R

Radical and Incremental Technical Change

Mary J. Benner
University of Minnesota, Carlson School of Management, Minneapolis, MN, USA

Abstract

Inspired by Schumpeter, researchers studying the influences of technical change on industries and firms have sought to classify types of technical change according to how dramatic the shift is in the technology underlying products in an industry. Such technical changes have been commonly classified as either radical or incremental. Radical technical change is a discontinuous shift in the base of scientific or technical knowledge underlying the products in an industry or product class, whereas incremental technical change is continuous refinement along an existing technological trajectory.

Definition Radical technical change is a discontinuous shift in industry technology to a new base of scientific or technical knowledge underlying the products in an industry, offering an improved price/performance frontier. Incremental technical change is the continuous refinement or elaboration of an existing technology.

Types of Technical Change

Initially spurred by Schumpeter's work (1934, 1942), the topic of ▶ technological change and its effects on industries and firms has gained prominence in research in the past several decades. A central idea in such work, with clear parallels to Thomas Kuhn's work on scientific paradigms, is that while technological change typically evolves 'normally' over time, focused on elaboration and refinement of an existing technology, technologies are periodically replaced by entirely new technological paradigms based on a new foundation of scientific or technical knowledge underlying the products in an industry (Dosi 1982). The shift to the new technological paradigm offers the promise of an improved performance/price frontier, and initiates a new trajectory of unfolding technological improvement. The continuous small changes along an existing technological trajectory have been characterized as incremental technical change, while the dramatic discontinuous shifts to a new technological paradigm have been considered radical technical change (e.g., Tushman and Anderson 1986; Henderson and Clark 1990; Henderson 1993; Green et al. 1995; Gatignon et al. 2002). For example, the shift from mechanical escapement technology to quartz technology in watches in the 1970s represents a radical technical change (Landes 1983), while the continuous small improvements in the precision of watches based on mechanical escapement technology over several decades

M. Augier, D. J. Teece (eds.), *The Palgrave Encyclopedia of Strategic Management*,
https://doi.org/10.1057/978-1-137-00772-8

illustrates incremental technical change. From the outset, quartz technology offered dramatically improved performance (precision) in watches at a lower cost. This promised improvement in the price/performance frontier was fulfilled as ongoing improvements unfolded in process ▶ innovation and production efficiencies in the quartz watch trajectory. Other examples of radical technical change include the shift from silver halide film to digital technology in photography, or the replacement of the entire ice harvesting and distribution industry by the advent of mechanical refrigeration (Utterback 1994).

The degree to which technical change is considered radical or incremental is from the perspective of a particular industry or product class. That is, a new technology may trigger radical technical change and fundamentally shift the price/performance frontier in a particular industry, yet the technology itself is not necessarily a new-to-the-world technology or a radical technical change for the products in other industries. For example, before its advent in photography, digital technology was already the underlying knowledge base for products in many other industries. Thus, at the time of the technological discontinuity, digital technology represented a radical technical change for the products in the photography industry, but may have been undergoing incremental technical change in other industries.

Technology Cycles and S-Curves

The classification of technical change in industries as radical or incremental is closely related to research that depicts technology change unfolding in patterns or cycles. A new technology cycle is triggered by a discontinuous shift or radical change in technology, indicated by the commercialization of the first product in the industry that is based on an entirely new foundation of technical knowledge. This phase ushers in an 'era of ferment' (Anderson and Tushman 1990) or a 'fluid' period (Utterback 1994), entailing rapid innovation and competition between multiple variants of products incorporating the new technology as well as products based on the old technology. This phase is further characterized by very high uncertainty about technological progress and the nature of demand for the product, as well as an increase in entry of new competitors. The period of ferment is generally followed by the emergence of a dominant product design, determined by a combination of technical, market, social and political forces (Anderson and Tushman 1990). In cases where one or a few dominant technological designs emerge, there is a shake-out, resulting in a small number of winning firms that possess the knowledge and skills to compete successfully in the new technical domain (Utterback 1994). With the emergence of a dominant design, the focus of technological change shifts to incremental refinement along the new technological trajectory, and increasing process innovation and production efficiencies (Abernathy and Utterback 1978).

These patterns in technological evolution and substitution have also been depicted as a series of S-shaped curves (Foster 1986), where effort and investment progress along an existing technological trajectory, result in ongoing improvement in the performance of the technology. The improvement in performance of a technology is steep early on, but then levels off (suggesting an 'S' shape) until the existing technology is replaced by a new technology with higher potential performance characterized by a new S-curve. The performance improvement offered by the new technology is so significant that no increase in scale, efficiency, or design can make older technologies competitive with the new technology (Tushman and Anderson 1986).

Radical and Incremental Technical Change and Organizational Innovation

Scholars in technical change and organization theory have further linked characteristics and patterns of industry technical change to implications for the appropriate types of technical innovation within firms (Ettlie et al. 1984). While incremental technical change builds upon or reinforces the accumulated knowledge and capabilities of firms associated with the existing technology, radical

technological changes, particularly to the extent they are 'competence destroying' (Tushman and Anderson 1986), can render a firm's prior accumulated knowledge and capabilities obsolete, and necessitate dramatic changes in organizational knowledge and competencies (Dosi 1982). These ideas have led to characterizing an organization's innovation as radical or incremental, echoing the classification of technical change at industry level. Incremental innovation within firms is likely to be appropriate in industries characterized by incremental change and elaboration of the existing technology, whereas radical innovation within a firm may be necessary to initiate or respond to discontinuous technical change. More recently, scholars have built on the work by March (1991) and March and Levinthal (1993) to characterize an organization's innovation in related ways, as either exploitation, building on the existing knowledge and competencies in a firm, or exploration, departing from the firm's current knowledge toward new technological or customer domains (Benner and Tushman 2003). A growing body of related research has examined the challenges for incumbent firms faced with threatening technical changes and the various forces that influence their responses (Reinganum 1983; Henderson and Clark 1990; Cooper and Smith 1992; Christensen and Bower 1996; Tripsas 1997a, b; Tripsas and Gavetti 2000; Gilbert 2005; Benner 2007, 2010).

See Also

- ▶ Information Technology and Strategy
- ▶ Innovation
- ▶ Technological Change

References

Abernathy, W.J., and J.M. Utterback. 1978. Patterns of industrial innovation. *Technology Review* 80: 40–47.

Anderson, P., and M.L. Tushman. 1990. Technological discontinuities and dominant designs: A cyclical model of technological change. *Administrative Science Quarterly* 35: 604–633.

Benner, M.J. 2007. The incumbent discount: Stock market categories and response to radical technological change. *Academy of Management Review* 32: 703–720.

Benner, M.J. 2010. Securities analysts and incumbent response to radical technological change: Evidence from digital photography and internet telephony. *Organization Science* 21: 42–62.

Benner, M.J., and M. Tushman. 2003. Exploitation, exploration and process management: The productivity dilemma revisited. *Academy of Management Review* 28: 238–256.

Christensen, C.M., and J.L. Bower. 1996. Customer power, strategic investment, and the failure of leading firms. *Strategic Management Journal* 17: 197–218.

Cooper, A., and C. Smith. 1992. How established firms respond to threatening technologies. *Academy of Management Executive* 6: 55–70.

Dosi, G. 1982. Technological paradigms and technological trajectories. *Research Policy* 11: 147–162.

Ettlie, J.E., W.P. Bridges, and R.D. O'Keefe. 1984. Organization strategy and structural differences for radical versus incremental innovation. *Management Science* 30: 682–695.

Foster, R. 1986. *Innovation: The attacker's advantage*. New York: Summit Books.

Gatignon, H., M. Tushman, W. Smith, and P. Anderson. 2002. A structural approach to assessing innovation. *Management Science* 48: 1103–1122.

Gilbert, C.G. 2005. Unbundling the structure of inertia: Resource versus routine rigidity. *Academy of Management Journal* 48: 741–763.

Green, S., M. Gavin, and L. Smith. 1995. Assessing a multidimensional measure of radical innovation. *IEEE Transactions on Engineering Management* 42: 203–214.

Henderson, R. 1993. Underinvestment and incompetence as responses to radical innovation. *RAND Journal of Economics* 24: 248–269.

Henderson, R.M., and K.B. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35: 9–30.

Landes, D. 1983. *Revolution in time: Clocks and the making of the modern world*. Cambridge, MA: Harvard University Press.

Schumpeter, J.A. 1934. *The theory of economic development*. Cambridge, MA: Harvard University Press.

Schumpeter, J.A. 1942. *Capitalism, socialism and democracy*. London: Routledge.

Tripsas, M. 1997a. Unraveling the process of creative destruction: Complementary assets and incumbent survival in the typesetter industry. *Strategic Management Journal* 18: 119–142.

Tripsas, M. 1997b. Surviving radical technological change through dynamic capability. *Industrial and Corporate Change* 6: 341–377.

Tushman, M.L., and P. Anderson. 1986. Technological discontinuities and organizational environments. *Administrative Science Quarterly* 31: 439–465.

Utterback, J. 1994. *Mastering the dynamics of innovation*. Boston: Harvard Business School Press.

RAND Corporation

Mie Augier
GSBPP, Naval Postgraduate School, Monterey, CA, USA

Abstract

This piece discusses some of the history and ideas present during the first decades at the RAND Corporation, the think tank which became significant during the post WWII years and became an important institution for the mentoring and development of strategic thinking in the US. It also had an important role for the development of important areas in economics and business school perspectives, such as game theory, evolutionary, behavioral and experimental economics, and others; and it also was a place where several key contributors to the strategic management field worked (including Herbert Simon, Sidney Winter, and Richard Nelson).

Introduction

RAND is an institution which has been significant for several developments relevant to strategy and strategic management, including the developments of the ideas of strategic thinkers (such as Herman Kahn, James Schlesinger and Andrew Marshall); and the institution was a place where several scholars relevant to the strategic management field worked early in their career (such as Richard Nelson and Sidney Winter). Intellectually, the institution also helped important developments such as game theory and evolutionary economics, as well as many others; and in many fields it became known as a key place for the development of fundamental knowledge relevant to understanding of strategic and practical problems. The name RAND was an acronym for Research and Development although some referred to it as "Research and No Development" because the organization didn't produce any hard outputs such as weapons (Kaplan 1983, p. 9).

RAND became known for its early work on multidisciplinary problem framing, mathematical social science and operations research and the application of intelligence and social sciences to new ways of understanding social and strategic problems (in the beginning a focus on problems of national security and later on spanning a wider range of areas of public concern, including health and education). Many of the intellectual developments at RAND became embedded in the changes in business schools at the time and there were significant overlap in terms of people, ideas and other postwar institutions as well. For example, an early Board member at RAND was Rowan Gaither who later went to the Ford Foundation and wrote the Gaither report which led to the Ford Foundations decision to support changes in business schools and management education, including their support of the Gordon-Howell report (Augier and March 2011, chapter 5). Others who were at RAND doing social science (including Hans Speier and Bernard Berelson) also were consultant to Gaither at the Ford Foundation. And a number of scholars whose work would form the 1950s and 1960s core of many of the fundamental contributions in areas of business school research, such as operations research and organization theory and economics, including Kenneth Arrow, Tjalling Koopmans, Duncan Luce, Jacob Marschak, Richard Nelson, Roy Radner, Howard Raiffa, Philip Selznick, Herbert Simon, Vernon Smith, Oliver Williamson and Sidney Winter.

Brief History

The creation of RAND was one result of the post WWII – recognition of the need for a group of scientists working full time on military matters in peace time. Many scientists had worked on military issues at all levels including the atomic bomb at Los Alamos; the MIT Radiation lab; the Office of Scientific Research and Development (headed by Vannevar Bush); the Statistical Research Group at Columbia University, and the Secretary of War's panel of expert consultants from science and industry. After the war, leaders in both

academic and policy making communities felt that collaboration of scientific and military minds was essential to stay ahead in the global competition and to expand knowledge related to future developments in war, peace and national security. A need was for something to be interdisciplinary from the beginning. As one early RAND staff member noted, the need was:

> for a mobilization of all the disciplines – physical sciences, engineering, mathematics whatever you have – to study and give advise to the military of the problems . . . of national defense that was going to come up in the future. (Bornet 1961, p. 2)

The victory in the war had illustrated the success of such wartime collaboration among scientists, the military, academia and industry; not just to win any future wars but also to understand and anticipate conflict and to try and maintain a strategic advantage. This in turn encouraged postwar discussions of ways to allow peace time scientists, who mostly had returned to universities, to continue to work on national security problems. Dwight Eisenhower was among the early enthusiasts. "The armed forces", he wrote in a memorandum from April 30, 1946, on 'Scientific and Technological Resources as Military Assets", "could not have one the war alone. Scientists and businessmen contributed techniques and weapons which enabled us to outwit and overwhelm the enemy" (quoted in Bornet 1961, p. 5).

A peacetime institution was indeed needed to sustain the partnership that had emerged during the war, and RAND was one result of such sentiments. Rather than simply negotiate separate contracts with individual university faculty members, the Scientific Advisory Board to the Army Air Force's chief of staff adopted a 'think tank' model for recruiting expertise and doing research.

At first it was assumed that to have stability and attract top scientists, the efforts should be housed in an industrial research facility. Thus in 1946, the Douglas Aircraft Company was persuaded to sign a contract with the Air Force to provide the facilities. RAND was created as a separate division of the Douglas Aircraft Company to produce long term scientific and technical planning for the Air Force. A special office in the Pentagon, headed by General Curtis LeMay, was established to handle the contract.

LeMay saw RAND as an experimental to explore ways of organizing the nation's scientific capabilities in the service of military and civilian policy making. As he wrote to Arthur Raymond in 1947:

> The Army Air Forces consider RAND one of its most important projects. We are interested not only in the end-results we hope it will achieve, but in the philosophy behind this type of contract. We believe it to be a distinct forward step in correlating the thinking and planning of the Army Air Forces and its scientific civilian partners. We cannot help but feel that if RAND should fail for any reason whatsoever, it would be a severe set-back to the Army Air Forces and to the nation. (quoted in Bornet 1961, p. 8)

In addition to research about winning the next war, a central purpose of RAND was to contribute to avoiding it:

> A central purpose of our external political policy must be to prevent major war. Prevention of war is only one facet of a basic policy of maintaining the defense, liberty and welfare of the U.S. populace. Yet the effectiveness of modern technical weapons has become so great that it is now (or shortly will be) within the power of at least the major nations to render large parts of the globe literally uninhabitable, and profoundly to change the nature of life on the earth, by measures taken to fight another large-scale war. The toll in waste, death ad misery would be tremendous. In a protracted war, it is conceivable that civilization itself would be seriously endangered. (Digby 1990, p. 4)

The initial Army Air Force – Douglas connection wasn't perfect. In particular, there was concern about protecting the independence of the research efforts in a commercial establishment beholden to the Air Force. The founders of RAND found it important that research had freedom to contribute to basic science and civilian sector research, without too much direction from the military. John Williams noted that "it is an unpleasant fact that on a number of occasions in recent years the Air Force, the Department of Defense, or the Executive Office of the President has shown a tendency to gag us". (Williams 1962, p. 17)

The model of setting up an alternative nonprofit organization was therefore explored, and

RAND was set up as an independent non-profit research corporation to help create a better research environment as well as better relations with the outside world and academic scholars (Goldstein 1961). A central person in the transformation was the Chairman of the Board, Rowan Gaither, who had worked at the MIT radiation lab during the war as well as an attorney specializing in the law of non profit corporations. He was an effective organizer and manager with an interest in both pure research and the application of ideas to practice and had met the then President of RAND, Frank Collbohm, in 1943 when they were both at the MIT Radiation Lab. Having also been a legal council to RAND since 1947, Gaither, along with Collbohm, raised funds (a loan from the Ford Foundation) for the reorganization and on November 1, 1948, RAND became a non profit organization (Gaither, although he later moved to work at the Ford Foundation, remained keenly interested in RAND until he died in 1961 Bornet 1961, p. 15).

Early Work at RAND

From the beginning, RAND had unusual independence, central to them being able to do objective research. The high degree of intellectual freedom was important both for their ability to attract academic talent as well as for them doing research focused on understanding and framing the central questions. It also enabled RAND to foster some innovative thinking which might have been otherwise suffocated in established intellectual and institutional bureaucratics, enabled not only by the resources and talents of the people, but also the ideas and visions of key leaders and managers at RAND; the relative small size of the organization; lack of bureaucracy and red tape and an emphasis on ideas; the culture of 'optimistic urgency' present at the time, and even the building design, intended to help facilitate chance encounters (Augier et al. 2015; Williams 1961). As Bob Spect noted with regards to the physical architecture at RAND early on: "As John Williams said in 1950, 'RAND represents an attempt to exploit mixed teams, and to the extent its facility can promote this effort it should do so'. That is, at RAND, much more than at a university, a physicist is apt to encounter the political scientist, the engineer to consort with the economist. This is true – and important – not only in the formal work of an interdisciplinary project team, but also in the many internal contacts, ones the building design should stimulate. An expert in international relations may write a book by himself, but he is a different man and it is a different book because he has been stimulated and educated by encounters with colleagues of many disciplines and varied experience".

The sense of urgency (realizing there were big and urgent problems such as intercontinental nuclear war) and optimism about science (and the ability of science to contribute to such problems) was focused through a problem driven way and confidence in two ay street communication between working on real problems and developing analytical frameworks (Simon 1986). This also inspired a sense of altruism in scholars who could have spent their life in academia but wanted to contribute to understanding of larger national and societal problems (Weaver 1970).

RAND had early visionary leaders (including the President, Frank Collbohn and after a few years also Charlie Hitch, head of the economics department) and founders who built the organization with high intellectual freedom, little or red tape, and focusing on stimulating problem oriented research which engaged scholars from different disciplinary background, using broad ways of thinking rather than narrow immediate practical instruments (in doing so they also created an institution which was flexible enough to later one evolve into something quite different). They also encouraged people to think outside the box, making RAND a place where innovative ideas could be nurtured.

In particular after its reorganization in 1948 as an independent research corporation, RAND's mandate began to expand beyond mere weapons planning for the Air Force, and it quickly became an institution involved in research on decision making and behavior under conditions of uncertainty. As more and more social scientists were hired, RAND scholars pioneered research across a broad range of social sciences, importing techniques from systems analysis, game theory,

and linear programming, and in many cases established the intellectual bases that continue to underpin the state of knowledge in these fields today. Several departments (numerical analysis, logistics, mathematics, economics) were subsequently set up at RAND to accommodate the growing diversity of scientists. However, because of the nature of the problems they were working on, departmental lines were frequently treated as arbitrary. It was a working premise that military problems didn't conform to disciplinary boundaries and didn't often fit a particular academic category very neatly. And often, once projects started, research projects would migrate through several departments, involving men of different skills from many different departments.

John Williams was an early spokesman for developing social science at RAND; a mathematician himself, and aware of the potential difficulties in attracting top talent from academia to leave their positions to come to RAND, he came up with the idea of organizing a conference specifically designed to be for social scientists, in order to attract the best talents in American (academic) intellectual social science. The conference took place in New York in 1947 and included a broad group including political scientists, sociologists, economists, and mathematicians. Thus, the realization that research in social science could have implications directly on national security and welfare was explicitly stated by Warren Weaver (head of the applied mathematics panel of the National Defense Research Committee during the war, board member at RAND), at this conference devoted to recruitment for the RAND Corporation. In particular, Weaver painted a picture of the academic landscape as inadequate to measure up to the challenges of the Cold War world:

> every piece of knowledge we have in sociology and in economics and in political science, everything we know about social psychology, everything we know about propaganda... Every piece of information of that sort, I say, is a weapon... since the last war there has been a change in the character of war, a change in the character of the inevitable amalgamation of all the intellectual and material resources of the country which are necessary to maintain our position in peace and to enable us to defend ourselves... There have also emerged some patters of working together, particularly among the biological, physical and social sciences, which seem to me to have great promise... the whole fields of the social sciences and of the physical sciences must be brought more closely together. (Weaver 194x)

Among the concerns of this conference was to discuss how social science could contribute to issues such as civilian and military policy; the costs of war, psychological warfare, and the economic war potential of USSR and US. And the conference is just one of many ways in which RAND served as the focal point for the development of what came to be known as behavioral science in the 1950s, accompanied by various satellite academic units established at MIT, Stanford, and, especially, Carnegie Tech. Many of the participants at the conference subsequently joined RAND (including Bernard Brodie, Charles Hitch and Herbert Goldhamer); and others became well known in other interdisciplinary movements. Thus, after this conference, and at least in part because of the high degree of intellectual freedom present, RAND did manage to foster Weaver's "working together" of researchers whom disciplinary boundaries would often keep isolated in an academic context. Some of the social scientists there also developed work later important in policy circles, including operational code analysis (George 1969; Leites 1951; Schlesinger 1989).

Some Examples of RAND Work and Their Competitive Advantage

In addition to recognition of the importance of social science, RAND became known for the use of quantitative analysis in analyzing defense related problems. An early emphasis on mathematics and statistics also meshed well with trends in the developments in economics which was becoming more and more based on mathematical statistics at the time. Thus many economic scholars found RAND to be an interesting place to work, including many who would later on win the Nobel Prize in economics (including John Nash, Kenneth Arrow, Vernon Smith, Oliver Williamson, and others).

RAND was important early on also in fostering the early developments in game theory, and von Neuman was part of a project early on, attracted to RAND in part because game theory wasn't widely accepted in economics in the 1940s; and at RAND, many strategists and operations analysts found it intriguing. Both von Neumann and Morgenstern were affiliated with RAND, as well as a host of other game theorists over the next decades (Poundstone 1992). For many researchers at RAND, game theory was a method to improve rational behaviour for people or for nations in competitive situations, such as the Cold War. For instance, in 1949 Olaf Helmer organized a conference on the application of game theory to tactics, with participants including Kenneth Arrow, Leonard Savage, Lloyd Shapley, and others. The conference was held to try and capture the recent progress in game theory. But there were also concerns at RAND of the limitations of game theory. For example, the psychologist Merrill Flood developed a program of experiments which raised doubts about the rationality assumptions underlying game theory (Flood 1958; Poundstone 1992). Other RAND researchers also found game theory overlooked human aspects of competition and decision making.

