Joint Discussion Paper Series in Economics

by the Universities of

Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 08-2021

Werner Güth and Hartmut Kliemt

Sustainable Procedures of Corporate Social Responsibility

This paper can be downloaded from
http://www.uni-marburg.de/fb02/makro/forschung/magkspapers

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

Sustainable Procedures of Corporate Social Responsibility

Werner Güth Hartmut Kliemt

February, 26 2021

Werner Güth, Max-Planck-Institut zur Erforschung von Gemeinschaftsgütern
Kurt-Schumacher-Str. 10, D 53113 Bonn, Germany

Hartmut Kliemt, Guest Professor Justus Liebig University Behavioral and Institutional Economics (VWL VI) Licher Str. 66, 35393 Giessen, Germany

Corresponding author e-mail-address: Hartmut.Kliemt@gmail.com

Abstract

Directly non-profit-seeking CSR strategies can be indirectly profitable and sustainable provided their presence can be detected by other actors with sufficient reliability at sufficiently low costs. This creates a niche for implementing CSR procedures in ways that are compatible with shareholder approaches to responsible corporate governance. Though ethical norms of equal interpersonal respect for all stakeholders are appealing, spelling them out in concrete procedural terms renders them ethically less attractive than they appear in theory. If ethical ideals of equal respect for all stakeholders in CSR do not have the extreme implications laid out here where are the alternatives?

Key words: Equal Ethical Respect, Corporate Social Responsibility, Obliquity, Shareholder and Stakeholder Conceptions

JEL Classification: Business Ethics, Welfare Economics, Evolutionary Economics

1. Introduction and overview

Milton Friedman maintains that under appropriate conditions the social responsibility of personal as well as corporate actors, when participating in the economy, demands that they act in line with "profit signals" as emerge on markets within a given political framework (Friedman 1970). According to normative welfare economics and in particular its so-called first theorem, following profit signals leads for a given *political* framework of law enforcement and given preferences to Pareto non-dominated results. Beneficence as an ethical principle demands doing good and this in turn demands that Pareto dominated results be avoided (because otherwise "doing better" would be possible in terms of what the individuals themselves want). Under competitive market conditions selfish profit seekers are unintentionally doing good to the extent that they avoid Pareto dominated alternatives when following market signals.[1] However,

[1] "Private Vices, Publick Benefits" (Mandeville 1924); social theorists from Cardinal de Reetz over Adam Smith, Adam Ferguson to Friedrich-August von Hayek have dwelled upon this insight. Well selected anthologies of the British in general and the Scottish Moralists in particular are (Raphael 1969),(Schneider 1967).

selfless moral actors who intend to do "good" by avoiding Pareto dominated results must show the same overt behavior and follow the same signals *if they respect the political framework of a free-market society* and the autonomy of individuals operating in it.[2]

In view of the conventional verdict against interpersonal comparisons of utility there can be no objective way to tell which of several Pareto efficient alternatives should be morally preferred when in a Pareto dominated state. There can be as many conceptions of the common weal and as many legitimate welfare functions as there are individuals. Singling out one of the many Pareto efficient results and imposing it against what would "spontaneously" emerge from autonomous individual decisions is disrespectful of the decision autonomy of at least one individual. Since it requires the application of the fundamental coercive power of the state it must be a responsibility of politics rather than of personal or corporate actors in their private capacities. In the private sphere the personal and corporate social responsibility to contribute to efficiency must not be imposed. It must be discharged by following market signals as provided in the process of market competition.

It may seem that according to the preceding line of argument, *"Corporate Social Responsibility", CSR*, implies an obligation to show directly profit-seeking behavior in all instances of market choice.[3] However, more often than not "our aims are best achieved indirectly" (Kay 2011). We reach them in an "oblique"[4] manner without directly intending it. Obliquity allows for foregoing proximate in favor of more remote advantages and for pursuing "higher order" preferences that acknowledge preferences among the objects of preference orderings.[5] This creates moral wiggle room for CSR. CSR aims can be pursued within a

[2] A still under-appreciated discussion of such arguments can be found in (Carens 1981); inspired by (Hayek 1945) and taken up later in (Cohen 2009).
[3] This rather obvious point is fully appreciated not often as well as in (Karnani 2010).
[4] It may not only be the case that we can reach our ends unintendedly (*weak obliquity*); quite often we can reach our ends only by intentionally committing to a strategy of not seizing opportunities to further our ends directly and intentionally (*strong obliquity).*
[5] Focusing merely on behavioral concepts like so-called revealed preference, economists try to get rid of higher order preference concepts. In particular, when dealing with (in)efficiency economists tend to avoid any distinction between what choice makers want and choose, on the one hand, and what they on reflection might want to choose, on the other hand.

justificatory framework that is compliant with the normative value pluralism and the skeptical meta-ethical framework of classical liberal economics.[6] Being meta-ethical skeptics ourselves we are aware that the substantive normative claim that plural values ought to be respected by a liberal legal order and the managers operating within it is not a "logical" consequence of a skeptical meta-ethical perspective. We happen to share the liberal normative ideal of interpersonal respect of plural values without. This is the premise from which we start without any claim to provide an ultimate justification for moral skepticism and the substantive moral pluralism implicitly underlying liberal neo-classical economics.

At the same time we will strongly diverge from neo-classical economics in emphasizing intrinsic besides extrinsic motivation and the indirectness of human pursuits that is related to this. Spelling out the basic intuitions underlying what we call an "indirect evolutionary approach" (to understanding competitive processes), on the one, and "corporate constitutional commitments" (as strategic competitive devices), on the other hand, we can put CSR into an evolutionary and a constitutional economics perspective:[7] The indirect evolutionary approach, models the conditions under which CSR-oriented corporate actors can in principle succeed in competition *provided* that internal corporate constitutional commitments to CSR can in principle be implemented such that their presence can be detected by other contract parties (or credibly revealed to them).

In the first main section (2.) we conceptualize companies as unitary actors whose overt behavior can be described *as if* they were pursuing an objective function. This corresponds to conventional economic ways of modeling inter-firm competition. But other than in more conventional approaches we rely on a dual objective function approach: one function represents direct or proximate substantive success and the other function represents *intrinsic motives* other

[6] See on this (Robbins 1935) and for further comments (Kliemt 1986).
[7] Even the implementation of stakeholder conceptions as a tool of strategic management may itself, if obliquely, contribute to ultimate shareholder interests; on stakeholder conceptions in general see, for instance, (Freeman et al. 2010).

than seeking proximate success (directly). The upshot of this application of our previous work on indirect evolution to CSR issues is that conditions under which intrinsic motives and directly non-profit-seeking behavior can ultimately be sustainable can render CSR sustainable as well.

