Summer 2011

The Coevolution of Transnational Corporations and Institutions

Sarianna M. Lundan
University of Bremen, Germany, s.lundan@uni-bremen.de

Follow this and additional works at: http://www.repository.law.indiana.edu/ijgls

Part of the Business Organizations Law Commons, International Law Commons, and the Transnational Law Commons

Recommended Citation
Available at: http://www.repository.law.indiana.edu/ijgls/vol18/iss2/3
The Coevolution of Transnational Corporations and Institutions

SARIANNA M. LUNDAN*

ABSTRACT

While economic theories of the firm have traditionally focused on the ownership of assets, the increasing use of contractual partnerships is beginning to challenge our conception of the firm by emphasizing its coordinating role. In structuring their contracts, as well as in managing the relationships governed by the contracts, firms try to mitigate uncertainties that could destroy the value-adding potential of such transactions. These uncertainties may be specific to the transaction partner, but they might also arise from the institutional context of the contracting parties, particularly in the case of transactions that cross borders. The coevolutionary process whereby firms both adjust to and shape the institutional constraints facing them results in new hybrid forms of governance, which contribute to the body of private transnational law. By studying how firms mitigate the uncertainties in their contractual relationships, empirical research can yield new insights about the emergence and impact of private law.

INTRODUCTION

The integrated industrial enterprises that emerged from the first global economy1 grew by horizontal and vertical integration, as well as by expanding abroad in order to take full benefit of the economies of scale and scope. To cope with the growing complexities created by the firm's expansion, emerging industrial giants such as DuPont, ICI, and Siemens started to experiment with the multidivisional (M-form) organization. This process began before World War I, but came into full

* Professor for International Management and Governance, Faculty of Business Studies and Economics, University of Bremen, Germany. Please send mail to s.lundan@uni-bremen.de.

1. The first global economy peaked on the eve of World War I, and the same level of outward direct investment stock to global GDP was not reached until the 1990s.
force in the interwar years. After World War II, with the help of the consultants at McKinsey, the M-form was subsequently transplanted to firms in Europe as part of the process of reconstruction.²

With fax and the Internet replacing the telephone, and with cheap air travel and container shipping replacing the railroads, the second global economy began to take shape in the early 1980s, reaching its peak before the onset of the current financial crisis. However, unlike the first global economy, the second global economy does not have a similarly emblematic organizational form associated with it. In the 1980s, transnational corporations (TNCs) underwent a period of experimentation with matrix organizational forms, only to return to variations of the multidivisional form.³ However, with the increase in outsourcing over the past decade or so, there is now a greater variety of organizational forms in use, which fall under the heading of network organizations.⁴

In broad terms, network organizations consist of an internal ownership-based network and an external contractual network of cooperative partners. At one end, this includes the conventional M-form divisional organizations that have simply responded to the greater availability of outsourcing partners by increasing the proportion of outside suppliers they use. At the other end are organizations, such as Nike or Apple, whose main value added at the headquarters consists of a combination of branding, design, and research and development (R&D) activities, and who contract for the manufacturing of their products by original equipment manufacturers (OEMs).

This shift in the governance forms adopted by large firms, and particularly those firms whose activities cross borders, has important implications for the theory of the firm, as well as for the institutional structure in which firms operate. Specifically, this article will put forward three interlinked propositions:


⁴. See generally D. Eleanor Westney & Srilata Zaheer, The Multinational Enterprise as an Organization, in THE OXFORD HANDBOOK OF INTERNATIONAL BUSINESS 349 (Alan M. Rugman & Thomas L. Brewer eds., 2001) (discussing the development of different forms of organization over the last few decades in response to changes in international business).

⁵. By governance I mean the tools used to induce order in exchange relations when potential conflicts due to information asymmetry or other reasons threaten the possibilities for cooperation.
1. The increasing use of contractual partnerships emphasizing the firm's coordinating role rather than its ownership of assets. More specifically, any theory of the firm that takes as its starting point the difference in the transaction costs over the market and inside the hierarchy has to account for the increasing use of external contracts in the coordinating function of the firm.

2. The firm that coordinates an increasing proportion of activities across the external market interface, rather than within the internal hierarchy, has external effects on other firms in the market by lowering the transaction costs of using the market.

3. The increasing amount of contractual activity opens up numerous opportunities for firms to shape the body of private transnational law. Consequently, by studying how firms mitigate the uncertainties in their contractual relationships, empirical research can yield new insights about the emergence and impact of private law.

This article is organized as follows. The first section discusses the implications of the changing structure of TNCs, particularly in reference to transaction cost economics. In this predominant theory of the TNC, the firm is seen as a nexus of contracts, where market transactions are replaced by employment relations inside the firm if the transaction costs of using the market are excessive. In this view, the internalization of market transactions in response to market failure is seen as the raison d'être of the firm. Since I will argue for the need to account for the implications of the increasing use of external contracts on our conception of the firm, I will offer an alternative view of the TNC as a system of coordinated activities. Such a system consists of both ownership-based and contractual activities and spans across multiple levels between subsidiaries, TNC headquarters, and the home and host markets.