RAND researchers pioneered research across a broad range of topics relevant to strategic decision making and strategy, using techniques from systems analysis, game theory, and linear programming. In many cases, scholars working or in conjunction with RAND established the intellectual bases that often underpins the state of knowledge in those fields today. In addition, 'systems analysis' became the foundation for much of McNamara's management of the Pentagon (and he hired Charlie Hitch and Alain Enthoven from RAND, among others).

RAND was also an organization which house central ideas in what became known as 'organization theory', in particular that part of studies of organizations that emphasized decision making. Olaf Helmer noted that "organization theory can be viewed as a very natural extension of game theory" (Helmer 1957), and Oskar Morgenstern was involved in a study designed to develop a mathematical theory of organizations, drawing considerably on game theory (Morgenstern 1944).

But RAND researchers recognized that organization studies was more than game theory and rational behavior. In 1951 they hosted a conference on organizations, with participants such as Arrow, Morgenstern, Flood and Allan Newell. Some early attempts also built on Simon's early work and Simon himself was a frequent visitor and consultant to RAND in the 1950s and 1960s. Others involved at RAND relevant to the development of organizations and strategy fields included Sid Winter and Richard Nelson; in fact their collaboration initiated at RAND [see entry on Sid]. Also, an emerging unhappiness with systems analysis (Hitch himself had pointed out some key limitations, Hitch 1955, 1956) and early warnings from Andrew Marshall, Herbert Goldhamer, Herman Kahn, James Schlesinger and others noted that systems analysis (and rational analysis in general) quite often ignores human factors that are essential for understanding real human decision making and behaviour in and outside organizations (Goldhamer 1950; Schlesinger 1967). Some also began exploring early gaming techniques and scenario analysis as more realistic frameworks for trying to understand and anticipate future developments relevant to the development of strategy.

The work on organizations also became the basis of work at RAND to apply organizational analysis to understanding countries, in particular Soviet behaviour. Rooted in conversations between Andrew Marshall and Joseph Loftus (a former Air Force analysis working on intelligence relating to Soviet nuclear programs), they started talking about understanding the Soviet as systems of organization and looked into the (very) early literature on organizations (in particular from the Carnegie school). Marshall also started a set of seminars around organizations resulting in the well known book by Graham Allison on the Cuban Missile Crisis, using different kinds of conceptions of decision making to understand the decision process during the crisis; the conceptual parts building on Marshall's work (Allison 1971; Marshall 1968; Allison and Marshall 1969). Marshall later one further developed this line of

thinking in his work on net assessment [see entry on Marshall].

Marshall also worked with Schlesinger, Sid Winter and Richard Nelson at RAND to develop the organizations theme even further. The intend was a long term project research all the existing contributions to organizational behaviour and to extend and adapt the ones suitable to understanding military organizations. The project did never materialize at RAND but the key people became central to the development of organizational and strategic management ideas (Nelson and Winter) as well as to US strategic thinking (Schlesinger and Marshall).

Closing

RAND as an organization has continued to evolve and grow. Many of the above mention scholars left but have often mentioned the influence of RAND on their work and thinking. Herbert Simon for instance mentioned that "for centrality to the postwar quantitative social sciences, the Cowles Commission and the RAND Corporation were definitely the places to see and be seen" (Simon 1991, p. x).

Especially during the early 'golden years', RAND successfully involved a group of academic scholars in physical, biological and social sciences; a list of RAND employees and consultants in the 50 and 60s include a number of leading economists and social scientists at the time, including Kenneth Arrow, Armen Alchian, Gary Becker, James Buchanan, Ronald Coase, Gerald Debreu, Merrill Flood, Tjalling Koopmans, Harold Lasswell, Nathan Leites, Charles Lindblom, James March, Oskar Morgenstern, John Nash, Richard Nelson, Allen Newell, Roward Raiffa, Paul Samuelson, Thomas Schelling, Lloyd Shapley, William Sharpe, Martin Shubik, Herbert Simon, Vernon Smith, Robert Solow, Hans Speier, Oliver Williamson, Albert Wohlstetter, and many others. Several of these became well known figures in business school and strategic management circles.

RAND was able to attract and retain such talents at least in large part by offering a combination of academic freedom; little bureaucracy; and interesting strategic problems. Allen Newell, one of the founders of the early work in Artificial Intelligence, recalled about RAND:

> RAND was a fairly free place at that time, bubbling with ideas and people who could sort of go off and . . . do everything, and they saw themselves as sort of experimenting with a whole way of life in terms of research, not in the university setting, with people who sort of consciously and rationality sort of decided what they were going to do and ideas could come up from the bottom of the organization and so forth. It was an exploration of the scientific way of life.

RAND overtime adapted to a changing strategic environment for the country and for the organization as well as a variety of other things. Although RAND had no direct role in the development of business schools or the field of strategic management, many of the key players and central ideas were nurtured by scholars who have been at RAND. RAND as an institution also influenced the thinking and strategic research and management around defense in particular during the McNamara years, and in the office of net assessment work.

Summary

RAND in the 1950s and 1960s was a premier example of an institution conducive (at least for a time) to the nurturing several ideas central to economics, organizations, and strategy, and it became grooming ground for many of the ideas and intellectual developments that became important to the field of strategic management as well as business school research in general (Augier and March 2011). While many of the scholars have since left, the ideas, frameworks and techniques they pioneered are still essential to strategy and strategic management today. The evolution of RAND as an organization itself is also an interesting an illustration of an organization adapting to changes it its strategic environment.

See Also

► Marshall, Andrew W. (Born 1921)
► Nelson, Richard R. (Born 1930)

▶ Simon, Herbert A. (1916–2001)
▶ Strategic Learning
▶ Winter, Sidney G. (Born 1935)

Acknowledgment Much of this entry builds on joint work with James March (esp. Augier and March, chapter 5) and Andrew Marshall (Augier et al. 2015; Augier and Marshall 2017), and I am grateful for their help and collaboration as well as many discussions on the topic.

References

Allison, G. 1971. *Essence of decision: Explaining the Cuban Missile crisis*. Boston: Little Brown.
Allison, Graham, and Andrew W. Marshall. 1969. *Explanation and prediction of governmental action: An organizational process model*. RAND RM-5897-PR.
Augier, M., and J. G. March. 2011. *The roots, rituals, and rhetorics of change: North American Business Schools after the Second World War*. Stanford: Stanford University Press.
Augier, M., and A. Marshall. 2017. The fog of strategy: Some organizational perspectives on strategy and the strategic management challenges in the changing competitive environment. *Comparative Strategy* 36 (4): 275–292.
Augier, M., J.G. March, and A. Marshall. 2015. The flaring of intellectual outliers: An organizational interpretation of the generation of novelty in the RAND Corporation. *Organizaton Science* 26 (4): 1140–1161.
Bornet, V. 1961. RAND: The first fifteen years. RAND document D-9461. Santa Monica: RAND Corporation.
Digby, J. 1990. Strategic thought at RAND, 1948–1963: The ideas, their origins, their fates. RAND Note, N-3096. June, 1990.
Flood, M. 1958. Some experimental games. *Management Science* 5 (1): 5–26.
George, A.L. 1969. The operational code: A neglected approach to the study of political leaders and decision making. *International Studies Quarterly* 13: 190–222.
Goldstein, J.R. 1961. *RAND: The history, operations, and goals of a non-profit corporation*. RAND Corporation Publications, P-2236-1.
Helmer, O. 1957. The Game-Theoretical Approach to Organization Theory. RAND P-1026, p. 1. RAND archives.
Hitch, C.J. 1955. An appreciation of systems analysis. *Journal of the Operations Research Society of America* 3 (4): 466–481.
Hitch, C.J. 1956. Fallacies in operations research. *Operations Research* 4 (4): 422–430.
Kaplan, F. 1983. *The wizards of Armageddon*. Stanford: Stanford University Press.
Leites, N. 1951. The Operational Code of the Politburo. New York: McGraw-Hill.
Marshall, A. 1968. Improving intelligence estimates through the study of organizational behavior. Paper presented to the RAND Board of Trustees, 1968. Santa Monica: RAND Corporation.
Poundstone, W. 1992. *Prisoners dilemma*. New York: Doubleday.
Schlesinger, J.R. 1967. "The uses and abuses of systems analysis". Testimony to Congress during confirmation hearings of James R. Schlesinger to become Secretary of Defense. In *Nomination of James R. Schlesinger to be Secretary of Defense Hearing, Ninety-Third Congress*. Washington, DC: US Government Printing Office.
Schlesinger, J.R. 1989. Nathan Leites: An old world figure in a new world setting. In *Remembering Nathan Leites. An appreciation*, ed. J. Goldhamer and C. Wolf, 55–62. Santa Monica: RAND Corporation.
Simon, H.A. 1986. Some design and research methodologies in business administration. In *La Production des Connaissances scientifique l'administration*, ed. M. Audet and J.-L. Malouin, 239–279. Quebec: Les Presses de L'university Laval.
Simon, H.A. 1991. *Models of my life*. Cambridge, MA: MIT Press.
von Neumann, J., and O. Morgenstern. 1944. Theory of Games and Economic Behavior. Princeton: Princeton University Press.
Weaver, W. 1970. *Scene of change: A lifetime in American science*. New York: Scribner.
Williams, J. 1962. An overview of RAND. RAND Document D-10053. Santa Monica: RAND Corporation.

Rational Expectations

William S. Neilson
The University of Tennessee, Department of Economics, Knoxville, TN, USA

Definition Rational expectations is the correct use of all publicly available information, including the appropriate model of the process that generates any random outcomes.

John Muth introduced the idea of rational expectations in 1961, and his argument can be explained with a simple story (Muth 1961). Suppose that farmers make planting decisions based on adaptive expectations formed from last year's prices. A year of bad weather yields high prices, which lead to overplanting, which then leads to low prices, which then leads to underplanting, which then leads to high prices, and so on.

A rational farmer would understand the underlying supply-and-demand model as well as the lack of serial correlation in weather shocks, and would therefore plant the same amount each year regardless of previous price outcomes. In essence, rational expectations dictate that when economic agents optimize and when the underlying economic system is in equilibrium, the agents' expectations about the future behaviour of the system are forward-looking and model-driven rather than backward-looking and data-driven.

Robert Lucas popularized Muth's idea in the 1970s by applying it to macroeconomics. In his models, firms and individuals are rational, forward-looking decision-makers with economic models of their own. This has two major implications. One is that policy changes not only impact on agents' behaviour but the models they use to forecast, and the resulting changes in belief impact on the effectiveness of the new policy. The second is that the impacts of policy changes cannot be estimated beforehand based on historical data, which cannot account for agents' revised expectations. This latter is the well-known Lucas Critique (Lucas 1976).

Lucas's use of rational expectations led to a revolution in macroeconomic theory that still continues. His findings suggest that an appropriate model of the economy must account for individual behaviour, that individuals form their own models of the economy, and that, in equilibrium, individuals' models must be correct. In other words, reduced-form approaches cannot be used to evaluate policy and must be replaced by structural approaches based on microeconomic foundations. Prominent outcomes of this line of research are real business cycle (RBC) and stochastic dynamic general equilibrium (SDGE) models of the economy.

The notion of rational expectations is closely related to the efficient market hypothesis (EMH), which in various forms says that asset prices reflect all available information, and therefore that one cannot use past prices to outperform the market. Economic theorists proved the existence of rational expectations equilibria for asset markets: prices fully reveal all information, individuals form their own models of the behaviour of the market, and, in equilibrium, these beliefs must be correct. However, these models beg the important question of how new information ever gets into the system, because if acquiring information is costly for agents and if prices fully reveal that information to the entire market, gathering information yields no return. Therefore, no one will collect it. It turns out that this logic holds when prices are determined by a Walrasian auctioneer but not when prices are determined by a real auctioneer. In the former case, the value of information is determined at the margin; in the latter case, the value is infra-marginal and determined by information rents.

Rational expectations play a central role in the theory of games with incomplete information. In these games, some players have information that the other players do not share, and the uninformed must infer the information from the behaviour of their informed opponents. Solution concepts in ▶ game theory are built around the notion of equilibrium, situations in which players have no reason to change their behaviour. If beliefs did not satisfy rational expectations – that is, if beliefs were inconsistent with all available information, including the equilibrium strategies of the informed players, then uninformed players would have a reason to change their beliefs and their strategies, so that the outcome could not be an equilibrium. Solution concepts such as Bayes–Nash equilibrium and sequential equilibrium implicitly incorporate rational expectations into their definitions.

Rational expectations theory runs contrary to behavioural economics, which is founded on the premise that humans are imperfect users of information. Behavioural agents are slowly making their way into theoretical models, where agents used to have rational expectations. Examples include noise traders in behavioural finance and rule-of-thumb price setters in new-Keynesian macroeconomics.

See Also

- ▶ Aspiration Levels and Learning
- ▶ Bounded Rationality
- ▶ Game Theory
- ▶ Prospect Theory and Strategic Decision-Making

References

Lucas, R.E. 1976. Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy* 1: 19–46.

Muth, J.F. 1961. Rational expectations and the theory of price movements. *Econometrica* 29: 315–335.

Real Options

Kalevi Kyläheiko
Lappeenranta University of Technology,
School of Business and Management,
Lappeenranta, Finland

Definition 'A *real option* is the investment in physical and human assets that provides the opportunity to respond to future contingent events' (Kogut and Kulatilaka 2001: 745). All the implicit managerial and operating *flexibilities* that are embedded in non-financial (real) assets and liabilities can be interpreted as real options.

Options reasoning has its roots in financial theory, where it is used for the valuation of derivatives for underlying financial instruments. The Nobel Prize laureates Scholes and Black provided, in 1973, the solution for the financial options that have to be closed at the end of the maturity: the value of an option positively depends on: (1) the price and (2) uncertainty of an underlying instrument, (3) the maturity of an option, (4) the risk-free rate on interest, and, negatively, on (5) the exercise price.

The real-options approach (ROA) can narrowly be viewed as 'an extension of financial theory to options on real (non-financial) assets' (Amram and Kulatilaka 1999). Not surprisingly, the first real-options applications used the direct analogy of financial options. However, in a strategic management context the valuation use of the ROA often proves to be hard, since the Black-Scholes model is based on several critical assumptions, such as that the asset price distribution is log-normal, the asset price follows the Markov process, there are no riskless arbitrage opportunities and no transaction costs (Hull 2003). Also, the markets for options and underlying assets are supposed to be efficient.

Options theory always copes with uncertainty: the greater the uncertainty, the greater the value of the (call) option. The problem, however, is that in the strategic management context uncertainty is often related to endogenous uncertainty (where uncertainties partly depend on manoeuvres taken by other players) that cannot be tackled with random walk-based options valuation models. Even harder is the use of valuation methods when the markets of underlying real assets are inefficient (thin) or even non-existent (Kyläheiko et al. 2002, 2008).

Because of the problems relating to valuation, we prefer to look at the use of the ROA in the context of strategic real investments in thin markets (as they typically are). As noticed by Trigeorgis (1996), the value of an asset is the net present value of expected cash flows plus the (positive or negative) value of all the (strategic) real options related to it.

Real options are often classified according to the flexibility they offer users. The most typical alternatives include: option to wait, option to abandon, option to alter operating scale (e.g., to expand, contract, shut down and restart), option to switch and learning-based growth options. Strategic investment projects often consist of various options called compound options. In multiple options both the options that enhance upside potential and protect against the downside risks work in combination.

In the global world it is no longer enough to have valuable, rare, inimitable and non-substitutable resources. A firm has to be able to proactively modify them as well. This brings us to the dynamic capability view of the firm (Teece et al. 1997; Teece 2007). According to this, the competitive advantage is primarily based on the ability to create, shape, extend and modify the existing resource base to quickly respond to changing preferences or new technologies that open up new strategic growth options. The firm's capacity to exercise these real options is based on ▶ dynamic capabilities that utilize learning and networking options. Sometimes, also, deferral,

abandonment, expansion, contraction or switching options are needed to create flexibility.

Strategies of firms and their networks can be interpreted as means to generate flexibility to overcome the problem created by the lack of complete sets of contingent-forward markets for dynamic capabilities (cf. Bowman and Hurry 1993; Sanchez 1993; Foss 1998; Kogut and Kulatilaka 2001; Foss and Roemer 2010). The ROA also helps dynamize existing retrospective theories of the firm (e.g., the resource-based view) and opens up new opportunities to use the firm boundaries as strategic tools by dynamizing the static transaction cost framework (cf. Kogut 1991; Sanchez 1993; Kyläheiko et al. 2002; Foss and Roemer 2010).

See Also

- ▶ Dynamic Capabilities
- ▶ Resource-Based Theories
- ▶ Risk and Uncertainty
- ▶ Strategic Risk Management

References

Amram, M., and N. Kulatilaka. 1999. *Real options: Managing strategic investment in an uncertain world*. Boston: Harvard Business School Press.

Bowman, E.H., and D. Hurry. 1993. Strategy through the option lens: An integrated view of resource investments and the incremental-choice process. *Academy of Management Review* 18: 760–782.

Foss, N.J. 1998. Real options and the theory of the firm. Working paper series 1998–3, Copenhagen Business School, Department of Industrial Economics and Strategy.

Foss, N.J., and E. Roemer. 2010. Real options, resources and transaction costs: Advancing the strategic theory of the firm. *International Journal of Strategic Management* 2: 73–92.

Hull, J.C. 2003. *Options, futures, and other derivatives*, 5th ed. Upper Saddle River: Prentice Hall.

Kogut, B. 1991. Joint ventures and the option to expand and acquire. *Management Science* 37: 19–33.

Kogut, B., and N. Kulatilaka. 2001. Capabilities as real options. *Organization Science* 12: 744–758.

Kyläheiko, K., J. Sandström, and V. Virkkunen. 2002. Dynamic capability view in terms of real options. *International Journal of Production Economics* 80: 65–83.

Kyläheiko, K., A. Jantunen, and J. Sandström. 2008. Valuing technological capabilities: Strategic options perspective. *International Journal of Business Excellence* 1: 71–91.

Sanchez, R. 1993. Strategic flexibility, firm organizations, and managerial work in dynamic markets: A strategic options perspective. *Advances in Strategic Management* 9: 251–291.

Teece, D. 2007. Explicating dynamic capabilities. *Strategic Management Journal* 28: 1319–1350.

Teece, D.J., G. Pisano, and A. Shuen. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18: 509–533.

Trigeorgis, L. 1996. *Real options: Managerial flexibility and strategy in resource allocation*. Cambridge, MA: MIT Press.

Realism

John Joseph
Duke University, Fuqua School of Business, Durham, NC, USA

Abstract

Critical realism is a philosophy of science, which claims that the social world exists independently of researchers and their investigations of it. Realism is distinct from both positivism and postmodernism. It rejects positivist beliefs that knowledge about the social world must be objectively identified and postmodernist beliefs that the world is entirely socially constructed. Realism recognizes that some causal mechanisms are unobservable and that causal relationships are affected by the openness of social systems and the role of human actors in the reproduction of social phenomena.

Definition Realism asserts that many generative or causal mechanisms in social science are unobservable and exist independently of their identification through experience or theorizing.

What Is Realism?

Realism is a philosophy of science, which provides the basic assumptions for building and

testing theory in the advancement of science. Realism in the social sciences focuses on identifying the generative mechanisms that explain events in social phenomena, even though such revelation is not a condition for assuming their existence (Ackroyd and Fleetwood 2000). The central doctrine of realism acknowledges that many causal or generative mechanisms are unobservable and exist independently of their identification through theorizing or experience (Boyd et al. 1991). From a realist perspective, we can believe a theory and its component causal mechanisms to be true, even if they are unobservable, as long as they explain a pattern of observable events.

In strategic management, these unobservables generally explain events within and between markets, industries, organizations and groups of individuals. Godfrey and Hill (1995) noted that constructs such as perceived opportunism in transaction cost economics, divergent interests in agency theory and tacit resources in the ▶ resource-based view are all unobservables. In the case of tacit resources, for example, they argue that 'the power of theory to explain performance persistence over time is based on the assumption that certain resources are by their nature, unobservable, and hence give rise to high barriers to imitation' (Godfrey and Hill 1995: 523). Hence, despite the difficulty in identifying tacit resources, they still may explain an empirical regularity about firm performance.

An important distinction for realists is that the absence of an observable pattern of events does not necessarily mean the generative mechanism does not exist or lacks causal powers (Tsang and Kwan 1999). As Tsang and Kwan (1999) suggest, to ascribe a power to an object is to say something about what it *will* or *can* do, in the appropriate conditions in virtue of its intrinsic nature (Harre and Madden 1975). Bhaskar (1978) makes this argument in his distinction between observable events in the *empirical* domain and the mechanisms capable of generating the co-occurrence of observable events in the *real* domain. The empirical domain refers to the domain of experienced events and the real domain refers to a separate domain where the generative mechanisms reside (Bhaskar 1978: 13). Mechanisms *may* be activated by certain conditions, move from the real to the empirical domains and produce patterns of events. But their causal power does not depend on their revelation in all circumstances.

Contrast with Positivism and Constructivism

Much of the work on critical realism in social science is attributed to the work of the contemporary philosophers Roy Bhaskar and Rom Harre (Harre and Second 1972; Harre and Madden 1975). Realism has gained in popularity since the early 2000s, particularly among scholars in management and strategy (Bacharach 1989; Tsoukas 1989; Ackroyd and Fleetwood 2000; Kwan and Tsang 2001; Tsang 2006; Van de Ven 2007). Consequently, strategy researchers have increasingly applied ontological and epistemological assumptions to the examination of phenomena in strategic management (Godfrey and Hill 1995; Tsang and Kwan 1999; Durand and Vaara 2009; Miller and Tsang 2010).