In the second main section (3.) we turn to the question of how corporate actors who are able to act opportunistically can at all credibly commit to not making use of that faculty when certain "temptations" (i.e. proximate extrinsic motives to the contrary) arise.[8] We ask by ways of simple illustrative examples what kinds of internal corporate governance processes may conceivably lead to credible commitments and constrained behavior of the corporate actor such that indirect evolution can favor committed over uncommitted actor types in situations that allow for type detection and discretionary decision-making based on it.

The final section (4.) puts our discussion of what makes the commitments underlying CSR proximately possible (3.) and ultimately sustainable (2.) into a broader economic and ethical perspective. CSR that goes beyond indirectly furthering business is morally precarious. Yet if we take the obliquity of all human endeavors seriously there is moral room for sustainable CSR. Such CSR is *ultimately* simply good business while *proximately* intending to do "good".

2. Sustainable *CSR*

A most general question about evolution in the animal kingdom (including, of course, the human species) concerns the possibility of altruism among separate competing individual units of selection.[9] Most solutions of this "riddle" depend in one way or other on the distinction

[8] This issue of corporate governance is quite analogous to the question whether a political constitution is not like a chastity belt to which the lady happens to hold the key; see (Jasay 1985).

[9] Darwin (and in a way (Hobbes 1651), § 10 ff. before him) already understood that self-sacrifice would be evolutionarily self-defeating unless there were some indirect benefits attached to it. For instance dealing with sterile castes of certain insect species Darwin came already up with a theory akin to kin selection to solve the problem. He talked about breeders who would get long-horned oxen by breeding long-horned cows and bulls, see (Darwin 1975). In response to Darwin's original work the astonishing amount of altruistic behavior, as vividly described in (Kropotkin 1919), had been pointed out by some so-called Social Darwinists (Hofstadter 1969) before modern evolutionary (game) theory solved the riddle; to name but a few examples of the more recent literature (Maynard-Smith 1982), (Sober and Wilson 1998) and more recently with respect to human behavior (Bowles and Gintis 2013). The upshot of the older literature between old and modern Darwinism has been presented conveniently in (R. Dawkins 1976).

between proximate and ultimate causes. For example, polar hares appear to be white because their fur reflects light of a certain wavelength (proximate cause of the appearance) while the emergence of the fur giving rise to the wavelength is explained by the fact that their predators, e.g. polar foxes, can identify dark objects in the Arctic environment more easily than lighter ones and therefore prey differentially on the darker individuals (ultimate cause of the appearance). Due to this selective force more of the relatively lighter than darker hares survive and have progeny in the next generation pushing average fur color towards the lighter end of the spectrum.

The selection of lighter rather than darker individuals in the example of arctic hares seems to suggest that ultimate success depends on the presence of direct, proximate advantage. If that were so, the chances for sustainable CSR would be dim. Since CSR activities are understood to be directly profit decreasing it has to be shown how, say, a company can develop a commitment to forego proximate profits and *thereby* increase its ultimate profit prospects.[10]

2.1 From trivial sacrifices to oblique profit-seeking

CSR investments can be normal profit-seeking activities in which opportunities are taken whenever they offer themselves and are evaluated not only in view of their proximate but also in view of their ultimate (full) consequences.[11] As long as CSR investments are made intentionally in the light of anticipated future consequences –however remote – they remain directly profit-seeking activities. Assuming non-myopic behavior within the constraints the corporate actor faces, the characteristic aspect of indirectly profitable CSR we are interested in is the foreclosing (intentional or unintentional) of options that an *un*boundedly rational actor would choose if he were still free to seize these opportunities. This leads to the question of

[10] It may be noted in passing that this not based on a "handicap" principle since the commitment as such is the property selected; see (Zahavi and Zahavi 1997).

[11] This in a way repeats the preceding argument that behaving according to a stakeholder approach in strategic management can ultimately be justified as contributing to shareholder value. In his elegant „economics in one lesson" Hazlitt describes the focus on all – including the most remote consequences of action – as the hallmark of rational economic behavior (Hazlitt 1988). No procrastination, no weakness of the will and no limits to the imagination of the future.

when the restriction of opportunistic behavior by commitments can be sustainable despite the presence of competitors not so restrained.

2.2 The sustainability of CSR commitments in competition

According to common sense a "responsible corporate citizen" is broadly speaking a corporate actor "bounded" to forego certain proximate opportunities. The "good" company acts in compliance with certain (self-)imposed restrictions internal to it. Like a virtuous individual who is by some intrinsic motivational process restricted from following extrinsic incentives the company is disposed by "firm commitments"[12] not to go for any opportunities as open up within the scope of legality. As far as CSR is concerned this means that the good corporate citizen adopts social responsibilities that seem directly unfavorable to its bottom line.

As in the case of natural selection of virtuous individuals who are restricted by inner commitments in their opportunity taking behavior it seems initially puzzling how it can ever be (even in a wide sense of the term) "profitable" for a fully rational corporate actor to lack or to forego the ability to realize opportunities. A rational actor endowed with "foresight and understanding" takes into account all future causal consequences of actions as foreseen by her in the instance of choice making. She does not know what she will know. However, as a rational actor she knows that she will never choose a dominated alternative. She does not suffer from weakness of the will and other imperfections of rationality. Therefore, she *will be better off with more rather than fewer options no matter what* and seems to have no use for commitments that restrict her future option sets.[13]

So, can it ever be advantageous for a non-myopic rational corporate actor to close off (part of) its ability to seize opportunities for (proximate) profit? or even more boldly: "If Homo

[12] Alluding to the nice pun of (Mayer 2013).
[13] See on such imperfections and rational self-management (Ainslee 2002), (Schelling 1984), (Thaler and Shefrin 1981).

Ecomomicus Could Choose his Own Utility Function, Would He Want One with a Conscience?" (Frank 1987)?

2.2.1 An elementary example of how inner commitments may pay off

Imagine a simple exchange in which one actor has to move first and another one second. Specifically, think of an exchange of apples and bananas. Person A has apples, *a* and person B has bananas, *b*.[14] A prefers *b* to *a* while person B prefers *a* to *b*. Let the apples be ripe before the bananas. A has to deliver his apples *a* first in the hope that B will deliver her bananas *b* afterwards. If A anticipates opportunistic behavior of B then he should not deliver his apples. A merely extrinsically motivated B is better off when holding apples *a* and bananas *b* rather than only apples *a*. Therefore, unless A could believe that B is committed to fairly reward it, trusting behavior of A should not be forthcoming. Both actors are stuck with their initial holdings even though they prefer the result of exchange to the status quo.