The second section will argue that the reflexive process whereby the TNC as a system both adjusts to and shapes its operating environment results in a coevolutionary trajectory that contributes to broader

---

6. When market failures occur, it is due to factors preventing international markets from functioning perfectly—including both structural market imperfections (e.g., cartels and government interventions) and such endemic market shortcomings as uncertainty, asymmetrical information, bounded rationality, and externalities.
transformations in the home and host countries. One important way in which TNCs are shaping the market context is through the different governance mechanisms they employ to deal with the uncertainties that arise in their external cooperative relationships. While such relationships are contractual, there is emerging evidence showing that the contracts covering such activities are not uniform, but quite heterogeneous, suggesting that firms are developing novel solutions to the specific uncertainties they encounter in a particular market context. As these solutions become articulated and codified, they become diffused within the firm itself and to its external partners. Over time, such practices may also become more widely diffused, whether deliberately or incidentally, and it is this process of diffusion that in the aggregate shapes the body of transnational private law that firms help create.

In the third section, I pay particular attention to the cross-border dimension of the firm's activities and the differences between developed markets, where private ordering takes place in the shadow of well-functioning public ordering, and emerging markets, where private ordering might be replacing or at least mitigating some of the deficiencies in public ordering. I conclude by briefly reviewing the implications of my argument, and by suggesting ways in which research combining legal and business perspectives can help to uncover the extent and forms of these extralegal activities conducted by firms. I begin by examining the changes in the theory of the TNC in response to increasing contractual activity.

I. TRANSNATIONAL CORPORATIONS AS COORDINATED SYSTEMS

Changes in the structure and form of TNCs are manifestations of the efforts of firms to identify practices that can mitigate the uncertainties presented by their environment. The transformation of the TNC from an ownership-based entity to a network firm has involved three types of changes that affect the firm boundary. First, the boundaries of firms have become more porous, consisting of a variety of equity-based and contractual interfaces with suppliers and customers. In addition to redefining the boundaries of traditional TNCs, we have seen the emergence (mainly in Asia) of new network-type TNCs.

Second, there has been a notable increase in the internal diversity of TNCs as a result of a continued devolution of responsibility to the subsidiary level.\(^1\) As a result of knowledge generation at the subsidiary level, the TNC parent is presented with new challenges for integrating the knowledge and material flows within the TNC network.\(^2\) Third, in addition to relationships between TNCs and government, TNCs have begun to engage in a greater variety of nonmarket interactions involving civil society and quasi-governmental institutions.\(^3\) In response to these changes, the theory of the TNC has had to come to terms with the fact that transactions in multiple markets (markets for assets, asset services, and firms) determine the governance form and that differential ownership advantages are likely to influence who internalizes what.\(^4\)

The seminal contributions of Buckley and Casson,\(^5\) Rugman,\(^6\) Teece,\(^7\) and Hennart\(^8\) led internalization theory to become one of the

---


primary theoretical foundations of the nascent field of International Business (IB). In addition to addressing the questions of the mode of entry and firm borders, these scholars put forward the idea that market internalization itself was sufficient to explain the existence of the transnational corporation.\textsuperscript{18} Internalization was thus presented as a general theory of the firm, capable of explaining the emergence of the TNC as a special case.

Internalization theory and the transaction cost economics (TCE) associated with Williamson\textsuperscript{19} developed independently of each other at around the same time and share a number of similarities since they both draw on the original work of Coase.\textsuperscript{20} One of the main differences between the two approaches is that while internalization theory focuses specifically on imperfections in the market for knowledge, TCE pays more attention to small numbers conditions and opportunism, which are likely to impede transactions between unaffiliated parties. However, as Verbeke and Greidanus suggest, the idea that considerable weight should be placed on designing governance forms that safeguard against the pervasive opportunism of one's partners does not sit well in a context where long-term collaborative relationships are becoming the norm and not the exception.\textsuperscript{21} These differences aside, both are comparative institutional theories that seek to explain why one form of governance substitutes for another.

When developing the eclectic or OLI paradigm, which became the preeminent theoretical framework to explain the pattern of TNC activity, Dunning placed considerably more emphasis on the need for

81 (highlighting the characteristics of the market for technology and the conditions for technology transfer).


18. I use the term "transnational corporation" (TNC) since it is commonly used in the legal literature, and it is also the nomenclature used by the United Nations. However, most contributions to the literature in economics and international business use the term "multinational enterprise" (MNE) since the word "transnational" is sometimes associated with a particular firm structure. See generally Christopher A. Bartlett & Sumantra Ghoshal, Managing Across Borders: The Transnational Solution 83-154 (2d ed. 2002).


the TNC to possess ownership advantages prior to the act of internalization. He argued that the pattern and extent of the cross-border activities of TNCs could only be explained by the interplay between the mobile ownership advantages of the firm (the O-factor), the immobile location-bound resources in the host country (the L-factor), and the degree to which these activities are internalized (the I-factor).

In the first instance, such advantages are likely to consist of little more than entrepreneurial judgment, which consists of the ability of the entrepreneur to acquire information and to make judgments about its usefulness. As the firm grows by hiring more employees and procuring additional inputs, a second kind of ownership advantage is called for that enables efficient organization within the firm. This requires an ability to set hierarchical incentives in a way that increases employee commitment and maximizes output.

As the firm continues to grow in scale and scope, it is likely that in addition to managing activities within the firm, there is a growing need for it to manage an increasing number of subcontractors. Indeed, in diverse fields such as construction, media, and software development, there is no compelling reason for the firm to internalize the output of the various input providers in the first place, since the same end result can be achieved contractually by specifying the characteristics of the output. Instead, extensive projects are carried out by teams consisting of a small core of salaried employees and a large number of external contractors. Thus the size of the firm is mitigated by its ability to effectively coordinate the inputs of the various independent subcontractors.