Realist philosophy began in the 1970s as an alternative to positivist and constructivist views of scientific enquiry (Bhaskar 1978). Positivism, which has its roots 18th-century philosophers including Locke, Berkeley and Hume (Russell 1946), has long held that theory should contain observable elements whose existence can be empirically verified. Propositions containing theoretical elements that cannot be verified should not be used for the purposes of scientific research. Over time, this stance against unobservables has been relaxed and replaced by an emphasis on the form and content of theory. This has shifted the positivists' emphasis to predictive ability rather than the underlying observables (Friedman 1953), and has elevated the status of mathematical form in generating and testing predictive theory.

Realism also stands in contrast to postmodernist or constructivist approaches, which focus on the highly interpretive nature of social phenomena (Mir and Watson 2000). Constructivists hold that scientific researchers are not separate from the phenomena and theories they seek to evaluate, measure and validate. For constructivists,

researchers, the process of theory-building and the phenomena of study interact to create a socially constructed reality. Moderates in this camp believe that scientific theories not only explain the real world, but are also encoded with the culture, categories, values and interests of researchers. Radical constructivists hold that individual agents structure and create the world as they attempt to discover it (Hess 1997; Kwan and Tsang 2001), and, unlike realists, they 'negate the causal power of generative mechanisms, demoting them to discourse with no material implications' (Durand and Vaara 2009: 1248).

Contentions of Realism

Realism has three foundational contentions (Tsang and Kwan 1999: 762): (1) understanding of the social world is based on understanding the causal mechanisms and social structures that govern social phenomena; (2) mechanisms and structures are only contingently related to observable empirical events; and (3) knowledge about events, mechanisms, structures and their contingent relationships may be acquired through verification and falsification, although our knowledge of objective reality is never definitive.

Causal Mechanism and Structures

Realists hold that the 'task of explanation in social science is to penetrate behind the surface of experiences and perceptions and to account for what occurs in terms of an understanding of connections at the level of structures' (Ackroyd and Fleetwood 2000: 13). Structure refers to a set of simultaneously constraining and enabling rules, which shape human interactions and are shaped by those same interactions (Giddens 1976; Sayer 1992). For example, social structures that guide how managers interact, decide, influence, or construct identity may all shape the conditions under which certain causal mechanisms have explanatory power (Tsang and Kwan 1999). At the same time, managers may draw upon and even alter the structures that govern their actions. Thus, realists believe that in order to understand the social world, it is necessary to incorporate the role of human agents in the reproduction and transformation of structures.

Structures and Contingent Associations

A second important distinction made by realists is that a pattern of events is neither a necessary nor a sufficient condition for a causal claim. This is because realists reserve a place for contingent conditions in theory-building and replication. Realists acknowledge that theories and their components are not isolated from interference and are therefore sensitive to the context in which events occur. Correspondingly, realists recognize that not all causal mechanisms and associations will occur under all conditions (Tsang and Kwan 1999). The causal power of mechanisms depends, in part, on alternative (and equally unobservable) causal mechanisms (Miller and Tsang 2010); this reflects the realist belief that organizations function as open systems (Sayer 1992) and are subject to diverse causal variations and state equifinality.

Correspondingly, realist assumptions of enquiry do not preclude the possibility that causal mechanisms may not function under certain conditions or be counteracted by countervailing or intervening mechanisms that lead to alternative outcomes. For example, the emergence of the M-form corporation is an event observed in the empirical domain, and the generative mechanisms, such the need for efficiency, perceived risk of opportunism and the calculation of *ex ante* transaction costs, reside in the real domain (Godfrey and Hill 1995). However, these generative mechanisms depend on contingent conditions as well, such as the institutional environments of the firm and particular historical background of the firms (Chandler 1962; Kwan and Tsang 2001).

Verification, Falsification and Replication

As with any science, the advancement of strategic management depends on the ability to build and

verify theories, their components and the causal mechanisms that describe their association. Realists assert that because structures and mechanisms exist in the world independently of researchers' knowledge of them, the critical evaluation of theories is possible. However, because realists also recognize the openness of social systems and the continent relationship between generative mechanisms and events, realists limit claims to objective knowledge. As Miller and Tsang (2010: 144) suggest, 'Lacking an indubitable basis for science, we can nevertheless, reasonably assert the veracity or falsity of scientific theories – albeit not definitively.'

It is for this reason that, in advancing theory, realists seek *explanations* rather than *predictions* for the conjunction of events. This focus on explanation demands that enquiry into strategy not be confined to the search for nomothetic regularities between cause and effect. Instead, realist enquiry focuses on the elaboration of the structures, processes, power and relationships and other countervailing mechanisms that undergird events and their association. As noted above, testing theory for its veracity should not require the measurement of observable variables (Godfrey and Hill 1995). Many generate mechanisms are not observable and even draw their causal power from the fact that they cannot be observed. Lastly, failure of exact replication should not be treated as a conclusive falsification of the theory. Since much of realist theory is based on the idea that the relationship between events and their causal mechanisms and structures is contingent, the inability to replicate findings may be indicative of countervailing mechanisms and conditions.

The method of enquiry is important in scientific progress as it must be able to advance the field, build on prior scholarship, repeat or falsify the predictions of various theories. To do so effectively, the method of enquiry must do it in a way that addresses the underlying ontological and epistemological assumptions of realism. Scholars from a variety of disciplines have argued for different approaches to test realist theory including statistical modelling (Mingers 2004), discourse analysis (Fairclough 2005), counterfactual causal analysis and case studies (Tsoukas 1989). Others call for a multi-method and multifocal approach (Tsang and Kwan 1999) that is suitable for addressing the complexities of social phenomena (Miller and Tsang 2010).

For example, case studies may be particularly useful for guiding and advancing strategic management research from a realists' perspective. Case studies are well suited to examining the real-world context in which the phenomena occur (Eisenhardt 1989; Graebner and Eisenhardt, cited in Siggelkow 2007). They can accommodate equal emphasis towards pre-existing structures, agency and the emergent phenomena arising from their interaction.

See Also

▶ Resource-Based View

References

Ackroyd, S., and S. Fleetwood. 2000. *Realist perspectives on management and organizations*. London: Routledge.

Bacharach, S.B. 1989. Organizational theories: Some criteria for evaluation. *Academy of Management Review* 14: 496–515.

Bhaskar, R. 1978. *A realist theory of science*. Leeds: Leeds Books.

Boyd, R., P. Gasper, and J.D. Trout (eds.). 1991. *The philosophy of science*. Cambridge, MA: The MIT Press.

Chandler, A.D. 1962. *Strategic and structure: Chapters in the history of the American enterprise*. Cambridge, MA: The MIT Press.

Durand, R., and E. Vaara. 2009. Causation, counterfacturals, and competitive advantage. *Strategic Management Journal* 30: 1245–1124.

Eisenhardt, K.M. 1989. Making fast strategic decisions in high-velocity environments. *Academy of Management Journal* 32: 543–576.

Eisenhardt, K.M., and M.E. Graebner. 2007. Theory building from cases: Opportunities and challenges. *Academy of Management Journal* 50: 25–32.

Fairclough, N. 2005. Discourse analysis in organization studies. *Organization Studies* 26: 915–939.

Friedman, M. 1953. *Essays in positive economics*. Chicago: University of Chicago Press.

Giddens, A. 1976. *The new rules of sociological method: A positive critique of interpretive sociologies*. London: Hutchinson.

Godfrey, P.C., and C.W.L. Hill. 1995. The problem of unobservables in strategic management research. *Strategic Management Journal* 16: 519–533.

Harre, R., and E.H. Madden. 1975. *Causal powers: A theory of natural necessity*. Oxford: Basil Blackwell.
Harre, R., and P.F. Second. 1972. *The explanation of social behavior*. Oxford: Basil Blackwell.
Hess, D.J. 1997. *Science studies: An advanced introduction*. New York: New York University Press.
Kwan, K., and E.W.K. Tsang. 2001. Realism and constructivism in strategy research: A critical realist response to Mir and Watson. *Strategic Management Journal* 22: 1163–1168.
Miller, K.D., and E.W.K. Tsang. 2010. Testing management theories: Critical realist philosophy and research methods. *Strategic Management Journal* 32: 139–158.
Mingers, J. 2004. Future directions in management science modeling. In *Critical realist applications in organization and management studies*, ed. S. Fleetwood and S. Ackroyd. London: Routledge.
Mir, R., and A. Watson. 2000. Strategic management and the philosophy of science: The case for a constructivist methodology. *Strategic Management Journal* 21: 941–953.
Russell, B. 1946. *A history of western philosophy*. New York: Simon & Schuster.
Sayer, A. 1992. *Method in social science: A realist approach*, 2nd ed. London: Routledge.
Siggelkow, N. 2007. Persuasion with case studies. *Academy of Management Journal* 50: 20–24.
Tsang, E.W.K. 2006. Behavioral assumptions and theory development: The case of transaction cost economics. *Strategic Management Journal* 27: 999–1011.
Tsang, E.W.K., and K. Kwan. 1999. Replication and theory development in organizational science: A critical realist perspective. *Academy of Management Review* 24: 759–780.
Tsoukas, H. 1989. The validity of idiographic research explanations. *Academy of Management Review* 14: 551–561.
Van de Ven, A.H. 2007. *Engaged scholarship: A guide for organizational and social research*. New York: Oxford University Press.

Recombination of Knowledge

Lee Fleming[1] and Gianna G. Giudicati[2]
[1]Harvard University, Business School, Boston, MA, USA
[2]University of Trento, Trento, Italy

Abstract

Many literatures suggest recombinant search as the source of invention, creativity, novelty and ▶ innovation. We discuss the history of this idea, the challenge of defining and operationalizing the idea, the social influences upon recombinant search, and the strategic implications.

Definition Recombinant search is a process (typically social but sometimes automated) that generates novelty. Novelty lacks normative connotation and only describes whether or not a combination is new to the world or a particular social context. New combinations must be assessed for their usefulness and in turn that usefulness, both the causes and consequences, is the primary interest of social scientists and managerial scholars.

History of the Idea and Application in Management Literature

The idea that all things (not just novelty) are the result of recombinations can be traced to the pre-Socratic philosopher Empedocles (*c*. 490–430 BC) (Kirk et al. 1983). Empedocles proposed that the 'components' or elements of recombination were air, fire, earth and water; similar ideas have been identified in other ancient societies (Russell 1945). The concept has continued to influence Western thought and can be recognized in the writings of John Locke and Adam Smith: new ideas are formed by the combination of various old ideas, which can be either empirical (that is, experienced) or theoretical. The 'growth of knowledge' principle proposed by Smith is based on the connections of existing elements (Loasby 2002: 1231): the generative process of recombinant search exploits the connections between elements rather than the elements per se (Smith 1982; Loasby 2002). Poincaré (1921) describes how mathematical creativity results from the cross-fertilization of idea pairs and Campbell (1960) explicitly adopts an evolutionary metaphor of recombinant variation, selection and retention.

The theme of recombination resurfaces has been a recurring feature of the last century of management literature. For instance, Schumpeter (1939) proposed three steps for the recombinant search process: identifying the simple

components of an existing complex system, separating them to form single independent entities and substantially recombining them in a new manner (Schumpeter 1939: 84–86). ▶ innovation, both at the broader economic and technical level and at the firm level, may be seen as the output of a recombinant search process of existing physical and non-physical materials (Nelson and Winter 1982). Baldwin and Clark (2000: 123–146) describe various types of recombinant novelty resulting from the possible combinations of six 'modular operators': splitting, substituting, augmenting, excluding, inverting and porting. By implementing and combining the previous operators, they suggest how changes in the process of invention, which they call 'design', can be classified and categorized. These same authors had earlier suggested that any form of new knowledge, novelty or technology, could be characterized as a change in the core concepts or relationships among them (Henderson and Clark 1990).

Challenges of Defining and Operationalizing Recombination

In contrast to the precise definition of the operators, the definition of components remains frustratingly imprecise. One open question is whether absolutely everything is a recombination of already existing components – in other words, questioning if there is nothing 'new' under the sun. This philosophy applies best to conceptualizing the invention of purely physical artefacts. For example, after the Big Bang when the periodic table settled out, was there an incredibly large – but finite, predefined and discrete – number of 'combinations' possible? If so, inventors explore the possible space of these finite combinations. However, the advent of nanotechnology, and its expansion of the periodic table as the properties of elements change when they get smaller, makes one question this concept of finite combinations. This definition also breaks down when applied to the area of knowledge and non-physical novelty – is knowledge a 'component' or is a 'component' only a physical thing? New 'components' of knowledge might come from science, and add to the stock of components and knowledge over time. This debate can degenerate into a religious argument, and is often best resolved by appeal to the particular context of the research question.

The authors' view is that the recombinant search process is more than purely artefactual, and is highly influenced and constrained by science, available knowledge and culture. Basalla (1988: 26–30, 56–63) defines the elements entering in the combination process as artefactual antecedents, which are then combined to create new artefacts or made things. Fleming and Sorenson (2004: 910) define components as 'any bits of knowledge or matter that inventors might use to build inventions' and are usually denoted as 'factors' (Schumpeter 1939) or 'chunks' (Simon 1991). More specifically, Henderson and Cockburn (1994) used the label 'component competencies' to identify those resources, skills, knowledge and technical systems which are effectively integrated in new and flexible ways to develop new component competencies. The components involved in the recombination process may be existing components, previously untried components, or new components created by the inventor.

Recombinant search perspectives also bring up another controversy between continuous and discontinuous views of technological change. Among the discontinuous perspectives have been Mokyr (1990) and Tushman and Anderson (1986). Such perspectives are motivated by the study of singular and extreme examples of high-impact inventions, for example, the transistor, or rapidly increasing process improvements, such as occurred in concrete production. Basalla (1988) provides the canonical counter-argument and carefully details the incredible minutiae which precede and follow breakthroughs. Examining a population of distributions of invention value would afford a test between continuous and discontinuous views: if breakthroughs are discontinuous, one would expect bimodal distributions; if continuous, one would expect monotonically distributed distributions.

The Actors of Recombination: Individuals and Groups or Organizations

Components do not just magically self-assemble –- they need some sort of possibly purposive combiner (though purely random invention machines, modelled on natural evolution and guided by selection criteria, have been invented, see Koza et al. 2003). This leads naturally to an investigation of the psychological and social influences upon the process of recombinant search. Intelligent and semi-intelligent search has been categorized under the rubric of exploitation (also called local search) and exploration (March 1991; Fleming 2001; Rosenkopf and Nerkar 2001). Because of the constraints and the fewer uncertainties characterizing the recombination process, local search is typically incremental and exploitative, using slightly modified components or slightly changing the relationships between them. Bounded rationality (March and Simon 1958; Simon 1991) implies that actors cannot simultaneously take into account an unlimited number of possible components and combinations. To ease the process, actors either use familiar components, or work within previously developed architectures. Local search, using previously successful components or architectures, minimizes uncertainty, variance and the possibility of failure (Fleming 2001).

While each recombinant trial occurs in the mind of a single inventor, ideas are communicated between creative workers, and hence the social context has a significant influence on recombinant search (Fleming et al. 2007). Inventors usually rely upon the knowledge and experience of their colleagues to improve the combination process and identify more useful components. Scholars underline how individual ideas cannot be derived from isolated experiences and that new ideas and individual creativity are highly influenced by the social environment (Amabile 1983). Social proximity, indeed, improves knowledge generation and diffusion together with other individual and organizational mechanisms such as personal movement, personal friendships, organizational mergers, conferences and strategic alliances.

Sociological research has analysed how networks influence the flow of information and ideas in recombinant search process (Burt 2004; Uzzi and Spiro 2005; Fleming et al. 2007). Despite these efforts to examine the dynamics of individual creativity and upcoming knowledge in social contexts, only recent contributions have started to take into account the collective dynamics of creativity (Kurzberg and Amabile 2000; Padgett and Powell 2011; Riccaboni and Frigotto 2011). Most importantly, the causality of idea creation, flow, and impact upon future creation remains quite tangled. Data and methodological innovation will be required to make progress in this area.

Strategy research has developed the implications of these ideas for firms. Starting from the idea that the ability to acquire external innovation is bounded by an organization's experience and expertise, Schumpeter (1939) pointed out that the commercialization of the invention process is more likely to be done by new, risk-seeking firms – though it is still possible, even likely, that the original breakthrough came from an incumbent (Fleming 2002). Penrose (1959) defined resource recombination as the main source of innovation in firms. The ability of a firm to synthesize and use existing and acquired knowledge has been defined as 'combinative capabilities' (Kogut and Zander 1992). The ability to integrate knowledge has been defined as 'architectural competences' (Henderson and Cockburn 1994). The ability to exploit existing competences and resources to deal with changing environments has been defined as 'dynamic capabilities' (Teece et al. 1997). Galunic and Rodan (1998) found that the likelihood of recombination is negatively related to the tacitness and the dispersion of knowledge and to the closeness of the boundaries and the institutionalization of competences. As with the sociological perspective, endogeneity and causality remain difficult challenges for the strategy literature, particularly in regard to innovation and ▶ technology strategy.

See Also

- ▶ Innovation
- ▶ Innovation Strategies
- ▶ Technology Strategy

References

Amabile, T.M. 1983. *The social psychology of creativity*. Berlin: Springer-Verlag.

Baldwin, C.L., and K.B. Clark. 2000. *Design rules*, The Power of Modularity, vol. 1. Cambridge, MA: The MIT Press.

Basalla, G. 1988. *The evolution of technology*. Cambridge: Cambridge University Press.

Burt, R.S. 2004. Structural holes and good ideas. *American Journal of Sociology* 110: 349–399.

Campbell, D.T. 1960. Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review* 67: 380–400.

Fleming, L. 2001. Recombinant uncertainty in technological search. *Management Science* 47: 117–132.

Fleming, L. 2002. Finding the organizational sources of technological breakthroughs: The story of Hewlett Packard's ink jet invention. *Industrial and Corporate Change* 11: 1059–1084.

Fleming, L., and O. Sorenson. 2004. Science as a map in technological search. *Strategic Management Journal* 25: 909–928.

Fleming, L., S. Mingo, and D. Chen. 2007. Collaborative brokerage, generative creativity, and creative success. *Administrative Science Quarterly* 52: 443–475.

Galunic, C.R., and S. Rodan. 1998. Resource recombination in the firm: Knowledge structures and the potential for Schumpeterian innovation. *Strategic Management Journal* 19: 1193–1201.

Henderson, M., and K.B. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Special Issue, Administrative Science Quarterly* 35: 9–30.

Henderson, R., and I. Cockburn. 1994. Measuring competence? Exploring firm effects in pharmaceutical research. *Special Issue, Strategic Management Journal* 15: 63–84.

Kirk, G.S., J.E. Raven, and M. Schofield. 1983. *The presocratic philosophers: A critical history*, 2nd ed. Cambridge: Cambridge University Press.

Kogut, B., and U. Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* 3: 383–397.

Koza, J.R., M.A. Keane, and M.J. Streeter. 2003. Evolving inventions. *Scientific American* 288: 52–59.

Kurzberg, T.R., and T.M. Amabile. 2000. From Guilford to creative synergy: Opening the black box of team-level creativity. *Creativity Research Journal* 13: 285–294.

Loasby, B.J. 2002. The evolution of knowledge: Beyond the biological model. *Research Policy* 31: 1227–1239.

March, J.G. 1991. Exploration and exploitation in organization learning. *Organization Science* 2: 71–87.

March, J.G., and H. Simon. 1958. *Organizations*. New York: Wiley.

Mokyr, J. 1990. *The lever of riches: Technological creativity and economic progress*. Oxford: Oxford University Press.

Nelson, R., and S. Winter. 1982. *An evolutionary theory of economic change*. Cambridge: Belknap Press.

Padgett, J.F., and W.W. Powell. 2011. *The emergence of organizations and markets*. Princeton: Princeton University Press.

Penrose, E.G. 1959. *The theory of the growth of the firm*. New York: Wiley.

Poincaré, H. 1921. *The foundations of science: Science and hypothesis, the value of science, science and method*. New York: Science Press.

Riccaboni, M., and M.L. Frigotto. 2011. A few special cases: Scientific creativity and network dynamics in the field of rare diseases. *Sociometrics* 89: 397–420.

Rosenkopf, L., and A. Nerkar. 2001. Beyond local search: Boundary-spanning, exploration, and impact in the optical disk industry. *Strategic Management Journal* 22: 287–306.

Russell, B. 1945. *A history of western philosophy, and its connection with political and social circumstances from the earliest times to the present day*. New York: Simon & Schuster.

Schumpeter, J. 1939. *Business cycles*. New York: McGraw-Hill.

Simon, H. 1991. Bounded rationality and organizational learning. *Organization Science* 2: 125–134.

Smith, A. 1982. *Lectures on jurisprudence*. Indianapolis: Liberty Fund.

Teece, D.J., G. Pisano, and A. Shuen. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18: 509–533.

Tushman, M.L., and P. Anderson. 1986. Technological discontinuties and organizational environments. *Administrative Science Quarterly* 31: 439–465.

Uzzi, B., and J. Spiro. 2005. Collaboration and creativity: The small world problem. *American Journal of Sociology* 111: 447–504.

Red Queen Among Organizations, the

William P. Barnett
Stanford University, Graduate School of Business, Stanford, CA, USA

Abstract

This article considers the 'Red Queen' effect, which was coined to describe the phenomenon whereby an organization embarks on a policy of continuous improvement simply to maintain its relative position within the industry. The extent to which this is

applicable is affected by the 'logics of competition', a phrase which refers to the technological and market structures and social settings in which an organization operates. This will vary both from industry to industry and also over time.

