Beyond The Grip of the Sovereign

BEYOND THE GRIP OF THE SOVEREIGN: DELIBERATION, 'WE-REASONING' AND THE LEGITIMACY OF NORMS AND STANDARDS IN A GLOBALISED WORLD

Juljan Krause (London School of Economics and Political Science) & Markus Scholz (University of Applied Science Vienna)[1]

ABSTRACT

Nation states are increasingly losing the power to regulate markets and corporations to the effect that many regulatory guidelines and principles are now either being worked out by corporations themselves (so-called business-driven programmes) or agreed upon in more inclusive processes in which social and environmental interest groups are involved (so-called multi-stakeholder initiatives). This paper aims to analyse and model the complex ways in which these new norms and standards emerge, in particular when the partners at the negotiating table have conflicting interests. Unlike recent contributions that do not consider power imbalances we argue that diverging interests are a key characteristic of negotiations. We present a team game-theoretical framework in which players can switch between two kinds of reasoning: an individual mode in which stakeholders aim for the best possible outcome for themselves and a 'we-mode' in which they are genuinely concerned with finding a standard that is optimal for the whole group. We show that a higher inclination towards 'we-mode' reasoning is beneficial overall and maximises individual pay-offs as well as the outcome for the entire group. We argue that cooperation is therefore in the rational self-interest of the stakeholders, which backs up the argument in favour of honest deliberation at the negotiating table with references to rationality. Based on the results of our model we conclude that in a world where national regulatory frameworks are losing their grip, only norms and standards that have been worked out by a sufficiently large number of 'we-reasoning' stakeholders can be called legitimate.

Keywords: Global business standards and norms, emergence of standards and norms, deliberative political theory, legitimacy, multi-stakeholder initiatives, collective intentionality, team game theory.

Introduction

In a globalised economy, rules, norms, standards and guidelines increasingly emerge outside the traditional democratic framework: rather than deliberated and passed in parliament, new forms of business standards are the result of inter-corporational efforts, which prompts urgent questions about their legitimacy. These new ways of establishing norms are often referred to in terms of multi-stakeholder initiatives (MSI), clubs, private authority regimes, private law or private regulation (see Brammer, Jackson and Matten 2012; Cashore, Auld and Newsom 2004; Coglianese and Nash 2001; Dingwerth 2007; Fransen 2012; Hahn 2011; Mattli and Büthe 2003; Potoski and Prakash 2009; Prakash and Potoski 2006; 2006b; 2007; Scherer, Palazzo and Baumann 2006; Windsor 2004). A key requirement for assessing the legitimacy of MSI guidelines and others is a thorough understanding of the processual character by which these norms emerge in the first place. The analysis of these processes is the focus of this paper. In particular the question how norms arise out of negotiations among a limited number of stakeholders is unclear. We draw on recent advances in social philosophy and argue that 'ethics guidelines' that emerge outside sovereign control are better understood if read against recent literature on 'collective intentionality' (see Gilbert 2006, Searle 1995, Tuomela 2007) and 'team game theory' (see Bacharach 1999; 2006 and Sugden 2000; 2008; cf. Hakli et al. 2010) .

[1] The authors appear in alphabetical order.

Building on Hakli et al. (2010), we present a team-game theoretical model in which actors switch between an 'I-mode' that is only concerned with what is best for the corporation, NGO or stakeholder they represent, and a 'we-mode' that is genuinely concerned with what is best for the MSI as a whole. There is a trade-off between maximising one's own expected pay-off, i.e. seeing a guideline agreed upon that is only optimal for oneself, or maximising the expected outcome of the group by agreeing on a standard that considers all positions at the table fairly. We show that actors who are not concerned with the pay-off of the whole group are indeed better off if they follow egoistic strategies. However, actors who are willing to reason along the lines of what is best for the group should cooperate and deliberate so that they eventually pass a guideline which has been subject to honest deliberation. We show that it is not sufficient for a new norm or standard to simply be passed in order for it to be called legitimate – legitimacy requires more than the simple fact that some kind of agreement has been reached. Only if the inclination towards 'we-reasoning' is sufficiently strong, the newly found standard or norm is a truly legitimate one. Hence, legitimacy turns into an endogenous factor in norm-setting processes which is intrinsic to the dynamics of the MSI in question. Only if important conditions of processual fairness and minimal coercive power are met, a standard that emerges outside traditional democratic frameworks can be labelled legitimate (Hahn 2011; Fransen 2012; Quack 2010).

This paper proceeds as follows. The next section discusses in some detail new forms of MSI and the ways in which the emergence of new standards and norms can be approached by deliberative political theory. It reviews some of the important work that has recently been done in this field and argues that the issue of legitimacy has not yet received the attention it deserves. The third section introduces a game-theoretical model that aims to capture the key issues in MSI negotiations. In our model, actors can switch between two kinds of reasoning, and the degree to which they are inclined to genuinely reason from the point of view of the group will have an impact on the kinds of agreements that can be reached. One of the insights the model generates is that it is in the rational self-interest of 'we-reasoning' stakeholders to cooperate and deliberate fairly, hence deliberation is not simply a moral issue.

The legitimacy of multi-stakeholder initiatives

In a globalised world, national governments are increasingly losing the power to regulate business practices, in particular those of trans-national corporations (TNC), e.g. in terms of labour standards that prevent human right abuses or environmental standards. The prospect of a democratically constituted global institution that would regulate TNC, i.e. a world government, is quite unlikely. This phenomenon is widely known as the 'governance gap' (Dingwerth 2007; Djelic & Quack 2010). It is, however, widely acknowledged that economic activity requires rules (e.g. property rights, contractual rights and obligations), and the necessary enforcement mechanisms in order for stable markets to be possible at all (see Scherer, Palazzo and Baumann 2006: 505). Importantly, such 'rules of the game', as Milton Friedman (1962; 1970) would have put it, are beneficial for corporations in the sense that they, for instance, render legal decisions less volatile, balance different levels of bargaining power and generally reduce transaction costs. In addition, TNC face increasing pressure from non-governmental organizations (NGO) to follow certain business practices. In an attempt to fill the governance gap, many regulatory guidelines, as well as respective enforcement mechanisms, are now being worked out and implemented by corporations themselves (Windsor 2004).

