

Regulating Corporate Virtue

Industry Self-Regulation and Institutional Change

Working Paper prepared for:
The Political Economy Doctoral Workshop
Department of Government, LSE

September 2004

Anthony D. Williams
Department of Government
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
England UK

adw@anthonydwilliams.com

REGULATING CORPORATE VIRTUE

AN INSTITUTIONAL ANALYSIS OF INDUSTRY SELF-REGULATION

In the past 10-15 years there has been a concerted effort on the part of many multinational firms, both individually and collectively, to improve their social and environmental performance standards. At first glance this behaviour seems paradoxical. Given the opportunities afforded by globalisation to do better (i.e., to lower costs and increase profits), one would not expect “socially responsible” behaviour to be high on the list of corporate priorities.¹ This surprising development requires a more thorough and theoretically grounded explanation than has yet been provided in the academic literature.²

The core proposition of the paper is that an analytical focus on economic institutions and institutional change provides the requisite theoretical framework for analysing recent regulatory change and explaining the contemporary proliferation of industry self-regulation.³ The analysis and defence of this proposition proceeds in three parts. Part one reviews some of the major contributions to the new institutional economics literature, including, but not limited to, the work of David Kreps, Douglass North, and Oliver Williamson, each of whom provide different lenses of the role of institutions in organizing economic life. In particular, three distinct, but related, theoretical domains are examined: the role of institutions and institutional change in economic analysis, the economics of organization and management, and the economics of reputation and corporate culture. Part two applies these theoretical perspectives to build a coherent story of the evolution toward industry self-regulation, thus making an original contribution to taking the literature forward. Part three offers a research agenda for studying industry self-regulation, focusing specifically on what this program of comparative institutional research can give back to the theory of institutions.

¹ For a description of recent events see (Haufler 2001) and for some further empirical evidence see the results of an OECD fact-finding mission that found 246 social and environmental codes of conduct alone in a survey of North American, European and Asian business sectors, which, by their estimate, is only a sample of the total number of codes in existence (OECD 2001).

² The main academic focus on business self-regulation is, not surprisingly, in the business management literature where a small group of scholars (e.g., Michael Lenox, Andrew King, Jennifer Nash, and John Erhenfeld) have proceeded to analyse these developments systematically, albeit from a business management perspective. With very few exceptions, industry self-regulation has received scant attention in the political science literature, and what has emerged (e.g., Haufler 2001, Fung 2003) has not produced much of theoretical interest.

³ A quick note on nomenclature: I prefer industry self-regulation to corporate social responsibility because self-regulation highlights the fact that corporations and industries are voluntarily controlling their individual and collective behaviour while “social responsibility” (CSR) is ambiguous to the point of having little objective meaning. Self-regulation denotes an objective process, while CSR denotes a subjective standard of behaviour.

PART I: THE NEW INSTITUTIONAL ECONOMICS

In 1960, Ronald Coase penned a now famous and foundational essay: “The Problem of Social Cost” (1960). Simply put, Coase proposed that when it is costless to transact, property rights and bargaining strength do not affect the efficiency of outcomes. This insight would later be formalized as the Coase Theorem. Initially, the widely cited theorem was given as a reason to ignore institutions in economic analysis because they were assumed to play no independent role in economic performance (North 1990). For many, it provided evidence that governments are unable to improve on the efficient outcomes that individuals themselves can obtain through comprehensive bargaining.

In a subtle twist of fate, the Coase Theorem would later give birth to the New Institutional Economics, as scholars such as North, Williamson, and Coase himself, began to consider what an assumption of positive transaction costs – simply defined as the costs of running an economy – would mean for economic analysis. The resounding answer was that “institutions matter.” Thanks to their work, institutions have begun to regain a prominent place in economic analysis.

Defining Institutions: Rules of the Game or Outcomes of the Game?

As testimony to the fact that we are in early days, there is still considerable debate within economics and political science about what institutions are. Most scholars agree the institutions are enduring regularities of human action. The differences concern the grounds on which these observed regularities rest. Using the analogy of a game, definitions vary according to whether the analyst perceives institutions as the rules of the game or the outcomes of the game (Aoki 2001).

Many analysts insist on separating the players and their behaviour from the rules of the game (North 1990, Ostrom 1990, Shepsle 1986). North, for example, defines institutions as the man-made constraints on human interaction that define the incentive structure of society (1990: 3). In his view, institutions consist of both formal constraints such as codified rules and laws, and informal constraints such as norms, customs, and shared beliefs. More generally, analysts within this tradition define institutions as contextual features that define constraints on, and opportunities for, individual behaviour (Diermeier and Krehbiel 2002).