These two types of advantages correspond quite closely to the two categories of ownership advantages outlined by Dunning. These are the asset-based ownership advantages (Oa) and the transaction-based ownership advantages (Ot). When these were introduced, the conception of the TNC was still prominently ownership based. Thus, Oa refers to advantages the firm owns, whether these are intellectual property rights in the form of patents and trademarks, stocks of tacit knowledge, or tangible property and equipment. The Ot advantages are essentially the economies achieved by a multiplant firm over a single plant firm.


and they also incorporate any advantages derived from transnationality per se. This formulation of ownership advantages thus separates the advantages based on the possession of superior competitive assets, whether intellectual or real assets, from those that are based on the cross-border coordination of these assets.

The transaction-based advantages (Ot) of the firm can be related directly to the ideas of Penrose on the managerial limits to growth. As firms grow, the ability of management to oversee an increasingly diverse set of operations declines and sets the limits to further growth, since top management cannot be expanded ad infinitum. Chandler expressed similar ideas in his analysis of the emergence of the multidivisional firm. He argued that as functionally organized firms were diversifying both in terms of their product range as well as their geographical reach, the multidivisional form became the preferred structural solution to reduce the complexity facing top managers.

However, from a contemporary perspective, this analysis ignores the increasing ability of firms to expand their value-adding activities through contractual means, which neither requires the resource commitment nor carries the problems of hierarchical coordination envisaged by Penrose and Chandler. Thus to the extent that the ability to contract for output simplifies the coordination task, managers should find it possible to control a firm of greater complexity than would be possible only hierarchically.

While some types of Oa advantages have salvage value if the firm were to be liquidated, any premium associated with its other Oa and Ot advantages would be likely to disappear. However, if the firm were acquired by another firm, part of the premium paid would be precisely for these types of ownership advantages, with the expectation and hope that subsequent to the merger, the merged entity would be able to maintain similar or even higher levels of Oa and Ot advantages. That such advantages do not always materialize reflects the fact that the disembodied value of such advantages is difficult to assess, and therefore the premium paid by any acquirer may overstate their usefulness when deployed in a different context. This is particularly the case if the acquired firm cannot be left intact, but is subject to


rationalization and cost cutting, which are often used to gain the desired management transformation.

Greater reliance on transactions that extend beyond hierarchical control poses a governance challenge. Such transactions, whether they involve OEM production or design contracts, or various types of outsourcing arrangements, incorporate a variety of ways to balance the interests on both sides. These may include sequential equity purchases, mutual hostages, and various other means to align the incentives of the contracting parties. It is likely that some firms develop superior skills at managing such relationships that provide them with an advantage, but such advantages are not the traditional multiplant economies envisaged by the Ot advantages. Similar challenges are posed by various kinds of nonmarket activities that may extend from the simple philanthropic activities undertaken by TNCs to multistakeholder partnerships involving governments and NGOs.

To address the capabilities required to manage such relationships, Dunning and Lundan introduced a third category of ownership advantages, namely institutional advantages (Oi). They argued that the ability of the TNC to grow requires effective management of both the market and nonmarket domains, and it necessitates the simultaneous deployment of asset-based advantages, advantages of common governance, and institutional advantages.

Oi advantages include the formal and informal institutions that govern the value-added processes within the firm and between the firm and its stakeholders. This might include instruments such as codes of conduct, norms, and corporate culture, and such advantages might also be incorporated into the incentive system of the firm. In this sense, some of the Oi advantages resemble intangible Oa advantages, but they relate specifically to the expectations and limits set by the institutional framework in the home and host countries. Because of this, the cases where Oi advantages are most visible are in connection with the exploitation and augmentation of the firm's ownership advantages using the nonmarket interface. Since increasing the geographical scope of the firm also increases the institutional complexity with which it is confronted, Dunning and Lundan argue that Oi advantages are in large


part responsible for the limited ability of many TNCs to expand beyond their regional market. We now turn to the evolving interface between TNCs and external institutions.

II. ADAPTATION AND COEVOLUTION

According to North, economic growth in the long run depends on the development of more complex institutions to deal with the uncertainties that arise from more complex forms of exchange, involving both market and nonmarket actors. These institutional aspects have also become increasingly interconnected over geographical space, and TNCs both contribute to, and are affected by, the contagion effects from the external environment that are now capable of being transmitted more rapidly and effectively from one location to another.

TNCs have responded to the more profound nature of uncertainty in part by shifting toward more open business network structures that provide greater flexibility in adapting to changes. In so doing, TNCs have developed solutions for dealing with the problems of impersonal exchange across multiple markets and the acquisition and recombination of dispersed knowledge. The methods of organizing and structuring relationships within and between firms, including in the conduct of R&D, the design and enforcement of contracts, and the management of cooperative relationships, evolve into routines or best practices that are gradually diffused to other firms and become institutionalized. Thus, while the institutional environment is external to any individual firm, the process through which new institutions are created is often initiated through the experimental actions of individual firms. Since firms (and especially TNCs) are increasingly interconnected with other market and nonmarket actors, a variety of alternative experimental directions across different network

31. See generally DOUGLASS C. NORTH, INSTITUTIONS, INSTITUTIONAL CHANGE AND ECONOMIC PERFORMANCE (1990) (explaining an analytical framework describing drivers of institutional change and the implications of this type of institutional development for economic theory and economic history); DOUGLASS C. NORTH, UNDERSTANDING THE PROCESS OF ECONOMIC CHANGE (2005) (developing and extending the argument that institutions, and particularly informal institutions, are important to achieving economic growth).
actors increases the system-wide extent of institutional experimentation and diversity, as well as generating novelty from within this complex system.\textsuperscript{34}