The literature distinguishes two kinds of approaches in this respect. Attempts to take over political responsibilities of public administration (namely regulatory issues), with no or only minor participation of other stakeholders such as NGOs or representatives from national governments are usually referred to as 'business driven programmes' (BDPs).[2] Since BDPs often systematically exclude a wider audience of stakeholders from developing new regulatory frameworks the legitimacy of these programmes is often questioned (Fransen 2012).[3] In this paper however, we focus on so-called multi-stakeholder initiatives. MSI bring together non-corporate stakeholders and businesses in order to discuss and eventually

[2] For an overview of BDP and MSI see Fransen, 2012: 167, Table 1.

[3] There are many reasons why corporations would participate in such programmes, ranging from self-interests, e.g. to reduce transaction costs, (thereby reducing legal and moral uncertainties), avoid calls for tougher regulation on a national level or pressure from NGOs (in the worst case consumer boycotts), or a new understanding of extended global corporate citizenship with its respective rights and responsibilities (corporate citoyens). For an overview see Matten & Crane 2005; Scherer et al 2006.

develop adequate regulatory frameworks in order to solve pressing social and environmental issues. Although MSI have received some scholarly attention - primarily from fields such as business ethics, political theory and international relations - the question of the legitimacy of these new regulatory frameworks remains an urgent one. These new regulatory frameworks have a great impact on both public perception of business behaviour as well as decision-making processes within corporations (Hahn 2011: 123-124; cf. Potoski and Prakash 2009; Prakash & Potoski 2006b). This is why a thorough analysis of the complex ways by which these guidelines emerge is needed.

An increasing number of authors suggest drawing on international political theory for investigating the legitimacy of these new norms and standards (see Djelic & Quack 2010; Dingwerth 2007; Hahn 2011; Klinke 2009; Quack 2010). In particular deliberative political theory (also known as discursive policy) informs much of the debate. Deliberative political theory considers authentic, sincere and honest deliberation as the relevant source of legitimacy. In its classic approach, the ideal form of deliberation is concerned with pursuing one shared token goal. 'Democratic processes are oriented around discussing [the] common good rather than competing for the promotion of the private good of each. Instead of reasoning from the point of view of the private utility maximizer, through public deliberation, citizens transform their preferences according to public-minded ends, and reason together about the nature of those ends and the best means to realize them' (Young 1996: 120-121; see Mansbridge et al. 2010: 66-67). The legitimacy of a norm or standard can be assessed by comparing the dynamics of its emergence to this idealised scenario. However, as we argue below, legitimacy does not exclude some degree of self-interest among the stakeholders involved.

Deliberative political theory is already being applied to the analysis of the legitimacy of forest certification schemes (Overdevest 2010) or the regulation of financial markets (Morgan 2010). Usually, the analysis of the processes by which new laws, rules and norms come about is broken down into three levels: the input, throughput and output dimensions (see Hahn 2011: 126-128; cf. Risse & Kleine 2007; Klinke 2009). The input dimension provides information about the participants in decision-making processes and whether these processes are inclusive, i.e. whether or not relevant stakeholders have been considered. The throughput dimension, which is the focus of this article, looks at the characteristic features of the decision-making process itself, namely the distinctive character of discursive processes (modes of communication, bargaining power, preference aggregation, transparency, underlying motives of the actors involved etc.). Finally, the output dimension investigates if the newly found laws, rules and norms provide efficient solutions to the problems at hand (Hahn 2011: 126-128).

Some TNC voluntarily participate in MSI. These programmes primarily aim to prevent negative externalities but create positive ones (Potoski and Prakash 2009: 2). Positive and negative externalities result from market failures, as market mechanisms alone do not promote the socially optimal level of the production of goods (Pigou [1920] 1960). Generally, market failures are due to asymmetric information among the actors involved which entails uneven levels of bargaining power. For instance, it might be the case that producers continue to use environmentally damaging input factors (e.g. palm oil), because they are unaware that their customers would be willing to pay more for less damaging input factors. Conversely, even if TNC ensure their customers that they use environmentally-friendly ingredients exclusively, these self-declarations may not be convincing if consumers are unable to determine whether or not the TNC's claims can be trusted. As Potoski and Prakash argue, '[i]nformation asymmetries coupled with the lack of assurance mechanisms prevent actors from transacting business. In such situations the market has "failed", and the gains from win-win exchanges have not been realized' (Potoski and Prakash: 2009: 4).

Typically, the solution for these kinds of market failures is for governments to intervene in market transactions (Pigou [1920] 1960; cf. Coase 1960). National governments could, for instance, put in place laws that prevent negative externalities (e.g. to protect the environment) or introduce labels and certificates that inform consumers about the origins of ingredients (e.g. non-genetically modified organisms). However, in a globalised world, the capacity of national governments to implement norms and standards (which could prevent the creation of negative externalities) or certify certain practices (that could inform customers about a producer's trustworthiness) is steadily decreasing. Hence, if society is burdened with negative externalities that result from corporative actions, and when corporations face

strong pressure from stakeholders (e.g. in the form of NGO protests) and when national governments are not powerful enough to provide adequate solutions to these problems, then a new form of governance is needed. This is why voluntary programmes such as MSI can be useful. As Potoski and Prakash explain: '[voluntary programmes] seek to reconfigure the institutional space between the potential exchange participants by providing new rules and new mechanism to facilitate exchange. In effect, they create a new market for corporate reputation—one could perhaps call it a "market for virtue"' (Potoski and Prakash 2009: 4).[4] Recent contributions in business ethics confirm this thesis: corporations seek for a license to operate, i.e. they seek legitimacy (Conzelmann 2012: 194; Palazzo and Scherer 2006; Scherer, Palazzo and Baumann 2006; Suchman 1995, Vogel 2005).