For many rational choice theorists, particularly game theorists, institutions are the “outcomes-of-the-game.” In other words, they are equilibriums or focal points that arise from mutually understood actor preferences and optimising behaviour in repeated game situations (Calvert 1995, Riker 1980, Schotter 1981). Such institutions are stable and enduring because the strategies players adopt in social, political, and economic exchange are best responses to the actions of other players. No player will independently deviate unless the underlying structure of the game changes.

The particular advantage of institutions-as-equilibria approach is that responsibility for the social order lies with individuals, rather than an external state or third-party enforcer (Crawford and Ostrom 1995). Hence, the origins of institutions and issues of enforcement are endogenous to the model because the institutions themselves are no more than regular patterns of behaviour sustained by mutual expectations about the actions others will take (Calvert 1995).⁴ The weakness is that the rules of the game are often regarded as historically given, which simplifies modelling but then limits our ability to focus on how the rules change (Kreps 1990a). This highlights a key methodological concern: it is important to maintain a distinction between institutions and the behaviour that transpires within them (Diermeier and Krehbiel 2002).

Structure, Agency, and Institutional Change

To build a theory of institutions we need to explain institutional change. Conventional game-theoretic accounts of institutions have difficulty accounting for change endogenously. Because the choice sets are fixed by the structure of the game – and the players themselves cannot alter those structures independently – there is generally no conceivable source of endogenous change. As Tsebelis explains, this paradox emerges because “... individual action is assumed to be an optimal adaptation to an institutional environment, and the interaction among individuals is assumed to be an optimal response to one another” (1990: 40). Theoretically, this leads to two conclusions: institutions are overwhelmingly stable and equilibriums change only when punctuated by external shocks (Aoki 2001), and/or they are nested in a structure of meta-games in which higher-order choices between institutional features are made (Diermeier and Krehbiel 2002).

⁴ This does not make formal rules irrelevant in the analysis of equilibrium institutions; it simply clarifies the source of these rules. That is, behaviour standards and shared beliefs are often formally articulated and codified within the formal rules sets of institutions. If, on the other hand, rules are not consistent with an equilibrium choice by agents, they will be largely unenforceable (Aoki 2001).

North has arguably developed the most cogent account of institutional change. In his view, institutional change proceeds through the interaction between organizations and the institutional matrix (1981, 1990). While stability derives from the fact that there are a large number of specific constraints that affect a particular choice set, the nexus of formal and informal constraints in society are not static and immutable. Clever organizational entrepreneurs will induce institutional change as they perceive new or altered opportunities. Changes in opportunities are sometimes the consequence of exogenous changes in the environment, but more often they are the result of endogenous competition among boundedly rational actors. As North points out, it is the bargaining strength of individuals and organizations that counts; only when it is in the interest of those with sufficient bargaining strength will there be changes in the institutional framework.⁵

Rational-choice institutionalism's insistence that institutions reflect the bargaining equilibria of games focuses analysis on why institutions exist and why they remain stable. North's path dependent institutionalism more fruitfully identifies the sources (i.e., changes in prices and perceptions) and mechanisms (i.e., bargaining and competition) of institutional change. A powerful explanation of industry self-regulation, and for regulatory change more broadly, lies in a synthesis of these views. In other words, we need a view of equilibrium institutions that conceptualises individuals as opportunistic and boundedly rational actors seeking advantages within an institutional framework.

This view restores the agency of individuals to alter the structure of the game through processes of competition, learning, and experimentation, and, importantly, through political conflict.⁶ At the same time, it allows that institutional arrangements can be endogenous and self-enforcing, while eliminating the requirement for an external enforcer to impose and enforce the rules of the game.⁷ Finally, it goes beyond the idea the

⁵ North maintains that institutional change is overwhelming incremental and path dependent. Change is incremental because large-scale change would harm the organizations whose viability and profitability depend on the institutional matrix. Institutional change is path dependent because the direction of change will be broadly consistent with the existing institutional matrix and governed by the kinds of knowledge and skills that the entrepreneurs of organizations have invested in. Abrupt revolutionary change will only occur in cases where "gridlock" among competing organizations is thwarting their ability to capture gains from trade.

⁶ Breakthroughs in game-theoretic modelling that enable analysts to relax assumptions of perfect information and hyper-rationality are moving in this direction. Kreps, for one, sees the potential for game theory to model institutions as "the product of long-term experiences of a society of boundedly rational and retrospective individuals" (Kreps 1990a: 183). Milgrom and Roberts examine the role of adaptive learning in enabling players to iterate their way toward stable equilibria when there are multiple equilibria in a given situation (1988). Aoki posits a view of institutional change whereby actors with subjective perceptions of the game structure and imperfect inferences about the choices of other agents experiment with their strategic choices, producing a spontaneous and decentralized shift in the institutional order (2001). Together, these developments bring the equilibrium view strikingly closer to North's view of institutional change as an incremental bargaining process.