In a recent paper, Cantwell, Dunning, and Lundan identify three types of engagement between TNC and institutions.\textsuperscript{35} The first type is institutional avoidance, in which TNCs take the external institutional environment as a given, but in which they are able to make choices between different institutional environments. Faced with a weak institutional environment, characterized by a lack of accountability, political instability, poor regulation, and deficient enforcement of the rule of law, the response of most TNCs is likely to be characterized by an “exit” rather than a “voice” strategy.\textsuperscript{36} However, exceptions may include natural resource-seeking investment and some forms of infrastructure investment, where the number of alternative investment locations is more limited. The prevalence of this type of behavior is evident in the results of many studies confirming that the more footloose forms of TNC activity are mostly concentrated in countries characterized by good governance.\textsuperscript{37}

The second form of engagement is institutional adaptation. As in the previous case, the TNC treats the institutional environment as essentially exogenous, but in this case it adjusts its own structure and policies to better fit the environment. The means to achieve this objective include the use of political influence and, in some cases, bribery, but it may also involve efforts by the TNC to intentionally emulate the behavior, commercial culture, and institutional artifacts that are most desirable in the host country context. At the extreme, the TNC may wish to “go native”—to become an insider in the host country market, possibly even hiding the aspects that make it appear foreign.

In contrast to the first two cases, in the third case, the institutional environment is assumed to be partly endogenous, and the TNC is engaged in a process of coevolution. In institutional coevolution, while firms may employ some of the same tactics they used under the adaptive scenario, their objective is no longer simply to adjust, but to effect change in the local formal and informal institutions. For example, a TNC might engage in political activities to advance specific kinds of regulation or market structure that give it an advantage over its

\textsuperscript{35} Id. at 574-78.
\textsuperscript{36} See ALBERT O. HIRSCHMAN, \textit{EXIT, VOICE AND LOYALTY} 3-4 (1970).
\textsuperscript{37} See DUNNING & LUNDAN, supra note 29, at 308-09.
competitors. In doing so, the TNC might also align itself with domestic firms in lobbying the government for economic protection or support.

While it could be argued that this last form of institutional response represents merely a difference in degree rather than in kind when compared to the adaptive scenario, there are a variety of forms of coevolution that may be observed, each involving a different balance between adaptation and coevolution. These include the introduction of new organizational routines and best practices that the affiliates of TNCs have either developed locally or that are transferred to them from elsewhere within the TNC network.\(^{38}\) They also embrace the transmission of home country institutional practices that are adopted by the TNC parent and transferred within the TNC network.\(^ {39}\) Coevolution may also involve activities in which the TNC engages to effect institutional change at the supranational level. These include various efforts to influence the standards being promoted or enforced by existing supranational bodies,\(^ {40}\) as well as proactive institutional entrepreneurship to self-regulate specific aspects of corporate behavior.\(^ {41}\)

Effective management of the firm's nonmarket environment is likely to require both adaptation and active agency on the part of the TNC. For example, NGOs and firms can become partners in creating and legitimatizing new standards.\(^ {42}\) Indeed, in the area of social responsibility, TNCs often prefer to preempt future regulation by proposing to regulate themselves by creating codes of conduct, or being active in endorsing other standards like the U.N. Global Compact or the Dow Jones Sustainability Index. In other cases, new institutions gain prominence and legitimacy through the process of imitation (isomorphism). For example, in the legal area, the concentration of large clients to a handful of New York and London law firms can be seen as a

---

38. See generally THE EMERGENCE AND IMPACT OF MNC CENTRES OF EXCELLENCE (Ulf Holm & Torben Pedersen eds., 2000) (presenting the results of a multi-country study on the emergence of subsidiary centers of excellence, which are subsidiaries that develop distinct competencies that are advantageous over competitors and are often subsequently adopted in other units of the TNC).


42. Teegan et al., supra note 12, at 473-76.
risk minimization response by firms in cases that involve very large financial stakes. Similarly, the spread of U.S.-style commercial arbitration has been argued to be an outcome of the use of these practices by TNCs.\textsuperscript{43}

These three forms of engagement are not mutually exclusive, and TNCs are likely to exhibit both adaptation and coevolution with institutions in different home and host countries, in different industrial sectors, and at different points in time. One would expect, however, that adaptation would be more likely to occur in relatively stable environments and in less innovative sectors, while more dynamic environments are likely to both allow and require innovation and a continuous coevolution for the firm to sustain a competitive advantage.

In a comparatively settled local environment, it is likely that, conforming to the expectations of the neo-institutionalist approach, subsidiaries would mainly adapt to their local surroundings. The variety the subsidiaries introduce is likely to be accepted by other local actors, but it is unlikely to have much impact on their behavior. By contrast, in a faster moving environment, perhaps associated with a process of political and institutional reform, the coevolution of TNC subsidiaries and the institutional environment is more likely, and the variety they introduce to the local environment is more likely to influence (as well as be influenced by) the behavior of other agents.

In particular, an institutional system that is in flux is likely to exhibit “institutional voids” that offer opportunities for institutional entrepreneurship and coevolution. Consequently, in emerging markets, TNCs might be welcomed in part because they introduce institutional elements that are missing in the local environment.\textsuperscript{44} By the same token, in such dynamic environments, TNCs are more likely to be seen as legitimate if they contribute to a transformational process that is already ongoing. New practices introduced by TNCs can become a part of the wider process of changing values and institutional structures in the host country, but even in such cases, institutional entrepreneurship by TNCs is likely to contribute only one element to the broader restructuring underway. The next section will explore the process of coevolution and the diffusion of new routines in more detail.