TNC seek a kind of moral legitimacy which 'refers to conscious moral judgments on the organization's output, procedures, structures and leaders. Moral legitimacy is socially constructed by giving and considering reasons to justify certain actions, practices, or institutions' (Palazzo and Scherer 2006: 73). In order to gain such moral grounding, the authors argue that corporations should engage in processes of public communication, i.e. in deliberative and inclusive processes that are conducive to building up a level of legitimacy which turns them into trustworthy actors. 'And the more active citizens in their different stakeholder roles become, the greater the need to deal with their demands in a discursive way. Stating this we are able to conceptualize the process by which corporate moral legitimacy must be reproduced: by placing corporations into public communication network [sic]'(81).

Hence a thoroughgoing analysis of the very processes by which new forms of legitimacy emerge in deliberative processes is a pressing issue: legitimacy has turned into an endogenous factor in corporate strategy. Rather than a desirable attribute that is awarded by an external credible source – e.g. a sovereign national government – the problem of legitimacy is now a constructive one in that it is generated from within. The issue of legitimacy no longer pertains to the output dimension of standard-setting procedures but is inseparable from input and throughput dimensions. As Fransen argues: 'Inclusiveness, expertise-based effectiveness and procedural fairness thus contribute in different ways to the notion that the institutional design of multi-stakeholder-governed voluntary programmes, in normative-theoretical terms, can be considered more legitimate than the institutional design of business-driven programmes that exclude societal stakeholders; or, in empirical social science terms, is likely to be granted more legitimacy by relevant actors' (2012: 169). Fransen concludes that 'the evidence presented suggests that there is indeed pressure in voluntary programme policy fields to organize governance with inclusion of various stakeholders in order to be seen as legitimate by an outside audience of governmental, intergovernmental and non-governmental organizations' (2012: 165). Moreover, there is an explicit 'normative pressure on voluntary programmes to adhere to what is understood *as appropriate*' (2012: 170, emphasis added). It is the emergence of this kind of endogenous normative pressure the following section focuses on.

What is to be understood as 'appropriate' may have been subject to deliberation – we follow Mansbridge et al. and define deliberation broadly as 'communication that induces reflection on preferences, values and interests in a non-coercive fashion' (2010: 65) – but often coercion will play a role. It is the sometimes subtle, sometimes obvious forms of guiding the formation of attitudes of others in the throughput dimension which requires more attention. The fact that a new standard or norm is agreed upon by an inclusive number of stakeholders does not in itself render it legitimate. Legitimacy requires coercion in the throughput dimension to be kept at a minimum. Only if there is an actual 'we-reasoning' among a sufficiently large number of stakeholders that takes into account the perspective of the entire group, the norm can be a legitimate one.[5] A dominant stakeholder (e.g. a powerful corporation or popular NGO) who gets others to sign up to her preferences may push through a norm that has been agreed to by all, but one would hesitate to call it legitimate. It is reasonable to assume that not all stakeholders enter negotiations from a neutral perspective. It is this interplay between vested interests and concerns for the entire group in MSI which is modelled in the following section.

'We-reasoning' in deliberative processes: a team game-theoretical approach

[4] See Vogel (2005) for the „market for virtue".

[5] A good example is ISO 26000 which brought a large number of stakeholders together (see Hahn 2011 for an analysis).

Business guidelines regulate business activity; they are a special case of what Searle (1995) calls 'regulative rules'. In our context, the norm or standard counts as a definite rule (or soft law as it is sometimes called) to follow in a specific business in certain countries and/or industries, e.g. the Forest Stewardship Council's deliberations aim to regulate the issue of deforestation in a global context; bringing together, among others, Kingfisher, Faber-Castell, the WWF and the ICCO (see Fransen 2012 for an overview of recent MSI initiatives).

It is a necessary but not sufficient condition for a new standard to pass that all stakeholders intend the standard to pass. This is by no means a trivial issue, as recent advances in social philosophy on this matter attest to. Searle's (1995) account has proved highly influential among social philosophers who aim to analyse the ways in which collectively shared intentions serve as preconditions for collective action. 'With consciousness comes intentionality, the capacity of the mind to represent objects and states of affairs in the world other than itself' (7). If intentionality is so intimately tied to consciousness, it is unlikely that there can be any bargaining situation in which intentions play no role. This is to say that intentions and intentionality will always be a factor in dynamic negotiations, whether or not they are subject to debate or self-reflection. It is extremely implausible to conceptualise negotiations as entirely free of individual, and perhaps egoistic intentions on the part of the actors involved.

MSI are no exception. Actors do not enter the arena of debate without having formed intentions on some baseline level – they do not sit down at the table without having formulated a notion of the outcome they would welcome given their preference sets. In short, there is no 'innocent' deliberative bargaining situation. As far as the regulation of business activity is concerned, the move towards increasingly endogenous forms of regulation complicates the relationship between deliberation, collective action and rule-making. In the following, we assume that status asymmetries and unbalanced levels of bargaining power are characteristic of negotiations. Actors enter negotiations with certain strategies in mind and in the participation of specific outcomes. They have preferences about the kinds of norms they would like to see implemented when they represent specific industries or corporations. In this sense, the arena of deliberation is no neutral space in which power is distributed evenly.