⁷ This doesn't mean that we eliminate governments from our analysis; rather we simply include them as another set of players in the game and work out what their strategic choices would be vis-à-vis the choices of other players.

institutions are merely constraints on behaviour to highlight the enabling capacities of institutions for making credible commitments, monitoring behaviour, and economizing on information processing for decision-making.

The Economics of Organization and Management

The economics of corporate organization and management provides a rich theoretical foundation for the new institutional economics. This growing body of knowledge, which has advanced mainly through the work of scholars such as Oliver Williamson (1985), and Paul Milgrom and John Roberts (1990, 1992), has led the economics profession toward a much deeper understanding of the structure and governance of economic organizations and the contractual relations among them.

In conventional economic models, firms are a collection of potential production plans together with a rule for selecting among them. These models, however, provide no insight into why some transactions are conducted efficiently in the market while others are conducted under the centralized authority of the firm. Moreover, they allow no explicit role for management activities (Milgrom and Roberts 1990: 59)

Beginning with Coase, transaction cost economics shifted the emphasis from production plans and the maximization of profit to transactions and the management of relationships (1937). The basic hypothesis proposed by Coase was that economic activities tend to be organized to economize on transaction costs. Decisions to acquire inputs in the market or produce them internally are based on the relative transaction costs of either alternative. In 1985, Williamson proposed a framework to specify the costs associated with different ways transacting and to discover how circumstances cause these costs to vary.

Williamson's central premise is that the differences among simple market contracting, complex contracting, and vertical integration, lie primarily in the institutions they specify for governing a relationship when unforeseen contingencies arise. To predict which of these contracting structures will govern a given relationship Williamson identifies the critical dimensions that favour one form of contracting over another. He concludes, "the principal dimensions ...are asset specificity, uncertainty, and frequency" (1985: 52).

While frequency and uncertainty are self-explanatory, asset specificity refers to the degree to which an asset's value depends on the continuation of a particular relationship. The more specific an asset is to a particular transaction, the more a party to a transaction stands to lose should the transaction be prematurely terminated.⁸ It follows that when assets are specific, and transactions subject to opportunism, parties will benefit if appropriate safeguards can be devised to enable credible commitments (1985: 48). Safeguards may take the form of incentive alignments or superior governance structures within which to organize transactions.

Given these defining attributes, the principal predictions of Williamson's transaction cost theory are as follows: Governance structure will be most complex and centralized for transactions with 1) the highest asset specificity, 2) the greatest uncertainty about performance conditions, and 3) the greatest frequency (1985: 55). Holding frequency and uncertainty constant, market contracting gives way to complex (bi-lateral) contracting, which in turn is supplanted by unified contracting (vertical integration) as asset specificity progressively deepens.

In the big picture, a key role of economic institutions – as Williamson sees them – is to facilitate credible commitments, thereby economizing on transaction costs, and permitting economic exchange in cases where it might otherwise have been too costly. Having now provided the bare bones of transaction cost economics we will leave Williamson's theory in the abstract and draw out the implications for industry self-regulation in part two.

Reputation and Corporate Culture

For our final piece of the puzzle we turn to Kreps' (1990b) economic model of reputation and corporate culture. Kreps' theory is interesting because he gives us a different perspective on the *raison d'être* for the firm than Williamson.

The starting point in Kreps' analysis of corporate culture is the same foundational observation of transaction cost economics. Many transactions will be too costly to undertake unless participants can rely on efficient and equitable adaptation to unforeseen contingencies. The problem is that each party to the transaction requires a source of faith that unforeseen contingencies will be appropriately resolved. The source of this faith, as

⁸ In Williamson's taxonomy, assets have several dimensions of specificity including human, dedicated, physical, site, and brand specificity (1985: 95-6).

Kreps sees it, is each party's reputation, which supplies a powerful (self-enforcing) mechanism for avoiding the costs of specifying and enforcing the terms of a contract in transactions where one side must trust the other.⁹ As Kreps puts it, "[Reputation] is the glue that permits mutually beneficial transactions to take place..." (1990b: 93).¹⁰

The importance of reputation to economic exchange highlights the role of corporate culture. Kreps argues that unforeseen contingencies are best met by the sort of principle that underlies what game theorists call a focal point – a principle or rule individuals use naturally (often as a product of experience) to select a mode of behaviour in a situation with many possible equilibrium behaviours (Schelling 1960). Firms will be judged according to the principles it selects and its diligence in embracing and applying these principles. Corporate culture, according to Kreps, is partly the principles themselves, and partly the vehicle for communicating the principles to future trading partners, and importantly, to the people the firm employs to uphold and apply them.¹¹

There are three important points to impart from Kreps' analysis. 1) Reputation provides a strong incentive for firms who wish to realize the value of repeating transactions to engage in "responsible" behaviour – it ties past actions to future trading opportunities. 2) Corporate culture establishes principles for dealing with unforeseen contingencies as they arise and mechanisms for ensuring that principles are applied in all cases, even when its application will *not* be optimal in the short-run. 3) The consistency with which these principles are applied establishes the strength of a firm's reputation, while visibly inconsistent behaviour is often sufficient to irreparably impair it. Together, these insights give us a unique explanation for what a firm is: an intangible asset carrying a reputation that is beneficial for efficient and enduring transactions (1990b: 95).