Only if individuals are in fact sufficiently committed to honest dealings (by natural retributive emotions and/or upbringing) *and* can be recognized as bearers of a commitment to deal honestly with sufficient certainty, they can (relative to their subjective evaluation and knowledge) *rationally* realize the exchange. If it is possible to recognize the presence of "virtue" in other actors, then those endowed with such traits would be sought as partners for mutually advantageous transactions. More chances of co-operation should be on offer to committed individuals and they should outcompete the uncommitted.[15] Individuals who are uncommitted should desire to be committed – as long as the additional opportunities opened up by commitments are not outweighed through the opportunities closed off by the commitments. If external mechanisms like explicit contract enforcement by third parties are lacking

[14] In principle one would have to think of a certain quality and quantity of apples and bananas at a certain time.
[15] Early examples of models that use the (Axelrod 1984) framework but allow for an exit option are (Schüssler 1990) and (Vanberg and Congleton 1992).

individuals would want to be endowed with internal commitments to the extent that their presence can be detected or signaled at sufficiently low costs with sufficient reliability.

To stay with the simple example, person B would have good reason to use an *additional* commitment option amounting to guaranteeing and to signaling that she will deliver the bananas *b* after receiving the apples *a*. If this commitment opportunity *in fact* exists and if its existence can be credibly signaled to A this will induce actor A to deliver the apples, *a*. And, in consequence, both actors will ultimately be better off. By becoming committed to doing "good" and by being able to signal this, B is made better off in *substantive* terms and so is A (i.e both are doing well or better than without B doing good and being able to commit to it).

2.2.2 Indirect evolutionary competition

Assuming that rational human individuals *can* be *internally* committed to forego direct material advantages and can signal the presence of such *intrinsic motivational constraints* of their *ability* to take opportunities it is obvious how committed individuals may in principle outcompete (proximate or direct) opportunists: Provided they can detect the presence of commitments potential co-operation partners can choose to act more co-operatively when encountering committed rather than uncommitted partners. Under certain conditions this can be of sufficient advantage for the less flexible (committed) to outcompete the more flexible (uncommitted). To a characterization of these conditions within a basically neo-classical "maximization under constraints" framework we turn next.

2.2.2.1 A dual objective function model of representing unitary actor choices

Assume that the behavior of the unitary actor is subject to two evaluation functions, one representing substantive external and one representing modifying factors internal to the actor.[16] For instance, on a competitive market the substantive evaluation will typically be in pecuniary while the internal one may be in (broadly) moral terms. Behavior is internally determined in a

[16] More details on explicitly indirect evolutionary models can be found for instance in (Güth and Kliemt 1994), (Berninghaus, Güth, and Kliemt 2003).

process subject to moral constraints (like, e.g. honesty). Market success is determined according to substantive (pecuniary) results of overt action.

In the simplest formal model the overt behavior of actors $i = 1, 2, \ldots, n$ is chosen according to subjective evaluations that are represented by $u_i, i = 1, 2, \ldots, n$. Choices are yielding objective proximate success according to $\xi_i, i = 1, 2, \ldots, n$. It is possible that for alternatives *a* and *b* we can have $\xi_i(a) > \xi_i(b)$ and $u_i(a) < u_i(b)$; that is, the alternative *b* is subjectively preferred to *a* and chosen, despite $\xi_i(a) > \xi_i(b)$. This models the commitment to forego substantive opportunities.

We can reformulate the question raised at the end of section 2.1 and ask how corporate actors with preference representations, u_i, that lead to deviations from what is substantively advantageous, ξ_i, can survive over the long haul. To translate this query into the CSR context: Assume that the CSR orientation of corporate actor i is represented by an, u_i, leading to corporate actions deviating from what is in substantive terms, ξ_i, proximately advantageous, can such an actor ultimately survive? Or, to put it slightly otherwise, is "doing well by doing CSR" a case of "doing well by doing good" and sustainable over the long haul?

In the next step we intend to illustrate in terms of an example – akin to the initial trade-and-trust setting of the exchange of apples *a* and bananas *b* – that being intrinsically motivated to deviate from precepts that lead to proximate substantive success can ultimately pay off in substantive terms. We will restrict ourselves to a graphical presentation of the basic ideas and touch on analytical details only briefly.[17]

2.2.2.2 A dual objective function model of exchange

The following graph taken from (Güth and Kliemt 2000) represents the basic trust interaction in the familiar stylized way

[17] The latter will neither be necessary nor particularly instructive for the transfer of the basic ideas to *CSR*; see for a rather comprehensive overview (Berninghaus, Güth and Kliemt, 2012).

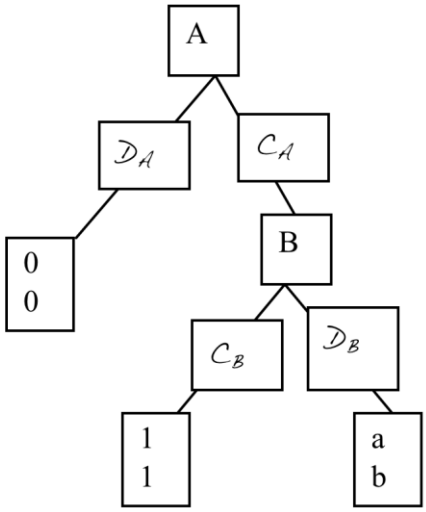

Figure 1, b>1>0>a

Interpreting Figure 1 as a game in the technical sense of the term the numerical values at the end-nodes of the tree are representations of preferences "all things considered". The scale factor and zero points of the representation have been chosen according to some appropriate positive linear transformation.[18] We have – left to right – at the origin of the tree two alternatives open to A. *First, D_A, or non-delivery of apples by first moving A.* Let both assign to this alternative a utility of "0". *Second, C_A, or delivery of apples a by the first mover A.* After C_A the second mover B decides: B delivers *b,* choosing move C_B – leading to a "subjective" overall payoff of "1" for, B, and of "1" for A – or the second mover chooses D_B keeping both, *a* and *b* – leading to a "subjective" overall payoff of "*b(>1)*" for B and of "*a(<0)*" for A.[19]

Assuming that having apples and bananas is substantively better for *B* than having apples only while having apples or bananas is substantively better for *A* than having neither the

[18] For broader up to date discussions of preference representations, (Hausman 2012), (Gilboa 2009).
[19] For the sake of simplicity we use the same characters for substantive and „utility" payoffs here.

substantive ordering of results $\xi_i, i = A, B$ is the same as the subjective ordering $u_i, i = A, B$. In that case the status quo will persist as result of a play of the game evaluated at (0, 0). A rational actor A who would foresee B's evaluation *b>1* would never deliver his apples in the first place. Due to backward induction he would choose D_A. Still, both could be better off – substantively and subjectively – if the actor in the second mover role were able to commit to behavior that goes against the extrinsic motives exerted by the substantive payoffs and could signal this commitment. In that case actor B would prefer the path $(\mathrm{C}_A, \mathrm{C}_B)$ to $(\mathrm{C}_A, \mathrm{D}_B)$, as indicated by $u_B(\mathrm{C}_A, \mathrm{D}_B) < u_B(\mathrm{C}_A, \mathrm{C}_B)$, even though, $\xi_B(\mathrm{C}_A, \mathrm{D}_B) > \xi_B(\mathrm{C}_A, \mathrm{C}_B)$.

