoe doet. We presenteren een model waarin actoren van het verleden leren en daarnaast met hun toekomstverwachtingen rekening houden bij het bepalen van hun gedrag in een huidige situatie.

VERTROUWEN EN DYADISCH SOCIAAL KAPITAAL: THEORIE

Hoofdstuk 3 stelt tegelijkertijd de effecten van het verleden en de toekomst op gedrag aan de orde. We proberen een antwoord te krijgen op de vraag welke effecten van temporele inbedding de productie van een surplus bevorderen dan wel in de weg staan. Net als in hoofdstuk 2 is de productie van een surplus geconceptualiseerd binnen een vertrouwensspel. Door het vertrouwensspel te herhalen kunnen we systematisch onderzoeken welke aspecten van het verleden en de toekomst van invloed zijn op de productie van een surplus. Onze hypothesen over het gedrag van de vertrouwengever zijn gebaseerd op een eenvoudig idee. Of de vertrouwengever wel of niet investeert in de productie van een surplus hangt af van *de hoeveelheid sociaal kapitaal in haar relatie met de vertrouwennemer.* We nemen aan dat sociaal kapitaal voortkomt door de interacties uit het verleden met de vertrouwennemer (het 'leer effect') maar ook de gemeenschappelijke toekomst van hun relatie weerspiegelt (het 'controle effect'). Daarbij wordt aangenomen dat het gedrag van de vertrouwennemer een simpele *vuist regel* volgt: hoe groter de belangen die op het spel staan, des te meer sociaal kapitaal is nodig om de vertrouwengever ertoe te brengen de noodzakelijke hulpbronnen te verschaffen voor de productie van een surplus. De grootte van belangen hangt af van hoeveel de vertrouwengever kan winnen of verliezen door het schenken van vertrouwen, maar ook van hoeveel de vertrouwennemer kan winnen door het schenden van vertrouwen. We gebruiken de uitbetalingsratio's RISICO := $\frac{P_1 - S_1}{R_1 - S_1}$ en VERLEIDING := $\frac{T_2 - R_2}{T_2 - S_1}$ als maten voor belangen in het spel.[5] Het RISICO kan uit het idee van subjectief verwacht nut worden afgeleid (Snijders 1996: h. 2; Coleman 1990: 97–102) terwijl de VERLEIDING volgt uit een schuld–model (Snijders 1996: h. 2). Snijders laat met een experiment met eenmalige vertrouwensspelen zien dat de kans dat de vertrouwengever vertrouwen schenkt afneemt naarmate RISICO en VERLEIDING groter worden.

De uitkomsten van interacties uit het verleden geven informatie die het mogelijk maakt *te leren.* Informatie uit eerdere ervaringen kan zowel van positieve als van negatieve aard zijn. Informatie is negatief als de vertrouwennemer eerder het gehele surplus 'opeiste'. Informatie is positief als het surplus zo werd verdeeld dat zowel de vertrouwengever als de vertrouwennemer ervan profiteerden. We nemen aan dat de voorraad sociaal kapitaal afneemt met negatieve uitkomsten maar toeneemt met positieve uitkomsten. We nemen ook aan dat de mate waarin de voorraad kapitaal krimpt of groeit afhangt van de eerdere situatie, namelijk, van hoe laag of hoe hoog de belangen waren die op het spel stonden. Laten we bijvoorbeeld aannemen dat de vertrouwen-

[5]Als in het vertrouwensspel de vertrouwengever geen vertrouwen geeft, verdienen zij en de vertrouwennemer respectievelijk P_1 en P_2. Als geschonken vertrouwen wordt gehonoreerd, verdienen ze respectievelijk R_1 en R_2. Als geschonken vertrouwen wordt geschonden, ontvangt de vertrouwengever S_1 terwijl de vertrouwennemer T_2 krijgt. De uitbetalingen zijn als volgt geordend: $R_1 > P_1 > S_1$ en $T_2 > R_2 > P_2$.

nemer het gehele surplus heeft gehouden en vertrouwen heeft geschonden. Hoe hoger het surplus was, hoe groter de verleiding voor de vertrouwennemer was en des te meer het door de vertrouwengever kan worden 'geaccepteerd' dat het surplus in zijn geheel door de vertrouwennemer is gehouden. Een uitkomst waar de vertrouwennemer een hoger surplus voor zichzelf houdt is dus minder verrassend voor de vertrouwengever. Het sociaal kapitaal neemt dan af, maar in mindere mate dan wanneer een kleiner surplus door de vertrouwennemer wordt geïncasseerd. We beschouwen een vertrouwennemer die een klein surplus voor zichzelf houdt dus als minder te vertrouwen dan een vertrouwennemer die een groot surplus voor zichzelf houdt. Op dezelfde manier draagt het delen van een groot surplus meer bij aan sociaal kapitaal dan het delen van een kleiner surplus. Daarnaast nemen we aan dat een vertrouwengever meer belang hecht aan uitkomsten uit het recente verleden dan aan die uit het verre verleden: hij verdisconteert dus uitkomsten uit het verleden.

Het tweede aspect van temporele inbedding is de toekomst die de vertrouwengever en vertrouwennemer samen hebben. Als de vertrouwengever en vertrouwennemer elkaar nog een keer zullen tegenkomen heeft de vertrouwengever de mogelijkheid te reageren op het huidige gedrag van de vertrouwennemer. Dit wordt *reciprociteit* genoemd. Goed gedrag van de vertrouwennemer kan worden beloond door het ter beschikking stellen van in de toekomst van nut zijnde hulpbronnen. Aan de andere kant kan slecht gedrag worden bestraft door het niet meer ter beschikking stellen van hulpbronnen in toekomstige interacties. De vertrouwennemer loopt dan zijn aandeel in toekomstige opbrengsten mis. Hoe langer de wederzijdse toekomst, des te groter het sociaal kapitaal van de vertrouwengever, aangezien er meer mogelijkheden zijn het gedrag van de vertrouwennemer te beantwoorden. Dit wordt het *controle effect* van de toekomst genoemd. De effectiviteit van het controle effect hangt af van de omvang van de mogelijke toekomstige opbrengsten. Bijvoorbeeld, een groot surplus aan zich voorbij zien gaan is een zwaardere sanctie voor de vertrouwennemer dan het mislopen van een kleiner surplus. Hoe vaker er transacties in de toekomst zullen zijn, des te groter het huidige sociaal kapitaal van de vertrouwengever.

De theorie van sociaal kapitaal zoals we die in hoofdstuk 3 hebben ontwikkeld voorspelt hoe het geven van vertrouwen afhankelijk is van temporele inbedding. Een toepassing van hoe actoren in werkelijkheid omgaan met vertrouwenskwesties is te vinden in het inkoopmanagement van bedrijven. In **hoofdstuk 4** staan daarom relaties tussen inkopers en leveranciers centraal. We nemen aan dat economische transacties gezien kunnen worden als (een variant van) een vertrouwensspel. In het algemeen is er in economische transacties sprake van een tijdsvertraging tussen het moment waarop de transactie geregeld wordt en de werkelijke levering. Deze tijdsvertraging creëert allerlei onzekerheden voor de inkoper ('probleempotentieel'). Voorbeelden zijn dat goederen te laat worden geleverd of in het geheel niet. Voor bepaalde transacties zal dit probleempotentieel hoog zijn, bijvoorbeeld als het transactievolume groot is of als de gevolgen van te laat leveren voor de inkoper groot zijn. Een inkoper die bereid is zulke onzekerheden te accepteren laat de controle over aan de leverancier en geeft in die zin vertrouwen. In hoofdstuk 4 wordt aangenomen dat inkopende ondernemingen proberen het noodzakelijke vertrouwen te verminderen door tijd en moeite te steken in *ex ante management* van de transactie. Zo investeren ze in het onderhandelen over contractuele overeenkomsten. Het treffen van veel voorzorgsmaatregelen voor transacties duidt op een gebrek aan vertrouwen in de vertrouwennemer. Anderzijds beschrijft Macaulay (1963) dat inkopers en leveranciers vaak zulke contractuele investeringen beperken ten gunste van vertrouwen in de partner. Wanneer de inkoper herhaaldelijk zaken doet met dezelfde leverancier (temporele inbedding), wordt de investering van de inkoper in ex ante management niet alleen bepaald door het probleempotentieel van de huidige transactie, maar ook door hun gemeenschappelijke verleden en hun verwachte toekomst.[6] In hoofdstuk 4 stellen we dat de theorie van sociaal kapitaal bij vertrouwen ook gebruikt kan worden voor het voorspellen van de investering van de inkoper in het ex ante management van de betreffende transactie. In hoofdstuk 4 worden de hypotheses uit hoofdstuk 3 getoetst in het specifieke geval van inkoper–leverancierrelaties.

Vertrouwen en dyadisch sociaal kapitaal: empirische resultaten

Hypothesen over het gedrag van de vertrouwengever in een temporele setting worden in eerste instantie getoetst met laboratorium experimenten (hoofdstuk 3). De proefpersonen waren hoofdzakelijk eerstejaars studenten sociologie van de Ludwig–Maximilian Universiteit in München. Zij kregen een aantal herhaalde vertrouwensproblemen voorgeschoteld en moesten daarbij in de rol van de vertrouwengever beslissingen nemen; ze moesten dus beslissen al dan niet vertrouwen te schenken. De beslissingen van de proefpersonen werden gemeten met de strategie–methode (zie Roth 1995). Voor elk mogelijk verloop van het spel moeten proefpersonen aangeven hoe ze zich zouden gedragen. Deze methode maakt het mogelijk informatie over het gedrag van de proefpersoon in verschillende condities van temporele inbedding te verzamelen.

Ten tweede bekijken we of onze hypotheses over vertrouwen in een temporele setting een toets buiten het sociologische laboratorium kunnen doorstaan. In hoofdstuk 4 worden daartoe gegevens over inkoper–leverancierrelaties in de Nederlandse IT–industrie gebruikt. ('The External Management of Automation', MAT, zie Buskens en Batenburg 2000). Zoals boven beschreven, kan de investering van de inkoper in het ex ante management van de transactie worden gezien als een benadering van het gebrek aan vertrouwen van de inkoper in de leverancier. De data bevatten informatie om te beo-

[6] Het probleempotentieel van de transactie kan worden geïnterpreteerd in termen van de uitbetalingen of verschillen in uitbetalingen in het vertrouwensspel. Het probleempotentieel van een transactie wordt geoperationaliseerd via het financiële volume van de transactie en het opportunismepotentieel van de leverancier. Het financiële volume dient als een algemene maat voor het verlies van de inkoper wanneer vertrouwen wordt geschonden (nl., $P_1 - S_1$ versus $R_1 - S_1$) maar ook als een maat voor de drijfveer van de leverancier om vertrouwen te misbruiken (nl., $T_2 - R_2$). Het opportunismepotentieel wordt afgeleid als een gewogen gemiddelde van het schadepotentieel en de toezichtproblemen (alsmede het financiële volume van de transactie). Het schadepotentieel geeft weer 'hoe erg het zou zijn' wanneer zou blijken dat het geleverde product niet productief kan worden gebruikt en reflecteert direct de S_1–waarde (zie paragraaf 4.4.3).

ordelen hoeveel tijd en moeite de inkoper heeft geïnvesteerd in het ex ante management van de betreffende transactie.

Tabel 6.1 in hoofdstuk 6 biedt een overzicht van de hypotheses en van de resultaten van de experimenten en het survey–onderzoek. De laatste twee kolommen van tabel 6.1 laten zien dat onze bevindingen uit beide toetsen redelijk met elkaar overeenkomen. De hypotheses over het negatieve effect van toenemend risico in een lopende transactie worden bevestigd in beide analyses. In de experimentele toetsing wordt gevonden dat een toename in de belangen die op het spel staan (Risico en Verleiding) de kans laat afnemen dat de vertrouwengever vertrouwen schenkt. Ook de surveydata laten zien dat een toename in zowel het volume als het opportunismepotentieel van een lopende transactie een positief effect heeft op de investering van de inkoper in ex ante management. Een meer risicovolle lopende transactie gaat samen met minder bereidheid van de inkoper om de leverancier te vertrouwen en dit zet hem aan tot meer investeringen in het ex ante management van de betreffende transactie.

