

D055 16793 93707



The Social Institutions of Capitalism

Evolution and Design
of Social Contracts

Edited by
Pursey Heugens
Hans van Oosterhout
Jack Vromen

The Social Institutions of Capitalism

Evolution and Design of Social Contracts

Edited by

Pursey Heugens

Assistant Professor of Organization Theory and Organizational Economics, Utrecht School of Economics, Utrecht University, The Netherlands

Hans van Oosterhout

Assistant Professor of Institutions and Institutional Design, Rotterdam School of Management, Erasmus University, The Netherlands

Jack Vromen

Associate Professor in the Philosophy of Economics, Faculty of Philosophy, Erasmus University, The Netherlands

Edward Elgar

Cheltenham, UK • Northampton, MA, USA

© Pursey Heugens, Hans van Oosterhout, Jack Vromen, 2003

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise without the prior permission of the publisher.

Published by
Edward Elgar Publishing Limited
Glensanda House
Montpellier Parade
Cheltenham
Glos GL50 1UA
UK

Edward Elgar Publishing, Inc.
136 West Street
Suite 202
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Cataloguing in Publication Data

The social institutions of capitalism: evolution and design of social
contracts/edited by Pursey Heugens, Hans van Oosterhout, Jack Vromen.
p. cm.

Includes bibliographical references and index.

1. Social contract. 2. Capitalism. I. Heugens, Pursey. II. Oosterhout, Hans van,
1965–. III. Vromen, Jack, 1958–.
JC336.S54 2004
320.1'1—dc22

2003049268

ISBN 1 84376 495 4 (cased)

Printed and bound in Great Britain by MPG Books Ltd, Bodmin, Cornwall

Figures and tables

FIGURES

2.1	Canonical one-shot games	25
2.2	Equilibrium outcomes in repeated games	27
8.1	Farmers' Dilemma	142
8.2	Prisoners' Dilemma	147
8.3	'Reversed' Prisoners' Dilemma	147
8.4	Assurance game based on $m > 1$: repetition of Farmers' or Prisoners' Dilemma	149
8.5	A justice-Platonist contemplates entering into an agreement	153
8.6	Four-move Centipede game	155
8.7	A Chain Store game	161

TABLES

4.1	Sources of trustworthiness	73
4.2	Modes of intentional trust production	84
4.3	Modes of competence trust production	84
4.4	Attribution of the trustworthiness of X	85

List of contributors

Ken Binmore, University College London

David Campbell, Cardiff University

Thomas Donaldson, University of Pennsylvania

Thomas Dunfee, University of Pennsylvania

Pursey Heugens, Utrecht University

Bart Nooteboom, Erasmus University

Hans van Oosterhout, Erasmus University

Theo van Willigenburg, Erasmus University

Peter Vanderschraaf, Carnegie Mellon University

Toon Vandavelde, Catholic University Leuven

Jack Vromen, Erasmus University

Contents

<i>List of figures and tables</i>	vi
<i>List of contributors</i>	vii
1 Social contract theories: <i>e pluribus unum?</i> <i>Pursey Heugens, Hans van Oosterhout and Jack Vromen</i>	1
2 Stable social contracts <i>Ken Binmore</i>	20
3 The relational constitution of contractual agreement <i>David Campbell</i>	38
4 The foundations of trust <i>Bart Nooteboom</i>	66
5 Economics and the social contract <i>Toon Vandevelde</i>	97
6 Social contracts, <i>sic et non</i> <i>Thomas Donaldson and Thomas Dunfee</i>	109
7 Sources of normativity: reflectivity versus social contracting <i>Theo van Willigenburg</i>	127
8 Justice-conventionalism, justice-Platonism and the social contract <i>Peter Vanderschraaf</i>	141
<i>Index</i>	165