Once B can signal her commitment or A can detect the presence of suitable intrinsic motives, backward induction will lead to co-operative first mover choices. These will be rewarded due to the commitment of B to actions that are *directly substantively disadvantageous* yet in turn will be *indirectly advantageous in substantive terms*.

The basic comparative statics and the adaptation process in which the disposition to "do good" will be sustainable by well-making objective consequences can be illustrated in terms of Figure 2. We present and interpret Figure 2 in the next section. Afterwards an obvious way of how to transfer the basic insights that apply to all trade to the special case of the role CSR in market oriented corporate actors is illustrated.

2.2.2.3 A graphical representation of some formal modeling results

When called upon to act in the second-mover or B-role in exchange relationships transaction partners can be either trustworthy (characterized by $u_B(\mathrm{C}_A, \mathrm{D}_B) < u_B(\mathrm{C}_A, \mathrm{C}_B)$) or untrustworthy (characterized by $u_B(\mathrm{C}_A, \mathrm{D}_B) > u_B(\mathrm{C}_A, \mathrm{C}_B)$). How risky it is to trust when being assigned the first-mover, A-role, depends on the general knowledge of the share $p, 0 \leq p \leq 1$, of trustworthy types in the population.

It is assumed that the population is "large", the matching random and that p is known to all who are potentially acting in the A-role. Beyond knowledge of p behavior in the A-role also depends on the reliability and extent of information concerning the type of the randomly

assigned transaction partner in the B-role. Let this information after random matching be available to actors in the A-role at cost $C, 0 \leq C \leq 1$.

In Figure 2 one finds (p, C) combinations as well as the dynamics that unfold from each such combination (p, C) by (dashed) arrows.

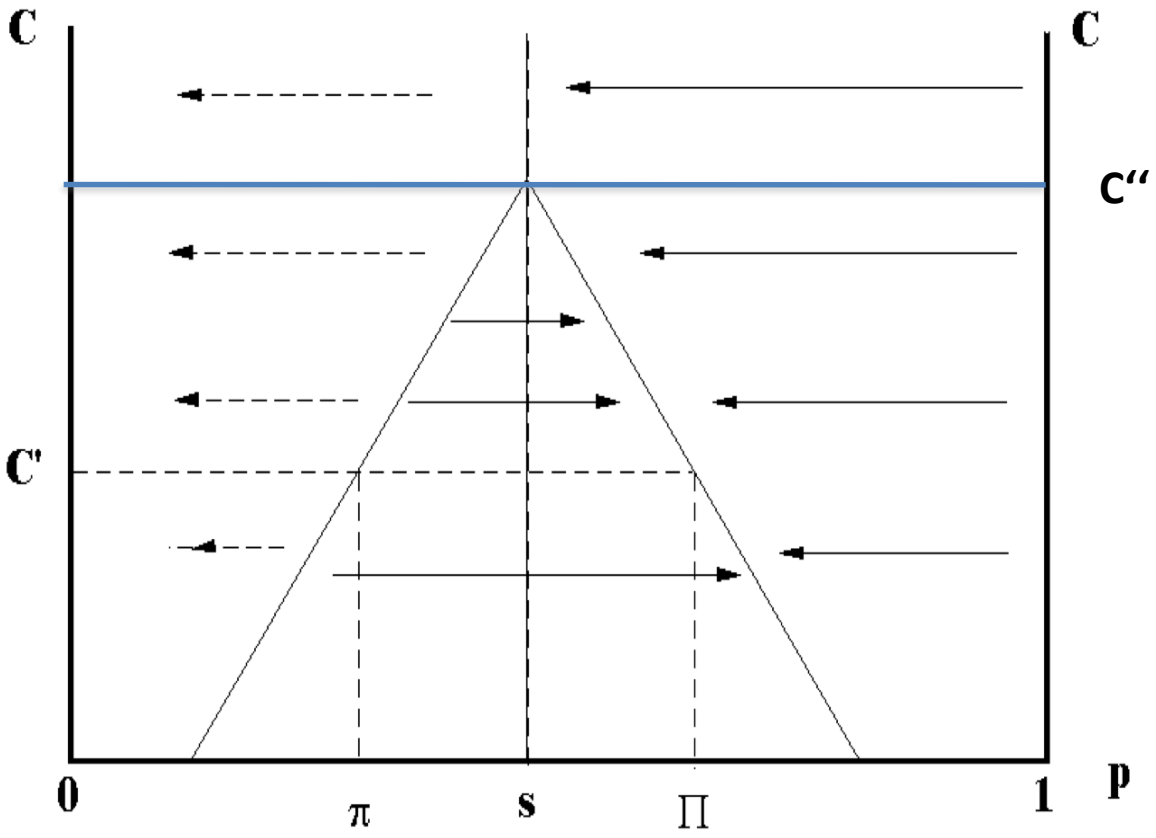

Figure 2 Evolution of population share *p* of committed actors for alternative levels of cost and given reliability of type detection

In Figure 2 values of *p* are indicated on the abscissa. On the ordinate we find alternative cost levels *C* at which a detection [signaling][20] technology of a given reliability may be available to the interacting parties. Horizontal lines in the triangle indicate for alternative values of *C* the intervals in which a detection technology of the given reliability will rationally be used; e.g. for *C*=*C'* the segment is ((π, C'), (∏, C')). The higher the costs of the detection technology

[20] It can be rather easily shown that detection and signaling technologies relevant here amount basically to the same thing.

the smaller the interval of p-values for which it actually is worthwhile to employ the costly technology of given reliability.[21]

On each round $t, t = 0, 1, ...$ of the process in which $p = p_t$ evolves, choices are made according to expected values. These are determined according to preference representations u_A, u_B and subjective probabilities of being matched with a trustworthy or non-trustworthy actor type in the B-role. Due to the dual nature of the conceptual framework underlying the indirect evolutionary approach the choices are made according to subjective expected utility u_A, u_B, yet the population share p_{t+1} is determined solely according to the substantive success ξ_A, ξ_B of these choices.

The width of the triangle base would be [0, 1] in case of the availability of a perfectly reliable type detection technology. With such a technology, in the limiting case of *C=0* obviously only and all trustworthy individuals would be trusted and would eventually crowd out the untrustworthy. If *C* increases towards *C''* the interval contains only a single point {(C'', s)} -- the top corner of the triangle -- and becomes empty beyond. The horizontal (dashed or bold) arrows indicate the evolution of the population share p at different cost levels for a given reliability of the signal of trustworthiness and non-trustworthiness, respectively.

The arrows indicate how the process will run its course whenever evolutionary success monotonically depends on direct substantive success (proximate fitness) on each round of interaction. It can be characterized for all initial (p, C) combinations where the process will ultimately be heading. To this we turn next.