We vinden in beide datasets ook evidentie voor het veronderstelde leer effect van het verleden. Uit de experimentele data blijkt dat de kans dat de vertrouwengever vertrouwen schenkt hoger is bij een succesvol verleden waarin vertrouwen werd gehonoreerd door de vertrouwennemer. Daarnaast vinden we evidentie voor een negatief effect op de kans dat de vertrouwengever vertrouwen schenkt als het ooit eerder is geschonden. In de surveydata kan helaas geen onderscheid worden gemaakt tussen positieve en negatieve ervaringen. Voor inkopers die eerder transacties hadden met dezelfde leverancier waren die transacties over het algemeen genomen positief. We kunnen aannemen dat negatieve ervaring inkopers ertoe heeft bewogen nieuwe leveranciers te zoeken. Met de surveydata kan daarom alleen de hypothese over het effect van een succesvol verleden getoetst worden. Een goed verleden met de huidige leverancier laat de management–investering van de inkoper afnemen. Voor de inkoper–leverancier relaties met een verleden bekijken we verder of het aantal eerdere transacties daarnaast nog een negatief effect heeft op de investering van de inkoper in het beperken van risico's. De data bevestigen dit inderdaad. Inkopers die aangeven een gemeenschappelijk verleden met de huidige leverancier te hebben investeerden significant minder in ex ante management wanneer die transacties frequenter hadden plaatsgevonden (zie tabel 4.6 in hoofdstuk 4).

Door beperkingen in de surveydata kunnen twee hypotheses over de effecten van het verleden alleen op de experimentele data worden getoetst. Ten eerste stellen we dat eerdere uitkomsten worden verdisconteerd. Hiervoor wordt ondersteuning gevonden. Ten tweede nemen we aan dat de kans dat de vertrouwengever vertrouwen schenkt, ceteris paribus, niet afhangt van of vertrouwen in een eerder spel niet was geschonken of dat dit spel überhaupt niet is gespeeld. De experimentele data bieden hiervoor enige, maar geen eensluidende evidentie. Wat betreft de inkoper–leverancier data zou dit betekenen dat zolang de inkoper geen interactie heeft gehad met de leverancier (en ook geen informatie over het gedrag van de leverancier heeft verkregen uit interacties met anderen), hij niets kan leren over de leverancier. Met de surveydata kan deze hypothese niet worden getoetst. De MAT–data bevat geen informatie of leveranciers in eerdere transacties reeds potentiële maar geen daadwerkelijke leverancier waren.

De hypothese over het controle effect van de toekomst kan echter wel met beide datasets worden getoetst. Deze hypothese stelt dat hoe langer de verwachte toekomst is, des te waarschijnlijker het is dat een vertrouwengever vertrouwen schenkt. De experimentele data bevestigen dit controle effect. De inkoper–leverancier data laten geen effect van de toekomst zien. Dit resultaat wijst echter mogelijk op een aanzienlijk verschil tussen het experiment zoals besproken in hoofdstuk 3 en de inkoper–leverancierrelaties van hoofdstuk 4. In het experiment toetsen we het 'werkelijke' controle effect dat werkt via het tit–for–tat mechanisme van reciprociteit. In de relatie van de inkoper met de leverancier zouden de verschillende mogelijkheden voor reciprociteit op dezelfde manier tot mindere investeringen in management moeten leiden. Daarnaast kan het inkoopmanagement echter ook afhangen van wat we het investeringseffect noemen. Zo'n effect suggereert meer investeringen bij een langere verwachte toekomst. Het idee is dat management–investeringen kunnen worden hergebruikt voor het veilig stellen van toekomstige transacties. Het controle effect zou op die manier afgezet moeten worden tegen het investeringseffect zodat de MAT–data mogelijkerwijs geen effect van verwachte toekomstige transactie op het ex ante management van de huidige transactie konden laten zien.

Tenslotte zijn er twee hypothesen waarvoor geen empirische ondersteuning wordt gevonden in de twee datasets. Enerzijds voorspelden we een interactie tussen uitkomsten in het verleden en de inzet van het betreffende spel in het verleden, anderzijds tussen de lengte van de verwachte toekomst en de inzet van de toekomstige spelen. De experimentele data laten noch een interactie tussen positieve uitkomsten in het verleden en Verleiding noch tussen negatieve uitkomsten in het verleden en verleiding zien. Ook blijken het eerdere volume en opportunismepotentieel geen significante interactie met succesvolle transacties in het verleden te hebben.[7] De interactie die we voorspelden over de lengte van de verwachte toekomst en de Verleiding in deze toekomstige spelen wordt ook niet ondersteund door de experimentele data. Aangezien de surveydata niet de noodzakelijke informatie bevat over toekomstige transacties, kan deze hypothese hiermee niet worden getoetst.

Samengevat duidt de empirische toets van onze hypothesen erop dat het schenken van vertrouwen door een vertrouwengever of inkoper theoretisch het best kan worden omschreven als een functie van drie verschillende kenmerken van temporele inbedding: de inzet van het huidige spel, de uitkomsten in het verleden en de lengte van de toekomst. Een sociaal kapitaal model dat deze drie kenmerken bevat wordt in hoofdstuk 3 een *simpel sociaal kapitaal model* genoemd. Bij een complex sociaal kapitaal model wordt ook rekening gehouden met kenmerken van transacties in het verleden en de toekomst. Onze data wijzen uit dat het simpele sociaal kapitaal model voldoet

[7] Bedenk hierbij dat de MAT–data geen informatie over onsuccesvolle zaken in het verleden met de huidige leverancier bevatten. De interactie tussen de belangen die op het spel stonden en een onsuccesvol verleden kan daarom niet worden getoetst.

voor de verklaring van waargenomen gedrag in experimentele vertrouwensspelen en IT–transacties.

Vertrouwen: suggesties voor verder onderzoek

Noch de experimentele noch de survey data verschaffen ons evidentie voor het belang van eerdere of latere spelinzetten. Zoals we uitvoeriger bespraken in hoofdstuk 4, bevatten de surveydata niet de noodzakelijke variabelen om op een goede manier de belangen die in het verleden en de toekomst op het spel staan of stonden te meten (nl., het volume en opportunismepotentieel van transacties in het verleden en in de toekomst). Met de experimentele data daarentegen kunnen inzetten in het verleden en de toekomst wel worden geconceptualiseerd, door middel van Risico en Verleiding. Hoe kan het nu dat we geen effect van inzetten in eerdere en latere spelen vinden? Bedenk dat er slechts twee verschillende uitbetalingsstructuren in de experimenten aan de orde kwamen. Toen we de experimenten ontwierpen waren er goede redenen om ons te beperken tot twee uitbetalingsstructuren. In verder onderzoek zou het effect van inzetten in het verleden en in de toekomst of gedrag in herhaalde vertrouwensspelen echter wel grondig moeten worden onderzocht. Hiervoor zijn experimenten nodig met herhaalde vertrouwensspelen en veel verschillende Risico– en Verleiding–ratio's. Wanneer we meer dan slechts lage en hoge inzetten kunnen onderscheiden zal het aantal experimentele condities echter enorm toenemen.[8] Al hoeft men niet elke uitbetalingscombinatie te gebruiken, voor een experimenteel design met een grotere variantie in inzetten is een hoger aantal proefpersonen nodig.

Ten tweede namen we in het kapitaalvoorraadmodel aan dat uitkomsten uit het verleden worden verdisconteerd. Zoals we hierboven beschreven volgt uit deze assumptie bijvoorbeeld dat vertrouwenschending vergeven kan worden of dat ook een goed lopende (vertrouwensvolle) relatie inspanningen nodig heeft om haar goede staat te behouden. De analyses in hoofdstuk 3 laten zien dat uitkomsten in het verleden inderdaad worden verdisconteerd. Het ligt voor de hand deze bevinding te generaliseren naar gevallen van meer dan twee tijdsperiodes en naar een sociaal kapitaalmodel met *exponentiële verdiscontering* van vroegere uitkomsten. Een model met exponentiële verdiscontering is echter slechts één mogelijkheid. Rationele–keuzemodellen nemen in het algemeen tijdsconsistent gedrag aan. Dit impliceert exponentiële verdiscontering (zie Becker 1996). Aan de andere kant wijst evidentie uit zowel experimenteel als survey onderzoek erop (bijv., Braun 1998: h. 4; Fehr en Zych 1998; Loewenstein en Elster 1992) dat gedrag soms tijdsinconsistent is. Zulk gedrag wordt voorspeld uit hyperbolische verdisconteringsmodellen. Zouden deze empirische bevindingen dan suggereren dat ook het verleden hyperbolisch in plaats van exponentieel wordt verdisconteerd? De experimentele resultaten van hoofdstuk 3 zijn gebaseerd op maximaal twee herhalingen van het vertrouwensspel. Een antwoord op hoe het verleden wordt

[8] Bedenk dat we al $2 \times 2 \times 4 \times 8 = 128$ verschillende uitbetalingscombinaties hadden die lage en hoge inzetten in het verleden hadden in combinatie met twee eenmalige spelen, een een–keer–herhaald spel, en een twee–keer–herhaald spel.

verdisconteerd kan daarom niet op basis van de ons ter beschikking staande experimentele data worden gegeven. Het is echter een interessante vraag en het zou de moeite waard zijn om onderzoek op te zetten waarin vertrouwensspelen vaker worden herhaald dan in onze studie. Uiteraard is 'een aantal keer' vaag. Het precieze aantal herhalingen van het vertrouwensspel waarbij een bevredigende toets van verschillende hypotheses over verdiscontering mogelijk wordt is moeilijk te geven. Een vier–keer–herhaald vertrouwensspel is waarschijnlijk een goed begin, maar tien keer herhalen is natuurlijk veel beter. Een bruikbaar design voor het simultaan meten van de effecten van inzetten in het verleden en in de toekomst zou een vier–keer–herhaald vertrouwensspel zijn met meer dan twee verschillende uitbetalingsstructuren.

Een laatste opmerking betreft de afleiding van de hypotheses in hoofdstuk 3 en hoofdstuk 4. In beide hoofdstukken werd aangenomen dat het gedrag van respectievelijk de vertrouwengever en de vertrouwennemer afhankelijk is van een sociaal–kapitaalvoorraad. Deze aanname wordt in grote mate ondersteund door de empirische bevindingen. Echter, om tot deze hypotheses te komen gebruiken we *verbale theorie*. Met het model zoals gepresenteerd in hoofdstuk 3 kunnen geen gedetailleerde voorspellingen over het gedrag van de vertrouwengever worden gemaakt, zoals voorspellingen over hoeveel de kapitaalvoorraad precies groeit na een positieve uitkomst in een spel met een inzet gerepresenteerd door een Verleiding van zeg 0.85. Een formeel model zou mogelijk maken het relatieve effect van leren ten opzichte van controle te voorspellen. Anders gezegd, men zou misschien in staat zijn een antwoord te krijgen op de vraag of het verleden dan wel de toekomst belangrijker is in een bepaalde situatie. Tenslotte zou in zo'n model de beslissing van de vertrouwengever om al dan niet vertrouwen te schenken moeten worden geconceptualiseerd als een gevolg van maximaliserend gedrag gegeven de specifieke temporele inbedding.

Vertraagde overeenkomst en onderhandelen: de verdeling van een surplus

In het tweede deel van het boek, **hoofdstuk 5**, wordt de verdeling van een surplus behandeld als een 'Alternating Offers Game' waarin twee actoren onderhandelen over hun aandeel in het surplus (Rubinstein 1982). In het spel wordt aangenomen dat ego en alter om beurten voorstellen en tegenvoorstellen doen over de verdeling van een surplus totdat ze tot een overeenkomst komen. Elke keer dat ego een voorstel doet heeft alter de keuze het voorstel of te accepteren of het af te wijzen om vervolgens op het daaropvolgende tijdstip met een eigen voorstel te komen. Ego kan dan op zijn beurt alter's voorstel accepteren dan wel afwijzen en later met een nieuw voorstel komen; enzovoort. Echter, het uitstellen van een overeenkomst levert de actoren kosten op. Deze kosten komen terug in de verdisconteringsfactoren van de actoren die een indicatie geven van hun geduld bij het uitstellen van de overeenkomst over de verdeling van een surplus.