As Krause (2011) argues, even in non-deliberative collective action problems in which agents do not come together to deliberate at length about a problem at hand, power and status issues do play a role: agents can be brought to give up on their own position and join in on collectively intending to pursue a certain action which is actually against their interests, all things considered. The coercion towards agreeing to a wage or pension cut if other employees have agreed to do the same is a prominent example. The fact that actors are empathic agents who are forward-looking means that some agents adopt the views of a dominant actor in the group even *prior to* deliberation: expecting their position to be untenable given the distribution of bargaining power, they conclude that in order to avoid punishment in the future, it is best to agree to contribute to an action which is actually not optimal for them. The puzzling insight is that in some situations agents actually act against their interests but do so voluntarily. It is precisely due to the anticipation of future losses that attitudes are being aligned in subtle processes in which lengthy deliberation and honest communication do not matter.

Bargaining power and status asymmetries matter even outside a deliberative discourse, e.g. when a well-experienced representative of a powerful corporation meets her inexperienced counterpart from a newly founded NGO. It is therefore highly implausible to assume that these asymmetries would not play a role in deliberations about collective action problems where the whole point of deliberating is to discuss the best norm or standard and try and convince other actors of one's point of view. Of course, this is not to say that actors were unwilling to compromise or let others convince them of a better norm. But it does mean that actors try and get their views across.

Following Bacharach (1999) and Hakli et al. (2010), we illustrate the key issues in strategic deliberation on a norm in the following team game-theoretic framework.[6] Actors have a choice between strictly

[6] We apply Hakli et al.'s (2010) insights about different kinds of modes of reasoning in collective action problems to the issue of formulating a norm. The following is an illustration of Bacharach's (1999) 'unreliable team interaction' (UTI). See Bacharach (1999) for a generalised approach to team game theory and the formal requirements for UTI equilibria to exist. We are grateful to Raul Hakli for helping with the calculation of the matrices.

following their own interests and only aim for what is best for the organisation they represent regardless of the outcome for others, or they can adopt a genuine concern with the outcome for the whole group. If they adopt the view of the group, the aim is to secure an outcome that is best for the group as a whole but perhaps not optimal for them as individual players. Essentially, players can switch between an 'I-mode' that aims to maximise individual pay-offs and a 'we-mode' that aims to work out what is best for the entire group. Hence the trade-off is between optimising a pay-off for oneself (as a representative of a specific cooperation, NGO or other stakeholder) and contributing to an optimal group pay-off that is the best plan for the entire MSI. If actors adopt the view of the group, they adopt its preference structure expressed as a share of it. This ensures that although groups have a preference as a group, a group only 'lives' through and is defined by its members. Hence there is no supra-human group agent or otherwise implausible entity.[7]

At the negotiating table, actors pursue either of two strategies. They can deliberate (D). This state is characterised by an honest and sincere approach to deliberation; one is willing to actively participate, listen to the points other players make, and compromise, if necessary. Alternatively, they can follow an entirely egoistic strategy (E) that is only concerned with pushing through one's own agenda. This is a strategy of skilful agents who may come across as actively and openly participating, but manage to get their way at the expense of other actors.

For simplicity, assume that there are only two agents who negotiate on a norm on behalf of an NGO (player 1) and a large cooperation (player 2). For the moment we ignore the possibility of taking a group's perspective, so players can only act in the I-mode. Both players can choose to either deliberate in good faith (D) or play egoistically (E). If they both deliberate, they find a genuine compromise which yields a result (pay-off) which is lower than what they could have achieved, had they pushed through their own agenda at the expense of the other player. If one player manages to skilfully persuade others to follow their lead, the pay-off of the dominant player is higher than what she could ever have achieved under a compromise as she did not have to give up on anything. It is zero for the other player who negotiated in good faith and did not realise that the dominant player actually followed her own agenda. In this case, a norm will be issued that only represents the preferences of the dominant player.

Players can avoid falling into this trap by insisting on their own individual, egoistic strategies. If both players behave egoistically, no genuine deliberation is taking place. In order to save their faces, the players quickly agree on a watered-down agreement which does not really do much justice to anybody but is of no harm either. This outcome still secures a positive pay-off for everybody as it is better than to leave with nothing at all, and actors can still defend their actions to the stakeholders they represent as having managed to agree on something. But the pay-off in this case is of course lower than what they could have achieved had they found a compromise that considered all positions in a balanced way.

This gives the following matrix which is functionally and logically equivalent to the well-studied Prisoners' Dilemma:

	D	E
D	3,3	0,4
E	4,0	2,2

[7] For an overview of recent debates on whether or not collective intentions constitute a collective subject see Krause (2011).

The rows of the table represent player 1's options, the columns represent player 2. If both players agree to deliberate, they take each other's views into account, compromise and reach an agreement. Assuming that they meet half way, they both secure a pay-off of 3 (DD in the upper left corner of the matrix). However, if player 1 plays egoistically and pushes her ideas through while player 2 believes that actual deliberation is taking place (the lower left cell of the table), player 1 secures the highest pay-off of the game as she manages to get player 2 to agree to a norm or guideline which actually only serves 1's interests (the pay-off for player 2 is therefore 0, as argued above). The game is symmetric, so the same rationale holds for the situation in which player 1 is willing to deliberate but player 2 behaves egoistically.

Such an outcome is typically the result of significantly uneven levels of bargaining power. Each player's bargaining power depends on a large number of factors. 'Every feature that could affect the cost of leaving or not entering the negotiation affects bargaining power, including one's best alternative to a negotiated agreement and the effects of the status quo of one's position. Possessing information and the resources to gather information, having ample time to wait, and many other factors all affect bargaining power' (Mansbridge et al. 2010: 91). Bargaining power can also entail that both players insist on their egoistic strategies: when they play egoistically and no one is willing to compromise, both players agree to pass a poor norm and leave with a pay-off of 2 (the lower right cell of the table).