The parameter *s* denotes the population share *p=s* at which, without access to type information, the expected values of trust and of no trust are equal. Knowing only the population share *p all* A-role actors will show trust if *p>s* and will show no trust if *p<s*. Without type information untrustworthy types in the second mover role will do better than trustworthy types.

[21] Figure 2 shows the relevant triangle for a technology of given reliability. This technology makes two signals, one of trustworthiness and one of non-trustworthiness available. The signals may differ in reliability yet since nothing substantial hinges on this for the present purposes we leave it out of account here.

The untrustworthy will crowd out the trustworthy if beyond knowledge of the population share *p* of trustworthy types no type information is available at sufficiently low costs: As long as *p>s*, untrustworthy B-role actors will exploit the rational bets placed by A-role actors. If *p<s*, untrustworthy B-role actors will be offered a chance to exploit first-mover trust only if A-role actors erroneously show trust. Since the chances to exploit "noise trust", *p<s*, are less frequent than under conditions in which trust is the better choice for all A-role actors, the noise-dynamic is slower – indicated by dashed arrows – than the choice-dynamic – indicated by bold arrows. Obviously if A-role actors can make use only of their knowledge of the population composition *p* then ultimately the share *p* will converge towards *p=0*.

The preceding may change if sufficiently reliable type information is provided at sufficiently low cost. For 0<C<C'' points (p, C) inside the "cost-reliability" triangle lead to a use of the detection technology by actors in A-roles. This will favor the trustworthy in the B-role until the right side of the triangle is reached at cost *C*.

For the exemplary cost level C' in Figure 2 the use of the technology will increase p within the range $\pi<p<\prod$. For any $p>\prod$ the technology will not be used at cost C' and p will decrease towards $p=\prod$ until ($\prod$, C') is reached. From this we learn that for costs C'<C'' the basin of attraction for $p=\prod$ is the interval $p\epsilon(\pi, 1]$. Not only (0, C') but also ($\prod$, C') forms an evolutionarily stable rest point if a sufficiently reliable detection technology is available at cost C'. *Depending on where the process starts the one or the other will be reached.*

In sum, from our modeling approach we know that for any *C<C''* not only *p=0* but also *p>0* may be an evolutionarily stable rest point and will in this sense be sustainable. There can be a "niche" for intrinsically motivated moral types even in large scale competitive interactions. *There is a niche for CSR if it can serve the same functions as intrinsic moral motivation.*

2.2.3 Sustainable CSR in the unitary actor framework

According to the preceding line of argument the thesis that market competition of necessity will drive out CSR investments is untenable. The indirect evolutionary model shows that it is at least conceivable that individuals who are committed internally to restrictions of their direct opportunism will not necessarily be crowded out by uncommitted competitors. Provided that individuals command the ability to detect committed actor-types with sufficient reliability at sufficiently low costs commitments to directly disadvantageous CSR choices may indirectly contribute to the ultimate success of committed types.[22]

We are in principle content to let it rest at the insight that commitments to CSR can be sustainable under certain conditions. The underlying insight carries over to all mutually advantageous contracts between personal or corporate actors.[23] Commitments to execute (incomplete) contracts not only according to the letter but in spirit – i.e. genuine trustworthiness – will be sought after in corporate actors as potential transaction partners. Yet, hunger is not bread and the fact that commitments would be desirable does not prove their feasibility. To demonstrate feasibility, we need to give up the unitary actor model.[24]

3. Corporate constitutions as "corporate conscience"

The dual function model underlying the indirect evolutionary approach remained externalist in focusing on overt behavior rather than on internal processes that generate behavior.[25] Acknowledging the role of internal processes we ask now how it can conceivably be possible: a. to internally commit the corporate actor to procedural restrictions on opportunistic choice

[22] Dawkins observation that in case of personal actors long term survival is the greatest ultimate success carries over to the case of corporate actors as well; see (Richard Dawkins 1995)

[23] Whether the argument will work beyond mutually advantageous exchange and contract is open. But even if the viability of "doing well by doing good" were restricted to exchanges and contracts that are directly advantageous to all involved it would still be a highly relevant result.

[24] For personal actors this has been discussed by philosophers like David Hume in terms of countervailing "calm" and "violent" passions (Hume 1739) and been revived for economics in (Frank 1988).

[25] In the standard approach there is no room for intrinsic motivation per se. Since only the choice behavior resulting from motivational processes matters is suffices to represent the choices made – or so the standard argument runs.

making and b. how the presence of internal commitments can become detectable with sufficient reliability at sufficiently low costs by potential interaction partners of a corporate actor.

3.1 Creating and signaling general corporate constitutional commitments

To the extent that commitments to act non-opportunistically can be substituted by a fully specified contract -- providing for external enforcement and sufficient extrinsic motivation at sufficiently low costs -- inner commitments cannot bestow a comparative advantage on actors endowed with them. Trust in an agent's (trustee's) intrinsic motivation to pursue the aims ends or values of the trustor (principal) can and will be substituted by explicit contract. But if complete contracts specifying substantive outcomes are either impossible or undesirable, potential interaction partners may place value on internal *procedures* of a corporate actor (generating the functional equivalent of a personal actor's trustworthiness). A simple example can illustrate how this may be accomplished.

3.2 Simple majority rule as commitment device

Imagine a group of five individuals operating as a corporate actor under simple majority rule in some realm of choice. One individual is of age up to 1, one up to 2, one up to 3, one up to 4, and one up to 5. Once an actor becomes a member, he will be able to go for at least five periods. At the beginning of each period the oldest actor presently on the team, upon turning 5, exits from the community and is substituted by an individual of age 0. There is no exit other than the regular one after five periods and no new member is admitted unless a regular exit has taken place.

In our stylized example of "procedural corporate governance" the age composition and the simple majority rule are fixed. The median voter is always aged less than three and has to go on for at least three periods of corporate choice making (including the one in which the decision is made). These aspects of the internal structure of the corporate actor may be fixed in externally enforceable contracts. But there is residual discretionary power in collective decision making of the corporate actor. All individuals who have a voice in the decisions retain their

powers of judgment and the freedom to vote for or against proposals, respectively, as seems fit to them. The decision procedure rather than decisions will be fixed. The procedure sees to it that when exercising their discretionary power, the median and all younger voters have always to take into account at least two future periods of interaction. Therefore, the collective decisions of the corporate actor are not suffering from the same "end-game" effects as decisions of individual actors.

Individuals who join such a corporate actor know that the group as a whole will be committed at every point in time to adopt a perspective that goes beyond the short term.[26] Personal or corporate actors who face such a corporate actor in external relations also understand that the actor is substantively unrestrained but will use its discretionary power in a way that reaches beyond the direct consequences for each single round of play taken separately. In that sense we have both, discretionary powers and a restriction on the nature of the opportunity set that will induce less short-term oriented, and in that sense, more "responsible" behavior.