⁹ Notice that while Kreps and Williamson identify the same problem, their interpretation of the solution is radically different. Whereas Williamson argues that hierarchical governance structures and "hostage exchanges" are often the only sufficient means by which to elicit credible commitments that constrain the tendency to act opportunistically, Kreps argues that the value of intangible assets such as reputation – in so far as a good reputation sustains future trading relationships – is often a sufficient disincentive to opportunistic behaviour.

¹⁰ Reputations are built on how parties have adapted to unforeseen contingencies in the past. Whether the party's past behaviour adds or detracts from their reputation will have consequences for the amount of faith future trading partners will have. As a result, most firms have a strong interest in preserving or even promoting a good reputation to allow for future beneficial transactions. This does not require that the transaction be repeated with the same player. It only requires that actions are observable, the firm endures to play future rounds, and that its opportunities in later rounds can be tied to behaviour in earlier rounds through the mechanism of reputation (p. 106).

¹¹ Communication of the principle is critical as Kreps points out. As he put it, "Potential trading partners are judging our decision-making entity on its faithful application of the principle; it is clearly important that the entity let other know just what that principle is" (p. 125).

PART II: NEW INSTITUTIONAL ECONOMICS AND INDUSTRY SELF-REGULATION

The major research challenge posed at the beginning of the paper was to explain why many multinational firms are ratcheting up their social and environmental performance standards, despite opportunities afforded by global economic integration to lower costs and increase profits. Deriving a satisfactory explanation requires a two-prong approach that pulls together the many threads of new institutional economics discussed above. First, I examine a series of institutional changes that have unfolded since the 1930s to explain why the relatively stable institutions of the Fordist era established after the Great Depression began to unravel in the 1980s, only to be partially replaced a decade later by a fledgling set of self-regulatory institutions. Second, I specify the pattern of constraints and incentives that underlie industry engagement in self-regulation, while making predictions about the contracting structures of self-regulation and the conditions under which self-regulating systems are likely to be successful. The result is a path dependent explanation of industry self-regulation based on solid micro-foundations.

Institutional Change and the Rise of Industry Self-Regulation

From 1914 through to 1945 the world endured several severe shocks that created the political and economic conditions in which a new institutional order was erected. Two world wars and the Great Depression considerably weakened big business, galvanized a powerful trade union movement, and legitimized big government as an authoritative social planner (Polanyi 1944, Ruggie 1998). The result, as Ruggie put it, was the “embedded liberalism compromise”: a new political and economic equilibrium based on implicit social contracts and consensual collective bargaining (2002).

From 1945 a stable equilibrium adheres until approximately the late 1970s. Given the bargaining the strength of the players and the set of contractual bargains that facilitated economic exchange, none of the players found it advantageous to devote resources to fundamentally restructuring the underlying institutional order.¹² In other words, maintaining the social bargain was the best-response strategy of each player.

¹² Indeed, impressive levels of economic growth and productivity gains were generating high rates of return for everyone: firms were increasing growth and profits, government were maximizing revenue and enlarging their bureaucracies, while workers were gaining higher wages and improved standards of living. But, as North points out, just because an institutional arrangement is in equilibrium does not mean that everyone is happy with the existing rules and contracts, but only that the relative costs and benefits of altering the game among the contracting parties does not make it worthwhile to do so (1990: 86).

Stability, however, did not imply complete institutional inertia. On the contrary, a great deal of incremental change and “recontracting” occurred within the institutional framework as firms, regulators, and unions sought to maximize their power and revenues (Stigler 1971, Buchanan, Tollison, and Tullock 1980). By the mid 1970s, the cumulative impact of these decentralized attempts to maximize wealth was exacting a heavy toll on economic performance (Olson 1982).¹³ Political support for the complex nexus of rules and regulations that governed economic exchange began to wither as its deadweight costs became too large (Becker 1983, 1985). Business interests in particular began to perceive that they could do much better with an altered agreement or contract – it had simply become too costly to transact within the prevailing institutional framework.

During the 1980s, the Fordist institutional equilibrium was punctuated. The unravelling was hastened by several prominent sources of disequilibrium, including changes in technology, manufacturing processes and corporate organization, ideology, international relations, and overseas investment opportunities.¹⁴ Altogether, these developments vastly strengthened the relative bargaining power of business interests, creating new, cost-efficient opportunities for business to restructure its relationships with the state and society. The results included a fairly sudden shift to deregulate, privatise, and liberalize the economies of advanced industrial countries, an erosion of the structural power of labour, and a general trend toward global economic integration that carried forward into the 1990s (Bhagwati 1995, Rodrik 1997, Ruggie 1998).