The combination (2,2) is the only equilibrium of the game. If the players act rationally so as to maximise their pay-offs (and we assume they do), then they will both decide against sincere deliberation and pursue an egoistic strategy. Consider the decision problem player 1 faces: if player 2 decides to deliberate (we are in the left column of the table) then it is optimal for player 1 to be egoistic (a pay-off of 4 is better than a pay-off of 3). But if player 2 decided to negotiate egoistically (now we have moved to the right column of the table), player 1 is also better off if she is egoistic (2 is better than 0). Hence regardless of player 2's choice, player 1 will choose E. Since the game is entirely symmetric, the same holds for player 2 who will also choose E irrespective of player 1's choice. This makes (E,E) the only equilibrium of the game. Both secure a pay-off of 2. The perplexing result is that both players would have been better off had they chosen (D,D). (D,D) is strictly better than (E,E) but not available to them – only a pseudo-norm passes which has not been subject to sincere deliberation.

Now consider the case in which both players can switch between an I-mode and a we-mode. In the I-mode, they reason according to what is best for them individually. In the we-mode they reason according to what is best for the group as a whole. This introduces a group preference to the game matrix (here to be taken the average of both players' individual pay-offs):

	D	E
D	3,3,3	0,4,2
E	4,0,2	2,2,2

Again, the first figure in each cell denotes the pay-off of player 1, the second figure that of player 2, and the third figure represents the group preference as the average share each player holds if the group is considered as a whole. So in the case of (D,D), the group pay-off is necessarily the same as in the individual cases. In both (D,E) and (E,D), one player takes it all (4) and the other player leaves with nothing. From a group's perspective, both players still hold an average share of 2 over the total pay-off for the group. It is obvious that (E,E) is still the best strategy if all players are in the I-mode (the game collapses to the standard form of the Prisoners' Dilemma). When the entire group is in the we-mode however and concerned with optimising the group pay-off, (DD), deliberation, is the best response: only (DD) secures a share of the group's pay-off of 3, all other forms of interaction (one actor deliberates and the other one behaves egoistically, or both insist on egoistic strategies) secure only a share of 2.

But if it is not known whether or not players operate in the we-mode, four different scenarios are possible: (1) Both players are in the we-mode and aim to maximise the group pay-off. (2) Player 1 is in the we-mode but player 2 is in the I-mode. (3) Player 1 is in the I-mode and player 2 is in the we-mode. (4) Both players are in I-mode. Let δ denote the inclination of each player to deliberate, expressed in terms of probability, i.e. $0 \leq \delta \leq 1$ which means that if $\delta=1$, both players deliberate, if $\delta=0$, both pursue egoistic strategies and play in the I-mode. If, for instance, $\delta=0.5$ then there is a 50 per cent chance that players will deliberate.

Given this, the expected utilities of each player and that of the group can be calculated as the sum of pay-offs across the matrix multiplied by the likelihood of their occurrence. The expected pay-off $E(P)$ of the whole group, $E(P_{1,2})$, for (DD) under the assumption that (E,E) is still the best response for players in the I-mode can therefore be calculated as follows:

(1):

$$E(P_{1,2)}(DD|E,E) = \delta^2 P_{1,2}(D,D) + \delta(1-\delta)P_{1,2}(D,E) + \delta(1-\delta)P_{1,2}(E,D) + (1-\delta)^2 P_{1,2}(E,E)$$

The first term captures the expected outcome for deliberating with a like-minded player, the second term denotes the expected pay-off of player 1 deliberating but player 2 is egoistic, the third term captures the chance of player 1 being egoistic but player 2 deliberates and finally, the fourth term captures the outcome of both players pursuing their egoistic strategies.

If the entire group is in the we-mode, four different cases emerge as far as possible and potential group strategies are concerned: (1) Both players choose to deliberate (this is the best response if $\delta=1$ as argued above). (2) Player 1 deliberates, player 2 is egoistic (DE). (3) Player 1 is egoistic, player 2 deliberates (ED). (4) Both players are egoistic (EE). In the following, we give the expected pay-offs for individual players for case (1); the calculations of the remaining pay-offs are easily done by manipulating the terms according to which scenario the group is in. The expected pay-off of player 1, $E(P_1)$, can therefore be calculated as follows. Under deliberation (DD), we obtain:

(2):

$$E(P_1)(D|E,DD) = \delta^2 P_1(D,D) + \delta(1-\delta)P_1(D,E) + \delta(1-\delta)P_1(D,D) + (1-\delta)^2 P_1(D,E)$$

In (2), player 1 deliberates regardless of what the second player does. When she is in the we-mode (the first two terms) she pursues D anyway. Similarly, for egoism follows:

(3):

$$E(P_1)(E|E,DD) = \delta^2 P_1(D,D) + \delta(1-\delta)P_1(D,E) + \delta(1-\delta)P_1(E,D) + (1-\delta)^2 P_1(E,E)$$

In this case, player 1 chooses E regardless of player 2's actions but plays D if in the we-mode.

The expected pay-off of player 1 for playing E if player 2 chooses D (and DD is the best response of the group) is:

(4):

$$E(P_1)(E|D,DD) = \delta^2 P_1(D,D) + \delta(1-\delta)P_1(D,D) + \delta(1-\delta)P_1(E,D) + (1-\delta)^2 P_1(E,D)$$

And finally, if player 1 pursues D given that player 2 does the same and the group plays DD:

(5):

$$E(P_1)(D|D,DD) = \delta^2 P_1(D,D) + \delta(1-\delta)P_1(D,D) + \delta(1-\delta)P_1(D,D) + (1-\delta)^2 P_1(D,D)$$

The following tables give calculations for δ=0.8, i.e. there is an 80 per cent chance that actors operate in the we-mode.