Definition 'Red Queen' competition is a dynamic among rival organizations, where competition triggers learning, which in turn increases competition in a self-accelerating process. Organizations in Red Queen competition are continuously changing in absolute terms, although they may be relatively stable.

When organizations compete, they initially make it difficult for each other to perform well. This pressure causes organizations to search for ways to improve performance. These improvements, in turn, make organizations stronger ▶ competitors, increasing the pressure on their rivals, who again search for improvements. So it is that competition causes organizational learning, which in turn intensifies competition in a self-accelerating process known as the 'Red Queen' effect. This term was coined by the evolutionary theorist Van Valen (1973) in reference to Lewis Carroll's Alice, who remarks to the Red Queen: 'Well, in our country, you'd generally get to somewhere else – if you ran very fast for a long time as we've been doing.' To this the Red Queen responds: 'A slow sort of country! Now, here, you see, it takes all the running you can do, to keep in the same place.' In an ecology of learning organizations, relatively stable performance masks ongoing, absolute development as organizations learn in response to rivalry, thereby increasing the strength of rivalry and so again triggering learning.

As applied to organizations, the theory of Red Queen competition is based on a synthesis of ▶ organizational ecology (Hannan and Freeman 1989) and organizational learning theory (March and Simon 1958; Levinthal and March 1993). Several premises are fundamental to the Red Queen theory (see Barnett 2008). First, ▶ competitiveness must be understood as a property of organizations rather than markets. Specifically, competition is defined as occurring whenever one organization reduces the viability of another organization. Competitiveness, then, can vary from organization to organization, and the strength of each organization's competitiveness can be estimated statistically based on its characteristics or experience. Second, the theory assumes that organizations respond to competition by searching for improvements. This search is not necessarily guided by managers; it is assumed to be the usual organizational response when performance falls below aspiration levels. Over time, such a search discovers improvements that cumulate as the capabilities of organizations. A third assumption is that an organization's capabilities, when aligned with its environment, make it a stronger competitor. Under these assumptions, organizations that have been exposed to competition are predicted to be more viable and to generate stronger competition – predictions supported in a variety of empirical tests (Barnett 2008).

The full implications of the Red Queen depend on the 'logic of competition' that exists in any particular context, defined as 'a system of principles in a given context that determines who can compete, how they compete, on what criteria they succeed or fail, and what are the consequences of success or failure' (Barnett 2008: 14). In a given setting, technologies, organizational architectures, market structures, social processes, political forces and various institutions all shape the prevailing logic of competition. For instance, time-to-market is rewarded in the semi-conductor and disk-drive industries due to the rate of technological change in those markets. In some other industries, by contrast, an advantage goes to national champions in any particular country due to political factors. Over time, logics of competition may change, as when the Internet disrupted the music recording industry, or when deregulation transformed retail banking in the US in the late twentieth century. When such changes occur, emerging logics of competition are ambiguous, so organizations must discover these logics by competing. Winning and losing in competition, organizations discover what works – and what does not work – and so come

to identify their strategies. In this way, the Red Queen is consistent with the 'emergent' perspective on strategy development.

Through the process of Red Queen evolution, an organization comes to be well aligned with the logic of competition in its context, and hence it beco mes a stronger competitor *in that context*. When the logic of competition changes, however, such a well-developed organization may find itself in a 'competency trap' (Levinthal and March 1993). Empirical estimates show that *outdated* competitive experience switches the direction of the Red Queen effect – making organizations *less* viable and *weaker* as competitors (Barnett 2008). Yet competitive experience also makes it especially hazardous for organizations to change (Barnett and Pontikes 2008). So because of organizational inertia, the Red Queen can cut both ways, making organizations stronger within a given logic of competition, but making them especially outdated when conditions change.

Sometimes alternative logics of competition are in contention, as when technology standards vie for prominence or when laws circumscribe the kinds of firms that can compete in a given industry or country. In these situations, whether an organization wins or loses depends on the larger question of which logic prevails. For instance, if environmental sustainability becomes important in the automobile or construction industries, then 'green' products and manufacturers will have an advantage. By contrast, luxury or quality might be emphasized, in which case a different set of firms will be favoured. In this way, meta-competitions among contending logics may determine who wins and who loses as industries evolve. Organizations then vie for legitimacy by framing their actions as appropriate given the prevailing logic of competition.

See Also

- ▶ Competitiveness
- ▶ Competitive Strategy
- ▶ Competitors
- ▶ Organizational Ecology

References

Barnett, W.P. 2008. *The Red Queen among organizations: How competitiveness evolves*. Princeton: Princeton University Press.

Barnett, W.P., and E.G. Pontikes. 2008. The Red Queen, success bias, and organizational inertia. *Management Science* 54: 1237–1251.

Hannan, M.T., and J. Freeman. 1989. *Organizational ecology*. Cambridge, MA: Harvard University Press.

Levinthal, D.A., and J.G. March. 1993. The myopia of learning. *Strategic Management Journal* 14: 95–112.

March, J.G., and H.A. Simon. 1958. *Organizations*. New York: Wiley.

Van Valen, L. 1973. A new evolutionary law. *Evolutionary Theory* 1: 1–30.

Regional Development

Sampsa Samila
National University of Singapore, Business School, Singapore, Singapore

Abstract

While great differences exist in the economic fortunes of different nations, the variations between regions within a nation can also be considerable. The study of regional development seeks to find the causes of these differences. The success of the Silicon Valley region in the technology industry has, in particular, attracted attention in the academic, business and policy worlds. Research has uncovered natural advantages, human capital, social capital, agglomeration economies and government as causal factors in explaining regional differences. These factors are interdependent, and often have strongly complementary effects.

Definition Whereas studies of economic growth generally focus on understanding the drivers of economic growth at national level and the differences in economic development between countries, the focus in studies of *regional development* is on understanding the sources of within-country variation in the rates and levels of economic development.

Why are some regions growing and some declining, even within the same country? For every Silicon Valley there is a Youngstown or Detroit, a once booming city that has endured a decline in fortunes, and regions that seem to perpetually lag behind others. The study of regional development seeks to understand the causes of these differences.

Research has identified five broad classes of causes for regional differences in economic development: natural advantages, human capital, social capital, agglomeration economies and government activity. All of them work by making economic activity more attractive, efficient and innovative in some regions over others, fostering faster growth and, ultimately, greater economic development.

Natural Advantages

Historically, many of the most important cities have been located within easy access to transportation by waterways, either through a protected sea harbour or a navigable river. Before the advent of railways, land transportation over longer distances moved barely faster than walking speed, while waterways provided significantly swifter methods of transport and trade. While the development of railways first and highways later has reduced the impact of natural transportation advantages, it has not made them irrelevant.

Similarly, some regions are inherently more suited for some economic activity than other regions. Agriculture requires a certain type of soil and climate, depending on the crop to be grown. Oil and minerals are found in certain regions and not others. To the extent that these are in short supply in the world markets, the economic impact of their presence can be considerable. For instance, the recent commercial viability of oil extraction from tar sands has transformed the economic fortunes of large parts of the province of Alberta in Canada.

Disease patterns can hinder economic activity significantly in some regions of Africa. The World Health Organization argues that disease extracts a considerable toll in sub-Saharan Africa (Sachs 2001). Weather can also affect the economic fortunes of regions. In the US, people have been migrating over time to regions with more pleasant weather (Rappaport 2007).

These natural factors still have a considerable impact, even in developed countries and in manufacturing industries. Ellison and Glaeser (1999) find that at least 20% of the agglomeration patterns of manufacturing industries within the US can be explained by a small set of natural factors.

Human Capital

One factor that seems to be consistently linked to regional economic growth in empirical studies is the level of human capital, measured as the average level of education (Glaeser et al. 1995).

Social Capital

One of the earliest attempts to define cultural differences as sources of economic differences was Weber's classic *The Protestant Ethic and the Spirit of Capitalism* (Weber 1958). Another influential early work was Banfield's examination of the role of trust in explaining the relative backwardness of the south of Italy compared with the north of Italy (Banfield 1958). More recently, Florida's ideas on the regional impact of the so-called 'creative class' of high technology workers, artists and musicians have attracted considerable public attention (Florida 2002). The proposed mechanisms through which culture works involve the affecting of individual motives (e.g., Weber), making cooperation with others easier (e.g., Banfield) and making the community more open to innovation (e.g., Florida). Recent evidence suggests that culture can indeed have a direct effect on regional differences in economic development (Tabellini 2010).

Agglomeration Economies

Agglomeration economies derive from the clustering of related businesses in one area, and can

lead to increasing returns where regions that are more developed develop faster than other regions. These were first analysed in detail by Marshall (1920), who noted that agglomeration economies derive mainly from three sources: input or market-sharing, labour-pooling and knowledge-sharing. By locating near firms in the same or related industries, firms can economize on access to inputs they need and on markets for their products. A region with multiple related firms provides many possibilities for employment and reduces the risk of developing focused skills, thus leading to greater division of labour and a more productive labour force. Likewise for firms, a region with similar firms provides access to a talented labour pool. And when many similar firms are concentrated in one area, knowledge about the business is likely to diffuse and be shared among the firms. The mysteries of trade become no mysteries; but are as it were in the air (Marshall 1920).

The empirical evidence supports these ideas. Knowledge flows between firms tend to be localized (Audretsch and Feldman 1996). While large firms can access national capital markets, for many smaller firms access to capital is localized (Sorenson and Stuart 2001; Becker 2007). Employees also tend to work harder in more concentrated settings (Rosenthal and Strange 2008). The outcome is that industries are often concentrated in certain areas, and the sharing of goods, labour and knowledge seem to be significant sources of that concentration in cross-sectional analyses (Rosenthal and Strange 2001; Ellison et al. 2010). The rates of entrepreneurship also seem to follow similar patterns, with access to knowledge and to capital as important factors (Sorenson and Audia 2000; Kerr and Nanda 2009; Samila and Sorenson 2011a).

Government

Government activity can have a significant effect on regional development through infrastructure, laws and regulations. Investments in infrastructure can improve regional development (Demurger 2001). While most laws and regulations tend to be national in scope, there still exists some regional variation. Of recent interest has been labour law, non-compete covenants in particular, which, in the US, is the responsibility of states. These clauses in labour contracts that prevent employees from leaving to a competing firm have been found to limit labour mobility, knowledge diffusion, entrepreneurship and growth (Gilson 1999; Marx et al. 2009; Samila and Sorenson 2011b).

The above-mentioned mechanisms do not operate in isolation, but instead can significantly reinforce each other. For instance, the access to finance suitable for start-ups can have a greater impact in areas with sources of innovations (Samila and Sorenson 2010). Likewise, social capital seems to improve the financial development of a region (Guiso et al. 2004). In addition, cultural amenities can attract high human capital employees (Falck et al. 2011). The social make-up of a region can also have an effect on the functioning of the government. In particular, diversity can reduce investment in public goods (Alesina et al. 1999). Government, educational institutions and industry all seem to play complementary roles (Etzkowitz and Leydesdorff 2000; Powell et al. 2002).

See Also

- ▶ Geography of Innovation
- ▶ Industrial Policy
- ▶ Innovation Policy
- ▶ Institutional Environment
- ▶ Knowledge Spillovers
- ▶ Marshall, Andrew W. (Born 1921)

References

Alesina, A.F., R. Baqir, and W. Easterly. 1999. Public goods and ethnic divisions. *Quarterly Journal of Economics* 114: 1243–1284.

Audretsch, D., and M. Feldman. 1996. R&D spillovers and the geography of innovation and production. *American Economic Review* 86: 630–640.

Banfield, E. 1958. *The moral basis of a backward society.* New York: Free Press.

Becker, B. 2007. Geographical segmentation of US capital markets. *Journal of Financial Econonomics* 85: 151–178.

Demurger, S. 2001. Infrastructure development and economic growth: An explanation for regional disparities in China? *Journal of Comparative Economics* 29: 95–117.
Ellison, G., and E.L. Glaeser. 1999. The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review* 89: 311–316.
Ellison, G., E.L. Glaeser, and W.R. Kerr. 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review* 100: 1195–1213.
Etzkowitz, H., and L. Leydesdorff. 2000. The dynamics of innovation: From national systems and 'mode 2' to a triple helix of university–industry–government relations. *Research Policy* 29: 109–123.
Falck, O., M. Fritsch, and S. Heblich. 2011. The phantom of the opera: Cultural amenities, human capital, and regional economic growth. *Labour Economics* 18: 755–766.
Florida, R. 2002. *The rise of the creative class*. New York: Basic Books.
Gilson, R.J. 1999. The legal infrastructure of high technology industrial districts: Silicon Valley, Route 128, and covenants not to compete. *New York University Law Review* 74: 575–629.
Glaeser, E.L., J. Scheinkman, and A. Shleifer. 1995. Economic growth in a cross-section of cities. *Journal of Monetary Economics* 36: 117–143.
Guiso, L., P. Sapienza, and L. Zingales. 2004. The role of social capital in financial development. *American Economic Review* 94: 526–556.
Kerr, W., and R. Nanda. 2009. Democratizing entry: Banking deregulations, financing constraints, and entrepreneurship. *Journal of Financial Economics* 94: 124–149.
Marshall, A. 1920. *Principles of economics*. London: Macmillan.
Marx, M., D. Strumsky, and L. Fleming. 2009. Mobility, skills, and the Michigan non-compete experiment. *Management Science* 55: 875–889.
Powell, W.W., K. Koput, J. Bowie, and L. Smith-Doerr. 2002. The spatial clustering of science and capital: Accounting for biotech firm-venture capital relationships. *Regional Studies* 36: 291–305.
Rappaport, J. 2007. Moving to nice weather. *Regional Science and Urban Economics* 37: 375–398.
Rosenthal, S., and W. Strange. 2001. The determinants of agglomeration. *Journal of Urban Economics* 50: 191–229.
Rosenthal, S.S., and W.C. Strange. 2008. Agglomeration and hours worked. *Review of Economics and Statistics* 90: 105–118.
Sachs, J.D. 2001. Macroeconomics and health: Investing in health for economic development. Report of the Commission on Macroeconomics and Health, World Health Organization.
Samila, S., and O. Sorenson. 2010. Venture capital as a catalyst to commercialization. *Research Policy* 39: 1348–1360.
Samila, S., and O. Sorenson. 2011a. Venture capital, entrepreneurship, and economic growth. *Review of Economics and Statistics* 93: 338–349.
Samila, S., and O. Sorenson. 2011b. Non-compete covenants: Incentives to innovate or impediments to growth. *Management Science* 57: 425–438.
Sorenson, O., and P.G. Audia. 2000. The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940–1989. *American Journal of Sociology* 106: 424–461.
Sorenson, O., and T.E. Stuart. 2001. Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology* 106: 1546–1588.
Tabellini, G. 2010. Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association* 8: 677–716.
Weber, M. 1958. *The protestant ethic and the spirit of capitalism*. New York: Scribner.

Regulation/Deregulation

Nydia Macgregor and Tammy L. Madsen
Santa Clara University, Leavey School of Business, Santa Clara, CA, USA

Abstract

Regulation establishes formal rules to govern organizational conduct in an industry, or multiple industries, with the force of law. Regulatory reform, either by imposing regulations or easing them (deregulation), changes the institutional rules governing competition and, in turn, the opportunity sets available to firms. Regulation constrains strategic choice, limits competition and, produces industry inefficiencies, whereas deregulation contributes to operating freedom, unfettered competition and improvements in efficiency. Sometimes regulatory reform also produces unintended consequences. Generally, shifts in regulatory regimes contribute to competitive dynamics via firm entry and exit, changes in industry structure and heterogeneity in competitive behaviour.

Definition A regulation is a formal rule, prescribed by a government institution such as an

administrative agency (e.g., Food and Drug Administration (FDA), Securities and Exchange Commission (SEC)) that governs organizational conduct in an industry or multiple industries (e.g., bank regulation versus antitrust policies) with the force of law. Deregulation is the process of removing formal rules from an industry or multiple industries in an effort to improve economic performance.

Overview

Regulation and deregulation are two possible outcomes of the same process. This process involves the creation of formal rules that govern the strategic actions of firms in a specific industry or across multiple industries. In general, these rules are intended to establish stable relationships among firms, workers and the state (Fligstein 2002) and, in turn, facilitate economic exchange. The traditional view was that, by controlling monopoly, regulation enhances welfare. Research discredited this view, showing instead that regulation limits competition among firms. The literature classifies regulations into two broad categories, social and economic; both categories may occur at different levels (e.g., local, state, federal, etc.). Social regulation targets non-economic activities across industries and covers areas such as environmental protection or health. For example, the World Health Organization (WHO) is 'the directing and coordinating authority for health within the United Nations system. It is responsible for providing leadership on global health matters, shaping the health research agenda, setting norms and standards, articulating evidence-based policy options, providing technical support to countries and monitoring and assessing health trends' (WHO 2011). In contrast, economic regulation guides competitive behaviour within an industry. Further, different types of regulation exist within each category and these types vary in their implications. For instance, cartel-like policies thwart competition and discourage industry concentration, whereas antitrust policies invigorate competition and encourage industry concentration.

States generally create regulatory regimes in an attempt to correct for market failures such as those associated with non-competitive market conditions, information asymmetries, externalities or public goods. As an example, efforts to regulate the US airline industry initially stemmed from the societal scepticism of the efficiency of markets that coincided with the Great Depression (Levine 1987). Further, shifts in societal values or understandings, or the power of key interest groups, can foster momentum for creating new regulatory regimes. In these instances, after the public and government officials affirm the existence of a problem, legislation is enacted. For instance, by the 1960s prominent economists converged upon the notion that the airline industry lacked perfect competition. In an effort to achieve this lofty status, the US Congress passed legislation to deregulate the industry in 1978. This example illustrates a typical pattern where, after a long period of stability in regulatory arrangements, significant transformation remains possible.

Regardless of the category or type, regulation fundamentally acts to solidify a predictable set of market and non-market relationships, enabling organizational actors to comprehend their opportunity set. The stability stemming from regulatory regimes may positively influence firm profits and viability, in contrast to the instability of environments devoid of any regulatory regime (Fligstein 2002). Nonetheless, a by-product of this 'stability' is that regulatory constraints limit competition among firms; the lack of competition reduces incentives for firms to search for, and invest in, efficient operating practices (e.g., Winston 1998). Moreover, under certain conditions, regulation forces firms to operate inefficiently. For example, regulation in the US trucking industry prohibited firms from using optimal trucking networks; in turn, operating efficiency suffered. Further, even in environments with predictable legal structures and free of war or invasion, strong regulatory constraints tend to increase transaction costs (Fligstein 2002). Last, regulations also may isolate firms from macroenvironmental disruptions and/or slow their adjustment to such shocks. As a result, regulation is associated with industry inefficiency.

In contrast, deregulation loosens the constraints circumscribed by extant regulatory regimes, shifting the balance of power between private well-being and social well-being (North 1990). A move towards a more deregulated environment thus reduces a state's intermediate role between a firm's principals and agents. As such, it increases the actions permissible to firms and, in turn, provides an opportunity for enhancing profitability. In other words, deregulation contributes to operating freedom and unfettered competition. These conditions typically allow firms to pursue novel ways of competing to increase operational efficiencies and to enhance their responsiveness to customer preferences and environmental disturbances. As a result, deregulation yields meaningful improvements in industry efficiency that endure and build over time (e.g., see also Winston 1998). For example, efficiency increased substantially in the US airline, trucking, railway, banking and natural gas industries under deregulation (see Winston 1998, for a summary). In most cases, however, the value created via such cost savings is transferred to consumers rather than captured by firms.

The transition from a regulated environment to a deregulated environment, while desirable from an efficiency viewpoint, is not necessarily predictable or complete. For one, multiple institutional actors (e.g., administrative agencies, powerful organizations, congress, legislators, etc.) may foster uncertainty regarding the content of emerging regulatory reform. For instance, in the US electric energy market between 1998 and 2002, uncertainty existed regarding whether regulations would increase or decrease, making adjustment plans difficult for firms (Delmas and Tokat 2005). These conditions typically lead incumbent firms to delay investing in operational changes until deregulation is formally introduced. In addition, regulatory change often has unintended consequences. For instance, in the US telecommunications industry, the break-up of AT&T ultimately benefited AT&T (see Haveman et al. 2001). The potential for unintended consequences combined with the fact that incumbent firms often inaccurately estimate how deregulation will unfold (e.g., Leone 1986) suggests that the anticipated consequences of regulatory reform are often partial, at best.

Source, Scope and Pace

Regulatory reform (institutional change) varies in source, scope and pace (e.g., Mahon and Murray 1981; Reger et al. 1992). To begin, the source of reforms may be exogenous, where actors external to an industry, such as regulatory agencies or legislative bodies, play a dominant role in initiating and defining the regulations. For instance, legislation gives birth to regulation but does not necessarily shape the affiliated policies; instead, legislators rely on professional bureaucrats within government agencies to implement the legislation by crafting rules and managing enforcement mechanisms; furthermore, in developed countries, these agencies are considered to be impartial arbiters (North 1990). Alternatively, as Stigler asserted as early as the 1960s (Peltzman 1993), regulatory reform might be more endogenous, where firms in an industry play a more dominant role than external institutional actors, in capturing the rules governing competition. From this view, regulatory regimes might reflect the conquest of producer interests over consumer interests, protecting the industries they regulate from competition (Stigler 1971; Demsetz 1982).

Regardless of the source, however, interactions among an industry's members and external institutional actors (e.g., public interest groups, regulatory agencies, legislative bodies) influence the scope, and pace of implementation, of institutional change (Winston 1998). In this process, and more often than not, regulation is informed by powerful organizational or institutional actors who successfully advocate for their interests above those of others (Fligstein 2002). As a result, regulatory reform typically has heterogeneous effects on organizations or industries, benefiting some and not others. For example, local or state regulations routinely favour taxis over other modes of transport to local airports, such as vans and limousines, and elevate fares and taxi profits (Winston 1998). As such, regulation must be

viewed as a matter of degree, not a binary construct where there is complete regulation or complete deregulation with homogeneous (equitable) implications for all.