1. Social contract theories: *e pluribus unum?*

**Purseuy Heugens,¹ Hans van Oosterhout
and Jack Vromen**

If the community of social contract theorists – past and present – had to select a single image as a symbol for their work, there is a fair chance that they would agree on picking the engraved title page of the first edition of Thomas Hobbes's *Leviathan* (1651). At the very least, it can be argued that it would be unreasonable for them to oppose this choice (Hampton, 1993: 379). After all, the *Leviathan* is one of the most important monographs in political theory, and Thomas Hobbes is one of the most influential thinkers in the history of social and political thought. The picture on the title page aptly expresses the gist of the theory. It depicts the sovereign power of a nation as being constituted by a compact of all its citizens, who voluntarily hand over most of their naturally obtaining autonomy to this 'mortal God' to escape the complete natural liberty which makes their lives intolerable (Lessnoff, 1990). The engraving by the French artist Abraham Bosse, now part of the collection of the British Museum in London, contains at least two unambiguous messages, aside from the many insightful subtleties it treasures. First, the illustration shows the upper half of a man's body. The man wears an imperial crown, and holds a sword in his right hand and a bishop's crosier in his left. He exercises his power over a walled city featuring a very prominent church, and his credo is: *non est potestas super terram quae comparetur ei* (There is no power on earth which can be compared to him; Job 41:24). This image leaves no room for doubt: Leviathan is the sole, indivisible, and absolute authority, both civil and ecclesiastical.

Upon closer inspection a second message unveils itself, which is even more telling than the first. The body of the sovereign comprises innumerable persons literally incorporated within him. The scene depicted here represents the compact through which individuals at once escape the threats of the state of nature and subject themselves unconditionally to the almighty sovereign. The engraving hence captures the essence of Hobbes's contractarian theory in a single image:

A multitude of men, are made *One* person, when they are by one man, or one Person, Represented; so that it be done with the consent of every one of that Multitude in

particular. For it is the *Unity* of the Representer, not the *Unity* of the Represented, that maketh the Person *One*. (Hobbes, 1968/1651: 220, original emphases)

The sovereign is therefore, in the words of Adams, Franklin and Jefferson, *e pluribus unum* – one from many.

Although Bosse's engraving is telling indeed, an important question is whether it really is an adequate symbol for the intertemporal community of contractualists alluded to in the opening paragraph of this introduction. It is difficult, after all, to overestimate just how many different argumentative uses of the notion of contract have appeared over the centuries (Hampton, 1993). Contractual arguments have been used to explain the state as well as its exact opposite – spontaneous (non-state induced) co-ordination amongst individual agents. Contractual arguments have also been deployed to justify authority, anchor the role of man in the divine plan of the universe, and justify conceptions of justice (to name but a few). Is it reasonable to presume that all these social contract theories have some non-trivial features in common, that is, that they all draw from one common conceptual source and as such form 'one from many' themselves?

Hampton's answer to this question, to be sure, is rather cynical. In her view 'the many differences and disagreements among [contractarians] show that although they are supposedly in the same philosophical camp, in fact they are united not by a common philosophical theory but by a common *image*. Philosophers hate to admit it, but sometimes they work from pictures rather than ideas' (Hampton, 1993: 379, original emphasis). Yet we do *not* subscribe to Hampton's cynical view, as we believe that there *is* a common thread that connects all contractarian theories despite their very diverse origins and articulations. In our view, all contractarian theory – and hence all contributions in this volume – branch out from a single core. The unity of contractarianism does not derive merely and shallowly from a shared use of the social contract metaphor, but rather from a single foundational idea that underlies all contractarian theory. Our conjecture is that this single foundational idea consists of a *basic argumentative structure*, which all theoretical manifestations of contractarianism have in common.

At first glance, the table of contents of this book already casts some serious doubts upon our thesis that all contractarian theories are united by a single basic argumentative structure. This book combines contributions from fields as diverse as economics, evolutionary game theory, contract law, business ethics, moral philosophy and anthropology. If we were to adopt a cynic's perspective, we would be tempted to admit that this variety alone makes it unlikely that we will find a common structure underlying these contributions. But even if we do succeed at making a convincing case for unity, as we certainly hope to do in this introduction, one can always maintain that this common thread is too permissive and lacks *sufficient* substance. This would be the case, for example, if some of