DD	D	E
D	3, 3, 3	2.4, 3.2, 2.8
E	3.2, 2.4, 2.8	**2.64, 2.64, 2.64**

DE	D	E
D	0.6, 3.8, 2.2	0, 4, 2
E	0.96, 3.36, 2.16	0.4, 3.6, **2**

ED	D	E
D	3.8, 0.6, 2.2	3.36, 0.96, 2.16
E	4, 0, 2	3.6, 0.4, **2**

EE	D	E
D	2.04, 2.04, 2.04	1.6, 2.4, 2
E	2.4, 1.6, 2	2, 2, **2**

The upper left cell in the corner denotes the strategy combinations the actors choose when they find themselves in the we-mode: DD means that both players deliberate, DE means that player 1 deliberates but player 2 plays egoistically etc. To find the equilibrium of the game, consider the outcome of (D,D,DD), the upper left cell of the first matrix. In this case, it would still be beneficial for player 1 to switch tactics and play E, as it yields a higher pay-off: 3.2>3. But the incentive to do so is now much lower. The same is true for player 2, which again leaves the lower right corner as the only possible equilibrium for individual strategies. But is it an equilibrium as far as the group strategy is concerned? Indeed it is. The group pay-off under DD (2.64), which is higher than any other group pay-off under different strategies (2): for instance, if actors are in the we-mode and follow a strategy (DE), meaning that player 1 deliberates and player 2 is egoistic, the group pay-off could only ever be 2. The same holds for the two remaining potential group strategies (ED) and (EE). Hence the strategy (DD) – 'if you are in the we-mode, then deliberate' – is the only equilibrium of the game even when δ is uncertain.

The central insight of this game which produces the strategy (E,E,DD) as the only equilibrium is the following: 'if you operate in the I-mode, be egoistic. But if you are in the we-mode, then really deliberate

with others'. The important point here is that this is not a *moral* message. To argue in favour of deliberation, and hence cooperation, is not an appeal to an ideal, perhaps utopian, state of deliberation actors should aspire to for the sake of a greater social good, e.g. human rights, social cohesion, or sustainable economic and environmental development. Such appeals may or may not fall on deaf ears. Rather, the model above illustrates that a call for honest and sincere deliberation can be made *solely by appealing to actors' rational self-interests*. If players aim to maximise their pay-off, it is in their own best interest to deliberate if they are in the we-mode. Importantly, the more actors are inclined to reason in the we-mode, the less beneficial individual egoistic strategies become. To illustrate this, consider the following tables that represent pay-offs for δ=0.9.

DD	D	E
D	3, 3, 3	2.7, 3.1, 2.9
E	3.1, 2.7, 2.9	**2.81, 2.81, 2.81**

DE	D	E
D	0.3, 3.9, 2.1	0, 4, 2
E	0.49, 3.69, 2.09	0.2, 3.8, **2**

ED	D	E
D	3.9, 0.3, 2.1	3.69, 0.49, 2.09
E	4, 0, 2	3.8, 0.2, **2**

EE	D	E
D	2.01, 2.01, 2.01	1.8, 2.2, 2
E	2.2, 1.8, 2	2, 2, **2**

If the we-mode is more likely (0.9>0.8), the relative advantage of deliberation as the group's strategy also increases: 2.81>2.64. If the chance of encountering actors who sincerely deliberate is high, then expected pay-offs are high. If I make a commitment to deliberation myself, then the chance that others at the negotiating table interact with a committed deliberator increases, which in turn increases the expected pay-off for the whole group as well as my share of it. This should be enough of an appeal to actors to commit to honest deliberation. Again, this is not a moral appeal, or an appeal to good conscience. It is an appeal to actors' own interest, which will probably not fall on deaf ears.

Importantly, even those stakeholders who still pursue their egoistic strategies because they are in the I-mode and egoism is their best response; have an interest in increasing the overall inclination towards honest deliberation. Under δ=0.8, the expected pay-off for individual egoists is 2.64. If δ=0.9, their pay-off

is 2.81. So even egoistic players who are committed to their I-mode have a rational interest in getting as many stakeholders to operate in the we-mode as possible. They are relatively better off following their egoism if more stakeholders are in the we-mode because then the relative gains from defecting from an optimal group scenario are higher.

Discussion

We believe that our illustration above can make a much stronger and robust case for sincere deliberation and cooperation in the throughput dimension. This kind of cooperation is by no means given, since it is rational for actors who do not care about the group to behave egoistically, and this fact cannot simply be 'appealed' away. But we also believe that our findings have important implications for the discussion of legitimacy:

The fact that a norm or guideline passes the throughput dimension and is therefore agreed upon does not render it legitimate. In line with the emerging body of literature with a critical focus on collective intentionality, the fact that agents agree to join in on collectively intending to a certain path of action does not necessarily make the action 'theirs' in that they would pursue it wholeheartedly as power and status asymmetries have been considered insufficiently. In our case, norms can be issued as the result of strategy combinations of the game that are not optimal: the lower right corners of the matrices above produce norms that constitute a suboptimal pay-off.

Hence the central insight of this model is the following: *only norms that are issued under the strategy combination (E,E,DD) are legitimate*. Only (E,E,DD) is an equilibrium of the game, and it captures the intuitive understanding that legitimacy requires a norm to have come out of a robust process of deliberation so that legitimacy actually 'can be understood as the conformation with social norms, values, and expectations' (Palazzo and Scherer 2006: 71). The higher actors' inclination towards sincere deliberation, the higher the expected pay-off of the entire group will be (and hence each actor's or representative's share of it). At the same time, the norm's degree of legitimacy will increase. Moreover, for $\delta \rightarrow 1$, the group pay-off approaches 3 and amounts to a possible yield that is unattainable in the I-mode scenario. For $\delta=1$, everybody reasons in the we-mode and egoistic concerns play no role.