\textsuperscript{43} Cf. Olaf Dilling et al., \textit{Introduction: Private Accountability in a Globalising World}, in \textit{RESPONSIBLE BUSINESS: SELF-GOVERNANCE AND LAW IN TRANSGNATIONAL ECONOMIC TRANSACTIONS} 2, 2-5 (Olaf Dilling et al. eds., 2008).

III. THE DIFFUSION OF NEW ROUTINES

A. Antecedents to Diffusion

Seen from an evolutionary perspective, all institutional systems represent adaptations to specific circumstances, with some features more deeply rooted than others and no expectation that the system is in any sense optimal. When new practices become adopted by local firms and embedded in the institutional fabric of particular countries, it is unlikely that such practices will remain unchanged. The ease with which indigenous firms in the countries hosting the affiliates of TNCs are able to absorb new practices and to fit them to the local context is analogous to the concept of absorptive capacity, which has been found to be an important determinant of technology creation and diffusion. As in the case of technology transfer, spillovers are likely to arise as a result of TNCs’ deliberate transfer of institutional knowledge. Such spillovers occur through demonstration and learning effects, the transfer of employees, and “leakage” from the training provided by TNCs to bring local suppliers or distributors in line with their global practices. The market structure in the host country has an impact on local firms’ interest in imitating new practices, but overall, such spillovers are an unavoidable consequence of the efforts undertaken by TNCs to make best practices transferable within the firm.

The centrality of the agency of TNCs to the cross-border transfer of technology is itself partly attributable to the need for the implementation of technology to be supported by the development of complementary institutions. Having already developed relevant organizational routines and the tacit knowledge that has become embedded in those routines, it is often easier for the TNC to extend these practices into the host country environment than it would be for an independent agent operating under license. By adapting its technology and associated routines of behavior for the local environment, the TNC also extends its own range of competence and hence its overall innovation potential.

At the same time, it should be acknowledged that, due to time and resource constraints, firms also frequently respond to uncertainties in an ad hoc manner. Engaging in institutional entrepreneurship incurs

substantial costs that are nonrecoverable and that are likely to result in path-dependent lock-in effects for the firm. While TNCs can learn to transcend the "imprinting" of their home country, due to the path dependence of their investment and search strategies, they are limited in their ability to adjust to the institutional systems of host countries. Thus, while the cross-border differences in institutions offer potential opportunities for TNCs to choose the most conducive environment for their activities, the legacy of home country institutions may result in firms preferring a similar (rather than superior) institutional context. Indeed, the findings of Rugman and Verbeke that some of the world's largest TNCs are regional rather than global in terms of their sales are in large part due to institutional differences that are a combination of formal market institutions and the societal norms and expectations that underpin those institutions. In the words of Kogut, "direct investment is the extension of the organizing principles and capabilities of the firm across countries." Without detailed knowledge of the market, and of the quality and quantity of information available to other firms and customers throughout the value chain, it is difficult to assess whether the business model the firm has adopted in its home country is transferable abroad. For example, Jacobides has argued that the failure of U.S. financial institutions (prior to the current crisis) to expand abroad was due to a lack of understanding of the extent to which their business models relied on a particular division of labor and the presence of specialized intermediaries within the domestic market. Furthermore, successful transfer presumes that a TNC has the ability to generate a sufficient variety of practices, to select those that are the most viable, and to codify them to the extent that they become transferable within the firm. Even if this is achieved, as a consequence of the increased autonomy gained by subsidiaries that evolve toward a higher share of competence-creating activities, problems of motivation and control may impede subsequent transfers of knowledge within the firm across disparate or potentially competing subunits.

49. See Rugman & Verbeke, supra note 30, at 14.
B. Effects on Market Structure

In fact, the contribution made by TNCs here is a dual one. The primary contribution of a TNC is its ability to assemble complementary assets when the input markets are thin. By internalizing transactions in cases where markets are thin or nonexistent due to cospecialized assets or information asymmetries, the TNC makes it possible for certain forms of economic activity to take place.\textsuperscript{52} However, in other cases, through its own activities, the TNC may make the markets thicker by transferring knowledge and providing training to its contractual partners. Indeed, there is a long history of TNCs providing such training to their suppliers, the impact of which has at times been quite significant, particularly in developing countries, as new technologies and organizational methods have become diffused into the local economy.\textsuperscript{53}

Contract research is a case in point. Traditionally, TNCs carried out R&D within the auspices of the firm and such activity was geographically concentrated in the home country.\textsuperscript{54} However, in recent years we have seen TNCs extending the range of the R&D activities that are conducted abroad and also engaging in market-based exchanges for contract research. Such markets do not come into existence automatically, but instead require a process whereby potential market participants create the standards that allow for successful transactions to take place.

In some cases the process of codification and subsequent training may help contribute to the emergence of modularity in the market for intermediate goods, which in turn allows for new entrants to enter the marketplace.\textsuperscript{55} Here we can think of technological modularity, as seen for instance in the emergence of standards such as the Global System for Mobile Communications that allowed for the explosive growth of mobile phone services.\textsuperscript{56} However, we might also think in terms of organizational modularity, i.e., the development of routines for


\textsuperscript{53} See DUNNING & LUNDA, supra note 29, at 444-50, 559, 573-77.