Net zoals in het eerste deel van het boek maken we bepaalde aannames over het gedrag van ego en alter. Ten eerste nemen we aan dat de actoren zich geen zorgen hoeven te maken over de totstandkoming van het surplus. Het te verdelen surplus is beschikbaar. Ten tweede nemen we wederom aan dat het nut van een actor wordt gegeven door de grootte van zijn of haar taartpunt en dat de actor liever een groot stuk heeft dan een kleiner stuk. Tenslotte nemen we ook aan dat de actoren zich gedragen volgens speltheoretische rationaliteit, dat wil zeggen: zij volgen Nash evenwichtgedrag.

Rubinstein (1982) laat zien dat onder deze aannames een evenwicht bestaat waarbij de verdeling van een surplus alleen afhangt van de ratio van de verdisconteringsfactoren van de actoren. Anders gezegd, het surplus wordt verdeeld naar rato van het geduld dat de actoren hebben bij het uitstellen van de overeenkomst over de verdeling van het surplus. Een actor die geduldiger is ondervindt lagere kosten bij vertraging en kan langer wachten, wat zijn onderhandelingspositie ten goede komt. We laten in hoofdstuk 5 zien hoe 'geduldig zijn' kan worden vertaald in de absolute *onderhandelingssterkte* ('bargaining power') van een actor.

De kosten van het uitstellen van de overeenkomst, en dus hun absolute onderhandelingssterktes, zijn in Rubinstein's model exogeen. Om tot voorspellingen te komen over de verdeling van het surplus en om die met experimentele evidentie te toetsen is informatie nodig over de onderhandelingssterktes van de actoren. We verrijken het scenario van dyadisch onderhandelen met een sociologisch relevante uitbreiding waaruit de onderhandelingssterktes van de actoren worden afgeleid. In hoofdstuk 5 volgen we Emerson (1972, 1981) door aan te nemen dat ruil tussen twee actoren niet onafhankelijk moet worden gezien van hun relaties met andere actoren. Dyadisch onderhandelen wordt beïnvloed door de inbedding van de actoren in een netwerk van andere onderhandelingspartners. Emerson volgend stellen we voor dat de netwerkinbedding van ego en alter kan worden gezien als de bron van hun onderhandelingssterkte. Vanuit het gezichtspunt van een actor bepalen netwerken de alternatieven voor dyadische ruil. Zodoende worden verschillen in opportuniteitsstructuren gedefiniëerd tussen actoren. Die komen terug in de onderhandelingssterkte van de actoren. De eerste onderzoeksvraag van het tweede deel van het boek is: *Hoe bepaalt de inbedding van een actor in een netwerk van onderhandelingspartners zijn of haar aandeel in de surplus in relaties met zijn of haar partners?*

Ingebed zijn in een netwerk biedt de actoren met meer dan één partner toegang tot verschillende onderhandelingspartners met wie een verdeling van een surplus mogelijk is. Zoals in hoofdstuk 5 in detail wordt uitgelegd kan de eigenlijke ruil in sommige netwerken tot één relatie beperkt zijn ten gevolge van exogeen gegeven regels. Een tweede onderzoeksvraag is daarom: *Met wie van zijn of haar partners verdeelt een actor een surplus wanneer beperkingen ervoor zorgen dat een actor slechts één surplus per keer kan verdelen?*

Onderhandelen en ruil: structureel ingebedde relaties

Als startpunt voor het bepalen van de absolute onderhandelingssterktes van de actoren richten we ons op een exogeen gegeven netwerk van actoren. Dit netwerk beschrijft onderhandelingsrelaties tussen actoren. Een positie in een netwerk kan worden 'omschreven' door drie afzonderlijke elementen. Alleen door tegelijkertijd met deze drie elementen rekening te houden kan de absolute onderhandelingssterkte van een actor worden bepaald. Ten eerste wordt elke positie in een netwerk gekarakteriseerd door het aantal beschikbare onderhandelingspartners. Relaties bieden ruilmogelijkheden. Een actor die veel onderhandelingsrelaties heeft is minder afhankelijk van een specifieke partner. Ten tweede heeft elk van deze onderhandelingspartners eigen onderhandelingspartners met wie zij ook kunnen onderhandelen. Het aantal partners van partners is daarom ook belangrijk. Hoe minder partners ego's partners hebben, des te afhankelijker ego's partners voor ruil van hem zijn. Tenslotte is het belangrijk of onderhandelingsverbindingen in netwerken negatief of positief zijn ('relational assessment'). Verbindingen zijn negatief wanneer de partners, gezien vanuit de eigen positie, alternatieven voor elkaar zijn (Emerson 1972: 70–71). Ruil in één van deze relaties verhindert dan ruil in de overige relaties. Emerson stelt dat verbindingen ook positief kunnen zijn. Dit betekent dat ruil in de ene relatie ruil in andere relaties stimuleert. In dat geval zijn partners wat ruil betreft complementair.

Wanneer het netwerk eenmaal is bekend en haar verbindingen zijn geklassificeerd als positief of negatief, kunnen we de onderhandelingssterkte van iedere actor bepalen. In een negatief verbonden netwerk neemt ego's onderhandelingssterkte toe met het aantal relaties waarover hij beschikt, maar neemt ze af met het aantal relaties waarover zijn partners beschikken. In een positief verbonden netwerk, daarentegen, zijn relaties complementair en ruil in de één bevordert die in andere. In zo'n netwerk is afhankelijk zijn van anderen een voordeel. Ego's absolute onderhandelingssterkte neemt dan dus toe met het aantal partners dat hij heeft.

De absolute onderhandelingssterkte bepaalt echter niet volledig ego's aandeel in het surplus in de relatie met alter. Ego's aandeel van het surplus in de relatie met alter is afhankelijk van zijn relatieve onderhandelingssterkte met betrekking tot alter: ego's absolute onderhandelingssterkte in relatie tot alter's absolute onderhandelingssterkte. Een hoge absolute onderhandelingssterkte verzekert niet automatisch een grotere taartpunt. Actoren met gelijke absolute onderhandelingssterktes (zoals structureel equivalente actoren) verdelen de taart in twee gelijke stukken.

Substitutie van de absolute onderhandelinsgssterktes in de evenwichtsoplossing van het afwisselende–voorstellenspel levert voorspellingen op voor de verdeling van een surplus in alle onderhandelingsrelaties in een netwerk. Om onze theoretische voorspellingen te toetsen maken we gebruik van gepubliceerde resultaten van sociologische netwerk–ruiltheorie (voor een overzicht zie bijv. Willer 1999). Verder confronteren we onze voorspellingen met die van prominente ruiltheorieën.

Onderhandelen en ruil: experimentele resultaten

We toetsen onze voorspellingen over de verdeling van een surplus door gebruik te maken van in de literatuur gepubliceerde experimentele resultaten (Burke 1997; Cook et al. 1983; Lovaglia et al. 1995; Markovsky, Willer, en Patton 1988; Markovsky et al. 1997; Skvoretz en Willer 1993; Yamagishi, Gillmore, en Cook 1988; Yamaguchi 1996). Bij laboratoriumexperimenten van netwerk–ruiltheorie wordt gewoonlijk een gestandaardiseerd protocol gebruikt. Proefpersonen worden toegewezen aan een specifieke positie binnen een netwerkstructuur die wordt bepaald door de experimentleider. De netwerkstructuur is bij iedereen bekend. Na een inleiding en een testronde kunnen proefpersonen onderhandelen over de verdeling van een surplus van 24 punten met elk van de buren waarmee ze zijn verbonden in het netwerk. In het algemeen worden zo 20 ronden gespeeld. Op één uitzondering na (Yamagishi, Gillmore, en Cook 1988) worden bij experimenten negatieve verbindingen verondersteld en in het algemeen wordt het aantal ruilmogelijkheden per actor en ronde beperkt tot één (‘one–exchange rule’).

De voorspellingen van ons model worden in drie stappen getoetst. Bij elke stap vergelijken we daarbij onze voorspellingen met voorspellingen van andere benaderingen: Lovaglia et al.’s (1995) GPI–RD, Yamaguchi’s (1996) power model, Skvoretz en Willer’s (1993) Exchange Resistance theory, Friedkin’s (1986, 1992) Expected Value theory, Cook en Yamagishi’s (1992) Equi–Dependence theory, and Burke’s (1997) Identity Simulation model. Eerst stellen we onze voorspellingen over de verdeling van een surplus in negatief verbonden netwerken tegenover de betreffende empirische resultaten. Dan vergelijken we onze voorspellingen over winstverdelingen in positief verbonden netwerken met het enig experiment over positief verbonden netwerken dat in de literatuur is beschrijven. Deze twee stappen hebben betrekking op de eerste onderzoeksvraag van hoofdstuk 5. Van sommige negatief verbonden netwerkstructuren met een enkele–ruil regel is empirisch waargenomen dat ze ‘breken’. Hiermee wordt bedoeld dat tenminste een actor voortdurend afziet van ruil met minstens één onderhandelingspartner. In de derde stap richten we ons daarom op de tweede onderzoeksvraag van hoofdstuk 5. We bekijken of ons model afwijkingen tussen onderhandelings– en ruilstructuren correct voorspelt.

Voorspellingen over de verdeling van een surplus in negatief verbonden netwerken worden naast experimentele bevindingen met betrekking tot vier netwerkstructuren gelegd: 4–Line, Stem, Kite, en 3–Branch.[9] De geobserveerde verdelingen van het surplus kloppen redelijk met de voorspellingen van alle theoretische benaderingen die we in onze analyse in beschouwing nemen. Ze verschillen echter in de mate van precisie. Niet alle theorieën doen het bij dezelfde netwerken even goed. Iedere theorie kent in ieder geval één netwerk waar de voorspellingen slecht bij de data passen. Onze theorie bijvoorbeeld klopt niet goed met de 4–Line structuur. We voorspellen dat van een surplus van 24, ego er 15.9 krijgt. In werkelijkheid blijkt dit 14.1 te zijn. Met betrekking tot de andere drie netwerken biedt onze theorie voorspellingen die allemaal binnen het

[9]Zie figuur 5.1 in hoofdstuk 5 voor een grafische voorstelling van deze netwerken.

95% betrouwbaarheidsinterval van de experimenteel waargenomen variabelen vallen. We maken gebruik van twee goodness–of–fit maten: de gemiddelde afwijking en de absolute afwijking. Op basis van beide maten passen onze voorspellingen beter by de data dan die van de andere theorieën.

In de tweede stap toetsen we onze voorspellingen over de verdeling van een surplus in positief verbonden netwerken. Empirische bevindingen over positief verbonden netwerken zijn schaars en we kunnen ons slechts beroepen op een enkel experiment over een 5–LINE structuur (Yamagishi, Gillmore, en Cook 1988). Uit het enorme aanbod van theorieën is alleen Yamaguchi's (1996) power model in staat tot voorspellingen te komen voor positief verbonden netwerken. Yamagishi et al. presenteren naar aanleiding van hun data alleen de winstaandelen van actoren. Dit betekent dat zij de winsten die een actor in al zijn dyadische onderhandelingsrelaties verkregen heeft optellen en dan het resultaat zo standaardiseren dat deze winstaandelen van alle actoren optellen tot 1. Men rapporteert winstaandelen voor de posities in de positief verbonden 5–LINE A–B–C–B–A als: A = 0.1503, B = 0.1931, en C = 0.3133. Voor alle posities voorspellen we winstsplitsingen die dicht tegen de data aanzitten: A = 0.1632, B = 0.1736, en C = 0.3264. Deze voorspellingen passen ook beter dan die van Yamaguchi: A = 0.0020, B = 0.2500, en C = 0.4961.