Indeed, more complete descriptions convey how the depth and breadth of regulatory reforms influence the opportunity set for firms. Along these lines, regulatory *scope* describes the extent to which industry competition is market-based. While scope refers to the breadth and depth of rule changes, the pace of regulatory reform highlights the rate of implementation of such reform. For instance, the scope of regulation ranges from incremental to more radical or transformative, whereas the pace of implementation may vary from stepwise to punctuated (see Kim and Prescott 2005). Punctuated and radical regulatory reform can disrupt an industry's course of evolution and, in turn, reset the industry's clock (e.g., Madsen and Walker 2007). In contrast, when new rules are rolled out at a slow pace, managers may be able to anticipate the new order, which reduces uncertainty and allows for gradual adjustment. However, even with a seemingly more benevolent slower pace of change, uncertainty may surround the content of regulation (Merton 1936; Winston 1998) and the uncertainty can last for some time (Delmas and Tokat 2005). This makes it difficult for firms to predict the best course of strategic adaptation. In addition, the successful creation of one regulatory arrangement does not preclude a transformation of this regime in the future (e.g., from regulation to deregulation to re-regulation), nor does it predict how quickly such a transformation may occur. In the short run, while firms are adjusting to the new conditions, deregulation may negatively impact efficiency (Delmas and Tokat 2005). Nonetheless, policymakers and the public often seek immediate benefits from deregulation; this type of short-run view leads policymakers to undervalue deregulation and, in turn, consider re-regulation and/or changes in the scope of deregulation (Winston 1998). In general, the benefits of regulatory reform typically emerge over the long run. Thus, a second dimension of pace lies with regulation's temporal, and potentially transitory, aspects.

Results: Highlights from Extant Work

Since conforming to industry regulation increases economic returns, firms attempt to align their policies and behaviours with the institutional rules governing an industry (North 1990). As a result, to avoid declining profits, firms respond to regulatory change with strategic adjustments (Reger et al. 1992; Peteraf and Reed 2008). This is true regardless of whether the scope of regulation is broad or narrow. In addition, the new rules may allow for the emergence of different types of competitors, such as *de novo* or *de alio* entrants, and, in turn, change the competitive landscape. In other words, new regulatory arrangements trigger new selection pressures (Haveman et al. 2001; Madsen and Walker 2007). As a result, regulation (deregulation) has significant implications for industry composition and competitive dynamics and for firms' performance, viability and rates of adjustment. The following sections highlight key findings and observations from the extant empirical work.

Industry Composition and Competitive Dynamics

When regulatory reform involves a radical and punctuated departure from existing rules governing industry competition, it destabilizes an industry. Typically, these conditions contribute to a massive shift in an industry's structure of competition (e.g., Winston 1998), largely due to a shift in the players in the industry (via entry and exit), the balance of power among institutional actors and the dimensions of competition. First, following such a disruptive change, two generic cohorts of competitors emerge: incumbents, which are firms that entered the industry before the change, and entrants, which started up after the change. Second, because the opportunity set is altered by the regulatory reform, both incumbents and entrants encounter challenges that differ from those that incumbents faced in the earlier regulatory regime. In order to adapt and survive, entrants must build capabilities de novo, while incumbents, in contrast, must replace or modify their traditional routines developed in the previous era. These conditions yield heterogeneous responses

from entrants and incumbents. Some entrants develop robust positions based on new ways of competing, such as a low-cost approach (e.g., low-cost airlines). Responses from incumbent firms vary but typical patterns observed involve incumbents adapting by building and leveraging scale to sustain viability. For instance, in trucking, some firms used mergers to facilitate rapid expansion and/or purely as a means of survival. Not surprisingly, this merger activity often accelerated exit rather than thwarting it. Despite these efforts, adaptation to new regulatory regimes is not straightforward and the survival rate for incumbent firms suffers. Indeed, in the trucking industry, less than half the incumbent population remained in the industry 10 years after deregulation ensued (e.g., Madsen and Walker 2007).

Implications: Competitive Heterogeneity

Shifts in regulatory regimes lead to variation in opportunity sets and competitive behaviours. Recent work relates these differences to issues fundamental to the strategy field: the heterogeneity in performance among firms and the durability of superior firm profits (e.g., Madsen and Walker 2002; Walker et al. 2002). The general conclusions from work in this area are that: (1) deregulation resets an industry's clock (to the early stages of industry development); (2) the patterns of performance heterogeneity among two cohorts of firms, entrants and incumbents differ under deregulation; and (3) the profit advantages developed by entrants under deregulation are more temporary than those held by incumbent firms. More specifically, in a study of the US airline industry during the 10 years before and after the deregulation of pricing and entry, Walker et al. (2002) found that the heterogeneity in performance among entrants was significantly greater than that of incumbents following deregulation but that the heterogeneity in performance among incumbents was relatively stable across regulatory regimes. The evidence suggests that, following changes to rules governing competition, the capabilities of entrants tend to be more powerful determinants of performance differences than the capabilities of incumbent firms. In addition, incumbents encounter difficulties in copying wide-ranging strategies, and instead, tend to retain commonly understood ways of operating. Considering the effects of price and entry deregulation in the US trucking industry, work shows that the profit advantages held by entrants are much more temporary than those held by incumbent firms (Madsen and Walker 2002). Similarly, whereas early work suggested that increased regulation in the US pharmaceutical industry contributed to a decline in R&D productivity and innovation (Henry et al. 1978; Wiggens 1979), more recent work shows that these conditions differentially affected the viability and profitability of small and large firms (Thomas 1990). Small firms suffered dramatically whereas large firms experienced sales gains due to reduced competition; these gains offset the large firms' moderate declines in research productivity (Thomas 1990). Taken together, the findings indicate that the distribution of profits observed in an industry and the duration of firms' profit advantages are strongly influenced by regulatory reform, cohorts' experiences under different regulatory regimes and heterogeneity among firms.

Adjustment and Adaptation

The preceding sections highlight some of the challenges different firms face in adapting to a new regulatory regime. Other studies explore adjustment and adaptation from different angles.

For example, controlling for productivity growth, work finds that large entrants are slower to adjust to a deregulated regime than large incumbent firms, whereas small incumbents are slower to adjust than small entrants (Madsen and Walker 2013). Thus, a firm's legacy of competing in a prior regulatory regime constrains its adjustment to the new regime; however, the imprint of competing in the prior regime is not necessarily debilitating for large firms. Further, as large incumbents endure under deregulation, they develop an increasing ability to compete successfully. Examining adaptation in the airline industry, Peteraf and Reed (2008: 99) find that firms adapt to deregulation by employing strategic choices in areas where they have substantial discretion in order to 'counteract the effects of constrained or

predetermined choices'. This adjustment process yields a tighter alignment between firms' operations and the deregulated environment.

See Also

- ▶ Competitive Heterogeneity
- ▶ Industry Transformation

References

Delmas, M., and Y. Tokat. 2005. Deregulation, governance structures, and efficiency: The US electric utility sector. *Strategic Management Journal* 26: 441–460.

Demsetz, H. 1982. The 1982 Nobel prize in economics. *Science* 218: 655–657.

Fligstein, N. 2002. *The architecture of markets: An economic sociology of twenty-first-century capitalist societies*. Princeton: Princeton University Press.

Haveman, H.A., M.V. Russo, and A.D. Meyer. 2001. Organizational environments in flux: The impact of regulatory punctuations on organizational domains, CEO succession, and performance. *Organization Science* 12: 253–273.

Henry, G., H.G. Grabowski, J.M. Vernon, and L.G. Thomas. 1978. Estimating the effects of regulation on innovation: An international comparative analysis of the pharmaceutical industry. *Journal of Law and Economics* 21: 133–163.

Kim, B., and J.E. Prescott. 2005. Deregulatory forms, variations in the speed of governance adaptation, and firm performance. *Academy of Management Review* 30: 414–425.

Leone, R.A. 1986. *Who profits: Winners, losers, and government regulation*. New York: Basic Books.

Levine, M.E. 1987. Airline competition in deregulated markets: Theory, firm strategy, and public policy. *Yale Journal on Regulation* 4: 393–494.

Madsen, T.L., and G. Walker. 2002. The evolution of heterogeneity in performance. *Academy of Management Best Papers Proceedings*, BPS V1-V6.

Madsen, T.L., and G. Walker. 2007. Incumbent and entrant rivalry in a deregulated industry. *Organization Science* 18: 667–687.

Madsen, T.L., and G. Walker. 2013. Heterogeneity between cohorts: Inertia and adjustment to deregulation. Working paper, Southern Methodist University and Santa Clara University.

Mahon, J.F., and E.A. Murray. 1981. Strategic planning for regulated companies. *Strategic Management Journal* 2: 251–262.

Merton, R.K. 1936. The unanticipated consequences of purposive social action. *American Sociological Review* 1: 894–904.

North, D.C. 1990. *Institutions, institutional change and economic performance*. New York: Cambridge University Press.

Peltzman, S. 1993. Georgie Stigler's contribution to the economic analysis of regulation. *Journal of Political Economy* 101: 818–832.

Peteraf, M., and R. Reed. 2008. Regulatory reform and managerial choice: An analysis of the cost savings from airline deregulation. *Managerial and Decision Economics* 29: 99–116.

Reger, R.K., I.M. Duhaime, and J.L. Stimpert. 1992. Deregulation, strategic choice, risk and financial performance. *Strategic Management Journal* 13: 189–204.

Stigler, G.J. 1971. The theory of economic regulation. *Bell Journal of Economics and Management Science* 2: 3–21.

Thomas, L.G. 1990. Regulation and firm size: FDA impacts on innovation. *Research Policy* 21: 497–517.

Walker, G., T.L. Madsen, and G. Carini. 2002. How does institutional change affect heterogeneity in performance among firms? *Strategic Management Journal* 23: 89–104.

WHO (World Health Organization). 2011. Available at http://www.who.int/about/en/. Accessed 13 Mar 2011.

Wiggens, S. 1979. Product quality regulations and new drug introductions: Some new evidence from the 1970s. *Review of Economics and Statistics* 63: 615–619.

Winston, C. 1998. U.S. industry adjustment to economic deregulation. *Journal of Economic Perspectives* 12: 89–110.

Repeated Games

Rajeev Bhattacharya
Berkeley Research Group, Boston, MA, USA

Abstract

A repeated game is where the stage game is repeated a number of times – the number of repetitions could be finite or infinite. We usually assume that (a) the stage game has a finite number of players, (b) for each player, the set of feasible actions for the stage game is finite, and (c) the stage game is simultaneous. For a finitely repeated game, the only subgame perfect equilibria of the repeated game are where a Nash equilibrium of the stage game is played in each period. For an infinitely repeated game, every feasible payoff vector that strictly

dominates the players' minmax values can be sustained in equilibrium for sufficiently high discount factors.

Definition A repeated game Γ^T is where the stage game Γ is repeated T times. When T is finite, the repeated game Γ^T is called a finitely repeated game, and when T is infinite, the repeated game Γ^∞ is called an infinitely repeated game.

Notations

Let $\Gamma = \left(\{1, 2, \ldots N\}, (A_n)_{n=1}^N, (\pi_n)_{n=1}^N\right)$, where $\{1, 2, \ldots N\}$ is the set of players; and for each $n = 1,2, \ldots, N$, A_n is the action space of player n, $\pi_n : \times_{n=1}^N A_n \to R$ is the payoff function of player n. In other words, a stage game is defined by the set of players, the action space for each player and the payoff function for each player, which provides the real-valued utility that a player receives for every combination of actions of all players.

We usually assume that (a) N is finite; (b) for each n, A_n is finite; and (c) Γ is a simultaneous move game. We also assume that if actions $(a^t)_{t=1}^T$, where $a^t = \left(a_1^t, a_2^t, \ldots, a_N^t\right)$, are taken by the players $n = 1, 2, \ldots, N$ in the stages $t = 1, 2, \ldots, T$, then player n's payoff function for the repeated game will be $\sum_{t=1}^T \delta_n^t \pi_n(a^t)$, where $\delta_n \in (0, 1)$ is the discount factor of player n.

Let $\delta = (\delta_1, \delta_2, \ldots, \delta_N)$–in order to make the dependence on the discount factors explicit, we will denote a finitely repeated game with T repetitions and with discount factors $\delta = (\delta_1, \delta_2, \ldots, \delta_N)$ by $\Gamma^T(\delta)$ and an infinitely repeated game with discount factors $\delta = (\delta_1, \delta_2, \ldots, \delta_N)$ by $\Gamma^\infty(\delta)$.

Histories and Subgames

Suppose that the game begins in period 1, with the null history h^1. For period $t \geq 2$, let $h^t = (a^1, a^2, \ldots, a^{t-1})$ be the period-t history of choices of actions before period t, and let H^t be the set of all such possible period-t histories. A strategy for player n is a sequence of functions $\left(s_n^t\right) : H^t \to \mathbf{A}_n$, where $\mathbf{A}_n$ is the set of probability distributions over A_n. Such a probability distribution is called a mixed strategy, so a pure strategy is a degenerate special case of a mixed strategy where the particular action takes place with probability one.

Note that for a strategy profile to be well-defined, it must specify probabilities over actions (mixed strategies) for all histories, not just the histories that would occur as a result of the strategy profile.

Every history begins a new subgame, and for a strategy profile to be subgame perfect, each such profile of functions needs to be a ▶ Nash Equilibrium for the corresponding subgame, on and outside the equilibrium path.

Discount Factor

A player's discount factor is a combination of a function of the player's own cost of capital (i.e., the rate at which the player discounts future utility) and the player's subjective probability of survival into the future.

Finitely Repeated Game

All players know the finite time at which the game will terminate.

In such a case, the only subgame perfect equilibria of the repeated game are where a Nash equilibrium of the stage game – not necessarily the same Nash equilibrium – is played in each period.

Consider the special case where the stage game Γ has a unique Nash equilibrium. Then, for a finitely repeated game, the only subgame perfect equilibrium is where the unique Nash equilibrium is repeated in every stage.

Infinitely Repeated Game

Consider the condition that for each player, at each point of time, there be a positive probability bounded away from zero that there will be another round of the game. Given the interpretation of

the discount factor mentioned earlier, such a condition is sufficient for an *indefinite* game to be modeled as an infinitely repeated game. In particular, if there is a constant probability that there will be another round of the game, the repeated game can be modelled as an infinitely repeated game.

Let player n's minmax value be $\underline{v_n} = \min_{a_{-n}}[\max_{a_n}\pi_n(a_n, a_{-n})]$. It is the payoff that player n can guarantee itself in each stage game.

The following "folk theorem" asserts that every feasible payoff vector that strictly dominates the players' minmax values can be sustained in Nash equilibrium for sufficiently high discount factors. It is called a folk theorem because for a time before it was actually proven it was assumed by folk wisdom to be true.

Theorem (*folk theorem*): For every feasible payoff vector v such that $v_n > \underline{v_n}, \forall n$, $\exists \underline{\delta} \in (0, 1)$ such that $\delta_n \in (\underline{\delta}, 1) \forall n \Rightarrow \exists$ Nash equilibrium of $\Gamma^{\infty}(\delta)$ with payoff vector v.

Let $V =$ convex hull of $\{v : \exists a \in A$ such that $\pi(a) = v\}$. In other words, V is the set of convex combinations of all utilities that can be achieved by actions of the players.

The following classic "Nash-threats folk theorem" shows that any payoff profile that strictly dominates a stage-game Nash equilibrium can be sustained in subgame perfect equilibrium, for sufficiently high discount factors.

Theorem (Friedman 1971): Let a^* be a Nash equilibrium of the stage game Γ with payoff vector v^*. Then, for any $v \in V$ such that $v_n > v_n^* \forall n, \exists \underline{\delta} \in (0, 1)$ such that $\delta_n \in (\underline{\delta}, 1) \forall n \Rightarrow \exists$ subgame perfect equilibrium of $\Gamma^{\infty}(\delta)$ with payoff vector v.

The following theorem asserts that if the feasible payoff space is of the same dimensionality as the number of players, than any payoff profile that strictly dominates the minmax values can be sustained in subgame perfect equilibrium, for sufficiently high discount factors.

Theorem (Fudenberg and Maskin 1986): Let dimension of V be equal to N, the number of players. Then, for every payoff vector $v \in V$ such that $v_n > \underline{v_n}, \forall n, \exists \underline{\delta} \in (0, 1)$ such that $\delta_n \in (\delta, 1) \forall n \Rightarrow \exists$ subgame perfect equilibrium of $\Gamma^{\infty}(\delta)$ with payoff vector v.

Repeated Prisoner's Dilemma

In each stage, the two firms Row Player (A) and Column Player (B) in a market are not price-takers. Each firm can either charge high price (cooperate) or charge low price (not cooperate). If they charge equal prices, they take equal shares of the market at that price. If they do not charge equal prices, the firm with the lower price takes the entire market at the lower price. The market revenues are \$200 m at the high price, and \$160 m at the low price. Neither firm can observe the other firm's decision when it has to make its decision, and costs are negligible relative to revenues. In the payoff matrix, it is conventional to put the row player's payoff first.

Stage Game Prisoner's Dilemma – Payoff Matrix

	Column player (B)		
		High price (cooperate)	Low price (not cooperate)
Row player (A)	High price (cooperate)	\$100 m, \$100 m	\$0, \$160 m
	Low price (not cooperate)	\$160 m, \$0	\$80 m, \$80 m

- Both firms are better off when they charge high prices than when they charge low prices (i.e., cooperation by both firms Pareto-dominates non-cooperation by both firms).
- Pricing low is the dominant strategy for each firm (i.e., no matter what the other firm does, it is in each firm's interest to not cooperate in the stage game).
- Both firms charging low is the unique Nash equilibrium (i.e., the only mutually self-enforcing pair of strategies in the stage game is where each firm does not cooperate). This is also the minmax outcome for each player.

From the previous discussions, non-cooperation in every stage is the only subgame perfect equilibrium of a finitely repeated ▶ Prisoner's Dilemma. For an oligopoly with a definite termination date, cooperation is not sustainable in

subgame perfect equilibrium. However, in experiments with known large but finite numbers of repetitions, we often find that players cooperate in the initial rounds.

The folk theorems imply that for the infinitely repeated Prisoner's Dilemma, any payoff vector higher than the non-cooperative payoffs is sustainable in equilibrium, for sufficiently high discount factors – in particular, the cooperative outcome is sustainable in equilibrium, for sufficiently high discount factors. For an oligopoly, these theorems mean that for low enough costs of capital and high enough probabilities of survival, tacit cooperation is sustainable in subgame perfect equilibrium. For example, the repeated interactions between Coca-Cola and Pepsi-Co in the beverage industry can be modelled as a repeated Prisoner's Dilemma. We can calculate the conditions under which implicit cooperation ("win-win") between the firms can be sustained in equilibrium.

Dynamic Strategy

A "good" dynamic strategy needs to score highly along the following dimensions:

- Clarity: it needs to be simple
- Niceness: it should not initiate deviating from a cooperative outcome
- Provocability: it should not let deviation from a cooperative outcome go unpunished
- Forgiving: it should not "hold a grudge" for too long.

Examples of Dynamic Strategy

The "grim trigger" or reversion strategy profile is:

- In every period t, price high (cooperate) if the other firm has charged high prices (cooperated) in each previous period
- Price low (do not cooperate) otherwise The grim trigger strategy is, therefore,
- Clear: absolutely
- Nice: only in the sense that it does not initiate non-cooperative behaviour
- Provocable: it punishes every single deviation from cooperative behaviour
- Completely unforgiving.

The grim trigger strategy profile is conceptually useful in the sense that it provides us with a bound of what is sustainable in subgame perfect equilibrium – in particular, if something is not sustainable through a grim trigger strategy profile, it is not likely to be sustainable in subgame perfect equilibrium through any other dynamic strategy. However, because of its draconian nature, it is unlikely to be used in a real-life situation. Consider the following punishment: if a driver is caught going even one mile per hour above the speed limit, the driver loses her/his licence for life. If this cannot stop speeding, very few other strategies can. However, such a draconian punishment is unlikely to be acceptable to society.

"Tit for tat" is another important dynamic strategy: cooperate in the first period and mimic the opponent's behaviour from the previous period. It scores highly on all the four criteria of clarity, niceness, provocability and forgiving. It is a robust strategy – it has performed well in competitive tournaments studied by Axelrod (2006). It manages to encourage cooperation, whenever possible, while avoiding exploitation. It can, however, start an escalation process.

A variant of the tit-for-tat strategy is the "What have you done for me lately?" strategy:

- Begin cooperating
- Continue cooperating
- Count how many times your opponent has not cooperated even though you have
- If the above proportion becomes "unacceptable", revert to tit for tat. Of course, determining what is "unacceptable" can be critical. Dixit and Nalebuff (1993) suggest the following: start cooperating, and continue to do so until one of the four tests below fails.
- First impression: non-cooperation on the first move is unacceptable; revert to tit for tat
- Short term: two non-cooperative moves in any three consecutive turns are unacceptable; revert to tit for tat

- Medium term: three non-cooperative moves out of the last 20 periods are unacceptable; revert to tit for tat
- Long term: five non-cooperative moves out of the last 100 periods are unacceptable; revert to tit for tat.

This strategy scores higher than tit for tat on the nicety and forgiving tests, but lower on the provocability and clarity tests.

See Also

- ▶ Multistage Games
- ▶ Nash Equilibrium
- ▶ Prisoner's Dilemma

References

Axelrod, R. 2006. *The evolution of cooperation*. New York: Basic Books.
Dixit, A., and B. Nalebuff. 1993. *Thinking strategically*. New York: Norton.
Friedman, J. 1971. A non-cooperative equilibrium for supergames. *Review of Economic Studies* 38: 1–12.
Fudenberg, D., and E. Maskin. 1986. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 54: 533–554.
Fudenberg, D., and J. Tirole. 1991. *Game theory*. Cambridge, MA: The MIT Press.