\textsuperscript{55} Teece, supra note 52, at 134-36.

organizing transactions, that become sufficiently clear and codified so that they too can become diffused in the marketplace and influence the prevailing standards of how transactions are conducted. Furthermore, the network structure of the TNC, which is a combination of transactions that are internal and external to the firm, might be one way the TNC can overcome the limits set by the availability of managerial resources and capabilities on the growth of firms, or the so-called “Penrose effect.”

Firms shape markets rather than the other way around, because firms introduce new routines, and because of the role they play in redefining the rules of selection in their operating environment. The need to invest in cospecialized and complementary assets together with the ability of the firm to cocreate the market makes the prospect of value creation more credible within the auspices of the firm than purely over the market.57 Indeed, it is organizational innovation and the development of new business models (rather than product or process innovation) that lie at the core of the development of so-called “dynamic capabilities.”

The combination of increasing interconnections between geographically dispersed markets and the greater use of market-based transactions has intensified the incentives to develop new routines. By virtue of the nature of their business, TNCs are leading the way in the development of new capabilities. In line with Jacobides and Winter,58 I suggest that as the transactions coordinated by TNCs come to involve more market-based than hierarchical coordination, this is likely to increase the transparency of the sources of firm capabilities within industry sectors. This is because in industry sectors where market-based transactions dominate, firms are likely to gain easier access to the innovations that have allowed other firms to gain lower transaction costs or to provide higher value. This process is likely to be self-sustaining, leading to more specialization and less internalization, as the innovations made by the pioneering firms allow others to reduce the costs of market transactions.

Conversely, in markets where integrated hierarchical solutions are dominant, such as in the period of the emergence of the large industrial firm described by Chandler,59 the opposite is the case. Since the pioneering firms have no incentive to increase the transparency of

59. See CHANDLER, supra note 26, at 8, 14.
competitive practices, organizational innovations remain hidden inside integrated firms. Consequently, other firms are likely to find it more difficult to identify and imitate such practices. This process is also likely to be self-sustaining, since little innovation can take place between firms, and any advances remain largely hidden inside integrated firms.

While the degree of internalization at the firm level is in part influenced by factors such as experience and reputation that predispose a particular firm to choose (or to persist with) particular forms of governance, the greater use of market-based transactions is also dependent on the existence and availability of suitable partners. It is precisely by identifying possible partners, and by transferring the relevant routines to carry out new types of transactions, that firms engage in institutional entrepreneurship and influence industry architectures.

Of course, not all industries are equally amenable to such restructuring. The incumbents have vested interests in maintaining the current architecture, and consequently the costs of changing industry structure, whether by attempting to combine highly specialized players or by inducing independent firms to take over parts of the value chain, can be substantial. For an entrenched industry, exogenous technological changes or major changes in antitrust regulation are likely to be required to force the industry structure to change from being dominated by integrated firms to one where market transactions are prevalent.

IV. DIFFERENCES IN INSTITUTIONAL ENVIRONMENTS: THE BALANCE BETWEEN PUBLIC AND PRIVATE ORDERING

As a result of the transfer of routines and the process of coevolution in market and nonmarket domains, TNCs shape the institutional environment in their home and host countries and contribute to mixing the boundaries between what has traditionally been the role of the state (the public domain) and what has belonged to the private domain. Since economic markets are not self-regulating but depend on nonmarket institutions for a stable economic structure and the

60. See John Hagedoorn & Geert Duysters, External Sources of Innovative Capabilities: The Preference for Strategic Alliances or Mergers and Acquisitions, 39 J. MGMT. STUD. 167, 172-73, 180-82 (2002).
62. See Cantwell, Dunning & Lundan, supra note 34, at 577.
THE COEVOLUTION OF TNCs AND INSTITUTIONS

The provision of a variety of public goods, the failures of nonmarket institutions also matter for the operation and performance of TNCs.

Thus, the transaction costs faced by firms are likely to be higher when property rights are not specified and protected by the polity, social disunity within the community hampers the recruitment and collaboration of personnel on account of ethnic and other social divisions, or cultural values discourage people from joining and participating in wealth-creating activities. While all countries are likely to suffer from some types of institutional failure, they are particularly evident in emerging markets, marked by what Khanna, Palepu, and Sinha called "institutional voids," and what Palazzo and Scherer named "deficits.

As with failing markets, TNCs will evaluate different governance types pertaining to the nonmarket and will consider the possibility of performing "in house" some nonmarket institutional functions related to public law and order, social cohesion, and moral engagement. The nonmarket institutional goods and services associated with these functions are akin to the "complementary assets" highlighted by Teece because, like economic intermediate products, they are necessary for the production and marketing of final products.

Legal theory distinguishes between "public" and "private" orderings to describe the two principal mechanisms employed to mitigate market and nonmarket failures. Public ordering rests on laws.

63. Public goods are defined here as those provided by nonmarket institutions—particularly, in terms of law and order, social cohesion and cultural engagement in production and innovation. In principle, they are nonrivalrous and nonexcludable—that is, everyone can use them without reducing their availability to others. However, it should be noted that many privatized public services do not meet this definition. This is the case, for example, with telecommunication networks, transport infrastructure, and power generation, which are neither nonexcludable nor nonrivalrous in use although the government as an owner may choose not to exclude specific users.

64. See generally Hoskisson et al., supra note 44 (discussing the challenges faced by multinational enterprises investing in emerging markets).


68. See, e.g., Larry Catá Backer, Multinational Corporations as Objects and Sources of Transnational Regulation, 14 ILSA J. INT'L & COMP. L. 499, 505 (2009).