the particular theories exhibiting the underlying structure turn out to be mutually incompatible when 'push comes to shove'. At the same time, it is not up to us to draw this conclusion, as the question whether the common conceptual underpinnings of contractarian thought are too permissive can only be answered by further substantive development of the contractarian paradigm. In the present state of affairs, we somehow have to live with the tension between being too constraining about the conceptual underpinnings on the one hand, and being too permissive on the other. This introductory chapter is divided into four sections. First, we provide some historical context by sketching a brief – yet obviously incomplete – exposé of the historical variety and robustness of contractarian thinking. Next, we sketch the contemporary variety in contractarian theorizing by introducing the contributions to this book. In combination, these first two sections demonstrate the pressing need to find some common conceptual ground for contractarian theory in general. Section three proposes a view on how all of these contributions branch out from a common core in the form of a basic argumentative structure. We demonstrate not only that this structure is discernible in all social contract theories discussed in this volume, but also that its key constitutive elements are typically interpreted differently across various social contract theories. We conclude with some further thoughts on the tension between being too constraining or too permissive in pinpointing common conceptual ground in contractarian theorizing.

THE INTERTEMPORAL SCOPE AND ROBUSTNESS OF SOCIAL CONTRACT THEORY

Social contract theories have been around for a long time and have proven remarkably robust over the centuries. McClelland (1996) reminds us that essential elements of such theories were already present in the biblical covenant between God and the Children of Israel. As the Children of Israel negotiated with God to keep his laws in return for his promise to give them the Land, they effectively created a nation-state by contract. McClelland also explains that since the book of Genesis is an extended account of Moses' struggles to make his people keep their side of the promise, 'Mosaic leadership is a perfect image of a prince keeping his unruly people to their faith in return for a promise of future benefits' (ibid.: 174).

Plato has also made use of a metaphorical social contract in his Socratic dialogues, mostly to explain the relationship between the state and its citizens. In one of these dialogues we find Socrates in jail awaiting his death sentence, when he is visited by his friend Crito. When Crito urges Socrates to flee in order to escape the death penalty, he refuses on the grounds that this would imply

a breach of the social contract he has tacitly consented to. He explains what the laws and the government of Athens would say if they would interrogate him:

he who disobeys us is, as we maintain ... wrong ... because *he has made an agreement with us that he will duly obey our commands*; and he neither obeys them nor convinces us that our commands are unjust; and we do not rudely impose them, but give him the alternative of obeying or convincing us; – that is what we offer, and he does neither. (Plato, 1993, 'Kriton': 43; our emphasis).

These two examples demonstrate that contractarian views have long coloured our ideas of social organization. Yet, more systematic contractarian theories did not emerge until the fifteenth century, in the context of reflection on the conditions of political obligation in feudal monarchy (Lessnoff, 1990). Early contributors, such as the Alsatian monk Manegold of Lautenbach, used contractarian theories to argue for the immanent equality of rulers and their people:

King is not a name of nature but a title of office: nor does the people exalt him so high above it in order to give him the free power of playing the tyrant in its midst, but to defend it from tyranny. So soon as he begins to act the tyrant, is it not plain that he falls from the dignity granted to him? Since it is evident that he has broken the contract by virtue of which he was appointed. If one should engage a man for a fair wage to tend swine, and he finds means not to tend but to steal them, would one not remove him from his charge? (Quoted in Ritchie, 1893: 203)

Views of this kind sent a warning to kings and princes that the people had only lent them their power, and that rebellion could be condoned if they were to violate the terms attached to that loan (Hampton, 1993).

No matter how conceptually convincing these early sparks were, they did not create much of a fire because in the fifteenth century there hardly ever was a serious case to be made for civil disobedience (McClelland, 1996). After all, every person in Western societies – ruler and subject – was bound by the common law of Christianity. Since rulers were widely seen as the embodiment of God's majesty on earth, disobedience would earn dissatisfied subjects the doubly unattractive titles of anarchist *and* heretic. All of this changed, although certainly not overnight, with the Reformation. When Luther, Calvin and their followers started questioning the authority of the Catholic Church, it became less and less obvious that God wanted us to obey a particular ruler and live up to a specified set of rules. When Christianity was split into two competing halves – one Catholic and one Protestant – it became necessary for any particular king or prince to *justify* his role as ruler and lawgiver. It almost goes without saying that many theorists and princes alike found that contractarian theories were best equipped for this challenging job.