Rather than settle on a new equilibrium however, globalisation upset many of the regularized, stability-inducing patterns of interaction between firms, governments, and labour unions. In many respects, the shift from national to international patterns of economic exchange left an “institution-free” void and a lack of mutual expectations about the actions that other players would take.¹⁵ Economic dislocation and growing insecurity dramatically increased uncertainty for many constituencies in society (Gilpin 2000). A new period of political conflict emerged, launching a popular backlash against globalisation and the so-called rise of “corporate power.”

¹³ As Becker explains, “deadweight costs of regulations and other policies often rise over time as labour and capital become more mobile, as substitutes develop for products that have been made more expensive, and as other costly methods of evading and avoiding the effects of particular regulations are discovered” (1985: 340).

¹⁴ For evidence of these transformations see Castells 1996, Gilpin 2000, Keohane and Nye 2001, and Ruggie 1997.

¹⁵ In actuality, new institutions were in the making as government and business associations crafted new formal rules to govern international economic exchange through bodies such as the WTO. As Ruggie has argued, what lags is a translation of the national social contracts into a complementary set of norms and rules to govern the social and environmental consequences of global economic exchange (Ruggie 2002).

None of this is surprising. North points out that “what makes up a stable choice theoretic context is the total package of formal and informal constraints and enforcement aspects” (1990: 87). The stable pattern of political bargaining and economic exchange that persisted during the Fordist era was due not just to the deeply nested institutionalised rules and structures provided by political institutions, but also to the informal norms, customs and ideological commitments to justice and social welfare that animated the consensus among firms, governments, and society. While the formal rules may have changed during the 1980s and 90s, the informal constraints remained the same. Indeed, the expectations that business would uphold its end of the “social bargain” were as strong as ever, if not stronger given the turbulence of the times.

Thus a key factor in explaining industry self-regulation is the tendency of informal constraints to persist. When changes in formal rules clash with informal constraints transaction costs can increase, particularly when people perceive these rules as unfair (North 1990: 76). This suggests that industry self-regulation is only partially new. Rather, it is best understood as an attempt to resolve political conflict over globalisation by re-embedding social norms and re-institutionalising the processes social bargaining in a more flexible and globally extensible institutional framework. Indeed, if the initial investment in negotiating agreements, and the ongoing costs of enforcement, are less than the costs of unmitigated conflict – or worse, the reversal of economic integration – then it is conceivable that industry self-regulation is a politically (if not economically) efficient means to create stable political foundations for global economic exchange.

Having described the institutional changes that preceded industry self-regulation lets return to the economic theories of organization, management, and corporate culture to briefly draw out some further implications of our theory.

Micro-Analytics: Incentives and Institutional Design in Self-Regulatory Systems

A transaction-cost theory of industry self-regulation suggests that complex forms of industry self-regulation will be viable, and indeed desirable, as long as the investment industries make in governance can be recouped in recurring gains from trade. Hence, the cost-benefit calculus for firms is not driven primarily by production costs, but rather by transaction cost considerations – particularly the desire to reduce risk and uncertainty. Governance structures that establish mutual expectations of “contract fulfilment” and facilitate credible commitments lower transaction costs. By creating stable, regularized

patterns of political exchange among firms and their stakeholders, they also enhance profitable opportunities. This explains why firms engage in what otherwise appears to be irrational behaviour. But, if so, why aren't all firms pursuing this course of action?

Transaction cost economics provides us with specific predictions along two dimensions of observed variability. First, the theory predicts that industries with high asset specificity are most likely to invest in self-regulation as insurance to safeguard value of those assets. Highly visible and valuable corporate reputations provide the greatest incentive for industry self-regulation because they tie past actions to future trading opportunities. In the terms specified above, reputation raises the cost of defecting from implicit social bargains that have persisted since at least the 1950s.

Not only does industry self-regulation substantiate the importance of reputation, it places emphasis to a range of other asset specific investments as well. For example, most extractive industries make highly site-specific investments. It is probably no coincidence that firms in the mining, forestry, and oil and gas industries are among those firms most actively engaged in self-regulation. Their inability to costlessly shift assets into other uses makes them particularly vulnerable to social pressure, particularly, but not exclusively, from the communities that live in close proximity to the site of extraction.

Second, transaction cost economics predicts that the contracting structures of self-regulation will become increasingly complex as asset specificity increases. As in Williamson's analysis, non-specific assets do not justify investments in complex contracting. At best, we are likely to see simple forms of market governance in which firms make marginal adjustments in behaviour to respond to social pressure. More complex contracting structures will arise as asset specificity increases. Firms will negotiate detailed codes of conduct. They will expend considerable resources to measure and report their performance. They may even engage third party monitors to publicly verify their compliance. All of these investments can be justified in relation to the value a firm's assets would lose if its transactions were prematurely terminated.