69. The nonmarket concept also includes nongovernmental (i.e., civil society) failures. For a discussion of the development of the nonmarket concept through four perspectives, see Jean J. Boddewyn, Understanding and Advancing the Concept of "Nonmarket," 42 BUS. & SOC'Y 297 (2003).
and regulations backed by a legal system of civil enforcement and criminal penalties. Internalization theory and Williamson's "institutional environments" assume such public ordering, which is predominant in most developed countries. The primary mean used by TNCs for obtaining nonmarket intermediate products under public ordering is financial (e.g., through the payment of taxes, fees, bribes, and donations) although political influence (e.g., the voting out of unfavorable politicians and the capture of regulators), trust, status, reputation, and legitimacy are also employed for that purpose, since exchanges are often "embedded" in social relations.

Private ordering is more likely to prevail in markets where the legal institutions that underlay public ordering are largely missing or deficient. While public ordering is a "top-down" process that rests on the threat of the use of sovereign force, private ordering is a consensual "bottom-up" practice that derives its efficacy and legitimacy from the continued support of its members. A private ordering system may allow for exchanges to take place even when pricing in the market is difficult, property rights are unclear or insecure, and the pursuit of self-interest is insufficient to guarantee orderly transactions free of malfausance and opportunism. In embedded relationships, cooperation is common among the exchanging parties who conform to shared rules, norms, and customs, thereby developing a "socialized rationality" and having the moral obligation to consider the appropriateness of their actions.

The basis of private ordering is contractual, and contract management and dispute settlement are dealt with directly by the parties involved or through mutually agreed-upon arbitration schemes. In lieu of relying on well-defined property rights and contract laws, exchanges may be embedded in a broad set of relationships with the same party, so that nonperformance in one exchange could be penalized in later ones. Reserves of social capital, performance bonds, and third-party guarantors can help solve contracting problems stemming from low levels of cohesion and solidarity, with trust playing an important role.

71. See generally Mark Granovetter, Economic Action and Social Structure: The Problem of Embeddedness, 91 AM. J. SOCIOl. 481 (1985) (presenting a seminal discussion on the importance of embedded relationships to the execution of economic transactions).
73. See id. at 1783-84.
Private ordering is used in emerging markets where exclusionary groupings of local companies, networks based on kinship ties, and even membership in criminal organizations predominate in order to compensate for the failures of public ordering.\textsuperscript{74} Private ordering may also exist in parallel with a strong government that exercises coercive force through state ministries and agencies that reward or punish particular firms through permits, purchases, credit, and protection, and that tend to favor state enterprises.\textsuperscript{75}

In terms of transaction costs, a rule-based governance system of the type found in developed countries involves large costs for drafting, interpreting, and implementing contract laws, but the marginal cost of enforcing an additional contract is negligible because the rules are explicit, impersonal, and standardized. By contrast, transacting in emerging markets is likely to involve fewer fixed costs but significant marginal ones because one needs to screen, test, and monitor every transaction partner. Thus, as the market expands from local to regional, national, and international levels, the number of business partners increases, and the marginal cost of relations could rise significantly.

Private ordering may be preferred even in societies where public ordering is of high quality—as Williamson acknowledged\textsuperscript{76}—because it is more flexible and faster in adjusting to changing circumstances. Thus, it has been used in connection with the emergence of electronic cross-border markets\textsuperscript{77} and with global challenges such as climate change, and its practice is also illustrated by the increasing popularity of private arbitration agreements among firms. In such cases, private ordering works in the shadow of a well-functioning public ordering system that provides a degree of continuity and reliability.

The extent to which firms might be willing to engage in some form of nonmarket internalization is influenced by changing economic and political ideologies. Social scientists have charted the changing balance between the public and private domains, following the transformation of

\textsuperscript{74} See, e.g., Tarun Khanna & Yishay Yafeh, Business Groups in Emerging Markets: Paragons or Parasites?, 45 J. ECON. LITERATURE 331, 339-41 (2007); Klaus E. Meyer & Mike W. Peng, Probing Theoretically into Central and Eastern Europe: Transactions, Resources, and Institutions, 36 J. INT'L BUS. STUD. 600, 612 (2005).

\textsuperscript{75} See generally Mike W. Peng, Institutional Transitions and Strategic Choices, 28 ACAD. MGMT. REV. 275 (2003) (discussing what firms can expect to encounter in terms of institutional transition as authoritarian governments transfer to more market-based systems).


the modern state in postnational times. This transformation has taken place along two main axes: one concerning a shift in the key actors from public to private (e.g., from governments to NGOs and TNCs) and the other a change in the level of governance from the national to the supranational or regional level. As a consequence, many nonmarket services are now produced or coproduced by combinations of private firms and NGOs.

For example, in emerging markets like Jordan, the workplace rules in the supplier agreements of Wal-Mart effectively constitute a private system of regulation, while in the context of Germany, the same rules were heavily contested and the private ordering instituted by Wal-Mart was overturned under the public ordering of German law, which considered Wal-Mart's policies on workplace fraternization to be overly intrusive. At the supranational level, TNCs have been increasingly involved in the development and drafting of international treaties, multilateral agreements, and codes of conduct.

In general, one would expect that in well-ordered societies, the complementary assets held by local owners are likely to be easier to transact, whereas the opposite would be the case in burdened societies. One would also anticipate that when institutional distance is low, the knowledge and other assets held by the TNC are easier to transact, while the opposite would be the case when institutional distance is high. What matters then in determining the governance form are the relative costs that the TNC would incur in acquiring complementary assets and the costs that the holders of complementary assets would bear to access the knowledge and other assets of the TNC.