Reputation

Violina P. Rindova
University of Texas at Austin, McCombs School of Business, Austin, TX, USA

Abstract

Management research on corporate reputation has grown exponentially in the last 20 years and has generated a proliferation of definitions and debates about what reputation is. These debates reflect often unacknowledged differences in the theoretical perspectives that have guided current research. This contribution provides a review of the different theoretical perspectives and their implications for the study and management of reputations, and highlights the recent trend towards developing integrative, multi-dimensional theoretical and empirical approaches.

Definition Corporate reputation refers to collective perceptions and beliefs held in a given social collective – stakeholder group, organizational field or society at large – about specific firm attributes, or the firm's general ability to create value for diverse stakeholders across multiple dimensions of performance.

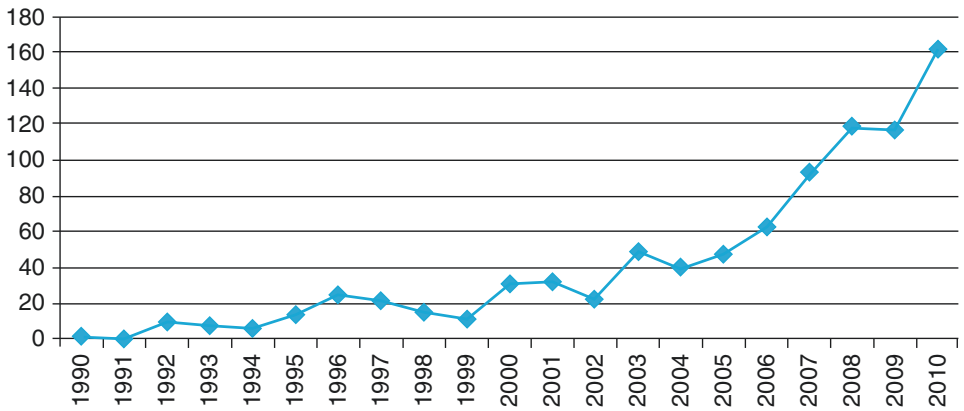
Research on corporate reputations has grown exponentially in the last two decades. In the field of strategic management the concept was introduced in the late 1980s in three articles that appeared around the same time. These articles represented the three core perspectives that have come to guide research in the area since. Weigelt and Camerer (1988) introduced a game-theoretic perspective on reputation, which they defined as 'a set of attributes ascribed to a firm, inferred from the firm's past actions' (Weigelt and Camerer 1988: 443). The game-theoretic perspective emphasizes that a firm's reputation enables observers to predict the likely future behaviour of a firm by observing its current actions and drawing inferences about its strategic type. In 1989, in a paper on the accumulation of ▶ intangible assets, Dierickx and Cool identified reputation as an intangible asset to the firm. They argued that reputation is an asset that is accumulated over time through consistent and sustained policies, practices and actions. In 1990, Fombrun and Shanley published, in the *Academy of Management Journal*, the first empirical study focusing on the effect of different types of information signals on the development of firm reputations. They conceptualized firm reputations as socially constructed assessments based on signals about firms' activities, achievements and prospects that are derived from diverse information sources, including the firms themselves, the media and other monitors.

The core ideas of these papers reflect three distinct conceptualizations of firm reputations rooted in three different theoretical perspectives: the game-theoretic view of strategy, the institutional social-constructivist view and the ▶ resource-based view. In the two decades since the publication of these papers, management research on reputations has grown exponentially, with scholars drawing on these different perspectives more or less explicitly. Figure 1 documents the exponential growth in research that features reputation as a 'main topic' published between 1990 and 2010 in the management discipline journals tracked in the *Web of Science* database.

This growth in reputation research has been accompanied by a proliferation of definitions (see Rindova et al. 2005, for a review) and debates about what reputation actually is. For example, debates exist about whether reputation consists of specific perceptions about specific unobservable attributes (e.g., the ability to deliver quality) (Deutsch and Ross 2003; Rhee and Haunschild 2006), or whether it encompasses broad collective knowledge and recognition of the firm (Fombrun 1996; Rindova and Fombrun 1999; King and Whetten 2008); about whether a firm has one reputation or many, that is, whether a firm's reputation integrates perceptions and beliefs across stakeholder groups, or whether distinct stakeholder groups, such as customers or alliance partners, hold distinct sets of perceptions and beliefs about firms; and about whether reputations are perceptions and beliefs in the minds of stakeholder audiences, or whether they exist as institutionalized assessments represented in reputational rankings. Recently, management scholars have begun pursuing more integrative approaches towards the study of reputation by considering the multi-dimensionality of the construct and its implications for understanding corporate reputations as strategic assets (Rindova et al. 2005; Lange et al. 2010; Rindova and Martins 2012). Below I review the three core theoretical perspectives that have guided management reputation research to date, and the recent developments in the area.

The Game-Theoretic View

Economics research from a game-theoretic perspective has identified reputations as important features of repeated competitive interactions with incomplete information (Milgrom and Roberts 1982, 1986; Shapiro 1983; Weigelt and Camerer 1988). Under conditions of incomplete information players are assumed to know their own type, but to be uncertain about the types of others. Determining other players' types is an important strategic issue for competitors who seek to predict each other's competitive actions.



Reputation, Fig. 1 Number of articles with reputation as a main topic (Adapted from Thompson ISI *Web of Science* database)

The observed past behaviours provide clues about underlying unobservable strategic types and serve as the basis of players' reputations. Further, under conditions of incomplete information players can use specific actions to signal specific desirable attributes, such as competitive toughness or ability to produce quality (Milgrom and Roberts 1982; Shapiro 1983; Weigelt and Camerer 1988). For example, Yamey (1972) suggested that firms can use predatory pricing to signal their likely reaction to future entry attempts. Kreps and Wilson (1982) and Milgrom and Roberts (1982) developed reputation-building models to formalize this idea. Milgrom and Roberts (1986) examine the signalling effects of pricing and advertising and show that a combination of these two strategic signals reliably identifies high- and low-quality producers. More generally, the patterns of strategic actions that firms undertake reveal information about the underlying unobservable attributes of firms because different attributes lead to different incentives and/or capabilities to take particular types of actions (Basdeo et al. 2006; Rindova et al. 2007).

The game-theoretic perspective therefore treats reputations as reliable mechanisms for signalling information about otherwise unobservable firm attributes. This information enhances the predictability of economic exchanges between the firm and a specific set of players, such as competitors, consumers or exchange partners. It enables these actors to predict the behaviour of the firm with regard to a specific attribute they value. This view of reputations has several implications for how reputations are understood and studied. First, because players have incentives to take actions that imitate a desirable strategic type, actions that contribute effectively to reputation-building have to meet certain criteria to be credible signals, such as costliness, and not being generally available. Second, observers must be in a position to observe and understand the strategic implications of a given type of action. For example, competitors in concentrated industries are in a better position to observe and interpret each other's actions than a set of consumers in fragmented markets. Third, depending on the type of actions being tracked, the same firm can have different reputations for different attributes with different observers, if different audiences assign differential value to different actions. This perspective, therefore, is particularly effective in analysing the effects of specific types of actions on the reputation of a firm with specific audiences. However, it has limited value for understanding a firm's overall reputational standing in a given organizational field or society, or for assessing the value of a firm's reputation as an intangible asset. These issues are the focal interest of the institutional-constructivist and resource-based views discussed next.

The Institutional Social-Constructivist View

In contrast to the game-theoretic view, which is interested in the specific reputational effects of specific types of signals on specific audiences, the constructivist view emphasizes the broad circulation and dissemination of diverse types of information about firms in markets (Fombrun and Shanley 1990; Rindova and Fombrun 1999). Because various market actors, and not only interested firms, create and disseminate information about firms, and because market audiences interact and take cues from each other's choices (Pollock et al. 2008), firm reputations are understood as outcomes of social construction processes (Rao 1994). As Fombrun and Shanley (1990: 234) explain:

> Publics construct reputations from available information about firms' activities originating from the firms themselves, from the media, or from other monitors. Publics use and propagate information they deem important for assessing firms' successes and failures at acquiring resource inputs, improving throughputs, and sustaining outputs. As signals about firms' activities, achievements, and prospects diffuse, individual interpretations aggregate into collective judgments.

This view therefore portrays reputations as aggregations of diverse perceptions and cognitions varying from simple awareness (Shamsie 2003), to general impressions (Rao 1994),

to attribute-specific knowledge (Rindova et al. 2005). From this perspective, firm reputation is not necessarily traceable to specific actions and observations, even if, in general, it tends to reflect the patterns in the history of a firm's actions (Rindova et al. 2007). Further, reputations are developed not only on the basis of the actions of firms, but on the basis of communications among diverse audiences who, despite their different interests and agendas, participate in a common information environment and share opinions and perspectives about firms. Some scholars have argued that, understood in those terms, reputation becomes decoupled from actual strategic actions and behaviours and becomes less rational and more similar to legitimacy (Rao 1994). Others argue, however, that even if these processes decouple reputation from specific actions, reputation preserves its uncertainty reducing function for firms' audiences through the collective verification processes that information exchanges among stakeholder audiences enable.

The institutional social-constructivist view therefore is concerned not as much with the evaluation of a firm by a specific individual or audience, but with its overall reputational standing in an organizational field. This perspective has highlighted two aspects of this reputational standing – relative prominence (or visibility) and favourability (Deephouse 2000; Rindova et al. 2005). The former describes the collective attention allocated to a given firm in its organizational field, whereas the latter describes the general favourability of collective perceptions. Barnett et al. (2006: 34) describe the general favourability as arising from 'assessments of the financial, social, and environmental impacts attributed to the corporation over time'. Some debate exists in the literature as to whether the favourability of collective evaluations includes an affective component (Cable and Graham 2000; Rhee and Valdez 2009), or whether the emotional responses to firms constitute an entirely different type of intangible asset termed 'firm celebrity' (see Rindova et al. 2006; Pfarrer et al. 2010 for a discussion of the differences between the two).

This perspective also has distinct implications for how reputations are understood and studied. First, reputations are understood not as individual assessments, but as accumulations of opinions and beliefs in the organizational field. As such, reputations must be studied either by aggregating individual opinions through the use of opinion polling techniques (e.g., Rindova et al. 2005), or from their representations in various reputational rankings (e.g., Martins 2005), which sometimes include opinion polls. Second, reputational rankings are seen not only as reflections of underlying collective opinions, but as representations that give collective opinions a quasi-objective status of 'social facts' (Rao 1994; Martins 2005). Consistent with the idea that reputational rankings objectify a firm's reputation, reputation researchers have extensively used reputational rankings as the measure of a firm's reputation (e.g., Fombrun and Shanley 1990; Basdeo et al. 2006; Roberts and Dowling 2002). However, because firms' positions in rankings depend on numerous measurement issues, reputations defined by reputational hierarchies have also been subject to controversy and critique (Gioia and Corley 2002; Dichev 1999; Martins 2005; Elsbach and Kramer 1996).

The Resource-Based View

The resource-based view of reputation as an intangible asset of the firm has been highly influential and has inspired a voluminous body of research investigating the effects of corporate reputations on firm performance (Deephouse 2000; Hall 1993; Jensen and Roy 2008; Podolny 2005; Rao 1994; Rindova et al. 2005; Roberts and Dowling 2002). Despite its importance in stimulating research on the reputation-performance relationship, the resource-based view of corporate reputations has remained somewhat theoretically underdeveloped. Resource-based theorists have stressed the importance of reputational assets for firm performance and sustainability of competitive advantage, but have not theorized the attributes of reputation that determine its value as an intangible asset. For example, in the paper

articulating a theory of how intangible assets accumulate, Dierickx and Cool (1989: 1506) argue that reputation, as a strategic asset, is 'the cumulative result of adhering to a set of consistent policies over a period of time'. Thus, they clarify that reputational assets are characterized by a level of accumulation which affects their ultimate impact on firm performance, but do not offer a general theoretical statement about the attributes of reputations that determine their value as intangible assets. Similarly, Barney (1991) argues that positive reputations may be rare and difficult to imitate, but does not discuss the nature of heterogeneity in reputational attributes that affect their rarity or ▶ imitability as resources.

The resource-based theory has advanced the idea that reputation is an intangible asset to the firm but has not theorized the attributes of reputations that define their asset quality and value. In fact, in a recent review of reputation research Lange et al. (2010: 162) observe that: 'most authors who make the case for organizational reputation being defined as an asset do so by detailing reputation's positive outcomes for the firm'. They further note that 'this practice seems to make the idea of organizational reputation as asset more of a description of the consequences of the concept than a definition of the concept', and call for considering 'what the idea of asset implies beyond reputation's positive outcomes for the firm'.

Rindova and Martins (2012) respond to this call and propose that the different theoretical perspectives discussed above provide the basis for articulating four dimensions, along which the composition of reputational assets can be characterized and analysed. Their approach builds on the study by Rindova et al. (2005), which demonstrates that prominence and perceived quality are two dimensions of reputation with distinct predictors and consequences. These findings are important because they imply that both the level of accumulation and content of perceptions need to be accounted for in studying how reputations create value and function as assets. Rindova and Martins (2012) extend these ideas to a multi-dimensional view of reputational assets that analyses them in terms of four dimensions: asset specificity defined as the extent to which reputations facilitate the interaction between a firm and a specific audience, as theorized by the game-theoretic view; asset accumulation defined as the level of collective attention a firm commands; asset breadth of appeal defined by the scope of the favourable assessments held about the firm; and asset codification: capturing the extent to which a firm's reputation is externalized and objectified in reputational rankings. Since each dimension is associated with different economic consequences, using a multi-dimensional approach should enable scholars and managers to form clear expectations regarding the effects of reputations with different characteristics on firm performance. Given the fact that a multi-dimensional approach yields both more complex and specific theory and empirical measurement, it is not surprising that, in reviewing the literature, Lange et al. (2011: 160) conclude that the development of integrative, multi-dimensional approaches to the study of reputation is 'a strong and emerging trend in management research'. This trend promises to increase the theoretical richness and empirical rigour of future studies on reputation, as the multi-dimensional analysis of reputations enables researches and managers to simultaneously take into account the multiple and diverse perceptions that contribute to corporate reputations in the marketplace.

See Also

- ▶ Game Theory
- ▶ Imitability
- ▶ Intangible Assets
- ▶ Intangible Resources
- ▶ Resource-Based View

References

Barnett, M.L., J.M. Jermier, and B.A. Lafferty. 2006. Corporate reputation: The definitional landscape. *Corporate Reputation Review* 9: 26–38.

Barney, J. 1991. Firm resources and sustained competitive advantage. *Journal of Management* 17: 99–120.

Basdeo, D.K., K.G. Smith, C.M. Grimm, V.P. Rindova, and P.J. Derfus. 2006. The impact of market actions on firm reputation. *Strategic Management Journal* 27: 1205–1219.

Cable, D.M., and M.E. Graham. 2000. The determinants of job seekers' reputation perceptions. *Journal of Organizational Behavior* 21: 929–947.

Deephouse, D.L. 2000. Media reputation as a strategic resource: An integration of mass communication and resource-based theories. *Journal of Management* 26: 1091–1112.

Deutsch, Y., and T.W. Ross. 2003. You are known by the directors you keep: Reputable directors as a signaling mechanism for young firms. *Management Science* 49: 1003–1017.

Dichev, I. 1999. How good are business school rankings? *Journal of Business* 72: 201–213.

Dierickx, I., and K. Cool. 1989. Asset stock accumulation and sustainability of competitive advantage. *Management Science* 35: 1504–1511.

Elsbach, K.D., and R.M. Kramer. 1996. Members' responses to organizational identity threats: encountering and countering the *Business Week* rankings. *Administrative Science Quarterly* 41: 442–476.

Fombrun, C.M. 1990. What's in a name? Reputation building and corporate strategy. *Academy of Management Journal* 33: 233–258.

Fombrun, C.J. 1996. *Reputation: Realizing value from the corporate image*. Boston: Harvard Business School Press.

Gioia, D.A., and K.G. Corley. 2002. Being good versus looking good: Business school rankings and the Circean transformation from substance to image. *Academy of Management Learning & Education* 1: 107–120.

Hall, R. 1993. A framework linking intangible resources and capabilities to sustainable competitive advantage. *Strategic Management Journal* 14: 607–618.

Jensen, M., and A. Roy. 2008. Staging exchange partner choices: When do status and reputation matter? *Academy of Management Journal* 51: 495–516.

King, B.G., and D. Whetten. 2008. A social identity formulation of organizational reputation and legitimacy. *Corporate Reputation Review* 11: 192–207.

Kreps, D.M., and R. Wilson. 1982. Reputation and imperfect information. *Journal of Economic Theory* 27: 253–279.

Lange, D., P.M. Lee, and Y. Dai. 2010. Organizational reputation: A review. *Journal of Management* 37: 153–184.

Martins, L.L. 2005. A model of the effects of reputational rankings on organizational change. *Organization Science* 16: 701–720.

Milgrom, P., and J. Roberts. 1982. Predation, reputation, and entry deterrence. *Journal of Economic Theory* 27: 280–312.

Milgrom, P., and J. Roberts. 1986. Price and advertising signals of product quality. *Journal of Political Economy* 94: 796–821.

Pfarrer, M.D., T.G. Pollock, and V.P. Rindova. 2010. A tale of two assets: The effects of firm reputation and celebrity on earnings surprises and investors' reactions. *Academy of Management Journal* 53: 1131–1152.

Podolny, J. 2005. *Status signals: A sociological study of market competition*. Princeton: Princeton University Press.

Pollock, T.G., V.P. Rindova, and P.G. Maggitti. 2008. Market watch: Information and availability cascades in the U.S. IPO market. *Academy of Management Journal* 51: 335–358.

Rao, H. 1994. The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the American automobile industry 1895–1912. *Strategic Management Journal* 15: 29–44.

Rhee, M., and P.R. Haunschild. 2006. The liability of good reputation: A study of product recalls in the US automobile industry. *Organization Science* 17: 101–117.

Rhee, M., and M.E. Valdez. 2009. Contextual factors surrounding reputation damage with potential implications for repuation repair. *Academy of Management Review* 34: 146–168.

Rindova, V.P., and C.J. Fombrun. 1999. Constructing competitive advantage: The role of firm-constituent interactions. *Strategic Management Journal* 20: 691–710.

Rindova, V.P., and L.L. Martins. 2012. Show me the money: A multi-dimensional perspective on reputation as an intangible asset. In *Oxford handbook of reputation management*, ed. M. Barnett and T. Pollock. Oxford: Oxford University Press.

Rindova, V.P., I.O. Williamson, A.P. Petkova, and J.M. Sever. 2005. Being good or being known: An empirical examination of the dimensions, antecedents, and consequences of organizational reputation. *Academy of Management Journal* 48: 1033–1049.

Rindova, V.P., T.G. Pollock, and M.L.A. Hayward. 2006. Celebrity firms: The social construction of market popularity. *Academy of Management Review* 31: 50–71.

Rindova, V.P., A.P. Petkova, and S. Kotha. 2007. Standing out: How new firms in emerging markets build reputation. *Strategic Organization* 5: 31–70.

Roberts, P.W., and G.R. Dowling. 2002. Corporate reputation and sustained superior financial performance. *Strategic Management Journal* 23: 1077–1093.

Shamsie, J. 2003. The context of dominance: An industry-driven framework for exploiting reputation. *Strategic Management Journal* 24: 199–215.

Shapiro, C. 1983. Premiums for high quality products as returns to reputations. *Quarterly Journal of Economics* 98: 659–679.

Weigelt, K., and C. Camerer. 1988. Reputation and corporate strategy: A review of recent theory and applications. *Strategic Management Journal* 9: 443–454.

Yamey, B.S. 1972. Predatory price cutting: Notes and comments. *Journal of Law and Economics* 15: 129–147.

Research and Development (R&D) Alliances

Fabio Pammolli[1], Massimo Riccaboni[1] and Gianna G. Giudicati[2]
[1]IMT Institute for Advanced Studies, Lucca, Italy
[2]University of Trento, Trento, Italy

Abstract

Innovation systems are characterized by a high level of inter-organizational collaboration in R&D activity. There is evidence that external collaborations and ▶ alliances have increased dramatically during the past decades. A significant proportion of such collaboration takes place between organizations of different type, size, age and location, which are endowed with complementary competences and resources. In this entry we examine the literature to first analyse bilateral contractual agreements. Second, we investigate the origin, structure and dynamics of R&D networks as a persistent trait of research arenas characterized by strong uncertainty and rapid technological change.

Definition R&D alliances are inter-organizational contractual relationships though which public and private research institutions subdivide innovative labour to perform R&D activities based on their resources and competences.

Introduction and History of the Concept

The organization of ▶ research and development (R&D) has been constantly reshaped by innovation and the evolution of knowledge bases. The dispersion of relevant knowledge among firms and other institutions worldwide implies that no single actor has all the necessary skills and competences to innovate (Powell et al. 1996; Ozman 2009). Therefore, the network of actors playing a role in the division of innovative labour has considerably expanded (Mowery 1999) and the organization of R&D has been radically transformed over the past few decades, especially in industries characterized by radical uncertainty, such as the life sciences (Arora and Gambardella 1990, 1994a; Pisano 1991; Powell et al. 1996; Orsenigo et al. 2001). R&D contracts and interorganizational ▶ alliances are needed to access various resources and competences, and to contribute to knowledge generation, recombination and transfer.