Now consider corporate culture. Recall that the role of corporate culture at the firm level is to establish principles for dealing with unforeseen contingencies. Industry codes of conduct perform the same function at the industry level. For self-regulation to work, all parties to the agreement must *ex ante* have some idea of the meaning of appropriate or equitable fulfilment of the agreement. The purpose of industry codes is to forge consensus on the principles that constitute responsible industry behaviour, while providing

mechanisms for ensuring these principles are applied in all cases and a vehicle for communicating them to internal and external stakeholders.

Finally, we come to the issues of institutional supply, effectiveness and enforcement. Most self-regulatory agreements provide collective goods – if one firm, or group of firms, opts to supply them they cannot feasibly exclude other firms in the industry from enjoying the benefits. Hence we should anticipate collective action problems. This leads many analysts to predict that unless there is a coercive power (i.e., government) to enforce a binding contract, the dominant strategy is always to not contribute. Yet, plenty functioning self-regulatory systems exist, so under what conditions will firms act collectively to supply new institutions?

Another way to put the question is: why do industry codes exist at all? Indeed, why don't firms simply address social responsibility issues at the firm level? A likely answer can be found in the collective nature of reputation and the fact that the behaviour of poor performers can tarnish the public image of an entire industry. This provides an incentive to address problems at the industry level, and well-established industry associations provide a clue as to how collective action problems might be solved. Nevertheless, there is much research to do to refine and test this proposition by examining nature of corporate reputation, as well as the structure of different industries.

Once an institution has been supplied, what makes it self-enforcing? Generally, a self-enforcing agreement between parties remains in force as long as each party believes himself to be better off by continuing the agreement than he would be ending it (Telser 1980: 27). The possibility of endangering future gains from trade by acquiring a reputation for irresponsible behaviour may compel firms to act benevolently in the present. But such agreements are inherently fragile, especially in a world of complex impersonal exchange. Any particular situation will depend on the structure of the pay-offs for each player: how substantial are the future gains of cooperation, how costly is it to comply, how often will violation be detected, or how credible is the threat of future harm. In order for self-regulatory agreements to take effect, at least two conditions must be met.

First, behaviour must be observable. This raises the problem of imperfect and asymmetric information. The problem is that the principal (this could include society, or simply the self-regulatory association) may not have perfect information about what type of agents (the industry participants) are voluntarily joining the self-regulatory regime, giving rise to the problem of adverse selection. Indeed, some firms may join to secure the

insurance and signalling benefits of membership, but have no actual intention of following through with the commitments. Once an agent has joined, the principal may not have perfect information about whether or not the agent is actually following through with their contractual obligation, giving rise to a moral hazard. To be sure, there are bound to be occasions when even well-intentioned firms will feel pressure to shirk, but will choose not to report this to the principal. These information asymmetries and conflicts of interest are important in determining the success of a self-regulatory regime.

Second, the gains from future trade must be greater than the costs of compliance and the expected utility of defection. Our discussion of reputation and asset specificity makes some specific predictions in this regard. But, in most cases, self-enforcing incentives will not be enough to deal with the hazards of opportunism discussed above. Indeed, the inherent difficulties in ensuring compliance and coping with principal-agent problems suggest that institutional remedies are a necessary component of effective self-regulation. Any successful self-regulatory system is therefore likely to require an organisational setting to enhance group interaction in order to secure and reinforce credible commitments to key principles; mechanisms to monitor and report industry performance on an ongoing basis (perhaps, with third party verification), and an appropriate and enforceable set of sanctions in place to punish free-riders and shirkers when necessary. There is insufficient scope to explore these institutional design issues further, but they form an important part of the subsequent research agenda.

PART III: AN AGENDA FOR THE STUDY OF INDUSTRY SELF-REGULATION

The new institutional economics provides a rich theoretical foundation for explaining self-regulation. In this paper I have developed an explanation of industry self-regulation grounded in an understanding of institutional change within advanced industrial countries over the past 75 years. I have also secured this account in the micro-foundations of corporate behaviour. In doing so, I have shown what can be accomplished by drawing together some distinct theoretical economic perspectives in the study of self-regulation. In conclusion I briefly propose a research strategy to elaborate, refine, and test this institutional theory of industry self-regulation.

Comparative Institutional Analysis

Institutional theories explain how particular institutional features constrain and enable the behaviour of individuals and organizations (Weingast 1996). A theory of institutions explains why some institutional features come into existence and persist, while others are non-existent or transient (North 1990). In building a theory of institutional change, some of the essential contextual features that were assumed to be constraining in foundational institutional theories become of the objects of choice within a more general theory of institutions (Diermeier and Krehbiel 2002). Comparative institutional analysis is the methodology for linking robust institutional theories with a theory of institutional change.