In de derde stap toetsen we onze voorspellingen over afwijkingen tussen onderhandelings– en ruilstructuren. In een aantal ruilexperimenten heeft de experimentleider bijgehouden hoeveel keer aan elkaar verbonden actoren met elkaar ruilden. Op basis van de resultaten analyseren enkele auteurs (bijv. Bienenstock en Bonacich 1993; Markovsky, Willer, en Patton 1988; Simpson en Willer 1999; Skvoretz en Lovaglia 1995) de afwijkingen tussen de onderhandelings– en de ruilstructuur. Een prominente kandidaat voor zulke 'breuken' is de T–SHAPE structuur (zie FIGUUR 5.2 in hoofdstuk 5) waarin twee actoren A_1 en A_2 ieder met een actor B verbonden is die zelf verbonden is met een actor C die op zijn beurt verbonden is met een perifere actor D. Markovsky, Willer, en Patton (1988: 227) doen melding van een breuk in deze structuur. In 100 onderhandelingsronden kwam slechts drie een keer ruil voor tussen B en C. Het netwerk viel dus uiteen in een 3–LINE A_1–B–A_2 en een DYAD C–D. Ons model voorspelt niet alleen op correcte wijze het ontstaan van deze breuk maar ook op precieze wijze de verdeling van de surplussen in de resulterende deelnetwerken. Van een surplus van 24 winstpunten bleek B er gemiddeld 19.14 te krijgen in relaties met de actoren op posities A.[10] Ons model voorspelt 19.86 winstpunten in het voordeel van B. C bleek 11.88 winstpunten te krijgen in de relatie met D—wij voorspelden een gelijke divisie van elk 12.0 winstpunten. Simpson en Willer (1999) doen melding van alternatieve netwerken waarvoor experimentele gegevens over het aantal keren ruil in alle relaties beschikbaar zijn: namelijk, een Y–TRIANGLE, XU–SHAPE, X4–LINE, V–TRIANGLE, en een H–SHAPE (zie FIGUUR 5.3 in het vorige hoofdstuk). Ons model

[10]De proefpersonen B realiseerden gedurende het experiment bijna indentieke winsten met zowel A_1 als A_2. De gerapporteerde 19.14 winstpunten vormen een gemiddelde van de winsten zoals gerealiseerd in relatie tot A_1 en A_2.

voorspelt breuken in al deze onderhandelingsnetwerken. Statistische toetsen op de data van Simpson en Willer ondersteunen onze voorspellingen. Voor het XU–Shape en het Y–Triangle netwerk hebben Simpson en Willer echter andere conclusies. Hun statistische toetsen wijzen uit dat deze netwerken niet breken.

Naast voorspellingen over de verdeling van een surplus en voorspellingen over netwerkbreuken stelt ons model ons tot verdere theoretische voorspellingen in staat. Ten eerste splitsen actoren in een Dyad hun surplus altijd in twee gelijke aandelen omdat ze structureel equivalent zijn en geen alternatieve ruilpartner hebben. Daarom is de Dyad een robuust netwerk (dwz., de ruilstructuur valt samen met de onderhandelingsstructuur en breekt dus niet). De beschikbare data ondersteunen deze voorspellingen. Ten tweede kunnen niet alle negatief verbonden netwerken breken. Wanneer we kijken naar die negatief verbonden netwerken waarin in alle relaties bedragen van gelijke waarde worden verdeeld, dan kan de structuur alleen breken wanneer aan de volgende voorwaarden wordt voldaan: (i) het netwerk moet minstens een actor met drie of meer ruilpartners bevatten, en (ii) deze ruilpartners moeten zich op tenminste twee structureel verschillende posities bevinden. Daaruit volgt dat netwerken met identieke bedragen waarin alle actoren ten hoogste twee ruilpartners hebben dus altijd robuust zijn. Zulke netwerken zijn óf ring– óf lijnstructuren. De beschikbare experimentele gegevens ondersteunen deze voorspelling. Ten derde zijn relaties in positief verbonden netwerken complementair en zulke structuren zijn daarom altijd robuust. Aangezien ruil in de ene relatie die in andere relaties stimuleert, zullen rationele actoren nooit één van hun mogelijke onderhandelingspartners buitensluiten. Wegens gebrek aan experimenten over positief verbonden netwerken kan hierover geen empirisch uitsluitsel worden gegeven. Ten vierde hangt het optreden van een breuk af van de relatieve onderhandelingssterkte van een actor in al zijn relaties. Echter, wanneer de te verdelen bedragen verschillen hangt de kans op een breuk ook af van de grootte van de te verdelen bedragen. In dergelijke onderhandelingsnetwerken kan het soms voordeliger zijn een kleiner deel van een groot surplus te krijgen dan een aantal grotere delen van kleinere surplussen. Een actor kan er soms dus voor kiezen uitgebuit te worden in de relatie die gekenmerkt wordt door een groot surplus. Er zijn geen empirische gegevens beschikbaar voor het toetsen van deze voorspelling. Ten vijfde bestaan er bepaalde onderhandelingsnetwerken waarin een actor een latente monopoliepositie bekleedt. Zo'n actor heeft volledige controle over twee of meer van zijn onderhandelingspartners—zij hebben geen andere onderhandelingspartner—maar is daarnaast verbonden met tenminste één partner waarover hij geen volledige controle heeft omdat de laatstgenoemde wel alternatieve onderhandelingspartners heeft. Een actor met zo'n latente monopoliepositie zal altijd minstens één van zijn minder afhankelijke onderhandelingspartners buitensluiten. Anders gesteld, zo'n actor zal proberen een pure monopoliepositie te verkrijgen. Experimenten die netwerkstructuren met latente monopolieposities onderzoeken laten inderdaad het voorspelde gedrag van de 'monopolie–actor' zien.

Onderhandelen en ruil: suggesties voor verder onderzoek

We zagen dat de in hoofdstuk 5 ontwikkelde theorie het in het algemeen beter doet dan de gevestigde theorieën van de sociologische netwerk–ruiltheorie. Deze conclusies rusten echter enkel op experimenten met kleine en eenvoudige netwerkstructuren. We komen met enkele voorspellingen voor grotere netwerken, met name in relatie tot onze resultaten over netwerkbreuken. Deze voorspellingen over de verdeling van een surplus moeten worden getoetst via experimenten met dergelijke complexere netwerkstructuren. Ook zouden zulke nieuwe experimenten een antwoord kunnen geven op de vraag of 'locale informatie' in plaats van volledige informatie over de hele netwerkstructuur inderdaad voldoende zijn voor precieze voorspellingen in grotere structuren.

Ten tweede rusten de voorspellingen van ons model op de onderliggende aanname dat onderhandelingssterkte een functie is van alleen de structuur van het netwerk. Deze aanname komt overeen met Emerson's (1972) basisidee en vervult ook de voorwaarden van de netwerk–ruiltheorie. Experimentele resultaten lijken er op te wijzen dat het redelijk is aan te nemen dat onze theorie ook de invloed van structuur op individueel gedrag goed beschrijft. Het is daarom verleidelijk een stap verder te gaan, zeker wanneer men probeert onderhandelingsgedrag in de werkelijkheid te verklaren. Eckel en Grossman (1998) vinden bijvoorbeeld dat mannen en vrouwen zich in onderhandelingen anders gedragen. Mannen accepteren lagere aanbiedingen van vrouwen dan van mannen. Vrouwen, daarentegen, accepteren lagere aanbiedingen van vrouwen. Op basis van Becker en Mulligan's (1997) resultaten over de endogenisering van tijdspreferentie kan worden verwacht dat bijvoorbeeld jongere of lager opgeleide personen minder geduldig zijn. Zij zouden zo tot minder gunstige overeenkomsten komen. Tenslotte lijkt het niet ondenkbaar dat ook welvaartseffecten een belangrijke rol kunnen spelen bij onderhandelingen. Rijke mensen zullen bijvoorbeeld minder scherp onderhandelen wanneer het een kleine taart betreft, terwijl arme mensen graag een klein stukje van een grotere taart zouden accepteren. Hieruit volgt dat onderhandelingssterkte een functie is van meer dan alleen structuur. Toekomstig onderzoek over netwerkruil zou daarom ook de effecten van zulke actorkenmerken aan de orde moeten stellen.

Als Emerson's (1972) voorstel dat ruiltheorie ook ruilstructuren verklaart serieus wordt genomen, zouden nieuwe theorieën in staat moeten zijn te verklaren hoe en wanneer actoren ruilrelaties met nieuwe partners aangaan buiten de gegeven onderhandelingsstructuur ('netwerkuitbreiding'). Cook en Whitmeyer (1992) stelden ook dat men bij netwerk–ruiltheorie meer de nadruk zou moeten leggen op het verklaren van veranderingen in netwerkstructuren in plaats van enkel winstverdelingen te voorspellen. In het geval van positief verbonden netwerken zou de aandacht voor netwerkuitbreidingen wel eens net zo belangrijk kunnen zijn als het bestuderen van netwerkbreuken in negatief verbonden netwerken. In positief verbonden netwerken zijn actoren complementair in ruil en het hebben van meer relaties kan de eigen onderhandelingssterkte laten toenemen. Een meer volledige theorie over netwerkruil zou daarom ook moeten toestaan dat nieuwe actoren een gegeven onderhandelingsstructuur binnenkomen en zou een gedragsconcept op individueel niveau moeten definiëren voor de acceptatie dan wel afwijzing van zulke nieuwe onderhandelingspartners. Of actoren in positief verbonden netwerken zich wel of niet volgens zulke aannames gedragen kan vervolgens in laboratoriumexperimenten worden getoetst.

References

Ahn, T. K., E. Ostrom, D. Schmidt, R. Shupp, and J. Walker (1998) "Cooperation in PD Games: Fear, Gread, and History of Play." Working Paper of the Workshop in Political Theory and Policy Analysis, Indiana University.

Ahn, T. K., E. Ostrom, D. Schmidt, R. Shupp, and J. Walker (1999) "Generating Distrust in PD Games: Fear, Gread, and History of Play." Working Paper of the Workshop in Political Theory and Policy Analysis, Indiana University.

Ahn, T. K., E. Ostrom, and J. Walker (1998) "Trust and Reciprocity: Experimental Evidence from PD Games." Working Paper of the Workshop in Political Theory and Policy Analysis, Indiana University.

Arrow, K. J. (1974) *The Limits of Organization*. New York: Norton.

Axelrod, R. (1984) *The Evolution of Cooperation*. New York: Basic Books.

Bandura, A. (1977) *Social Learning Theory*. Englewood Cliffs, NJ: Prentice Hall.

Batenburg, R. S. (1997a) "The External Management of Automation 1995: Fieldwork, Response, and Non–response." ISCORE Paper No. 59, Utrecht University.

Batenburg, R. S. (1997b) "The External Management of Automation 1995: The Codebook of MAT95." ISCORE Paper No. 58, Utrecht University.

Batenburg, R. S., W. Raub, and C. Snijders (2000) "Contacts and Contracts: Temporal Embeddedness and the Contractual Behavior of Firms." ISCORE Paper No. 107, Utrecht University.

Batenburg, R. S. and A. van de Rijt (1999) "The External Management of Automation 1998." ISCORE Paper No. 155, Utrecht University.

Becker, G. S. (1996) *Accounting for Tastes*. Cambridge, MA: Harvard University Press.

Becker, G. S. and C. B. Mulligan (1997) "The Endogenous Determination of Time Preference." *Quarterly Journal of Economics CXII*, 729–758.

Becker, G. S. and K. M. Murphy (2000) *Social Economics: Market Behavior in a Social Environment*. Cambridge, MA: Belknap Press of Harvard University Press.

Berg, J., J. Dickhaut, and K. A. McCabe (1995) "Trust, Reciprocity, and Social History." *Games and Economic Behavior 10*, 122–142.

BIENENSTOCK, E. J. AND P. BONACICH (1992) "The Core as a Solution to Exclusionary Networks." *Social Networks 14*, 231–243.

BIENENSTOCK, E. J. AND P. BONACICH (1993) "Game-Theory Models for Exchange Networks: Experimental Results." *Sociological Perspectives 36*, 117–135.

BIENENSTOCK, E. J. AND P. BONACICH (1997) "Network Exchange as a Cooperative Game." *Rationality and Society 9*, 37–65.

BINMORE, K. G. (1985) "Bargaining and Coalitions." In A. E. ROTH (Ed.), *Game–theoretic Models of Bargaining*, pp. 269–304. Cambridge: Cambridge University Press.

BINMORE, K. G. (1987) "Nash Bargaining Theory II." In K. G. BINNMORE AND P. DASGUPTA (Eds.), *The Economics of Bargaining*, pp. 61–76. Oxford: Blackwell.