Artificial and simple as the example may be, it nevertheless illustrates how procedures of collective decision-making "within" a corporate actor can serve as a commitment device. Of course, the workings of the mechanism depend on being committed to playing by the rules. To the extent that people have reasons to believe in rule of law and rely on the legal order they can, however, have good reason to rely on the external enforcement of procedural rules internal to the corporate actor. They can *rely* on the fact that the collective actor will have the internal age structure described before, decides by simple majority, and has a potentially infinite life as long as it can sustain itself. Thereby the externally monitored internal corporate governance structure can induce external interaction partners of the corporation to expect corporate social responsibility in the sense of its bounded rather than unrestricted opportunism.

[26] For the simple modeling idea due to Brennan see (Brennan and Kliemt 1994), for more general and sophisticated treatments, see eg (Kandori 1992), (Salant 1991).

Internal commitments to certain procedures of corporate governance are not only relevant for external relations of the corporate actor. They are also relevant for its relations to its own members. An extremely stylized simple example may, again, serve as an illustration of the potential advantages of such forms of *procedurally embodied* corporate social responsibility.

3.3 Unanimous co-determination of decisions

Other than shareholder, so-called stakeholder approaches argue that managers and the companies they run should be treating all potential interaction partners of the company respectfully and fairly in an ethically non-discriminatory manner.[27] Whether this can plausibly hold for external relations with suppliers and customers -- as stakeholder conceptions often insinuate -- seems doubtful. With respect to internal relations with employees or shareholders such an ideal of non-discriminatory treatment must seem more plausible and can indeed be spelled out in terms of unanimous decision-making.

The unanimity rule must in any event seem very attractive to adherents of stakeholder conceptions of corporate social responsibility since it implies that shareholder interests do not command ethical priority vis a vis stakeholder concerns. It *procedurally* expresses the stakeholder conceptions' core ethical value of equal moral concern for all stakeholder interests in concrete institutional rule terms.

Though it is natural for an ethicist to subscribe to norms of equal ethical respect this subscription rests uncomfortably with the privileges of private property and ownership in Western societies. It also disregards the fact that contracts with other individuals on behalf of the corporation in the last resort derive their authority from the legal entitlements of the residual claimants or owners. If we are not willing to brush these rather obvious ethical objections aside then following the precepts of a stakeholder approach, to the extent that it can be justified at all, must be justified indirectly: to adopt a stakeholder perspective in human resource management

[27] On Freeman's influential variant of such an approach see (Freeman 1984), (Freeman et al. 2010).

must pay off over the long haul for the corporation and its stakeholders while not violating the respect for shareholder entitlements.

Indeed, in particular, as far as relations to employees are concerned a corporate culture in which the values of fairness and equal respect for each and everybody are guiding principles of interaction may seem to have great promise. No wonder that members of senior management expresses support for such a corporate culture in their speeches. However, potential employees may demand something more reliable than speeches in which principles of corporate social responsibility in dealing with stakeholder employees are endorsed merely as lofty ideals. To win (potential) employees over, the implementation of corporate governance procedures that *embody* equal respect may be necessary.[28]

To amount to anything real "stakeholder-talk" – which is notoriously cheap – must be translated into stakeholder rules[29]. Majority rule though giving an equal vote to everyone does not express equal respect sufficiently in procedural terms. As has been observed for ages majorities may overrule minorities the same way as bosses in a hierarchy overrule their underlings. Taking stakeholder conceptions at their word only the strict unanimity rule could express guarantees of equal concern procedurally and in institutional terms.[30]

Though in some circles of contractarian political philosophers and political economists merely *conceivable* unanimity is regarded as a useful tool of conceptual analysis it is certainly not a real-world procedure.[31] Nobody would suggest in earnest that all decisions in real world affairs should be made unanimously by all affected individuals. Yet for certain fundamental

---

[28] Acquiring firm specific human capital can be conceived as analogous to showing trust in the interaction of Figure 1. Putting the employer in a role analogous to the B-role actor of Figure 1 the A-role employee has to be aware that his exit option may decrease in value and the employer's bargaining position may be strengthened (in particular the threat point of negotiations may alter in the course of time in ways unfavorable to the specializing employee).

[29] Implementing deferred bonuses would be another obvious way to credibly signal a commitment of the corporation and its management to the longer-term

[30] Though protection by substantive contractual claims and rights may be useful they can only regulate what is anticipated.

[31] Even the most ardent adherents of unanimity try to cope with this by making its operation more indirect, for the seminal analysis see (Buchanan and Tullock 1962) and the still devastating criticism of classical contractarianism in "of the original contract" (Hume 1985).

decisions veto power may be procedurally granted to a certain group of stakeholders (or to representatives of homogenous stakeholder groups). Procedurally fair bidding mechanisms form the case in point that we shall sketch next.[32]

3.4 The example of procedurally fair bidding mechanisms

It can quite easily be imagined that the course of action a corporate actor should take is determined not by voting or voice but rather sold to the highest bidder. If this were to be done without any further proviso it would seem to be the polar opposite of a stakeholder conception. It would seem as if money could buy another "thing". Wealth rather than equal concern would be decisive. On the other hand, even if we want that there should be "what money can't buy" (Sandel 2012), the typical alternative to negotiation with monetary compensation is imposition by power (hierarchically and/or democratically). As compared to force and collective coercion the option to buy and sell based on the willingness to agree seems much preferable provided a proper use of the measuring rod of money can be made.

Assume that everybody is free to bid positive sums - indicating a positive willingness to pay - or negative sums - indicating a demand for compensation when a decision of the corporate actor has to be made. Assume that positive and negative bids aggregated over all bidders must be larger than the external costs of a proposed project if it is to go through. Let a project for which the sum of bids does not exceed the sum of external costs simply be dropped as internally "unprofitable" – leading to zero transfer payments to each and every one. If the sum of the bids for a project generates a surplus over costs and compensation payments – if it is internally and externally profitable in this sense – its non-negative surplus can be split equally among all bidders. With respect to bids *substantively* equal treatment and in this sense equal respect can thereby "procedurally" be guaranteed.

If this procedure may seem somewhat peculiar – at least initially – it expresses in concrete institutional terms the basic values to which stakeholder theories normally pay only

[32] For some additional details see (Güth and Kliemt 2013).

lip service. The procedure bestows veto power on each participant. Everybody can bid as negatively as seems fit to her or him. If a proposed project, then does not go through and the status quo prevails nobody has to pay anything. If the project goes through, then compensation plus a share in the overall surplus will be paid after external costs are borne -- in case of negative bids -- or contributions will be demanded minus an equal share in the surplus – again after external costs are borne. Relative to bids or overt behavior everybody gains substantially the same monetary amount and is in this sense treated with equal respect also – on top of granting equal veto power. [33]

Those who are willing to accept norms of equal respect as propagated by stakeholder conceptions can hardly hold it against a procedure that it grants equal veto power to everybody and guarantees that everybody with respect to her or his freely submitted bids will be treated equally in terms of substantive gains. The objection that equal respect is something that should manifest itself in moral judgment rather than procedures does not seem convincing at all in a corporate governance context. If adherents of stakeholder theories are not willing to let the ideal of equal respect be expressed in terms of procedures that embody veto power this can only mean that we are not taking the ideal seriously. For, on some level they would be willing to impose on others what *they* think is right and in this sense would not be willing to respect them.