R&D alliances are inter-firm relationships through which actors subdivide innovative labour to perform subsets of R&D activities based on their resources and competences. Innovation emerges through the ▶ recombination of knowledge in formal or informal relations. The relationship between technological progress and division of labour among firms can be traced back to Adam Smith. Smith was the first to demonstrate that specialization and division of labour are the main sources of economic and technological progress (Smith 1776). In Smith's view, efficiency in specialization is based on dynamic learning by doing (Kogut 2000). Increased specialization allows problems to be decomposed and solved (Simon 1962): emerging forms of division of labour reflect the structure of the research spaces that are explored by economic agents and the evolution of problem-solving strategies in innovation contexts (Orsenigo et al. 2001; Riccaboni and Pammolli 2002). Division of labour also gives rise to a class of 'meer men of speculation', capable of acquiring general-purpose skills and recombining knowledge to create innovation. This 'power to recombine distant and dissimilar objects' (Smith 1776: 11) can take advantage of both technical specialization and knowledge recombination (Young 1928; Stigler 1951; Helpman 1998). Breakthroughs and architectural and radical innovations emerge by recombining knowledge bases and reconfiguring the knowledge space. As the market grows, 'general specialties' are introduced and used across different domains of application (Young 1928; Stigler 1951; Arora and Gambardella 1994; Bresnahan and Trajtenberg 1995; Arora et al. 2001; Gambardella and McGahan 2010).

The Austrian School has also focused on the intellectual aspect of the division of labour among economic agents. It introduced the notion of 'catallaxy' to emphasize the properties of a market as a coordination device among individual capabilities and goals (Hayek 1945; Mises 1998). New perspectives for future research come from the application of '*game semantics*' to R&D alliances between verifiers and falsifiers (Hintikka 1998).

R&D Alliances, Uncertainty and Contracts

R&D alliances may be implemented through the stipulation of contracts. Such R&D contracts usually involve at least two parties, one of which (the agent) typically performs research, while the other (the principal) usually has a broader scope (Arrow 1962, 1983; Arora and Gambardella 1990; Orsenigo et al. 2001). Recent research has also investigated the factors affecting the perception, access and participation in alliances by firms. Research at firm level has amply examined the risks and uncertainties of alliance formation, the extent to which actors collect information on potential partners, performance divergence across firms due to knowledge differentiation and resource heterogeneity, and the influence of previous networks on the effectiveness of new alliances (Gulati 1999; Gulati and Gargiulo 1999).

R&D alliances are preceded by a pre-contractual phase, aimed at defining motives and choosing partners. A firm's decision to enter an alliance is highly influenced by the need to minimize risks and moral hazard concerns due to the unpredictability of the behaviour of partners and possible opportunism and free-riding. Because of information asymmetry, ambiguities and risks, R&D alliances are often considered risky, despite the increasing number of partnerships (Kogut 1989; Gulati 1994). Grossman and Hart (1986) pointed out the importance of R&D contracts for the allocation of control and property rights over the R&D process through incomplete contracts (Pisano 1990). Transaction costs motivate the variety and complexity of contractual solutions for collaborative versus inhouse R&D (Aghion and Tirole 1994). R&D contracts allow knowledge to be transferred among partners, and also allow access to other complementary capabilities and improve organizational learning (Ozman 2009). Alliances also facilitate entry into new markets through the spread of risk among partners and the sharing of R&D costs (Mowery 1988). Powell et al. (1996) showed that R&D alliances are 'ambidextrous' organizational solutions for exploring new knowledge and exploiting existing competences and technologies (March 1991). Oliver (1990) also noted other motives, such as the need to conform with legal requirements or to adapt to environmental uncertainties, and the possibility of controlling partners and improving reputation and image.

In choosing R&D partners, firms consider similarities in know-how, the stage of each firm in its life-cycles, and the interdependence of resources as factors affecting the choice of a partner (Gulati and Singh 1998; Owen-Smith et al. 2002; Ozman 2009). Mimetic isomorphism and degree of similarity are additional explanations of the structure of alliances (DiMaggio and Powell 1983). Collaborations with domestic and foreign competitors are also formed to recombine knowledge and improve efficiency in technological and innovation processes in the industry in which firms operate (Mowery 1999).

Lastly, in the post-contractual phase, adequate incentive schemes must be structured in order to improve the productivity and efficiency of alliances (Aghion and Tirole 1994). Concerns as regards how parties adapt to changing circumstances or adjust the terms of the relationship to improve efficiency and knowledge-sharing are still underexplored (Powell et al. 1996). The termination of a relationship does not necessarily mean that the alliance has failed, but it can be a sign of the conclusion of collaborative activity, partly because of the achievement of partners' goals. In addition, other alliances can be formed with other external subjects to exploit dialectic relationships and critically analyse emerging issues (Freeman 1991). However, conflicting views appear in the literature on the advantages of external collaboration with respect to vertical integration. Pisano (1997) found potential 'lemon

problems', although whether collaborative agreements have a lower success rate than in-house R&D projects is still debated.

R&D Alliances as Networks of Innovators

Systems of inter-organizational contracts have been investigated as networks. By examining networks from a strategic management perspective, the literature has focused on the influence of inter-firm networks in various industries on the degree of innovation and performance. The variety of ties in which actors are involved in R&D alliances creates 'networks of innovators' (De Bresson and Amesse 1991; Freeman 1991; Orsenigo et al. 2001). However, networks of actors and firms have been labelled in various ways: business groups (Granovetter 1998), cooperative inter-organizational relationships (Oliver 1990), networks of learning (Powell et al. 1996), network organizations (Miles and Snow 1986) and inter-firm networks (Grandori and Soda 1995). The model of innovator networks has been analysed in many industries (Duysters and Hagedoorn 1995; Powell et al. 1996; Ahuja 2000) and developed over the past two decades as a way to reduce R&D costs, save time and increase flexibility. Networks also allow faster access to more diverse information and competences, increase the level of innovation inside firms (Powell et al. 1996) and improve efficient innovation appropriability (Dhanaraj and Parkhe 2006). The role of ▶ innovation networks in improving know ledge flows in R&D alliances and, in general, in collaboration alliances, has been widely investigated, together with the importance of network ties as key vehicles by means of which firms obtain access to external knowledge (Powell et al. 1996; Gulati 1999; Ahuja 2000). Networks are formed either by formal ties, such as subcontracts or participation in research consortia, or by informal ties, such as affiliations in technological or professional communities and associations. Recent studies have proved that the larger the number of ties and actors, the more the benefits increase, due to the decreased cost of network maintenance associated with direct ties (Burt 1992). Different types of ties positively influence the benefits derived from alliances and innovation performance (Powell and Owen-Smith 1999) due to the transfer of tacit knowledge. Powell et al. (1996: 66) suggested a 'cycle of learning process' in which R&D networks generate attention which attracts other collaboration partners by adding experience and competence in developing new ideas. A recent debate has also arisen on how to design an effective and efficient network structure aimed at increasing organizational benefits and innovative performances (Ahuja 2000). On one hand, a first stream of literature affirmed the need for closed, dense ties to improve the benefits of networking (Walker et al. 1997). On the other hand, other network scholars preferred an open structure, improving the benefits of brokerage behaviour (Coleman 1988; Walker et al. 1997).

In the last two decades, most of the research in economics, sociology and organizational sciences has focused on the relationships between the technological competences of firms, their capability to add innovation to the industry, and their role and position in R&D networks (Powell et al. 1996; Stuart and Podolny 1996; Stuart 1998; Orsenigo et al. 2001; Riccaboni and Pammolli 2002; Pammolli and Riccaboni 2004). The effect of the structure of R&D networks on firm performance has been examined to clarify the role of the various elements of the network structure in the innovation process. However, this effect also depends on the evolution of underlying technological processes.

A growing body of literature has examined the birth of networks of innovators, the characteristics of network structures, and the power of social network analysis in studying inter-organizational networks and strategic alliances (Freeman 1991; Powell et al. 1996; Granovetter 1998). Differing opinions emerge when future influences and the role of networks are investigated. One stream of literature identified networks as transitory and temporary phenomena, constantly overwhelmed by the power of the market and of single firms as the main sources of economic and technological progress (Pisano 1991). However, a description of the co-evolution of technological processes and organizations as a transition does not

acknowledge the long-term capacity of networks to accommodate both exploration and exploitation throughout different waves of technological change. Recent research has indeed proved that networks of innovators and external alliances through systems of contracts are long-term phenomena, fuelled over time by the entry of new firms which sustain high rates of specialization (Orsenigo et al. 2001; Riccaboni and Moliterni 2009).

See Also

- ▶ Alliances
- ▶ Innovation Networks
- ▶ Recombination of Knowledge
- ▶ Research and Development (R&D) Organization

References

Aghion, P., and J. Tirole. 1994. The management of innovation. *Quarterly Journal of Economics* 109: 1185–1209.

Ahuja, G. 2000. Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly* 45: 425–455.

Arora, A., and A. Gambardella. 1990. Complementarity and external linkages: The strategies of the large firms in biotechnology. *Journal of Industrial Economics* 38: 361–379.

Arora, A., and A. Gambardella. 1994. Evaluating technological information and utilizing it: Scientific knowledge, technological capability and external linkages in biotechnology. *Journal of Economic Behavior & Organization* 24: 91–114.

Arora, A., A. Fosfuri, and A. Gambardella. 2001. *Markets for technology: The economics of innovation and corporate strategy*. Cambridge, MA: The MIT Press.

Arrow, K.J. 1962. Economic welfare and the allocation of resources for invention. In *The rate and direction of inventive activity: Economic and social factors*, ed. H.M. Groves. Cambridge, MA: National Bureau of Economic Research.

Arrow, K.J. 1983. Innovation in large and small firms. In *Entrepreneurship*, ed. J. Ronen. Lexington: Lexington Books.

Bresnahan, T.F., and A. Trajtenberg. 1995. General purpose technologies: Engines of growth? *Journal of Econometrics* 65: 83–108.

Burt, R.S. 1992. *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.

Coleman, J.S. 1988. Social capital in the creation of human capital. *American Journal of Sociology* 94: 95–120.

De Bresson, C., and F. Amesse. 1991. Networks of innovators: A review and introduction to the issue. *Research Policy* 20: 363–379.

Dhanaraj, C., and A. Parkhe. 2006. Orchestrating innovation networks. *Academy of Management* 31: 659–669.

DiMaggio, P., and W. Powell. 1983. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review* 48: 147–160.

Duysters, G., and J. Hagedoorn. 1995. Strategic groups and inter-firm networks in international high-tech industries. *Journal of Management Studies* 32: 359–381.

Freeman, C. 1991. Networks of innovators: A synthesis of research issues. *Research Policy* 20: 499–514.

Gambardella, A., and A.M. McGahan. 2010. Business-model innovation: General purpose technologies and their implications for industry structure. *Long Range Planning* 43: 262–271.

Grandori, A., and G. Soda. 1995. Inter-firm networks: Antecedents, mechanisms and forms. *Organization Studies* 16: 183–214.

Granovetter, M. 1998. Coase revisited: Business groups in the modern economy. In *Technology, organization and competitiveness*, ed. G. Dosi, D.J. Teece, and J. Chytry. New York: Oxford University Press.

Grossman, S., and O. Hart. 1986. The costs and benefits of ownership: A theory of lateral and vertical integration. *Journal of Political Economy* 94: 691–719.

Gulati, R. 1994. Unilateral commitments and the importance of process in alliances. *Sloan Management Review* 35: 61–69.

Gulati, R. 1999. Network location and learning: The influence of network resources and firm capabilities on alliance formation. *Strategic Management Journal* 20: 397–420.

Gulati, R., and M. Gargiulo. 1999. Where do interorganizational networks come from? *American Journal of Sociology* 104: 1439–1493.

Gulati, R., and H. Singh. 1998. The architecture of cooperation: Managing coordination costs and appropriation concerns in strategic alliances. *Administrative Science Quarterly* 43: 781–814.

Hayek, F.A. 1945. The use of knowledge in society. *American Economic Review* 35: 519–530.

Helpman, E. 1998. *General purpose technologies and economic growth*. Boston: The MIT Press.

Hintikka, J. 1998. *The principles of mathematics revisited*. Cambridge, MA: Cambridge University Press.

Kogut, B. 1989. The stability of joint ventures: Reciprocity and competitive rivalry. *Journal of Industrial Economics* 38: 183–198.

Kogut, B. 2000. The network as knowledge: Generative rules and the emergence of structure. *Strategic Management Journal* 21: 405–425.

March, J. 1991. Exploration and exploitation in organizational learning. *Organization Science* 2: 71–87.

Miles, R., and C.C. Snow. 1986. Organizations: New concepts for new forms. *California Management Review* 28: 62–73.
Mises, L. 1998. *Human action: A treatise on economics*. Auburn: Ludwig von Mises Institute.
Mowery, D.C. 1988. *International collaborative ventures in US manufacturing*. Cambridge: Ballinger.
Mowery, D. C. 1999. America's industrial resurgence? An overview. In *U.S. Industry in 2000: Studies in competitive performance*. Washington, DC: National Academy Press.
Oliver, C. 1990. Determinants of interorganizational relationships: Integration and future directions. *Academy of Management Review* 15: 241–265.
Orsenigo, L., F. Pammolli, and M. Riccaboni. 2001. Technological change and network dynamics: Lessons from the pharmaceutical industry. *Research Policy* 30: 485–508.
Owen-Smith, J., M. Riccaboni, F. Pammolli, and W.W. Powell. 2002. A comparison of U.S. and European university–industry relations in the life sciences. *Management Science* 48: 24–43.
Ozman, M. 2009. Inter-firm networks and innovation: A survey of literature. *Economics of Innovation and New Technology* 18: 39–67.
Pammolli, F., and M. Riccaboni. 2004. Technological competences in networks of innovators. In *The economics and management of technological diversification*, ed. J. Cantwell, A. Gambardella, and O. Granstrand. New York: Routledge.
Pisano, G. 1990. The R&D boundaries of the firm: An empirical analysis. *Administrative Science Quarterly* 35: 153–176.
Pisano, G. 1991. The governance of innovation: Vertical integration and collaborative arrangements in the biotechnology industry. *Research Policy* 20: 237–249.
Pisano, G. 1997. *R&D performance, collaborative arrangements, and the market-for-know-how: A test of the 'lemons' hypothesis in biotechnology*. Working paper, Harvard Business School.
Powell, W.W., and J. Owen-Smith. 1999. Network position and firm performance. In *Research in the sociology of organizations*, ed. S. Andrews and D. Knoke. Greenwich: JAI Press.
Powell, W.W., K.W. Koput, and L. Smith-Doerr. 1996. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly* 41: 116–145.
Riccaboni, M., and R. Moliterni. 2009. Managing technological transitions through R&D alliances. *R&D Management* 39: 124–135.
Riccaboni, M., and F. Pammolli. 2002. On firm growth in networks. *Research Policy* 31: 1405–1416.
Simon, H.A. 1962. The architecture of complexity. *Proceedings of the American Philosophical Society* 106: 467–482.
Smith, A. 1776. *The wealth of nations*. New York: Modern Library.
Stigler, G.J. 1951. The division of labor is limited by the extent of the market. *Journal of Political Economy* 59: 185–193.
Stuart, T.E. 1998. Network positions and propensities to collaborate: An investigation of strategic alliance formation in a high-technology industry. *Administrative Science Quarterly* 43: 668–698.
Stuart, T.E., and J.M. Podolny. 1996. Local search and the evolution of technological capabilities. *Strategic Management Journal* 17(Evolutionary perspectives on strategy supplement): 21–38.
Walker, G., B. Kogut, and W. Shan. 1997. Social capital, structural holes and the formation of an industry network. *Organization Science* 8: 109–125.
Young, A.A. 1928. Increasing returns and economic progress. *Economic Journal* 38: 527–542.

Research and Development (R&D) Investment

Anu Wadhwa
Swiss Federal Institute of Technology, Lausanne, Switzerland

Abstract

R&D investment reflects an organization's willingness to invest in discovery and commercialization of new technologies in the form of products and processes as well as refinement of existing technologies. Investments in R&D have historically been made within firm boundaries and have been associated positively with firm innovation and performance. The uncertainty associated with returns to R&D investments and the increasingly distributed nature of knowledge has created an impetus for firms to invest in external sources of innovation to supplement their internal R&D efforts.

Definition R&D investment reflects organizational investment in 'creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications' (OECD 1963), to derive future benefits for the organization.

Traditionally, corporate R&D investment has been associated with monetary investments made in the R&D laboratories of large firms for the purpose of discovering new knowledge and applying existing knowledge to develop and refine new or existing products and processes. Firm-level R&D expenditures, which are typically composed of labour, capital and material costs, have usually been measured either as the total amount of money spent annually by the firm on R&D activity, or as R&D intensity (the ratio of annual R&D expenditure to annual sales). In a recent survey (Booze and Co 2010), the world's top 1,000 public firms spent a total of $503 billion on R&D investment. The vast bulk of this R&D expenditure came from firms in industries such as automotives, computing and electronics, and healthcare ($322.79 billion) and from regions such as the United States, Europe and Japan ($469.33 billion). While firms spend less than 5% of their revenues on R&D on average, each of the top ten R&D spenders incurs expenses in the range of 2–16% of their sales and invests in excess of $5 billion on R&D.

Returns to R&D Investment

Firms continue to invest in R&D because it is expected to create value for the firm. Extant research suggests that R&D investment has a strong relationship with firm growth and profitability. R&D investment has been shown to have a positive impact on firm innovation (Pakes and Griliches 1984; Ahuja and Katila 2001), facilitate absorption of external new knowledge (Cohen and Levinthal 1990), and has been associated positively with different indicators of firms' economic performance such as market value, profitability, sales and growth (Griliches et al. 1991; Grandi et al. 2008; Oriani and Sobrero 2008). However, the observed rate of return of R&D investment to firm performance has been decreasing and has exhibited increasing volatility over time and across industries (Grandi et al. 2008).

The trend of decreasing returns to R&D investment is partly due to the high degree of uncertainty inherent in the process of engaging in research and innovation (Mansfield et al. 1977). The resultant volatility in profitability or payoffs from R&D investments can arise from different sources of environmental uncertainty, two of which are particularly salient for technology-related investment – market uncertainty and technological uncertainty. Market uncertainty exists for firms when, depending on exogenous factors such as economic cycle, demographic changes, consumer preferences and institutional factors, they experience volatility in the expected level of demand for their products (Huchzermeier and Loch 2001). Firms face technological uncertainty when it is unclear which of the competing technologies in their industry are likely to emerge as the dominant technology (Anderson and Tushman 1990). These two types of uncertainties associated with the market and with the technology can influence patterns of R&D investment in significant ways. Faced with market uncertainty, firms can choose to wait before investing additional resources or commit to incremental pattern of investment to avoid potential losses (Folta and O'Brien 2004). Faced with technological uncertainty, firms may decide to wait for the technology to evolve further or invest in alternative technologies (McGrath 1997).

As firm size increases, the corresponding rise in the number of different R&D projects requiring investment and managerial attention creates substantial market and/or technological uncertainty and compounds the problem of resource allocation to these projects. For firms with multiple projects in their R&D portfolios, arriving at the most optimal mix is the key challenge confronting them.

Moving to 'Open' R&D

The increasing amount of R&D investment required to merely keep up with the competition, the increasing level of uncertainty associated with R&D activities, and the problem of effectively appropriating the benefits has begun to change the way firms think about how to replenish their innovation pipelines. This has implications for *where* they invest their R&D funds.

The *not-invented-here* syndrome has historically been the reason that firms have underinvested in external R&D or rejected investment in external technologies in favour of investing in their internal R&D efforts (Katz and Allen 1982). As the knowledge underlying innovation becomes more heterogeneous, complex and distributed, firms are now embracing the 'open' model of innovation as opposed to the closed paradigm to avoid missing out on opportunities that are either outside their current businesses and expertise, or which need to be combined with external technologies to unlock their potential (Chesbrough 2003). Thus, **open innovation** rests on the idea that firms should invest in external as well as internal research, and has been defined as 'the use of purposive inflows and outflows of knowledge to accelerate internal innovation, and expand the market for external use of innovation, respectively' (Chesbrough et al. 2006: 1). This trend towards open innovation is not completely new, but has been gathering steam in the last decade. An OECD study found that while three-quarters of the firms surveyed spent the bulk of their R&D investment on in-house R&D projects, as many as 51% of the firms allocated about 5% of their R&D investment to external projects while 31% of the firms allocated more than 10% outside (OECD 2008).

Types of External R&D Strategies

In their quest for innovative ideas beyond firm boundaries, companies channel their external R&D investments into external initiatives such as corporate venture capital (CVC) investments, alliances and acquisitions. Embeddedness in open, collaborative and interconnected innovation networks allow firms to create value by sourcing and leveraging external knowledge and technologies to supplement their internal R&D effort (Chesbrough 2003). In general, while there is some evidence that investments into new technologies and ideas via external initiatives influence firm innovation and economic performance positively (Ahuja 2000; Stuart 2000; Ahuja and Katila 2001; King et al. 2004; Man and Duysters 2005; Dushnitsky and Lenox 2005; Wadhwa and Kotha 2006; Allen and Hevert 2007), the strength of this relationship and the conditions under which it holds is still subject to much debate. In such an 'open' world, it has become extremely important for firms to re-evaluate and rethink how they can achieve returns on their investments in external R&D. From a value creation perspective, firms that choose an 'open' innovation strategy have to pay careful attention to how they can effectively stimulate their internal innovation process by integrating external knowledge and technology into new products and/or processes. From a value appropriation perspective, firms utilizing knowledge developed beyond their boundaries also have to ensure that they can claim a significant share of the value they create. Thus, despite the different paradigms adopted by firms to become more innovative, reaping the returns from their investments in R&D, whether external or internal, remains the key challenge confronting them.

See Also

- ▶ Collaborative Innovation
- ▶ Corporate Venturing
- ▶ Open Innovation
- ▶ Research and Development (R&D) Alliances
- ▶ Research and Development (R&D) Organization

References

Ahuja, G. 2000. Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly* 45: 425.