In the case of foreign direct investment between developed countries, where institutional environments are fairly similar and public ordering predominates, investors rely on the delivery of public goods by the key host institutions, and they expect that nonmarket intermediate products are readily obtainable—mainly through the use of money to pay taxes, make donations, and finance other transactions at low costs. In such circumstances, TNCs have many credible governance options and internalizing nonmarket activities is only moderately necessary—if at all.

78. See generally Transformations of the State? (Stephan Leibfried & Michael Zürn eds., 2005) (presenting a framework for analyzing state transformation in terms of the devolution of responsibilities both geographically and in terms of the level of governance).
80. Backer, supra note 72, at 1762-73.
81. E.g., Ramamurti, supra note 40, at 348-49.
82. See Dean Xu & Oded Shenkar, Institutional Distance and the Multinational Enterprise, 27 Acad. MGMT. Rev. 608, 610 (2002).
THE COEVOLUTION OF TNCs AND INSTITUTIONS

In the case of investment between two emerging markets, the institutional voids in the host country require expertise of operating under private and state orderings—an ability that investing TNCs from other emerging markets already have. Since the foreign investor might not be able to contract for the public goods it needs, the TNC must be able to transfer to a host country at a low cost the knowledge developed at home, and only limited internalization may be necessary. The relatively robust flows of investment into emerging markets by TNCs from other emerging countries testify to the feasibility of this pattern.

The two patterns that involve investment from developed to developing countries or vice versa are likely to prove more difficult. For investment originating in developing countries, the problem is the lack of experience required to establish effective nonmarket relationships in developed countries. However, to the extent that the investor possesses the financial resources necessary to engage in acquisitions, the latter provide an indirect way of accessing the relevant nonmarket resources in the host country, although this may be met with political opposition from the host country.

In the case of investment by established TNCs in emerging markets, the problem for the investor is again the low applicability of home experience when dealing with institutional voids in host countries. The prevalence of private and public orderings pose notable problems for outsiders, but to the extent feasible, acquisitions and joint ventures in the host country can be used to negotiate the nonmarket institutional environment. Furthermore, TNCs are likely to have broader capabilities and resources to support counterparties and nonmarket exchange mechanisms than the other way around. Only if attempts to forge such cooperation fail are TNCs likely to consider stepping into the nonmarket themselves, by either procuring the necessary assets or by taking over the providers of those assets.

V. THE WAY FORWARD: SOME SUGGESTIONS FOR FURTHER RESEARCH

While internalization theory provides a compelling narrative to answer the question of how firms (or TNCs) arise, the predicted

83. See Alvaro Cuervo-Cazurra & Mehmet Gene, Transforming Disadvantages into Advantages: Developing-Country MNEs in the Least Developed Countries, 39 J. INT'L BUS. STUD. 957, 975-76 (2008).


85. For example, this has been the case with respect to Chinese and Russian investment in the United States and Europe.
outcomes become less compelling as the firm matures. This is because in any particular market, full internalization might be the only conceivable alternative for some firms, while others may find it possible to access the same resources by means of partial equity purchases, contracts, or a combination of these. While in some cases the characteristics of what is being transacted are such that no firm would find it feasible to transact over the market (e.g., due to information asymmetry), the growing variety of different types of contractual market interfaces, and the multiplicity of different types of markets where TNCs transact, suggest that the TNCs themselves may play a role in generating new governance alternatives.

The evolutionary process whereby firms adjust their structure to the uncertainties confronting their value-adding activities, particularly across borders, is likely to lead to the emergence of new types of firms, including those that represent hybrids of some elements of public and private governance. As a consequence, there is a need for a better understanding of the factors that cause the balance between public and private governance to shift over time and that enable or encourage TNCs to internalize more of the kinds of functions normally assumed by nonmarket actors. While this is most evident in the cases where TNCs actually engage in the coproduction of public services such as safety, health, or education, the embedded contractual relations TNCs have formed with their suppliers and customers can also impact indirectly on the delivery of services in the nonmarket domain.

Furthermore, such contractual relationships can be unequal in terms of the balance of power, and various types of costs can be shifted to the weaker party (for example, this may happen in connection with contract farmers, suppliers in the production networks of electronics, or textile firms). The use of extensive contracting to govern the value-adding activities of the firm raises questions about accountability and transparency, which at least to some extent can be mitigated by the engagement of an impartial third party verification system. While large TNCs may not be beholden to territorially circumscribed law, they are beholden to their shareholders and customers, whose threat to withdraw support provides an incentive for firms to adjust their behavior in order to maintain legitimacy, although the effectiveness of this tool varies across industries.

From a policy perspective, there is thus a need to consider a much broader political role of TNCs than is typically envisaged in the literature, which has tended to view the relationship between governments and TNCs as bilateral and adversarial, thereby reflecting
the “obsolescing bargains” struck several decades ago. Here, TNC participation in various kinds of public-private partnerships has the potential to bring about more inclusive forms of governance to the investor-state relationship. Finally, the rise of different forms of contractual relationships should prompt scholars from both the legal and economic traditions to make greater use of contracts as sources of data concerning the evolution of governance forms. The practical difficulties of obtaining access to such sources aside, empirical research using contracts as a means to chart the evolving responses of firms to the uncertainties arising from their expanded social and political role could yield new insights both in terms of firm strategies and the formulation of public policy.