The first thinkers to take up the challenge presented by the Reformation were the radical contractarians Hobbes and Pufendorf. Hobbes is undoubtedly the

more influential of the two, but both share a rather bleak image of the nature of man as a point of departure and used this image to legitimate a subordination contract with a ruler that would allow them to abandon the miserable state that one man creates in the presence of another. Hobbes's description of the state of nature is so oft quoted that we will not repeat it here, but Pufendorf's account of man's natural state and the solution he proposes is well worth repeating. He states that:

There are two principal faults in human nature which prevent a number of independent men who are not subordinate to one another from achieving durable cooperation in a common end. One is the diversity of inclinations and judgments in deciding what is most conducive to that end. ... The other fault is indolence, and disinclination to do what is useful ... The first fault is countered by a perpetual union of the wills of all; the second by constituting some power, which shall be directly before their eyes, capable of inflicting suffering on those who oppose the common interest. (Pufendorf, 1991/1673: 136)

What makes both Hobbes and Pufendorf *radical* contractarians is, first, that they both agree that the intolerability of the pre-political state of nature makes it practically unjustifiable to ever resist or challenge the sovereign. Secondly, both thinkers insist that the supreme political authority is not accountable to, or punishable by, the individual citizens it governs (Lessnoff, 1990).

Unsurprisingly, the rather rabid version of contractarianism that Hobbes and Pufendorf proposed invoked critique from more moderate contractarians. It first and foremost came from John Locke, who could hardly believe that rational persons could consent to a scheme wherein the ultimate authority would literally be above the law. In critique of Hobbes's radical views he wrote:

As if when men, quitting the state of nature, entered into society, they agreed that all of them but one should be under the restraint of laws, but that he should still retain all the liberty of the state of nature, increased with power, and made licentious with impunity. This is to think that men are so foolish that they take care to avoid what mischiefs may be done to them by polecats or foxes, but are content, nay think it safety, to be devoured by lions. (Locke, 1993/1689: 308)

Locke's alternative to the unconditional surrender of power proposed by Hobbes is to construe the affiliation between the ruler and the people as an agency relationship, whereby the former is the representative while the latter are the principals (Hampton, 1986). His argument is that persons enjoy certain unalienable rights, such as those to life, liberty and estate. Hence, rather than abstaining from these rights for good, rational persons will only 'lend' these rights to government on condition that it will not abuse them. Pufendorf joined Locke in his criticism of Hobbes in that he also insisted that rational persons would not institute a sovereign without at least getting a promise of protection

in return (Lessnoff, 1990). Yet, whereas Pufendorf was hesitant to accept the consequences of his own views and plead for complete subjection to the sovereign nevertheless, Locke explicitly stated that the people are 'absolved from any further obedience' (Locke, 1993/1689: 374) if the state attempts to grasp an absolute power over their lives, liberties and estates.

Perhaps the most compelling explanation for the remarkable intertemporal scope of social contract theory is that the approach has always been able to morph itself to the social and political circumstances of the time. In the seventeenth century, the contractarian theories of Hobbes, Pufendorf and Locke had wide appeal because they harboured the promise of social stability when most of Europe was immersed in wars, revolutions and civil wars. At the dawn of the nineteenth century, over a hundred years after the publication of Hobbes's *Leviathan*, Europe had become so used to social living that stable government was almost taken for granted (McClelland, 1996). Thanks to Jean-Jacques Rousseau, social contract theory was transformed from an approach explaining the stability of societies to a vehicle for institutional redesign.