We find support for our findings in recent contributions to deliberative theory. Mansbridge et al. consider the classic definition of deliberation above as entirely free of coercive power and vested self-interests untenable. Rather, they argue that even if agents have their own agendas, their discussions can still meet the criterion of good deliberation (and hence legitimacy) if they adhere to standards of processual and procedural fairness. What is required for what they term 'fully cooperative distributive negotiations' is that 'participants enter a deliberation with conflicting interests, but, after deliberating on the content of those interests and on the principles of fairness appropriate to adjudicating them, adopt a distributive agreement that all consider fair. Distributive negotiations differ from integrative negotiations in having zero-sum outcomes. That is, although the parties reach an agreement that is better for each than the status quo or the other alternatives available, they give up some part of what they want to get an agreement' (Mansbridge et al 2010: 71-72).

This is exactly what happens in (E,E,DD). Players who operate in the we-mode 'adopt a distributive agreement that all consider fair' and genuinely deliberate about a new standard or norm. This new standard or norm is legitimate. They 'reach an agreement that is better for each than the status quo or the other alternatives available'. Indeed, if stakeholders operate in the we-mode then their best choice is deliberation: the expected pay-off under DD is larger than that of either ED, DE or EE. At the same time, however, 'they give up some part of what they want to get an agreement' as they could secure, at least theoretically, a pay-off of 4 in the individual mode when they play egoistically with actors who negotiate in good faith. The equilibrium norm under (E,E,DD) is legitimate but it does still allow for diverging interests to come into play as actors in the I-mode behave egoistically. Hence legitimacy does not require the utopian state of perfect deliberation amongst all stakeholders involved as the classical approach suggests. But it does require a sufficiently large number of stakeholders to reason in the we-mode and choose sincere deliberation as their approach to negotiations.

This is why the strategy combinations (E,D) and (D,E), regardless of what players choose in the we-mode, are unable to produce a norm that can be called legitimate: these pay-offs do not reflect what one would like to call good deliberation. Again, we find the results of the model above confirmed in recent deliberative theory where it is precisely argued that 'the use of threats in addition to promises moves the negotiations in question from "fully cooperative" distributive negotiations, a form of *deliberative* negotiation in which parties stand up for their interests and seek a fair outcome that incorporates those interests fairly, to distributive negotiations among "cooperative antagonists", a form of *non-deliberative* negotiation in which the parties deploy threats, a form of coercive power , to achieve their ends' (Mansbridge et al. 2010: 90, emphasis in original).

Interestingly, recent studies observe an increasing alignment or similarity in procedural reasoning by which new guidelines emerge (Bernstein and Cashore 2007; Overdevest 2010, Fransen 2012). It seems that participants develop a 'feel for the game' (Bourdieu 1977) and that indeed δ is increasing.

Conclusion

With the demise of government institutions as actors in the formulation of norms and guidelines, legitimacy has turned into an endogenous factor that must be determined within the deliberative process itself; legitimacy can no longer be granted ex post by an outside authority. Our approach models legitimacy as such an endogenous factor. It establishes a direct link between the issues of legitimacy and rationality in collective action: *it is rational for pay-off maximising agents to seek as much legitimacy as possible*.

We hope that our analysis will help steer the discussion away from a focus on idealised discursive practices towards a realist approach that considers norms and guidelines as fundamentally shaped by asymmetric interests and statuses. The question of power cannot simply be ignored. For French philosopher Michel Foucault, power is an ontological principle of collectivity and sociality as such: 'power produces knowledge' and 'power and knowledge directly imply one another' (Foucault 1991: 27).

If power is constitutive of the state of affairs in negotiations, it needs to be reckoned with. Our argument is exactly that power and sincere, honest deliberation are anything but mutually exclusive. Stakeholders can switch between modes of reasoning and their bargaining power in negotiations notwithstanding be persuaded that fair deliberation is beneficial for all. Neither do we claim that representatives from NGOs and corporations enter negotiations with eternally fixed sets of preferences. As Manin et al. argue in an early paper on political decision-making, individuals who participate in such processes 'only partially know what they want' (1987: 364). We believe the same is true for decision-making in MSI. Representatives partially know what they want and hope to satisfy their preference sets, in particular when they represent large organisations (e.g. NGOs or corporations) that have given them a specific mandate to negotiate. But the representatives involved are not necessarily complete antagonists. They can be – this is the case when each and every player operates in the I-mode. It is more likely, however, that there is a variety of interests and strategies. Some wish to cooperate and deliberate sincerely from the start, some will be persuaded to do so as negotiations continue and others will stick to their principles and only do what is best for them.

However, if they do wish to maximise their expected pay-off, a strong inclination towards 'we-reasoning' is to be recommended. For corporations, this means to renew their 'licence to operate' thanks to participating in a process that is truly legitimate. Importantly, although deliberation infuses standards and norms with 'moral legitimacy', legitimacy is *not just* a moral question. As our analysis above suggests, legitimacy is first and foremost in the rational self-interest of the stakeholders involved.

Bibliography

Bacharach, M. 1999. Interactive Team Reasoning: A Contribution to the Theory of Co-operation. Research in Economics. 53: 117-147.