Those who believe that it is unavoidable that some impose their views on others would typically try to strengthen free contracts and the option to exit from the corporate contract nexus. Those who believe that companies might command stronger means to express equal respect for stakeholders and certain forms of corporate social responsibility should think, however, along procedural lines. Appealing to the public at large and whispering in the ears of more or less benevolent CEO's and managers is no substitute for corporate governance proposals. Of course, such proposals need not at all be of the kind used here for illustrative purposes. However, cheap

[33] The procedure is obviously not under-bidding proof. But the risk that collective projects might not be realized even though they would be to everybody's advantage may put a cap on the proclivity to underbid.

talk will not do either. The CSR agenda will have to come up with governance procedures that approximate the well-intentioned appeals in practice.

4. Final observations

Due to the elegance of the revealed preference approach economists became almost oblivious of the fact that stenographical representation of the results of decision-making will work well only for contexts in which the internal springs of action do not matter much.[34] This will be the case if observed behavior is – as in case of some markets – mostly determined by the *structural* conditions of *inter*action[35] and the substantive or material payoffs that characterize them (typically in monetary terms). In such cases of market competition actors who decide against what substantive payoffs require will be swiftly eliminated and economists can feel safe in ignoring behavior that is not substantively advantageous.

It should be noted well that the unitary actor model will suffice only under the special conditions described before. Only then can psychology and in case of corporate actors organizational theory be taken out of economics without losing much. But if we consider more complex forms of behavior that at least allow for trade-offs between the short and the long term we cannot avoid taking into account how behavior is generated. Looking at internal processes of corporate actors we can treat their internal decision procedures as equivalent to mental processes of personal actors who are cognitively and emotionally bounded (committed) to some extent.

---

[34] Some of the most interesting work in making economic models of individual behavior more adeaquate make use of such techniques. To name but two examples one might think of (Bolton and Ockenfels 2000) as well as (Fehr and Schmidt 1999) which both build distributional and other regarding concerns into „individual utility function". We have no objection against devoting attention to motivational processes inernal to the actor. We resent, however, insiting that human behavior must by all means be described as if maximizing some function. This can be justified only if one has – as in a context of evolutionary competition – good reason to assume that inconsistent choice making that cannot be so represented will be selected against strongly.

[35] It is not by chance that the models in which economists tend to take the greatest pride "predict" certain outcomes of interaction mostly independently of specific assumptions on the behavioral level; e.g. evolutionary models as again in (Schumpeter 1959), (Alchian 1950), experimental markets (Smith 1962), or chess board models like in (Schelling 1971), (Hegselmann 2012).

Both, corporate and personal actors need not be utilitarian consequentialists. They might as well start from premises of interpersonal respect of a broadly speaking Kantian type. Whether such views are ethically reasonable or not is contested in ethics. But it seems beyond reasonable debate that Corporate Social Responsibility can be a serious corporate governance strategy only if it can be institutionalized: First, it must be shown how CSR can be embodied in institutional rules of corporate governance. Second, institutionalized CSR commitments must be shown to be sustainable under market competition. The theoretical framework laid out here shows how this could be done in principle.

To the extent that we can by means of corporate governance prepare at all for future decision situations that cannot yet be anticipated it should not be in terms of prescriptions of substantive choices. [36] Neither personal nor corporate actors will want to be operating according to a fully specified explicit strategy leaving no discretionary power to future manifestations of the actor. Commitments therefore must concern the procedures rather than the substance of choices. Why CSR should be an exception to this is as hard to see as why it should be impervious to the strictures of sustainability in competition.

In view of the fact that firms may in a wide sense of the term be conceptualized as "moral tribes" (Greene 2013) identification with the tribe and the moral motivation to "pull together" may be very much enhanced by corporate governance procedures that are perceived to be procedurally fair in the preceding sense. If so, they may bestow an advantage on the firms possessing them and they may be sustainable despite certain disadvantages of following complicated bidding procedures. What will prevail, will in the end be an empirical matter.

In any event those who endorse ideals of corporate social responsibility should seek for governance structures that render what they regard as socially responsible behavior to some

[36] The enumeration of several mutually exclusive sufficient action triggers can always be made collectively exhaustive by an „else clause" and specifying the corresponding action. But no person in her right mind would believe in her ability to anticipate "everything else" sufficiently to warrant a substantively fixed response, no matter what.

extent *independent of the good intentions of individual members of that corporate actor.*[37] Popular stakeholder conceptions of corporate social responsibility typically appeal merely to good intentions. This is not good enough. To make good on their own claims they should spell out institutional ways to realize their lofty ideals. If they cannot come up with anything other than what we suggested and express in institutional terms, then this is a major deficiency of these theories. The burden of proof is on them not on us who are skeptical about the ethical (not the empirical) claims of stakeholder theories of corporate governance. If our explication of "equal respect" is correct, if the ideals should indeed translate into such institutional terms, this should be regarded perhaps as a reason for being skeptical about stakeholder conceptions of corporate social responsibility altogether. Nevertheless, as we have shown, there is in principle room for such and other forms.

References

Ainslee, George. 2002. *Break Down of the Will*. Princeton: Princeton University Press.

Alchian, Armen A. 1950. "Uncertainty, Evolution, and Economic Theory." *Journal of Political Economy* Vol. 58: 211–21.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Berninghaus, Siegfried, Werner Güth, and Hartmut Kliemt. 2003. "From Teleology to Evolution. Bridging the Gap between Rationality and Adaptation in Social Explanation." *Journal of Evolutionary Economics* 13 (4): 385–410.

Berninghaus, Siegfried, Werner Güth, and Hartmut Kliemt. 2012. "Pull, Push or Both?: Indirect Evolution in Economics and Beyond." In *Evolution and Rationality: Decisions, Co-Operation and Strategic Beahviour; (Eds) Samir Okasha and Kenneth G. Binmore*, Cambridge: Cambridge University Press, 110–32.

Bolton, Gary, and Axel Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity and Competition." *American Economic Review* 90: 166–93.

Bowles, Samuel, and Herbert Gintis. 2013. *A Cooperative Species: Human Reciprocity and Its Evolution*. Reprint. Princeton, N.J.; Oxford: Princeton Univers. Press.

Brennan, Geoffrey, and Hartmut Kliemt. 1994. "Finite Lives and Social Institutions." *Kyklos* 47 (4): 551–71.

Buchanan, James M., and Gordon Tullock. 1962. *The Calculus of Consent*. Ann Arbor: University of Michigan Press.