For Rousseau, social contracts are not needed for explaining the origins of society, as he believed persons to be naturally sociable. This does not mean, however, that our sociability will prevent us from running into all kinds of social problems. On the contrary, Rousseau has argued that human history eventually but unavoidably leads to fraudulent and deceitful societies, as 'progress' requires exploitation of the masses and appropriation of public resources. His description of this process is famous:

The first man who, having enclosed a piece of land, thought of saying 'this is mine' and found people simple enough to believe him, was the true founder of civil society. How many crimes, wars, murders; how much misery and horror the human race would have been spared if someone had pulled up the stakes and filled in the ditch and cried out to his fellow men: 'Beware of listening to this impostor. You are lost if you forget that the fruits of the earth belong to everyone and that the earth itself belongs to no one!' (Rousseau, 1984/1755: 109)

In his quest for a just society, Rousseau initiated the use of the social contract as an intellectual tool for social and political *redesign* that started not from some hypothetical state of affairs, but from the world as we know it to be. The key problem to which such an effort must provide an answer is to 'Find a form of association which will defend and protect, with the whole of its joint strength, the person and property each associate, and under which each of them, uniting himself to all, will obey himself alone, and remain as free as before' (Rousseau, 1994/1762: 54–5). Rousseau denies in the strongest possible terms that the answer to this problem could lie in securing safety through a (partial) surrender of freedom. Instead, he proposes a social contract on the basis of complete equality and reciprocity, in which each person is an essential and inseparable

part of the sovereign body governing society. With this answer, Rousseau managed to bring the notion of social contract in line with the requirements and expectations of modern times.

THE CONTEMPORANEOUS DIVERSITY OF SOCIAL CONTRACT THEORIES

Versatility has always been one of social contract theory's main assets, and its persistence in the twentieth century clearly testifies to this quality. What was unprecedented even for a successful approach like contractarianism, however, was that the second half of the century brought forward a renaissance in which social contract theories unexpectedly started to appear in almost every area of the social sciences. In public choice, or in what is now referred to as *positive* political theory, for example, the notion of social contract has become the foundation for the sub-discipline of constitutional political economy. In normative political theory, the social contract idea was famously revived by John Rawls (1971) in his work on social and political justice. Gauthier (1986) used a contractarian theoretical structure to demonstrate that substantive ethical norms can be derived from assumptions on the rationality and self-serving behaviour of individuals. Rational choice theorists in general, and game theoreticians in particular (Binmore, 1994; 1998; Skyrms, 1996; Sugden, 1986), have used the notion of social contract in their efforts to explain how social co-ordination can evolve spontaneously, that is, without being initiated by state intervention or by concerted individual action. In organization studies and business ethics, finally, the idea of social contract has been elaborated on in conceiving the ethical norms that should guide and constrain practices in business and organization (Donaldson, 1982; Donaldson and Dunfee, 1999; Keely, 1988). The seven contributions to this book both exemplify this renaissance and illustrate the remarkable interdisciplinary scope that characterizes contemporaneous social contract theorizing.

Chapter 2 is entitled 'Stable social contracts', a contribution by evolutionary game theorist Ken Binmore. Skyrms (1996) notes that game theorists are less likely to be inspired by Hobbesian questions about what sort of contract rational decision-makers would agree to in the state of nature than by Rousseauist puzzles concerning the origin and continued evolution of the existing social contract. Binmore seems to fit this profile as he sets out to explore why some social organizations survive when others collapse and die. His answer is that stable social contracts are not the ones that are dependent on the contractors abiding by the terms of the agreement because of some intrinsic commitment to honour their duties and obligations. Instead, social con-

ventions become durable when it is in the interest of each individual participant with the power to uproot the scheme not to do so. As long as societal members are only asked to play those roles that are compatible with their own individual interests, a collection of Mr Hydes can be expected to co-operate like a collective of Dr Jekylls.

Contract law scholar David Campbell contributes the third chapter of this volume, 'The relational constitution of contractual agreement'. Goodhart (1997: 1) once wrote that the field of law and economics is characterized by 'too much one-way traffic', in that economic tools are frequently used for evaluating legal structures, whilst few assessments are made of how the law facilitates or affects market transactions. Campbell sets out to address this problem by proposing that insights from contract law can help new institutional economists attain a better grasp on the fundamental notion of transaction costs. In his view, economists are mistaken if they see transactions costs simply as 'friction' that has to be minimized. He urges us that activities like negotiating, information gathering and making mutual adjustments are not simply costly frictions that have to be reduced to the absolute minimum, but vital parts of the underlying social relationships without which economic exchange would not even be possible in the first place. In Campbell's own words: 'If one really took away *all* the costs of exchanging, the exchange would not take place cost free; it would not take place.'