BINMORE, K. G. (1992) *Fun and Games: A Text on Game Theory.* Lexington, MA: D.C. Heath.

BINMORE, K. G. (1998) *Game Theory and the Social Contract, Volume 2: Just Playing.* Cambridge, MA: MIT Press.

BINMORE, K. G., J. GALE, AND L. SAMUELSON (1995) "Learning to be Imperfect: The Ultimatum Game." *Games and Economic Behavior 8*, 56–90.

BINMORE, K. G., A. SHAKED, AND J. SUTTON (1985) "Testing Noncooperative Bargaining Theory: A Preliminary Study." *American Economic Review 75*, 1178–1180.

BLALOCK, H. M. (1972) *Social Statistics* (2nd ed.). New York: McGraw-Hill.

BLAU, P. M. (1964) *Exchange and Power in Social Life.* New York: Wiley.

BLUMBERG, B. F. (1997) *Das Management von Technologiekooperationen: Partnersuche und Verhandlungen mit dem Partner aus empirisch–theoretischer Perspektive.* Amsterdam: Thesis Publishers.

BOLTON, G. E. (1991) "A Comparative Model of Bargaining: Theory and Evidence." *American Economic Review 81*, 1096–1136.

BOLTON, G. E. AND A. OCKENFELS (1998) "Strategy and Equity: An ERC–Analysis of the Güth–van Damme Game." *Journal of Mathematical Psychology 42*, 215–226.

BOLTON, G. E. AND A. OCKENFELS (2000) "ERC: A Theory of Equity, Reciprocity and Competition." *American Economic Review 90*, 166–193.

BONACICH, P. (1998) "A Behavioral Foundation for a Structural Theory of Power in Exchange Networks." *Social Psychology Quarterly 61*, 185–198.

BONACICH, P. (1999) "An Algebraic Theory of Strong Power in Negatively Connected Networks." *Journal of Mathematical Sociology 23*, 203–224.

Bonacich, P. and E. J. Bienenstock (1995) “When Rationality Fails: Unstable Exchange Networks with Empty Cores.” *Rationality and Society 7*, 293–320.

Bonacich, P. and N. E. Friedkin (1998) “Unequally Valued Exchange Relations.” *Social Psychology Quarterly 61*, 160–171.

Boon, S. D. and J. G. Holmes (1991) “The Dynamics of Interpersonal Trust: Resolving Uncertainty in the Face of Risk.” In R. A. Hinde and J. Groebel (Eds.), *Cooperation and Prosocial Behavior*, pp. 190–210. Cambridge: Cambridge University Press.

Boone, R. T. and M. W. Macy (1999) “Unlocking the Doors of the Prisoner’s Dilemma: Dependence, Selectivity, and Cooperation.” *Social Psychology Quarterly 62*, 32–52.

Boyd, R. and P. J. Richerson (1985) *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.

Boyle, R. and P. Bonacich (1970) “The Development of Trust and Mistrust in Mixed–Motive Games.” *Sociometry 33*, 123–139.

Braun, N. (1998) “Rationalität und Drogenproblematik.” Habilitationsschrift. University of Berne: Department of Sociology.

Braun, N. and T. Gautschi (2000) “Who Gets How Much in Which Relation? A Non–Cooperative Bargaining Approach to Exchange Networks.” ISCORE Paper No. 132, Utrecht University.

Braun, N. and T. Gautschi (2001) “Who Gets How Much in Which Relation? A Flexible Theory of Profit Splits and Exchange Ties in Networks.” Working Paper, University of Munich and Utrecht University.

Burke, P. J. (1997) “An Identity Model for Network Exchange.” *American Sociological Review 62*, 134–150.

Burt, R. S. (1992) *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Burt, R. S. (2000) “The Network Structure of Social Capital.” *Research in Organizational Behavior 22*, 345–423.

Burt, R. S. and M. Knez (1995) “Kinds of Third–Party Effects on Trust.” *Rationality and Society 7*, 255–292.

Bush, R. R. and F. Mosteller (1955) *Stochastic Models for Learning*. New York: Wiley.

Buskens, V. (1999) *Social Networks and Trust*. Amsterdam: Thesis Publishers.

Buskens, V. and R. S. Batenburg (2000) “The External Management of Automation: Codebook for the Combined Data from the Netherlands and Germany.” ISCORE Paper No. 175, Utrecht University.

CAMERER, C. F. (1997) "Progress in Behavioral Game Theory." *Journal of Economic Perspectives 11*, 167–188.

CAMERER, C. F. AND K. WEIGELT (1988) "Experimental Tests of a Sequential Equilibrium Reputation Model." *Econometrica 56*, 1–36.

CHAMBERLAIN, G. (1980) "Analysis of Covariance with Qualitative Data." *Review of Economic Studies 47*, 225–238.

COASE, R. H. (1937) "The Nature of the Firm." *Economica 4*, 386–405.

COLEMAN, J. S. (1973) *The Mathematics of Collective Action*. London: Heinemann.

COLEMAN, J. S. (1988) "Social Capital in the Creation of Human Capital." *American Journal of Sociology 94*, S95–S120.

COLEMAN, J. S. (1990) *Foundations of Social Theory*. Cambridge, MA: Belknap Press of Harvard University Press.

COLMAN, A. M. (1995) *Game Theory and its Applications*. Oxford: Butterworh–Heinemann.

COOK, K. S. (Ed.) (2001) *Trust in Society*. New York: Russell Sage.

COOK, K. S. AND R. M. EMERSON (1978) "Power, Equity, and Commitment in Exchange Networks." *American Sociological Review 43*, 721–739.

COOK, K. S., R. M. EMERSON, M. R. GILLMORE, AND T. YAMAGISHI (1983) "The Distribution of Power in Exchange Networks: Theory and Empirical Results." *American Journal of Sociology 89*, 275–305.

COOK, K. S. AND J. M. WHITMEYER (1992) "Two Approaches to Social Structure: Exchange Theory and Network Analysis." *Annual Review of Sociology 18*, 109–127.

COOK, K. S. AND T. YAMAGISHI (1992) "Power in Exchange Networks: A Power–Dependence Formulation." *Social Networks 14*, 245–265.

DASGUPTA, P. (1988) "Trust as a Commodity." In D. GAMBETTA (Ed.), *Trust: Making and Breaking Cooperative Relations*, pp. 49–72. Oxford: Blackwell.

DAVIS, D. D. AND C. A. HOLT (1993) *Experimental Economics*. Princeton, NJ: Princeton University Press.

DEUTSCH, M. (1958) "Trust and Suspicion." *Journal of Conflict Resolution II*, 265–279.

DUGATKIN, L. A. (1997) *Cooperation among Animals: An Evolutionary Perspective*. Oxford: Oxford University Press.

ECKEL, C. C. AND P. J. GROSSMAN (1998) "Chivalry and Solidarity in Ultimatum Games." Working Paper, Virginia Polytechnic Institute and State University.

EMERSON, R. M. (1972) "Exchange Theory, Part II: Exchange Relations and Network Structures." In J. BERGER, M. ZELDITCH, JR., AND B. ANDERSON

(Eds.), *Sociological Theories in Progress*, Volume 2, pp. 58–87. Boston, MA: Houghton Mifflin.

Emerson, R. M. (1981) "Social Exchange Theory." In M. Rosenberg and R. H. Turner (Eds.), *Social Psychology: Sociological Perspectives*, pp. 30–65. New York: Basic Books.

Engels, F. (1909 [1845]) *Die Lage der arbeitenden Klasse in England: nach eigner Anschauung und authentischen Quellen* (3rd ed.). Stuttgart: Dietz.

Falk, A., E. Fehr, and U. Fischbacher (1999) "On the Nature of Fair Behavior." Working Paper No. 17, Institute for Empirical Research in Economics, University of Zurich.

Fehr, E. and P. K. Zych (1998) "Do Addicts Behave Rationally?" *Scandinavian Journal of Economics 100*, 643–661.

Feld, S. L. (1991) "Why Your Friends Have More Friends Than You Do." *American Journal of Sociology 96*, 1464–1477.

Fischer, G. H. (1997) "Unidimensional Linear Logistic Rasch Models." In W. J. Van der Linden and R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory*, pp. 225–243. New York: Springer.

Flap, H. D. (1988) *Conflict, Loyalty, and Violence*. Frankfurt am Main: Peter Lang.

Flap, H. D. (1999) "Creation and Returns of Social Capital. A New Research Program." *The Tocqueville Review 20*, 1–22.

Flap, H. D. and E. Boxman (1999) "Getting a Job as a Manager." In S. M. Gabbay and R. T. A. J. Leenders (Eds.), *Corporate Social Capital*, pp. 197–216. Amsterdam: Kluwer.

Flap, H. D. and N. D. de Graaf (1986) "Social Capital and Attained Occupational Status." *Netherlands' Journal of Sociology 22*, 145–161.

Forsythe, R., J. Horowitz, N. E. Savin, and M. Sefton (1994) "Fairness in Simple Bargaining Experiments." *Games and Economic Behavior 6*, 347–369.

Friedkin, N. E. (1986) "A Formal Theory of Social Power." *Journal of Mathematical Sociology 12*, 103–126.

Friedkin, N. E. (1992) "An Expected Value Model of Social Power: Predictions for Selected Exchange Networks." *Social Networks 14*, 213–229.

Friedkin, N. E. (1993) "An Expected Value Model of Social Exchange Outcomes." In E. J. Lawler, B. Markovsky, K. Heimer, and J. O'Brien (Eds.), *Advances in Group Processes*, pp. 163–193. Greenwich, CT: JAI Press.

Friedkin, N. E. (1995) "The Incidence of Exchange Networks." *Social Psychology Quarterly 58*, 213–221.

Friedman, J. W. (1986) *Game Theory with Applications to Economics*. New York: Oxford University Press.

Fudenberg, D. and D. K. Levine (1998) *The Theory of Learning in Games*. Cambridge, MA: MIT Press.

Fudenberg, D. and J. Tirole (1993) *Game Theory* (3rd ed.). Cambridge, MA: MIT Press.

Fukuyama, F. (1995) *Trust. The Social Virtues and the Creation of Prosperity*. New York: Free Press.

Gautschi, T. (1999) "A Hostage Trust Game with Incomplete Information and Fairness Considerations of the Trustee." ISCORE Paper No. 131, Utrecht University.

Gouldner, A. W. (1960) "The Norm of Reciprocity: A Preliminary Statement." *American Sociological Review 25*, 161–178.

Granovetter, M. S. (1985) "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology 91*, 481–510.

Gravelle, H. and R. Rees (1992) *Microeconomics* (2nd ed.). London and New York: Longman.

Greene, W. H. (1993) *Econometric Analysis* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Güth, W., R. Schmittberger, and B. Schwarze (1982) "An Experimental Analysis of Ulimatum Bargaining." *Journal of Economic Behavior and Organization 3*, 367–388.

Güth, W. and R. Tietz (1988) "Ultimatum Bargaining for a Shrinking Cake — An Experimental Analysis." In R. Tietz, W. Albers, and R. Selten (Eds.), *Bounded Rational Behavior in Experimental Games and Markets*, pp. 111–128. Berlin: Springer.

Hardin, R. (2001) "Conceptions and Explanations of Trust." In K. S. Cook (Ed.), *Trust in Society*, pp. 3–39. New York: Russell Sage.

Harsanyi, J. C. and R. Selten (1988) *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.

Hayashi, N., E. Ostrom, J. Walker, and T. Yamagishi (1999) "Reciprocity, Trust, and the Sense of Control: A Cross–Societal Study." *Rationality and Society 11*, 27–46.

Heckathorn, D. D. (1983) "Extensions of Power–Dependence Theory: The Concept of Resistance." *Social Forces 61*, 1206–1231.

Herkner, W. (1991) *Lehrbuch Psychologie* (5th enlarged ed.). Bern: Huber.

Hess, H. (1986) *Mafia* (2nd ed.). Tübingen: Mohr.

Hobbes, T. (1651) *Leviathan, or the Matter, Forme & Power of a Common–Wealth.* Revised Student Edition edited by Richard Tuck. Cambridge: Cambridge University Press, 1996.