In the study of industry self-regulation, both types of analysis are necessary. That is, an institutional theory of industry self-regulation will compare a variety of self-regulatory systems to explain and predict how various institutional features such as systems for learning, monitoring, and sanctioning affect the ability to make credible commitments, control free-riding, and manage principal agent problems. A higher-order theory of institutional change will seek to explain why industry self-regulation emerges in some industries and not in other, or in some nations or regions more than others, or more generally, in the presence of certain contextual features rather than others. More specifically, a theory of institutions approach to self-regulation seeks to identify the endogenous and exogenous features that drive change and enable institutional supply.

Propositions

With a deductive theory-building approach I have already developed a number testable hypotheses to explore institutional dynamics of industry-self-regulation, including, but not limited to, the following propositions:

1. Self-regulation is a product of the institutional environment. Levels of self-regulation will tend to be highest in those countries in which institutional history has established high expectations for social responsibility and where globalisation has created the highest levels of economic dislocation and insecurity.
2. The issues addressed in self-regulatory systems will be those for which there is a high concentration of organized interest group activity that increases the level of public concern and the observability of corporate behaviour.

3. Self-regulation will be greatest in those industries with the highest levels of asset specificity. Further distinctions could be made here by delineating different forms of specificity and testing to see which is most significant.
4. Governance structures will be increasingly complex and centralized for those industries with the highest levels of asset specificity.
5. Given collective action problems in arranging institutional supply, many of the conventional dynamics of group behaviour apply. Industry self-regulation will be greatest in concentrated industries with a small number of large firms with deep resources, established associations, and overlapping fortunes given the collective, non-excludable nature of industry reputation.
6. Self-regulation is not sufficiently self-enforcing. The highest rates of compliance, therefore, will be achieved in systems that have robust systems for learning, monitoring, and enforcement. Those that don't will suffer from free riding, adverse selection, and moral hazard.

Operationalization

Testing these hypotheses requires a comparative institutional analysis across nations, industries, issues, and self-regulatory systems. At each level we seek to explain dependent variables such as the supply of self-regulation, the level of commitment within the industries (i.e., what percentage of firms join), the presence or absence of institutional features, and the levels of compliance (to the degree that this is observable). I will quickly specify what each might entail without getting too granular at this stage.

Nations: A comparative study of self-regulation across nations addresses proposition one that there is a causal association between “national experiences” of globalisation, institutional history, and the prevalence of national systems of industry self-regulation. In addition to looking at recent histories of political economy across nations, this might entail looking at national measures of global economic integration, public opinion surveys, levels of interest group organization, and various economic indicators to assess whether national contextual features are correlated with our response variables.

Industries. Of course, many self-regulatory systems transcend the nation-state, so we must look beyond national dynamics to explain them. Indeed, the theory suggests that transaction cost considerations, and other factors endogenous to the industry, are of primary importance. Propositions 3, 4, and 5 point to three key explanatory variables:

asset specificity, industry structure, and reputation. As above, a comparative analysis of industries will examine a wide sample of industries (to ensure variability in the dependent variables) to measure the association between, for example, high levels of asset specificity and the presence of self-regulation, the level of commitment, and particularly, the complexity of the governance structures that emerge.

Issues. A comparative examination of patterns of industry self-regulation across various issues provides a methodology for testing proposition two. Proposition two suggests that the presence of NGO activity, which generates public concern and higher levels of transparency, is a key explanatory variable for explaining the widely variable issue coverage of self-regulatory agreements. Thus, issues with a high level of NGO activity generate higher levels in our response variables.

Self-regulatory systems. Finally, a comparative analysis of self-regulatory systems will examine a sample of self-regulatory systems with different institutional design features. In this case we move away from explaining the supply of self-regulation in which case institutional supply and the presence or absence of certain institutional features was dependent on exogenous contextual factors. Instead, we make the institutional features exogenous and attempt to explain the behaviour that transpires within the institutional constraints. Thus, the units of analysis become the levels of compliance, free-riding, and principal agent problems, as we seek to reach conclusions about how effective institutional design enables credible and enforceable commitments.

Conclusion

There is no doubt that the time has come for more extensive research into industry self-regulation. As John Ruggie remarked in a recent lecture addressing the UN Global Compact, “There are no comprehensive surveys or even inventories of the exploding universe of such arrangements, or even a common vocabulary to describe them, and no good understanding yet of when or where they produce what effects” (2002). Indeed, too much recent rule-making activity is transpiring in the private sphere for political scientists to remain oblivious to these developments. It is conceivable that the private ordering strategies of multinational corporations will come to dominate many areas of regulation that have recently been the preserve of nation-states. When, or if, that will happen (and with what consequences) are questions we can only resolve with a theoretically and empirically grounded study of industry self-regulation.