Chapter 4, 'The foundations of trust', comes from organization theorist Bart Nooteboom. This contribution adds a novel perspective to the continuing debate in new institutional economics concerning the validity of transaction cost theory's behavioural assumptions, especially on the question whether or not individuals will opportunistically seek to advance their own interests at the expense of their exchange partners. Nooteboom sets out to find new answers to an old question raised by Oliver Williamson (1985): is trust that goes beyond calculative self-interest blind? His answer is that it is not, and that there are two good reasons to assume that opportunism is often absent even when there are good opportunities and strong incentives for it. First, the correct observation that trust is not unconditional or unbound does not imply that trust cannot be real. Trustworthiness is generally believed to have its limits, and it is not rational to expect individuals to live up to their commitments when their life or liberties are at stake. But as long as those limits are not transgressed, trust can still be genuine and lead to forbearance, reciprocity, predictability and benevolence. Secondly, to trust someone can still be rational – and hence not be blind – even if it is not calculative. Nooteboom reminds us that all individuals have their cognitive limitations, and that we often have to make the rational decision to economize on calculative deliberations in order to free up mental resources for more important decisions in life. If we must rely on other individuals to make certain decisions for us because of our lack of decision-making powers, it can

be a rationally justifiable decision to let only those individuals whom we trust make decisions on our behalf.

Philosopher of economics Toon Vandevelde is responsible for the fifth chapter of this book: 'Economics and the social contract'. Vandevelde is inspired by the question whether contractarianism should be seen as a normative device for guiding institutional redesign attempts at creating more just societies or as an explanatory theory that could help social scientists explain all exchange-oriented aspects of social life. He is rather sceptical about the explanatory power of contractarianism as a social science theory. Just like rational choice economics has a tendency to explain all human action in terms of the self-interests of the actors involved – even deeds that suggest great altruism or that require great sacrifice – it is not difficult to see a hypothetical contract underlying every form of social exchange for those who *want* to see one. Vandevelde expects more from contractarianism as a normative device. His argument is that contractarianism is an important cognitive tool that can help us understand the requirements for and consequences of pro-social norms like the insurance principle and norms of solidarity. Because contractarianism is a powerful tool that can inform us about the effects that alterations to present-day conventions will have in the future, it can aid policy-makers in their attempts to deal with matters of societal unjustness.

Chapter 6, 'Social contracts, *sic et non*', is written by business ethicists Tom Donaldson and Tom Dunfee. Donaldson and Dunfee are long-time advocates of theoretically rigorous yet practically feasible approaches to business ethics, and their own 'integrative social contracts theory' has been praised as 'an important contribution to both moral philosophy and the field of management that amply illustrates how ethical theory can yield practical business results' (Boatright, 2000: 452). The theory is a novelty in that it clarifies the relationship between two different types of social contracts. At the most basic level, there are micro-social contracts: norm-setting agreements between members of a particular community. The theory does not prescribe or predict the content of these contracts, and in fact allows community members the 'moral free space' to adopt and maintain the norms which they prefer. This is not the same as saying that anything goes, however, because sooner or later these lower-level agreements will be tested against the background of what Donaldson and Dunfee call the macro-social contract. This second type of contract consists of a set of 'hypernorms' which guarantee persons a set of unalienable rights, such as those to life, fair treatment and non-discrimination. When a certain norm promoted in a micro-social contract is incongruent with these hyper-norms, the contract cannot be accepted as legitimate, even though it may be an authentic norm recognized and acted upon by members of the contracting community.