Hoffman, E., K. A. McCabe, K. Shachat, and V. L. Smith (1993) "Preferences, Property Rights and Anonymity in Bargaining Games." Mimeo. University of Arizona.

Holland, J. H., K. Holyoak, R. Nisbett, and P. Thagard (Eds.) (1986) *Induction, Processes of Inference, Learning and Discovery.* Cambridge, MA: MIT Press.

Hollis, M. (1998) *Trust Within Reasons.* Cambridge: Cambridge University Press.

Homans, G. C. (1961) *The Human Group.* New York: Harcourt.

Kagel, J. H. and A. E. Roth (Eds.) (1995) *Handbook of Experimental Economics.* Princeton, NJ: Princeton University Press.

Kahneman, D. (1994) "New Challenges to the Rationality Assumption." *Journal of Institutional and Theoretical Economics 150*, 18–36.

Kahneman, D., J. L. Knetsch, and R. Thaler (1986) "Fairness and the Assumption of Economics." *Journal of Business 59*, S285–S300.

Kelley, H. H. and J. W. Thibaut (1978) *Interpersonal Relations: A Theory of Interdependence.* New York: Wiley.

Keren, G. and W. Raub (1993) "Resolving Social Conflicts through Hostage Posting: Theoretical and Empirical Considerations." *Journal of Economic Psychology: General 122*, 429–448.

Knack, S. and P. Keefer (1997) "Does Social Capital Have an Economic Payoff? A Crosscountry Investigation." *Quarterly Journal of Economics 112*, 1251–1288.

Kohlberg, E. and J. F. Mertens (1986) "On the Strategic Stability of Equilibria." *Econometrica 54*, 1003–1037.

Kollock, P. (1994) "The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust." *American Journal of Sociology 100*, 313–345.

Kreps, D. M. (1990a) "Corporate Culture and Economic Theory." In J. Alt and K. Shepsle (Eds.), *Perspectives on Positive Political Economy*, pp. 90–143. Cambridge: Cambridge University Press.

Kreps, D. M. (1990b) *A Course in Microeconomic Theory.* New York: Harvester Wheatsheaf.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982) "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." *Journal of Economic Theory 27*, 245–252.

LAHNO, B. (1995a) "Trust and Strategic Rationality." *Rationality and Society 7*, 442–464.

LAHNO, B. (1995b) "Trust, Reputation, and Exit in Exchange Relationships." *Journal of Conflict Resolution 39*, 495–510.

LAWLER, E. J. AND J. YOON (1996) "Commitment in Exchange Relations: Test of a Theory of Relational Cohesion." *American Sociological Review 61*, 89–108.

LAZARIC, N. AND E. LORENZ (Eds.) (1998a) *Trust and Economic Learning*. Cheltenham: Edward Elgar.

LAZARIC, N. AND E. LORENZ (1998b) "Trust and Organizational Learning During Inter–Firm Cooperation." In N. LAZARIC AND E. LORENZ (Eds.), *Trust and Economic Learning*, pp. 209–226. Cheltenham: Edward Elgar.

LENSKI, G. E. (1966) *Power and Privilege. A Theory of Social Stratification*. New York: McGraw–Hill.

LEWICKI, R. J. AND B. BENEDICT BUNKER (1995) "Trust in Relationships: A Model of Development and Decline." In B. BENDEDICT BUNKER AND J. Z. RUBIN (Eds.), *Conflict, Cooperation, and Justice: Essays Inspired by the Work of Morton Deutsch*, pp. 133–174. San Francisco, CA: Jossey–Bass.

LIEBRAND, W. B. G. (1984) "The Effect of Social Motives, Communication and Group Size on Behavior in an N–Person Multi–Stage Mixed–Motive Game." *European Journal of Social Psychology 14*, 239–264.

LIN, N. (2001) *Social Capital. A Theory of Social Structure and Action*. Cambridge: Cambridge University Press.

LOEWENSTEIN, G. AND J. ELSTER (1992) *Choice over Time*. New York: Russell Sage.

LONG, J. S. (1997) *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Russel Sage.

LOVAGLIA, M. J., J. SKVORETZ, D. WILLER, AND B. MARKOVSKY (1995) "Negotiated Exchanges in Social Networks." *Social Forces 74*, 123–155.

LUCE, R. D. AND H. RAIFFA (1957) *Games and Decisions: Introduction and Critical Survey*. New York: Wiley.

LUHMANN, N. (1979) *Trust and Power*. Chichester: Wiley.

MACAULAY, S. (1963) "Non-Contractual Relations in Business: A Preliminary Study." *American Sociological Review 28*, 55–67.

MACY, M. W. (1989) "Walking out of Social Traps. A Stochastic Learning Model for Prisoner's Dilemma." *Rationality and Society 1*, 197–219.

MACY, M. W. (1993) "Social Learning and the Structure of Collective Action." In E. J. LAWLER, B. MARKOWSKY, K. HEIMER, AND J. O'BRIAN (Eds.), *Advances in Group Processes. Volume 10*, pp. 1–35. Greenwich, CT: JAI Press.

Macy, M. W. (1996) "Natural Selection and Social Learning in Prisoner's Dilemma. Coadaption with Genetic Algorithms and Artifical Networks." *Sociological Methods and Research 25*, 103–137.

Malinowski, B. (1922) *Argonauts of the Western Pacific. An Account of Native Enterprise and Adventure in the Archipelagoes of Melanesian New Guinea.* London: Routledge.

Markovsky, B., J. Skvoretz, D. Willer, M. J. Lovaglia, and J. Erger (1993) "The Seeds of Weak Power: An Extension of Network Exchange Theory." *American Sociological Review 58*, 197–209.

Markovsky, B., D. Willer, and T. Patton (1988) "Power Relations in Exchange Networks." *American Journal of Sociology 53*, 220–236.

Markovsky, B., D. Willer, and T. Patton (1990) "Theory, Evidence and Intuition." *American Sociological Review 55*, 300–305 (Reply to Yamagishi and Cook 1990).

Markovsky, B., D. Willer, B. Simpson, and M. J. Lovaglia (1997) "Power in Exchange Networks: Critique of a New Theory." *American Sociological Review 62*, 833–837 (Comment on Yamaguchi 1996).

Marx, K. (2000 [1867]) *Das Kapital.* Köln: GLB Parkland.

Matthews, B. A., W. M. Kordonski, and E. Shimoff (1983) "Temptation and the Maintenance of Trust." *Journal of Conflict Resolution 27*, 255–277.

Mauss, M. (1954) *The Gift. Forms and Functions of Exchange in Archaic Societies.* Glencoe, IL: Free Press.

Maynard Smith, J. (1982) *Evolution and the Theory of Games.* Cambridge: Cambridge University Press.

McCabe, K. A., S. J. Rassenti, and V. L. Smith (1996) "Game Theory and Reciprocity in Some Extensive Form Experimental Games." *Proceedings of the National Academy of Science 93*, 13421–13428.

McClintock, C. G. (1972) "Social Motivation — A Set of Propositions." *Behavioral Science 17*, 438–454.

McClintock, C. G. (1978) "Social Values: Their Definition, Measurement and Development." *Journal of Research and Development in Education 12*, 121–137.

McKelvey, R. D. and T. R. Palfrey (1992) "An Experimental Study of the Centipede Game." *Econometrica 60*, 803–836.

Misztal, B. A. (1996) *Trust in Modern Societies.* Cambridge: Polity Press.

Mlicki, P. (1996) "Hostage Posting as a Mechanism for Co–operation in the Prisoner's Dilemma Game." In W. B. G. Liebrand and D. M. Messick (Eds.), *Frontiers in Social Dilemma Research*, pp. 165–183. Berlin: Springer.

Muthoo, A. (1999) *Bargaining Theory with Applications*. Cambridge: Cambridge University Press.

Nash, J. F. (1950) “The Bargaining Problem.” *Econometrica 18*, 155–162.

Nash, J. F. (1951) “Non–Cooperative Games.” *Annals of Mathematics 54*, 286–295.

Nash, J. F. (1953) “Two–Person Cooperative Games.” *Econometrica 21*, 128–140.

Neelin, J., H. Sonnenschein, and M. Spiegel (1988) “A Further Test of Noncooperative Bargaining Theory: Comment.” *American Economic Review 78*, 824–836.

Nelson, W. R. (2001) “Incorporating Fairness into Game Theory and Economics: Comment.” *American Economic Review 91*, 1180–1183.

Neumann, M. (1996) *Theoretische Volkswirtschftslehre I. Makroökonomische Theorie: Beschäftigung, Inflation und Zahlungsbilanz* (5th enlarged ed.). München: Franz Vahlen.

Ochs, J. and A. E. Roth (1989) “An Experimental Study of Sequential Bargaining.” *American Economic Review 78*, 824–836.

Oliver, P. (1984) “Rewards and Punishments as Selective Incentives: An Apex Game.” *Journal of Conflict Resolution 28*, 123–148.

Osborne, M. J. and A. Rubinstein (1990) *Bargaining and Markets*. San Diego, CA: Academic Press.

Putnam, R. D. (1993) *Making Democracy Work*. Princeton: Princeton University Press.

Putnam, R. D. (1995) “Turning In, Turning Out: The Strange Disappearance of Social Capital in America.” *Political Science and Politics 24*, 664–683.

Rabin, M. (1993) “Incorporating Fairness into Game Theory and Economics.” *American Economic Review 83*, 1281–1302.

Rabin, M. (1998) “Psychology and Economics.” *Journal of Economic Literature XXXVI*, 11–46.

Raftery, A. E. (1995) “Bayesian Model Selection in Social Research.” In P. V. Marsden (Ed.), *Sociological Methodology*, pp. 111–163. Oxford: Blackwell.

Raith, W. (1994) *Das neue Mafia–Kartell. Wie die Syndikate den Osten erobern.* Reinbeck bei Hamburg: Rowohlt.

Raith, W. (1995) *Organisierte Kriminalität*. Reinbeck bei Hamburg: Rowohlt.

Rapoport, A. (1988) “Experiments with N–Person Social Traps II. Tragedy of the Commons.” *Journal of Conflict Resolution 32*, 473–488.

Rapoport, A. and A. M. Chammah (1965) *Prisoner’s Dilemma. A Study in Conflict and Cooperation.* Ann Arbor, MI: University of Michigan Press.

Rapoport, A., M. J. Guyer, and D. G. Gordon (1976) *The 2 × 2 Game.* Ann Arbor, MI: University of Michigan Press.

Rasmusen, E. (1994) *Games and Information. An Introduction to Game Theory* (2nd ed.). Cambridge: Blackwell.

Raub, W. (1997) *Samenwerking in duurzame relaties en sociale cohesie.* Amsterdam: Thesis Publishers.

Raub, W. and G. Keren (1993) "Hostages as a Commitment Device: A Game-Theoretic Model and an Empirical Test of Some Scenarios." *Journal of Economic Behavior and Organization 21*, 43–67.

Raub, W. and C. Snijders (2001) "A Reluctant Match: Models for the Analysis of Trust in Durable Two Party Relations." In C. Castelfranchi and Y. Tan (Eds.), *Trust and Deception in Virtual Societies*, pp. 195–220. Amsterdam: Kluwer.

Raub, W. and J. Weesie (1990) "Reputation and Efficiency in Social Interactions: An Example of Network Effects." *American Journal of Sociology 96*, 626–654.

Raub, W. and J. Weesie (1993) "Symbiotic Arrangements: A Sociological Perspective." *Journal of Institutional and Theoretical Economics 149*, 716–724.

Raub, W. and J. Weesie (2000a) "The Management of Durable Relations." In J. Weesie and W. Raub (Eds.), *The Management of Durable Relations. Theoretical Models and Empirical Studies of Households and Organizations*, pp. 1–32. Amsterdam: Thela Thesis.

Raub, W. and J. Weesie (2000b) "Reliable Signals via Hostages." *Analyse und Kritik 22*, 19–43.

Riedwyl, H. (1992) *Angewandte Statistik* (2nd ed.). Bern: Haupt.

Rivaud-Danset, D. (1998) "Trust and Collective Action in Financial Markets." In N. Lazaric and E. Lorenz (Eds.), *Trust and Economic Learning*, pp. 78–93. Cheltenham: Edward Elgar.