Bacharach, M. 2006. Beyond Individual Choice: Teams and Frames in Game Theory. Princeton NJ: Princeton University Press (ed. by Gold, N. and Sugden, R.).
Bernstein, S./ Cashore, B. 2007. Can Non-State Global Governance be Legitimate? An Analytical Framework. Regulation and Governance. 1: 247–371.
Bourdieu, P. 1977. Outline of a Theory of Practice. Cambridge: Cambridge University Press.
Brammer, S./ Jackson, G./ Matten, D. 2012. Corporate Social Responsibility and Institutional Theory: New Perspectives on Private Governance. Socio-Economic-Review, Special Issue: Corporate Social Responsibility and Institutional Theory: New Perspectives on Private Governance. 10: 3-28.
Cashore, B./ Auld, G./ Newsom, D. 2004. Governing through Markets: Forest Certification and the Emergence of Non-State Authority. New Haven, CT: Yale University Press.
Coase, Ronald H. 1960. The Problem of Social Cost. Journal of Law and Economics. 3: (1): 1–44.
Coglianese, C./ Nash, J. (eds.) 2001. Regulating from the Inside: Can Environmental Management Systems Achieve Policy Goals? Washington, DC: Resources for the Future.
Conzelmann, T. 2012. A Procedural Approach to the Design of Voluntary Clubs: Negotiating the Responsible Care Global Charter. Socio-Economic Review, 10: 193–214.
Cutler, C./ Haufler, V./ Porter, T. (eds.) 1998. Private Authority and International Affairs. Albany: State University of New York Press.
Djelic, M.L./ Quack, S. (eds.) 2010. Transnational Communities. Shaping Global Economic Governance. Cambridge: Cambridge University Press.
Dingwerth, K. 2007. The New Transnationalism: Transnational Governance and Democratic Legitimacy. Basingstoke: Palgrave Macmillan.
Foucault, M. 1991. Discipline and Punish: the Birth of the Prison. (transl. by Alan Sheridan) London: Penguin Books.
Fransen, L. 2012. Multi-stakeholder Governance and Voluntary Program Interactions: Legitimation Politics in the Institutional Design of Corporate Social Responsibility. Socio-Economic Review. 10: 163–191.
Friedman, M. 1962. Capitalism and Freedom. Chicago: University of Chicago Press.
Friedman, M. 1970. The Social Responsibility of Business is to Increase its Profits, New York Times Magazine. Sep. 13.
Gilbert, M. 2006. Rationality in Collective Action. Philosophy of the Social Sciences. 36: 3-17.
Hahn, R. 2011. International Standardisation and Global Governance. Die Betriebswirtschaft. 71(2): 121-137.
Hakli, R./ Miller, K./ Tuomela, R. 2010. Two Kinds of We-Reasoning. Economics and Philosophy. 26: 291-320.
Klinke, A. 2009. Deliberative Politik in transnationalen Räumen – Legitimation und Effektivität in der grenzüberschreitenden Umweltpolitik zwischen Kanada und USA. Politische Vierteljahreszeitschrift. 50: 774-803.
Krause, J. 2011. Collective Intentionality and the (Re)Production of Social Norms: The Scope for a Critical Social Science. Philosophy of the Social Sciences. (available online from publisher's website, forthcoming in print).
Manin, B./ Stein, E./ Mansbridge, J. 1987. On Legitimacy and Political Deliberation. Political Theory. Vol. 15, 3: 338-368.
Mansbridge, J. et al 2010. The Place of Self-Interest and the Role of Power in Deliberative Democracy. The Journal of Political Philosophy. Vol. 18, 1: 64-100.
Matten, D./ Crane, A. 2005. Corporate Citizenship: Toward an Extended Theoretical Conceptualization. Academy of Management Review, 30, 166–179.
Mattli, W./ Büthe, T. 2003. Setting International Standards: Technological Rationality or Primacy of Power? World Politics. 56, 1: 1-42.
Morgan, G. 2010. Legitimacy in Financial Markets: Credit Default Swaps in the Current Crisis. Special Issue on Law and Legitimacy in Transnational Governance Socio-Economic Review. 8(1): 17-45.
Overdevest, C. 2010. Comparing Forest Certification Schemes: The Case of Ratcheting Standards in the Forest Sector. Special Issue on Law and Legitimacy in Transnational Governance Socio-Economic Review. 8(1): 47-76.
Palazzo, G./ Scherer, A. 2006. Corporate Legitimacy as Deliberation: A Communicative Framework. Journal of Business Ethics. Vol.66, 1: 71-88.
Pigou, A. [1920.] 1960. The Economics of Welfare. 4th ed. London: Macmillan.

Prakash, A./ Potoski, M. 2006. Racing to the Bottom? Trade, Environmental Governance, and ISO 14001. American Journal of Political Science. 50, 2: 350-364.
Prakash, A./ Potoski, M. 2006b. The Voluntary Environmentalists: Green Clubs, ISO 14001, and Voluntary Environmental Regulations. Cambridge: Cambridge University Press.
Potoski, M./ Prakash, A. 2009. Voluntary Programs - A Club Theory Perspective. Cambridge, MA: The MIT Press.
Tuomela, R. 2007. The Philosophy of Social Reality: The Shared Point of View, Oxford: Oxford University Press.
Quack, S. 2010. Law, Expertise and Legitimacy in Transnational Economic Governance. Special Issue on Law and Legitimacy in Transnational Governance Socio-Economic Review. 8(1):3–16.
Risse, T. / Kleine, M. 2007. Assessing the Legitimacy of the EU´s Treaty Revision Methods. Journal of Common Markets Studies. 1: 69-80.
Scherer, A. /Palazzo, G. /Baumann, D. 2006. Global Rules and Private actors: Toward a New Role of the Transnational Corporation in Global Governance. Business Ethics Quarterly. 16(4): 505-532.
Searle, J. R. 1995. The Construction of Social Reality. Penguin Press.
Suchman, M.C. 1995. Managing Legitimacy: Strategic and Institutional Approaches. Academy of Management Review. 20: 571–610.
Sugden, R. 2000. Team Preferences. Economics and Philosophy. 16:175-204.
Sugden, R. 2008. Nash Equilibrium, Team Reasoning, and Cognitive Hierarchy Theory. Acta Psychologica. 128: 402-404.
Vogel, D. 2005. Market for Virtue. Washington, DC: Brookings Press.
Quack, S. 2010. Law, Expertise and Legitimacy in Transnational Economic Governance: An Introduction. Socio-Economic Review. 8: 3–16.
Young, I. M. 1996. Communication and the other: beyond deliberative democracy. In: S. Benhabib (ed.). Democracy and Difference. Princeton, NJ: Princeton University Press.
Windsor, D. 2004. The Development of International Business Norms. Business Ethics Quarterly. Vol. 14, 4: 729-754.