The seventh chapter of this book, 'Sources of normativity: reflectivity versus social contracting', comes from moral philosopher Theo van Willigenburg. For

radical contractarians, the normativity of any set of principles derives from the fact that they are mutually agreed upon by freely consenting individuals (Gauthier, 1986). More moderate contractarians, like Donaldson and Dunfee (1999), may believe that these principles then still need to be tested against a background contract consisting of higher-order norms. Yet even these temperate contractarians agree that their macro-social contract should only be used as a marginal test, and that within these limits contractors enjoy ample 'moral free space' to consent to norms they deem appropriate. Van Willigenburg is very critical of this approach, and believes that what he calls rational choice contractarianism can never be considered a satisfactory source of normativity because this approach is hampered by a number of insurmountable problems. He argues that contractarian theory offers no convincing response to a free-rider who has effectively found a loophole in the social contract allowing her to evade her responsibilities. In his Kantian view, the most authoritative source of normativity we should accept to guide our behaviours is practical reflection on compelling and conclusive reasons by autonomous individuals. Because of our capacity for critical reflection, we cannot simply ignore compelling reasons speaking for or against what is attractive or repulsive once we know they exist. The normativity of any reason, then, stems from a reflective individual's own insight into the normative force of this reason and one's *taking* it as a reason. In social situations, reasons do not derive their normativity from the fact that they are mutually agreed upon by a group of rational actors with their own interests in mind, but from their universality; from the fact that every rational individual in similar circumstances would also have to admit to the compellingness of that reason.

Political philosopher Peter Vanderschraaf has written the eighth and final chapter of this book, 'Justice-conventionalism, justice-Platonism and the social contract'. In his view, justice-conventionalism and justice-Platonism represent two alternative approaches for vindicating the rationality of keeping one's promises. Justice-Platonists argue that there is something inherently desirable about promise-keeping, and that some of the goods that might accrue to any person are inseparable from them being just. Justice-conventionalists do not believe that there are goods internal to the practice of justice, and propose that promise-keeping can only be defended when it helps generate and maintain a system of reciprocal expectations that enable members of society to reach higher material gains than the ones they could achieve working independently. Vanderschraaf presents a careful analytical account of what we may expect when justice-conventionalists interact with justice-Platonists in the context of promise-making and promise-keeping. The implications of his analyses for the empirical world are even more telling, however. They follow from the observation that in any society we will find justice-Platonists (who will always keep their promises), justice-conventionalists (who will keep their promises as long as

this yields more benefits than defecting) and ‘Fooles’ or ‘Mutants’ (who can never be trusted to keep their promises). The gist of this observation is that the rationality of entering into and keeping agreements varies according to one’s beliefs about the relative sizes of these groups (which determine the perceived probability of encountering a promise-breaker or a promise-keeper in a social exchange situation with an unfamiliar actor) as well as one’s personal commitment to justice-Platonism or justice-conventionalism. Hence, empirical beliefs matter when one considers entering into social contracts.

THE BASIC ARGUMENTATIVE STRUCTURE OF SOCIAL CONTRACT THEORY

The historical and contextual robustness of social contract theory indicates that there must be something compelling about the approach. At the same time, its remarkable intertemporal and interdisciplinary scope also raises serious doubts as to whether it really represents an unambiguous and unified body of work. It may very well be that the social contract tradition merely constitutes a family of theories lacking a common substantial underpinning. In such a ‘family’ of approaches all the individual family members are related, but not in any definite sense because no single assumption or basic proposition is shared by all members of this family (Wittgenstein, 1967). It is not easy to see, after all, which fundamental assumptions and propositions are shared by ‘contractualists’ like Rousseau, Rawls and Scanlon (1982; 1998) on the one hand, and ‘contractarians’ such as Hobbes, Gauthier and Binmore on the other.² Yet on closer inspection, and with some presumptive goodwill, something of a common *argumentative structure* can be discerned underneath the variety of theories denoted by the term ‘social contract theory’. In any case, one could always stipulate that theories that diverge too much from this basic structure are not really social contract theories. But it is clear that such a move should only be made as a last resort. Let us first examine how inclusive the basic argumentative structure we propose actually is.

In our view, *the gist of all social contract theory is that only those institutional and social constraints to human action have (a) normative appeal and (b) will be acknowledged and conformed to in everyday behaviour, that rational or reasonable agents could, and arguably would agree to if they had the choice* (cf. Hampton, 1993; Rawls, 1971; Scanlon, 1998). In order for this basic argumentative structure to do the job we have in mind for it, which is to uncover the common identity of all branches of the social contract tradition, we need to be rather liberal in interpreting the individual terms of the formula given above. Our priority at this point is not to further investigate the true meaning of concepts