Rooks, G. (2002) *Contract en Conflict. Strategisch Management van Inkooptransacties.* Amsterdam: Thesis Publishers.

Roth, A. E. (1995) "Bargaining Experiments." In J. H. Kagel and A. E. Roth (Eds.), *The Handbook of Experimental Economics*, pp. 253–348. Princeton, NJ: Princeton University Press.

Roth, A. E. and I. Erev (1995) "Learning in Extensive–Form Games: Experimental Data and Simple Dynamic Models in the Medium Term." *Games and Economic Behavior 8*, 164–212.

Rubinstein, A. (1982) "Perfect Equilibrium in a Bargaining Model." *Econometrica 50*, 97–109.

SAKO, M. (1998) "The Information Requirements of Trust in Supplier Relations: Evidence from Japan, Europe and the United States." In N. LAZARIC AND E. LORENZ (Eds.), *Trust and Economic Learning*, pp. 23–47. Cheltenham: Edward Elgar.

SANDEFUR, R. L. AND E. O. LAUMANN (1998) "A Paradigm for Social Capital." *Rationality and Society 10*, 481–501.

SAVAGE, L. J. (1954) *The Foundations of Statistics*. New York: Wiley.

SCHELLING, T. C. (1960) *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

SCHÜSSLER, R. A. (1989) "Exit Threats and Cooperation Under Anonymity." *Journal of Conflict Resolution 33*, 728–749.

SELTEN, R. (1965) "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit." *Zeitschrift für die gesamte Staatswissenschaft 121*, 301–324, 667–689.

SELTEN, R. (1967) "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments." In H. SAUERMANN (Ed.), *Beiträge zur experimentellen Wirtschaftsforschung*, pp. 136–168. Tübingen: Mohr.

SIMPSON, B. AND D. WILLER (1999) "A New Method for Finding Power Structures." In D. WILLER (Ed.), *Network Exchange Theory*, pp. 270–284. Westport, CT: Praeger.

SKVORETZ, J. AND T. J. FARARO (1992) "Power and Network Exchange: An Essay toward Theoretical Unification." *Social Networks 14*, 325–344.

SKVORETZ, J. AND M. J. LOVAGLIA (1995) "Who Exchanges with Whom: Structural Determinants of Exchange Frequency in Negotiated Exchange Networks." *Social Psychology Quarterly 58*, 163–177.

SKVORETZ, J. AND D. WILLER (1991) "Power in Exchange Networks: Setting and Structural Variations." *Social Psychology Quarterly 54*, 224–238.

SKVORETZ, J. AND D. WILLER (1993) "Exclusion and Power: A Test of Four Theories of Power in Exchange Networks." *American Sociological Review 58*, 801–818.

SMELSER, N. J. AND R. SWEDBERG (Eds.) (1994) *The Handbook of Economic Sociology*. Princeton, NJ: Princeton University Press.

SNIJDERS, C. (1996) *Trust and Commitments*. Amsterdam: Thesis Publishers.

SNIJDERS, C. AND G. KEREN (1997) "Trust and Reciprocity." ISCORE Paper No. 112, Utrecht University.

SNIJDERS, C. AND G. KEREN (2001) "Do you Trust? Whom Do you Trust? When Do you Trust?" In S. R. THYE, E. J. LAWLER, M. W. MACY, AND H. A.

Walker (Eds.), *Advances in Group Processes. Volume 18*, pp. 129–160. Amsterdam: JAI, Elsevier Science.

Snijders, T. A. B. (1998) "A Measurement–Orientated Definition of Social Capital." Working Paper, University of Groningen.

Ståhl, I. (1972) "Bargaining Theory." Economic Research Institute, Stockholm.

StataCorp (2001) *Stata Statistical Software: Release 7.0.* College Station, TX: Stata Corporation.

Swinth, R. L. (1976) "The Establishment of the Trust Relationship." *Journal of Conflict Resolution XI*, 335–344.

Taylor, M. (1987) *The Possibility of Cooperation.* Cambridge: Cambridge University Press (revised edition of *Anarchy and Cooperation*, London: Wiley, 1976).

Thibaut, J. W. and H. H. Kelley (1959) *The Social Psychology of Groups.* New York: Wiley.

Thye, S. R., M. J. Lovaglia, and B. Markovsky (1997) "Responses to Social Exchange and Social Exclusion in Networks." *Social Forces 75*, 1031–1049.

Trivers, R. L. (1971) "The Evolution of Reciprocal Altruism." *Quarterly Review of Biology 46*, 35–57.

Ultee, W., W. Arts, and H. Flap (1996) *Sociologie: vragen, uitspraken, bevindingen* (2nd ed.). Groningen: Wolters–Noordhoff.

Von Neumann, J. and O. Morgenstern (1944) *Theory of Games and Economic Behavior.* Princeton, NJ: Princeton University Press.

Voss, T. (1982) "Rational Actors and Social Institutions: The Case of the Organic Emergence of Norms." In W. Raub (Ed.), *Theoretical Models and Empirical Analysis: Contributions to the Explanation of Individual Actions and Collective Phenomena*, pp. 101–128. Utrecht: Explanatory Sociology Publications.

Voss, T. (1985) *Rationale Akteure und soziale Institutionen. Beitrag zu einer endogenen Theorie des sozialen Tausches.* München: Oldenbourg.

Walker, S. (1987) *Animal Learning. An Introduction.* London: Routledge.

Wasserman, S. and K. Faust (1994) *Social Network Analysis: Methods and Applications.* Cambridge: Cambridge University Press.

Weesie, J. (1994) "Social Orientations in Symmetric 2 × 2 Games." ISCORE Paper No. 17, Utrecht University.

Weesie, J. and W. Raub (1996) "Private Ordering: A Comparative Institutional Analysis of Hostage Games." *Journal of Mathematical Sociology 21*, 201–240.

Weesie, J. and W. Raub (Eds.) (2000) *The Management of Durable Relations. Theoretical Models and Empirical Studies of Households and Organizations.* Amsterdam: Thesis Publishers.

Wilde, G. J. S. (1970) *Neurotische labiliteit gemeten volgens de vragenlijstmethode*. Amsterdam: F. van Rossen.

Willer, D. (1981) "Quantity and Network Structure." In D. Willer and B. Anderson (Eds.), *Networks, Exchange, and Coercion: The Elementary Theory and its Application*, pp. 108–127. Oxford: Elsevier.

Willer, D. (Ed.) (1999) *Network Exchange Theory*. Westport, CT: Praeger.

Willer, D. and J. Skvoretz (1999) "Network Connection and Exchange Ratios: Theory, Predictions, and Experimental Tests." In D. Willer (Ed.), *Network Exchange Theory*, pp. 195–225. Westport, CT: Praeger.

Williamson, O. E. (1985) *The Economic Institutions of Capitalism*. New York: Free Press.

Williamson, O. E. (1996) *The Mechanisms of Governance*. Oxford: Oxford University Press.

Yamagishi, T. and K. S. Cook (1990) "Power Relations in Exchange Networks: A Comment on 'Network Exchange Theory'." *American Sociological Review 55*, 297–300.

Yamagishi, T., M. R. Gillmore, and K. S. Cook (1988) "Network Connections and the Distribution of Power in Exchange Networks." *American Journal of Sociology 93*, 833–851.

Yamagishi, T. and M. Yamagishi (1994) "Trust and Commitment in the United States and Japan." *Motivation and Emotion 18*, 129–166.

Yamaguchi, K. (1996) "Power in Networks of Substitutable and Complementary Exchange Relations: A Rational–Choice Model and an Analysis of Power Centralization." *American Sociological Review 61*, 308–332.

Yamaguchi, K. (1997) "The Logic of the New Theory and Misrepresentations of the Logic." *American Sociological Review 62*, 338–841 (Reply to Markovsky et al. 1997).

Yamaguchi, K. (2000) "Power in Mixed Exchange Networks: A Rational Choice Model." *Social Networks 22*, 93–121.

Author Index

Subject Index

Curriculum Vitae

Thomas Gautschi was born on April 22, 1969 in Thun, Switzerland. After completing his secondary education, he started his undergraduate study in social sciences at the University of Bern. In 1994, he became a research and teaching assistant at the Department of Sociology of the University of Bern. In 1997, he took his master's degree in social sciences with a thesis in sociology and a thesis in economics. His thesis in sociology addressed questions of rationality in persons' environmental behavior while his thesis in economics was concerned with possibilities to implement a personnel development concept and its impact on a firm's performance. Thomas Gautschi joined the Ph.D. program at the Interuniversity Center for Social Science Theory and Methodology (ICS) in 1997. The research reported in this book is the result of his work at the ICS. He is currently an Assistant Professor at Eindhoven University of Technology.

ICS Dissertation Series

The ICS–series presents dissertations of the Interuniversity Center for Social Science Theory and Methodology. Each of these studies aims at integrating explicit theory formation with state–of–the–art empirical research or at the development of advanced methods for empirical research.

The ICS was founded in 1986 as a cooperative effort of the universities of Groningen and Utrecht. Since 1992, the ICS expanded to the University of Nijmegen. Most of the projects are financed by the participating universities or by the Netherlands Organization for Scientific Research (NWO). The international composition of the ICS graduate students is mirrored in the increasing international orientation of the projects and thus of the ICS–series itself.

1. Cees van Liere (1990). *Lastige Leerlingen. Een empirisch onderzoek naar sociale oorzaken van probleemgedrag op basisscholen.* Amsterdam: Thesis Publishers.
2. Marco H. D. van Leeuwen (1990). *Bijstand in Amsterdam, ca. 1800–1850. Armenzorg als beheersings- en overlevingsstrategie.* Amsterdam: Thesis Publishers.
3. Ineke Maas (1990). *Deelname aan podiumkunsten via de podia, de media en actieve beoefening. Substitutie of leereffecten?* Amsterdam: Thesis Publishers.
4. Marjolein I. Broese van Groenou (1991). *Gescheiden Netwerken. De relaties met vrienden en verwanten na echtscheiding.* Amsterdam: Thesis Publishers.
5. Jan M. M. van den Bos (1991). *Dutch EC Policy Making. A Model–Guided Approach to Coordination and Negotiation.* Amsterdam: Thesis Publishers.
6. Karin Sanders (1991). *Vrouwelijke Pioniers. Vrouwen en mannen met een 'mannelijke' hogere beroepsopleiding aan het begin van hun loopbaan.* Amsterdam: Thesis Publishers.
7. Sjerp de Vries (1991). *Egoism, Altruism, and Social Justice. Theory and Experiments on Cooperation in Social Dilemmas.* Amsterdam: Thesis Publishers.
8. Ronald S. Batenburg (1991). *Automatisering in bedrijf.* Amsterdam: Thesis Publishers.
9. Rudi Wielers (1991). *Selectie en allocatie op de arbeidsmarkt. Een uitwerking voor de informele en geïnstitutionaliseerde kinderopvang.* Amsterdam: Thesis Publishers.
10. Gert P. Westert (1991). *Verschillen in ziekenhuisgebruik.* Amsterdam: Thesis Publishers.
11. Hanneke Hermsen (1992). *Votes and Policy Preferences. Equilibria in Party Systems.* Amsterdam: Thesis Publishers.

12. Cora J. M. Maas (1992). *Probleemleerlingen in het basisonderwijs.* Amsterdam: Thesis Publishers.

13. Ed A. W. Boxman (1992). *Contacten en carrière. Een empirisch–theoretisch onderzoek naar de relatie tussen sociale netwerken en arbeidsmarktposities.* Amsterdam: Thesis Publishers.

14. Conny G. J. Taes (1992). *Kijken naar banen. Een onderzoek naar de inschatting van arbeidsmarktkansen bij schoolverlaters uit het middelbaar beroepsonderwijs.* Amsterdam: Thesis Publishers.

15. Peter van Roozendaal (1992). *Cabinets in Multi–Party Democracies. The Effect of Dominant and Central Parties on Cabinet Composition and Durability.* Amsterdam: Thesis Publishers.

16. Marcel van Dam (1992). *Regio zonder regie. Verschillen in en effectiviteit van gemeentelijk arbeidsmarktbeleid.* Amsterdam: Thesis Publishers.