Bibliography

- Alt, James and Kenneth Shepsle. 1990. *Perspectives on Positive Political Economy*. Cambridge: Cambridge University Press.
- Aoki, Masahiko. 2001. *Towards a Comparative Institutional Analysis*. Cambridge: MIT Press.
- Becker, Gary. 1983. "A Theory of Competition Among Pressure Groups for Political Influence." *Quarterly Journal of Economics*. 98: 371-99.
- Becker, Gary. 1985. "Public Policies, Pressure Groups, and Deadweight Costs." *Journal of Public Economics*. 28: 330-47.
- Buchanan, James, Robert Tollison and Gordon Tullock. 1980. *Toward a Theory of Rent-seeking in Society*. College Station: Texas A&M Press.
- Calvert, Randall. 1995. "The Rational Choice Theory of Social Institutions: Cooperation, Coordination, and Communication." In, Jeffrey Banks and Eric Hanushek, eds. *Modern Political Economy: Old Topics, New Directions*. 216-28. New York: Cambridge University Press.
- Coase, Ronald. 1937. "The Nature of the Firm." *Economica* 4: 386-405.
- Coase, Ronald. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3(1): 1-44.
- Crawford, Sue and Elinor Ostrom. "A Grammar of Institutions." *American Political Science Review* 89 (3): 582-600.
- Diermeier, Daniel and Keith Krehbiel. 2002. "Institutionalism as a Methodology." Unpublished manuscript.
- Fung, Archon. 2002. "Making Social Markets: Dispersed Governance and Corporate Accountability." In, John Donahue and Joseph Nye, eds. *Market-Based Governance: Supply Side, Demand Side, Upside, and Downside*. Washington D.C.: Brookings Institution Press.
- Gilpin, Robert. 2000. *The Challenge of Global Capitalism*. Princeton: Princeton University Press.
- Haufler, Virginia. 2001. *The Public Role of the Private Sector: Industry Self-Regulation in a Global Economy*. Washington: Carnegie Endowment for International Peace.
- Kreps, David. 1990a. "Corporate Culture and Economic Theory." In, Alt, James and Kenneth Shepsle. 1990. *Perspectives on Positive Political Economy*. Cambridge: Cambridge University Press.
- Kreps, David. 1990b. *Game Theory and Economic Modelling*. Oxford: Oxford University Press.
- Milgrom, Paul and John Roberts. 1992. *Economics, Management and Organization*. New Jersey: Prentice Hall.
- Milgrom, Paul and John Roberts. 1990. "Bargaining costs, influence costs, and the organization of economic activity." In, Alt, James and Kenneth Shepsle. 1990. *Perspectives on Positive Political Economy*. Cambridge: Cambridge University Press.
- Milgrom, Paul and John Roberts. 1988. "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities." *Econometrica* 58: 1255-1277.
- North, Douglass. 1990. *Institutions, Institutional Change, and Economic Performance*. Cambridge: Cambridge University Press.
- North, Douglass. 1981. *Structure and Change in Economic History*.
- Olson, Mancur. 1982. *The Rise and Decline of Nations*. New Haven: Yale University Press.
- Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

- Peltzman, Sam. 1976. "Toward A More General Theory of Regulation." *Journal of Law and Economics* 19: 211-40.
- Riker, William. 1980. "Implications from the Disequilibrium of Majority Rule for the Study of Institutions." *American Political Science Review* 74: 432-46.
- Ruggie, John. 2002. "Taking Embedded Liberalism Global: The Corporate Connection." Speech given at the London School of Economics, June 6 2002.
- Ruggie, John. 1998. *Constructing the World Polity: Essays on International Institutionalization*. London: Routledge.
- Schelling, Thomas. 1960. *The Strategy of Conflict*. New York: Oxford University Press.
- Schotter, Andrew. 1981. *The Economic Theory of Social Institutions*. Cambridge: Cambridge University Press.
- Shepsle, Kenneth. 1986. "Institutional Equilibrium and Equilibrium Institutions." In, Herbert Weisberg. *Political Science: The Science of Politics*. New York: Agathon Press.
- Stigler, George. 1971. "The Theory of Economic Regulation." *Bell Journal of Economics*, 2:3-21.
- Telser, L. G. 1980. "A Theory of Self-enforcing Agreements." *Journal of Business*. Vol. 53, No. 1: 27-44.
- Tsebelis, George. 1990. *Nested Games*. Berkeley: University of California Press.
- Weingast, Barry. 1996. "Rational Choice Institutionalism." In Robert Goodin and Hans-Dieter Klingemann, eds. *A New Handbook of Political Science*. Oxford: Oxford University Press.
- Williamson, Oliver. 1985. *The Economic Institutions of Capitalism: Firms Markets, Relational Contracting*. New York: The Free Press.
- Williamson, Oliver. 1993. "Calculativeness, Trust, and Economic Organization." *Journal of Law and Economics*. 34: 453-500.