Carens, J. 1981. *Equality, Moral Incentives, and the Market. An Essay in Utopian Politico-*

[37] As in politics in which according to an old wisdom "everybody should be supposed to be a knave …" (Hume 1985) essay, iv, we may want to focus in economics also on structures and incentives for good behavior rather than selecting individuals who are intrinsically motivated to do good; in a general vein (Robertson 1956). Still extrinsic motivation and incentives work only to some extent.

Economic Theory. Chicago and London: The University of Chicago Press.
Cohen, G. A. 2009. *Why Not Socialism?* Princeton University Press.
Darwin, Charles R. 1975. *On the Origin of Species*. Facsimile of 1 ed. Cambridge, Mass.: Harvard University Press.
Dawkins, R. 1976. *The Selfish Gene*. Oxford: Oxford University Press.
Dawkins, Richard. 1995. *River out of Eden. A Darwinian View of Life*. Science Masters. London: Weidenfeld & Nicolson.
Fehr, Ernst, and Klaus Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114: 817–68.
Frank, Robert. 1987. "If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience?" *The American Economic Review* 77/4: 593–604.
———. 1988. *The Passions within Reason: Prisoner's Dilemmas and the Strategic Role of the Emotions*. New York: W. W. Norton.
Freeman, R. Edward. 1984. *Strategic Management: A Stakeholder Approach*. Boston: Pitman.
Freeman, R. Edward, Jeffrey S. Harrison, Andrew C. Wicks, Bidhan L. Parmar, and Simone de Colle. 2010. *Stakeholder Theory: The State of the Art*. Cambridge ; New York: Cambridge University Press.
Friedman, Milton. 1970. "The Social Responsibility of Business to Increase Its Profits." *The New York Times Magazine*, September 13.
Gilboa, Itzhak. 2009. *Theory of Decision under Uncertainty*. 1st ed. Cambridge ; New York: Cambridge University Press.
Greene, Joshua. 2013. *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. New York: Penguin Press HC, The.
Güth, Werner, and Hartmut Kliemt. 1994. "Competition or Co-Operation: On the Evolutionary Economics of Trust, Exploitation and Moral Attitudes." *Metroeconomica* 45 (2): 155–87.
———. 2000. "Evolutionarily Stable Co-Operative Commitments." *Theory and Decision* 49: 197–221.
Güth, W., Kliemt, H. 2013. Fairness that Money Can Buy: Procedural Egalitarianism in Practice, Rationality, Markets and Morals Vol. 4, S. 20-30.
Hausman, Daniel M. 2012. *Preference, Value, Choice, and Welfare*. Cambridge ; New York: Cambridge University Press.
Hayek, Friedrich August von. 1945. "The Use of Knowledge in Society." *American Economic Review* XXXV (4): 519–30.
Hazlitt, Henry. 1988. *Economics in One Lesson: The Shortest and Surest Way to Understand Basic Economics*. New York: Crown Business.
Hegselmann, Rainer. 2012. "Thomas C. Schelling and the Computer: Some Notes on Schelling's Essay 'On Letting a Computer Help with the Work.'" *Journal of Artificial Societies and Social Simulation* 15 (4): 9.
Hobbes, Thomas. 1651. *Leviathan*. Harmondsworth: Penguin.
Hofstadter, Richard. 1969. *Social Darwinism and American Thought*. New York.
Hume, David. 1739. *A Treatise of Human Nature*. Oxford: Clarendon.
———. 1985. *Essays. Moral, Political and Literary*. Indianapolis: Liberty Fund.
Jasay, Anthony de. 1985. *The State*. Indianapolis: Liberty Fund.
Kandori, Michihiro. 1992. "Repeated Games Played by Overlapping Generations of Players." *The Review of Economic Studies* 59 (1): 81–92.
Karnani, Aneel G. 2010. "Doing Well by Doing Good: The Grand Illusion." *California Management Review*, 53(2), DOI: 10.1525/cmr.2011.53.2.69
Kay, John. 2011. *Obliquity. Why Our Goals Are Best Achieved Indirectly*. London: Profile Books.
Kliemt, H. 1986. "Individualism, Libertarianism and Non-Cognitivism." *Analyse und Kritik* 8,

2: 211–28.
Kropotkin, Peter. 1919. *Mutual Aid: A Factor of Cooperation*. New York: A.A. Knopf.
Mandeville, Bernard. 1924. *The Fable of the Bees: Or, Private Vices, Publick Benefits: The Fable of the Bees. Pt. II. Appendixes: Mandeville's Family (with Genealogical Table) Description of the Editions. Criticisms of the Fable. A List, Chronologically Arranged, of References to Mandeville's Works. Index to Commentary*. Vol. 2. The Clarendon press.
Mayer, Colin. 2013. *Firm Commitment: Why the Corporation Is Failing Us and How to Restore Trust in It*. Oxford University Press.
Maynard-Smith, John. 1982. *Evolutionary Game Theory*. Cambridge.
Raphael, David-Daiches. 1969. *British Moralists*. Oxford: Oxford University Press.
Robbins, Lionel. 1935. *An Essay on the Nature and Significance of Economic Science.* London: Macmillan.
Robertson, D. H. 1956. *Economic Commentaries*. London: Staples Press.
Salant, David J. 1991. "A Repeated Game with Finitely Lived Overlapping Generations of Players." *Games and Economic Behavior* 3 (2): 244–59. doi:10.1016/0899-8256(91)90025-A.
Sandel, Michael. 2012. *What Money Can't Buy: The Moral Limits of Markets*. Open Market edition. Allen Lane.
Schelling, T.C. 1984. *Choice and Consequence*. Cambridge, MA.: Harvard University Press.
Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *The Journal of Mathematical Sociology* 1 (2): 143–86. doi:10.1080/0022250X.1971.9989794.
Schneider, Louis. 1967. *The Scottish Moralists on Human Nature and Society.* Chicago und London.
Schumpeter, Joseph Alois. 1959. *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.
Schüssler, R. 1990. *Kooperation unter Egoisten.* München: Oldenbourg.
Smith, Vernon L. 1962. "An Experimental Study of Competitive Market Behavior." *Journal of Political Economy* 70:2 (April): 111–37.
Sober, Elliott, and David Sloan Wilson. 1998. *Unto Others. The Evolution and Psychology of Unselfish Behavior*. Cambridge and London: Harvard University Press.
Thaler, R. H., and H. M. Shefrin. 1981. "An Economic Theory of Self-Control." *Journal of Political Economy* 89/2: 392 ff.
Vanberg, Viktor J., and Roger Congleton. 1992. "Rationality, Morality and Exit." *American Political Science Review* 86 (2): 418 ff.
Zahavi, Amotz, and Avishag Zahavi. 1997. *The Handicap Principle. A Missing Piece of Darwin's Puzzle*. New York, Oxford: Oxford University Press.