17. Tanja van der Lippe (1993). *Arbeidsverdeling tussen mannen en vrouwen.* Amsterdam: Thesis Publishers.

18. Marc A. Jacobs (1993). *Software: kopen of kopiëren? Een sociaal–wetenschappelijk onderzoek onder PC–gebruikers.* Amsterdam: Thesis Publishers.

19. Peter van der Meer (1993). *Verdringing op de Nederlandse arbeidsmarkt. sector- en sekseverschillen.* Amsterdam: Thesis Publishers.

20. Gerbert Kraaykamp (1993). *Over lezen gesproken. Een studie naar sociale differentiatie in leesgedrag.* Amsterdam: Thesis Publishers.

21. Evelien Zeggelink (1993). *Strangers into Friends. The Evolution of Friendship Networks Using an Individual Oriented Modeling Approach.* Amsterdam: Thesis Publishers.

22. Jaco Berveling (1994). *Het stempel op de besluitvorming. Macht, invloed en besluitvorming op twee Amsterdamse beleidsterreinen.* Amsterdam: Thesis Publishers.

23. Wim Bernasco (1994). *Coupled Careers. The Effects of Spouse's Resources on Success at Work.* Amsterdam: Thesis Publishers.

24. Liset van Dijk (1994). *Choices in Child Care. The Distribution of Child Care among Mothers, Fathers and Non–Parental Care Providers.* Amsterdam: Thesis Publishers.

25. Jos de Haan (1994). *Research Groups in Dutch Sociology.* Amsterdam: Thesis Publishers.

26. Kwasi Boahene (1995). *Innovation Adoption as a Socio–Economic Process. The Case of the Ghanaian Cocoa Industry.* Amsterdam: Thesis Publishers.

27. Paul E. M. Ligthart (1995). *Solidarity in Economic Transactions. An Experimental Study of Framing Effects in Bargaining and Contracting.* Amsterdam: Thesis Publishers.

28. Roger Th. A. J. Leenders (1995). *Structure and Influence. Statistical Models for the Dynamics of Actor Attributes, Network Structure, and their Interdependence.* Amsterdam: Thesis Publishers.

29. Beate Völker (1995). *Should Auld Acquaintance Be Forgot...? Institutions of Communism, the Transition to Capitalism and Personal Networks: The Case of East Germany.* Amsterdam: Thesis Publishers.

30. Anneke Cancrinus–Matthijsse (1995). *Tussen hulpverlening en ondernemerschap. Beroepsuitoefening en taakopvattingen van openbare apothekers in een aantal West–Europese landen.* Amsterdam: Thesis Publishers.

31. Nardi Steverink (1996). *Zo lang mogelijk zelfstandig. Naar een verklaring van verschillen in oriëntatie ten aanzien van opname in een verzorgingstehuis onder fysiek kwetsbare ouderen.* Amsterdam: Thesis Publishers.

32. Ellen Lindeman (1996). *Participatie in vrijwilligerswerk.* Amsterdam: Thesis Publishers.

33. Chris Snijders (1996). *Trust and Commitments.* Amsterdam: Thesis Publishers.

34. Koos Postma (1996). *Changing Prejudice in Hungary. A Study on the Collapse of State Socialism and Its Impact on Prejudice against Gypsies and Jews.* Amsterdam: Thesis Publishers.

35. Jooske T. van Busschbach (1996). *Uit het oog, uit het hart? Stabiliteit en verandering in persoonlijke relaties.* Amsterdam: Thesis Publishers.

36. René Torenvlied (1996). *Besluiten in uitvoering. Theorieën over beleidsuitvoering modelmatig getoetst op sociale vernieuwing in drie gemeenten.* Amsterdam: Thesis Publishers.

37. Andreas Flache (1996). *The Double Edge of Networks. An Analysis of the Effect of Informal Networks on Cooperation in Social Dilemmas.* Amsterdam: Thesis Publishers.

38. Kees van Veen (1997). *Inside an Internal Labor Market: Formal Rules, Flexibility and Career Lines in a Dutch Manufacturing Company.* Amsterdam: Thesis Publishers.

39. Lucienne van Eijk (1997). *Activity and Well–Being in the Elderly.* Amsterdam: Thesis Publishers.

40. Róbert Gál (1997). *Unreliability. Contract Discipline and Contract Governance under Economic Transition.* Amsterdam: Thesis Publishers

41. Anne–Geerte van de Goor (1997). *Effects of Regulation on Disability and Duration.* Amsterdam: Thesis Publishers.

42. Boris Blumberg (1997). *Das Management von Technologiekooperationen. Partnersuche und Verhandlungen mit dem Partner aus empirisch–theoretischer Perspektive.* Amsterdam: Thesis Publishers.

43. Marijke von Bergh (1997). *Loopbanen van oudere werknemers.* Amsterdam: Thesis Publishers.

44. Anna Petra Nieboer (1997). *Life–Events and Well–Being: A Prospective Study on Changes in Well–Being of Elderly People Due to a Serious Illness Event or Death of the Spouse.* Amsterdam: Thesis Publishers.

45. Jacques Niehof (1997). *Resources and Social Reproduction: The Effects of Cultural and Material Resources on Educational and Occupational Careers in Industrial Nations at the End of the Twentieth Century.* Amsterdam: Thesis Publishers.

46. Ariana Need (1997). *The Kindred Vote. Individual and Family Effects of Social Class and Religion on Electoral Change in the Netherlands, 1956–1994.* Amsterdam: Thesis Publishers.

47. Jim Allen (1997). *Sector Composition and the Effect of Education on Wages: An International Comparison.* Amsterdam: Thesis Publishers.

48. Jack B. F. Hutten (1998). *Workload and Provision of Care in General Practice. An Empirical Study of the Relation Between Workload of Dutch General Practitioners and the Content and Quality of Their Care.* Amsterdam: Thesis Publishers.

49. Per B. Kropp (1998). *Berufserfolg im Transformationsprozeß. Eine theoretisch–empirische Studie über die Gewinner und Verlierer der Wende in Ostdeutschland.* Amsterdam: Thesis Publishers.

50. Maarten H. J. Wolbers (1998). *Diploma–inflatie en verdringing op de arbeidsmarkt. Een studie naar ontwikkelingen in de opbrengsten van diploma's in Nederland.* Amsterdam: Thela Thesis.

51. Wilma Smeenk (1998). *Opportunity and Marriage. The Impact of Individual Resources and Marriage Market Structure on the When and Whom of First Marriage in the Netherlands.* Amsterdam: Thela Thesis.

52. Marinus Spreen (1999). *Sampling Personal Network Structures: Statistical Inference in Ego–Graphs.* Amsterdam: Thela Thesis.

53. Vincent Buskens (1999). *Social Networks and Trust.* Amsterdam: Thela Thesis.

54. Susanne Rijken (1999). *Educational Expansion and Status Attainment. A Cross–National and Over–Time Comparison.* Amsterdam: Thela Thesis.

55. Mérove Gijsberts (1999). *The Legitimation of Inequality in State–Socialist and Market Societies, 1987–1996.* Amsterdam: Thela Thesis.

56. Gerhard G. Van de Bunt (1999). *Friends by Choice. An Actor–Oriented Statistical Network Model for Friendship Networks Through Time.* Amsterdam: Thela Thesis.

57. Robert Thomson (1999). *The Party Mandate: Election Pledges and Government Actions in the Netherlands, 1986–1998.* Amsterdam: Thela Thesis.

58. Corine Baarda (1999). *Politieke besluiten en boeren beslissingen. Het draagvlak van het mestbeleid tot 2000.* Amsterdam: Thela Thesis.

59. Rafael Wittek (1999). *Interdependence and Informal Control in Organizations.* Amsterdam: Thela Thesis.

60. Diane Payne (1999). *Policy Making in the European Union: an Analysis of the Impact of the Reform of the Structural Funds in Ireland.* Amsterdam: Thela Thesis.

61. René Veenstra (1999). *Leerlingen – Klassen – Scholen. Prestaties en vorderingen van leerlingen in het voortgezet onderwijs.* Amsterdam: Thela Thesis.

62. Marjolein Achterkamp (1999). *Influence Strategies in Collective Decision Making. A Comparison of Two Models.* Amsterdam: Thela Thesis.

63. Peter Mühlau (2000). *The Governance of the Employment Relation. A Relational Signaling Perspective.* Amsterdam: Thela Thesis.

64. Agnes Akkerman (2000). *Verdeelde vakbeweging en stakingen. Concurrentie om leden.* Amsterdam: Thela Thesis.

65. Sandra van Thiel (2000). *Quangocratization: Trends, Causes and Consequences.* Amsterdam: Thela Thesis.

66. Rudi Turksema (2000). *Supply of Day Care.* Amsterdam: Thela Thesis.

67. Sylvia E. Korupp (2000). *Mothers and the Process of Social Stratification.* Amsterdam: Thela Thesis.

68. Bernard A. Nijstad (2000). *How the Group Affects the Mind: Effects of Communication in Idea Generating Groups.* Amsterdam: Thela Thesis.

69. Inge F. de Wolf (2000). *Opleidingsspecialisatie en arbeidsmarktsucces van sociale wetenschappers.* Amsterdam: Thela Thesis.

70. Jan Kratzer (2001). *Communication and Performance: An Empirical Study in Innovation Teams.* Amsterdam: Thela Thesis.

71. Madelon Kroneman (2001). *Healthcare Systems and Hospital Bed Use.* Amsterdam: Thela Thesis.

72. Herman van de Werfhorst (2001). *Field of Study and Social Inequality. Four Types of Educational Resources in the Process of Stratification in the Netherlands.* Amsterdam: Thela Thesis.

73. Tamás Bartus (2001). *Social Capital and Earnings Inequalities. The Role of Informal Job Search in Hungary.* Amsterdam: Thela Thesis.

74. Hester Moerbeek (2001). *Friends and Foes in the Occupational Career. The Influence of Sweet and Sour Social Capital on the Labour Market.* Amsterdam: Thela Thesis.

75. Marcel van Assen (2001). *Essays on Actor Perspectives in Exchange Networks and Social Dilemmas.* Amsterdam: Thela Thesis.

76. Inge Sieben (2001). *Sibling Similarities and Social Stratification. The Impact of Family Background across Countries and Cohorts.* Amsterdam: Thela Thesis.

77. Alinda van Bruggen (2001). *Individual Production of Social Well–Being. An Exploratory Study.* Amsterdam: Thela Thesis.

78. Marcel Coenders (2001). *Nationalistic Attitudes and Ethnic Exclusionism in a Comparative Perspective: An Empirical Study of Attitudes Toward the Country and Ethnic Immigrants in 22 Countries.* Amsterdam: Thela Thesis.

79. Marcel Lubbers (2001). *Exclusionistic Electorates. Extreme Right–Wing Voting in Western Europe.* Amsterdam: Thela Thesis.

80. Uwe Matzat (2001). *Social Networks and Cooperation in Electronic Communities. A theoretical–empirical Analysis of Academic Communication and Internet Discussion Groups.* Amsterdam: Thela Thesis.

81. Jacques Janssen (2002). *Do Opposites Attract Divorce? Dimensions of Mixed Marriage and the Risk of Divorce in the Netherlands.* Amsterdam: Thela Thesis.

82. Miranda Jansen (2002). *Waardenorinties en partnerrelaties. Een panelstudie naar wederzijdse invloeden.* Amsterdam: Thela Thesis.

83. Anne Rigt Poortman (2002). *Socioeconomic Causes and Consequences of Divorce.* Amsterdam: Thela Thesis.

84. Alexander Gattig (2002). *Intertemporal Decision Making.* Amsterdam: Thela Thesis.

85. Gerrit Rooks (2002). *Contract en Conflict. Strategisch Management van Inkooptransacties.* Amsterdam: Thela Thesis.

86. Károly Takács (2002). *Social Networks and Intergroup Conflict.* Amsterdam: Thela Thesis.

87. Thomas Gautschi (2002). *Trust and Exchange. Effects of Temporal Embeddedness and Network Embeddedness on Providing and Dividing a Surplus.* Amsterdam: Thela Thesis.