Ahuja, G., and R. Katila. 2001. Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study. *Strategic Management Journal* 22: 197–220.

Allen, S.A., and K.T. Hevert. 2007. Venture capital investing by information technology companies: Did it pay? *Journal of Business Venturing* 22: 262–282.

Anderson, P., and M.L. Tushman. 1990. Technological discontinuities and dominant designs: A cyclical model of technological change. *Administrative Science Quarterly* 35: 604–633.

Booze & Co. 2010. Global innovation 1000 report.

Chesbrough, H.W. 2003. The era of open innovation. *MIT Sloan Management Review* 44: 35–41.
Chesbrough, H., W. Vanhaverbeke, and J. West (eds.). 2006. *Open Innovation: Researching a New Paradigm*. Oxford: Oxford University Press.
Cohen, W.M., and D.A. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* 35: 128–152.
de Man, A.P., and G.M. Duysters. 2005. Collaboration and innovation: A review of the effects of mergers, acquisitions and alliances on innovation. *Technovation* 25: 1377–1387.
Dushnitsky, G., and M.J. Lenox. 2005. When do incumbents learn from entrepreneurial ventures? Corporate venture capital and investing firm innovation rates. *Research Policy* 34: 615–639.
Folta, T.B., and J.P. O'Brien. 2004. Entry in the presence of duelling options. *Strategic Management Journal* 25: 121–138.
Grandi, A., B.H. Hall, and R. Oriani. 2008. R&D and financial investors. In *R&D evaluation and performance measurement*, ed. V. Chiesa and F. Frattini. Cheltenham: Edward Elgar.
Griliches, Z., B.H. Hall, and A. Pakes. 1991. R&D, patents, and market value revisited: Is there a second (technological opportunity) factor? *Economics of Innovation and New Technology* 1: 183–201.
Huchzermeier, A., and C.H. Loch. 2001. Project management under risk: Using the real options approach to evaluate flexibility in R&D. *Management Science* 47: 85–101.
Katz, R., and T. Allen. 1982. Investigating the not invented here (nih) syndrome: A look at the performance, tenure and communication patterns of 50 R&D project groups. *R&D Management* 12: 7–19.
King, D.R., D.R. Dalton, C.M. Daily, and J.G. Covin. 2004. Meta-analyses of post-acquisition performance: Indications of unidentified moderators. *Strategic Management Journal* 25: 187–200.
Mansfield, E., J. Rapoport, A. Romeo, S. Wagner, and G. Beardsley. 1977. Social and private rates of return from industrial innovations. *Quarterly Journal of Economics* 91: 221–240.
McGrath, R.G. 1997. A real options logic for initiating technology positioning investments. *The Academy of Management Review* 22: 974–996.
OECD (Organisation for Economic Co-operation and Development). 1963. *Frascati manual of the organization for economic cooperation and development*. Paris: OECD.
OECD (Organisation for Economic Co-operation and Development). 2008. *Policy brief*. Paris: OECD.
Oriani, R., and M. Sobrero. 2008. Uncertainty and the market valuation of R&D within a real options logic. *Strategic Management Journal* 29: 343–361.
Pakes, A., and Z. Griliches. 1984. Patents and R&D at the firm level: A first look. In *R&D, patents, and productivity*, ed. Z. Griliches. Chicago: University of Chicago Press.
Stuart, T.E. 2000. Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. *Strategic Management Journal* 21: 791–811.
Wadhwa, A., and S.B. Kotha. 2006. Knowledge creation through external venturing: Evidence from the telecommunications equipment manufacturing industry. *Academy of Management Journal* 49: 819–835.

Research and Development (R&D) Organization

Nicholas Argyres
Washington University in St Louis, Olin Business School, St Louis, MO, USA

Abstract

This article briefly reviews research on the determinants and impacts of firms' internal R&D structures. The focus is on the relative advantages of centralization and decentralization of R&D activities within large firms. Other important topics related to R&D organization are mentioned as well.

Definition R&D organization relates to the manner in which a corporation organizes its research and development (R&D) activities, which includes the reporting relationships involving executives and technical personnel, the ways in which information about R&D is designed to flow within the organization, and the incentives provided to employees producing and using R&D within the organization.

Technological knowledge plays a vital role as a basis for competitive advantage in many industries. The internal organization of corporate research activities is in turn a key determinant of the direction and impact of technological innovation. Research on R&D organization, however, has been rather thin. This article focuses on perhaps the most prominent question in this small literature: that regarding the trade-offs between centralization and decentralization of the R&D function. Towards the end of the article other important topics are mentioned, including

(1) informal R&D networks with the organization, (2) the relationships between internal R&D organization and strategic alliances, and (3) the relationships between R&D organization and R&D location choices.

The emphasis on centralization vs decentralization of R&D in part reflects the influence of Hounshell and Smith's (1988) history of DuPont's successful R&D organization. As a pioneer in adopting the multi-divisional ('M-form') form of organization, DuPont decentralized virtually all research activities to the divisional level. The firm later restored the funding authority and capacity of its central research unit after realizing the merits of the centralized form. Since these two shifts in R&D organization in the 1920s, DuPont has maintained a hybrid research structure. During its history, DuPont experienced several of the trade-offs between the various R&D structures that are discussed next.

Centralization vs Decentralization of R&D

Decentralization has in general, and with respect to R&D in particular, been traditionally associated with efficiency advantages, owing to improved information processing and reduced information and time demands on top management (Williamson 1975, 1985; Galbraith 1977). With respect to information processing, the decentralized structure can efficiently link research effort to divisional needs, via divisional staff who are better equipped to make decisions that affect their unit (Jensen and Meckling 1992). This advantage is particularly evident where successful innovation depends on close understanding of user needs (Von Hippel 1988). With respect to scope of managerial opportunism, first, decentralization of research establishes clear lines of authority and responsibility to the divisional manager, thus reducing the scope for non-cooperation by R&D personnel; second, the decentralized organization also facilitates the measurement of R&D performance, thereby mitigating opportunism by R&D personnel; finally, decentralization of research improves the credibility of corporate-level management's promises not to intervene in the divisions' operations, thus enhancing incentives of divisional managers (Williamson 1985).

Centralization of R&D, on the other hand, is thought to enable firms to better exploit economies of scale and scope in R&D activities relative to decentralized R&D (Galbraith 1977; Daft 1989). This is because business units within a decentralized firm may not take cross-unit 'spillovers' into account when making their R&D investment decisions (Arrow 1962). Centralization is therefore argued to be more efficient when research is 'non-specific' (Kay 1988). Non-specific research is research whose fruits are applicable beyond the confines of a specific business unit, and is generally associated with great uncertainty. It is more likely to generate innovations with greater and wider technological impact (Kuznets 1962), because non-specific research involves the selection of research problems that are more likely to lead to the discovery of fundamental or generic knowledge (Nelson 1990).

Centralized R&D supports the pursuit of non-specific research because it avoids the negotiation and haggling costs that decentralized divisions can impose on the firm as they attempt to work out an arrangement to share rights to, and responsibilities for, the inputs and outputs of non-specific research – costs that are magnified in the presence of high uncertainty (Argyres 1995). Centralized R&D and/or corporate-level funding of R&D, however, can overcome these problems by either directly funding non-specific research or concentrating R&D activities in a single cost centre, which replaces the 'high-powered' incentive system with a 'low-powered' one (Williamson 1985; Milgrom and Roberts 1992). Thus, transaction cost logic indicates that centralized research will favour investment in non-specific R&D. In terms of welfare economics, a centralized R&D function helps internalize the externalities involving R&D investments with the firm, and thereby generate innovation that transcends the business unit (Kay 1988). Information-processing approaches to organization (Thompson 1967; Egelhoff 1991) also imply that non-specific R&D would tend to be centralized.

Centralized R&D is also thought to facilitate non-local search for solutions to problems more effectively than does decentralized research, for two reasons. First, in decentralized R&D divisional managers and engineers are typically asked to adopt a customer-centric orientation (Kay 1988) and are less likely to pursue opportunities in new markets (Galunic and Eisenhardt 2001). A focus on current customer demands trades off opportunities to develop more radical technologies (Christensen and Bower 1996; Jaworski et al. 2000). Centralized R&D, on the other hand, allows for more freedom for researchers to explore broader and non-local research projects. Second, because researchers in centralized R&D labs are less deeply engaged in local communication channels, they are less subject to the associated information filters. Therefore, research staffs are more likely to appreciate and explore (broader) architectural innovations (Henderson and Clark 1990) with superior capacity to recognize the value of new knowledge (Cohen and Levinthal 1990).

Empirical Evidence

Two surveys conducted by the Industrial Research Institute (IRI) in 1994 and 2001 show that large companies typically feature one of three distinct R&D structures; centralized, decentralized and hybrid. Each accounts for approximately 30%, 10% and 60% of the total number of firms in the surveys, respectively. In the centralized structure, a single executive is in charge of the firm's research activities, and reports directly to a corporate-level executive such as the CEO or president. In the decentralized structure, individual divisions or business units conduct research exclusively within themselves, and R&D directors directly report to their division general managers. In the hybrid structure, research is conducted both within a centralized function and within the firm's divisions or business units function. The R&D directors of the two functions report to corporate management and division general manager respectively.

By decoupling research effort from the immediate demands of divisions and reducing the transaction costs associated with internal R&D coordination, centralized R&D theoretically can generate innovations of a greater impact on future technological developments and of a broader impact upon technological domains. Argyres and Silverman (2004) provide empirical evidence to support these predictions using the IRI survey data combined with patent citation and other data.

Lerner and Wulf (2007) attribute the impact of centralized R&D to the influence of corporate R&D leaders over research decisions, and find that innovation has a larger impact when longer-term incentives (stock options and restricted stock) are granted to corporate R&D leaders of a centralized research organization. However, this association disappears for firms with decentralized R&D organizations.

Firms tend to search for knowledge 'close to' their existing knowledge (Cohen and Levinthal 1990; Helfat 1994), because searching is path-dependent and constrained by organizational routines (Nelson and Winter 1982). The ability to move beyond local search is crucial for gaining competitive advantage (Kogut and Zander 1992; Levinthal and March 1993; Henderson and Cockburn 1994). Rosenkopf and Nerkar's (2001) study of patented innovations in the optical disk industry demonstrates that broader search efforts contribute to a broader impact of innovation.

Argyres and Silverman (2004) find that firms with centralized R&D appear to conduct technological search outside their organizational boundaries more widely than do decentralized R&D firms. This implies that by facilitating more distant ('capabilities-broadening') search, centralized R&D can generate innovations that draw on previous innovations developed in a wider range of organizations and technological domains.

Separate from the authority relations in R&D, the source of research funding within large firms can be the business units, corporate headquarters or some combination of the two. Whereas, similar to line of authority, funding authority is also one of the many instruments that firms use to influence the R&D decisions of their managers and

technical staff. The arguments above regarding the advantages and disadvantages of centralization vs decentralization of R&D apply to either instrument of allocating authority independently.

Argyres and Silverman (2004) show evidence that the marginal differences in the centralization of R&D funding are large when the R&D reporting relationship is already highly centralized. This suggests that R&D structure and budget authorities complement each other in affecting innovative impact (Siggelkow 2002), and increases in the degree of corporate-level control of R&D funding associated with bigger increases in innovative impact when R&D decision-making authority is centralized.

As noted in the IRI surveys above, the preponderance of hybrid R&D structures is striking. Hybrid organizations in general may be able to combine the advantages and disadvantages of centralized and decentralized structures in terms of coordination, control and information processing, while suffering the consequence of greater role ambiguity than other structures (Daft 1989). Tushman and O'Reilly (1996) conceptualize 'ambidextrous organizations' as able to simultaneously pursue radical and incremental innovation. Hybrid R&D organizations may therefore be seen as a type of ambidextrous organization, producing innovation that is intermediate in the breadth of its impact. However, Argyres and Silverman (2004) do not find consistent evidence that hybrids tend to produce innovation with such intermediate-level impact. Hybrid R&D organizations remain poorly understood.

Other Topics: Informal Networks, Alliances and Location

While formal R&D structure is an important topic that has been studied to some extent, informal R&D organization has long been known to be important as well. Recently, scholars have begun to study R&D organization using social network approaches. Scholars have studied how knowledge flows among researchers within the firm (e.g., Reagans and Zuckerman 2001; Nerkar and Paruchari 2005; Singh 2005) and shown how informal communication among scientists and engineers within the firm can stimulate innovation. Relatively little is known, however, about the interaction between formal R&D organization structures and informal networks. How the two mechanisms affect each other, and how they combine to influence innovative outcomes, remain intriguing questions for future research.

The relationship between intra-firm organization of R&D and inter-firm alliances in research is important for understanding how the two mechanisms interact with each other. Bercovitz and Feldman (2007) examine the relationship between internal R&D organization and a particular type of alliance – the industry–university partnership. Although this study does not find a direct link between internal organization of research and firm–university collaboration, it does show that firms make greater use of university-based research when they both focus on exploration internally and have more centralized R&D operations.

Questions regarding the determinants and impacts of the geographic distribution of R&D are conceptually distinct from those about the causes and consequences of R&D organization structure. Several findings in the literature on R&D geography do carry organizational implications, however. Singh (2008), for example, shows that geographically distributed R&D is associated with worse innovation outcomes. This is presumably because the potential gains from access to diverse ideas and expertise from different locations are outweighed by the increased cost and difficulty in achieving integration of knowledge across multiple locations. However, when cross-regional integration of knowledge does occur, the effect on innovative outcomes is positive. Kuemmerle (1999a, b) earlier studied the determinants of location choices for international R&D investments by large firms.

Conclusion

Intra-firm organization of R&D activities affects the process of technological knowledge creation and, consequently, the formation of competitive advantage. Since each type of R&D organization

structure promotes a different type of R&D, there are trade-offs between these structures, and firms find their own efficient matches between R&D organization structure and the type of R&D they choose to pursue. Understanding the interactions between R&D organization structure and informal social networks, strategic alliances and R&D location choices is also important but still little studied.

See Also

- ▶ Corporate Strategy
- ▶ Exploration and Exploitation
- ▶ Federative Multinational Enterprise (MNE)
- ▶ Firm Size and Boundaries, Strategy
- ▶ M-Form Firms
- ▶ Organizational Design

References

Argyres, N. 1995. Technology strategy, governance structure and interdivisional coordination. *Journal of Economic Behavior & Organization* 28: 337–358.

Argyres, N., and B. Silverman. 2004. R&D, organization structure, and the development of corporate technological knowledge. *Strategic Management Journal* 25-(winter special issue): 929–958.

Arrow, K. 1962. Economic welfare and the allocation of resources to invention. In *The rate and direction of inventive activity: Economic and social factors*, ed. National Bureau of Economic Research. Princeton: Princeton University Press.

Bercovitz, J., and M. Feldman. 2007. Fishing upstream: Firm innovation strategy and university research alliances. *Research Policy* 36: 930–948.

Christensen, C., and J. Bower. 1996. Customer power, strategic investment and the failure of leading firms. *Strategic Management Journal* 17: 197–218.

Cohen, W., and D. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* 35: 128–152.

Daft, R. 1989. *Organization theory and design*, 3rd ed. St. Paul: West Publishing.

Egelhoff, W. 1991. Information-processing theory and the multinational enterprise. *Journal of International Business Studies* 22: 341–368.

Galbraith, J. 1977. *Organization design*. Reading: Addison-Wesley.

Galunic, D., and K. Eisenhardt. 2001. Architectural innovation and modular corporate forms. *Academy of Management Journal* 44: 1229–1249.

Helfat, C. 1994. Evolutionary trajectories in petroleum R&D. *Management Science* 40: 1720–1747.

Henderson, R., and K. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35: 9–30.

Henderson, R., and I. Cockburn. 1994. Measuring competence? Exploring firm effects in pharmaceutical research. *Strategic Management Journal* 15(winter special issue): 63–84.

Hounshell, D., and J.K. Smith. 1988. *Science and corporate strategy: DuPont R&D 1902–1980*. New York: Cambridge University Press.

Jaworski, B., A. Kohli, and A. Sahay. 2000. Market-driven vs. driving markets. *Journal of the Academy of Marketing Science* 28: 45–54.

Jensen, M., and M. Meckling. 1992. Specific and general knowledge, and organization structure. In *Contract economics*, ed. L. Werin and H. Wijkander. Oxford: Blackwell.

Kay, N. 1988. The R&D function: Corporate strategy and structure. In *Technical change and economic theory*, ed. G. Dosi, C. Freeman, R. Nelson, G. Silverberg, and C. Soete. London: Pinter.

Kogut, B., and I. Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* 3: 383–397.

Kuemmerle, W. 1999a. Foreign direct investment in industrial research in the pharmaceuticals and electronics industries – Results from a survey of multinational firms. *Research Policy* 28: 179–193.

Kuemmerle, W. 1999b. The drivers of foreign direct investment into research and development: An empirical investigation. *Journal of International Business Studies* 30: 1–24.

Kuznets, S. 1962. Inventive activity: Problems of definition and measurement. In *The rate and direction of inventive activity: Economic and social factors*, ed. National Bureau of Economic Research. Princeton: Princeton University Press.

Lerner, J., and J. Wulf. 2007. Innovation and incentives: Evidence from corporate R&D. *Review of Economics and Statistics* 89: 634–644.

Levinthal, D., and J. March. 1993. The myopia of learning. *Strategic Management Journal* 14: 95–112.

Milgrom, P., and J. Roberts. 1992. *Economics, organization and management*. Englewood Cliffs: Prentice Hall.

Nelson, R. 1990. Capitalism as an engine of progress. *Research Policy* 19: 193–214.

Nelson, R., and S. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Belknap.

Nerkar, A., and A. Paruchari. 2005. Evolution of R&D capabilities: The role of knowledge networks within a firm. *Management Science* 51: 771–786.

Reagans, R., and E.W. Zuckerman. 2001. Networks, diversity, and productivity: The social capital of corporate R&D teams. *Organization Science* 12: 502–517.

Rosenkopf, L., and A. Nerkar. 2001. Beyond local search: Boundary-spanning, exploration, and impact in the

optical disk industry. *Strategic Management Journal* 22: 287–306.
Siggelkow, N. 2002. Misperceiving interactions between substitutes and complements: Organizational consequences. *Management Science* 48: 900–917.
Singh, J. 2005. Collaborative networks as determinants of knowledge diffusion patterns. *Management Science* 51: 756–770.
Singh, J. 2008. Distributed R&D, cross-regional knowledge integration and quality of innovative output. *Research Policy* 37: 77–96.
Thompson, J. 1967. *Organizations in action: Social science bases of administrative theory.* New York: McGraw-Hill.
Tushman, M., and C. O'Reilly. 1996. Ambidextrous organizations: Managing evolutionary and revolutionary change. *California Management Review* 38: 8–30.
Von Hippel, E. 1988. *The sources of innovation*. New York: Oxford University Press.
Williamson, O. 1975. *Markets and hierarchies*. New York: Free Press.
Williamson, O. 1985. *The economic institutions of capitalism*. New York: Free Press.

Research Consortia

Mariko Sakakibara
University of California Los Angeles, Anderson School of Management, Los Angeles, CA, USA

Abstract

Research consortia refers to the associations of a group of firms or other types of organizations which agree to share the costs and results of a research project before the execution of that project. They typically involve a large membership and undertake pre-competitive research, but this term has been used more broadly. Government often uses research consortia to promote innovation, but their effectiveness depends on a country's innovation systems. Cost-sharing and skill-sharing research consortia have different implications for a participant's R&D spending and productivity. Researchers need to cope with the endogenous nature of participation when conducting performance evaluation studies.

Definition Associations of a group of firms or other types of organizations which agree to share the costs and results of a research project before the execution of that project.

Research consortia refers to the associations of a group of firms or other types of organizations which agree to share the costs and results of a research project before the execution of that project. Research consortia between firms can take a variety of forms. They may be a ▶ joint ventures, formed by two or more partners as a separate company with shared equity investments. They can be a partnership or ▶ strategic alliances, linking firms on the basis of continuing commitment to shared technological objectives without equity sharing. They may take the form of 'innovation networks', combinations of firms and research organizations that share research agendas (Sakakibara and Dodgson 2003). While research consortia typically involve a large membership and undertake pre-competitive research of common interest to all participants (Vonortas 1997), the terms research consortia, research joint ventures, research ▶ alliances and strategic research partnerships are often used interchangeably. Collaborating firms can have vertical relationships throughout a value chain, while they can be direct competitors in the product market. Research consortia have become increasingly important vehicles for developing new innovations because firms often face resource and time constraints to develop innovations on their own. Firms collaborate in their research activities for cost-sharing or skill-sharing purposes (Sakakibara 1997a, b). There are many other motives, such as to share and/or reduce risks or uncertainties (Hagedoorn 1993), and to affect competitive positions through standard-setting, competitor exclusion or locking in key players.

Government often uses research consortia as a policy tool to promote ▶ innovation because they are an important feature in the generation and diffusion of technology and, by extension, industrial development. They are also an important feature of the research environment and industry in most industrialized and industrializing nations. Government can fund research projects conducted

by research consortia, affect their membership and organizational decisions, and set rules of antitrust enforcements. The 1984 US National Cooperative Research Act is an example which relaxed antitrust regulations in order to allow the formation of research joint ventures. The degree and type of government involvement depend on industry structures, business systems and research infrastructure within national innovation systems. Similarly, the effectiveness of research consortia as a means to promote innovation varies by country.

Theoretical Arguments on Research Consortia

From an economic perspective, research consortia are considered as a means to set cost-sharing and/or output-sharing rules for the participants in an R&D project in order to correct market failures. Spence (1984) argues that the existence of R&D spillovers makes it difficult for innovators to capture the full social benefits of their innovative activity, which depresses the incentives to conduct R&D. Through R&D cooperation, firms internalize the externality created through spillovers, thus restoring the incentive to conduct R&D. However, Katz (1986) shows that if a higher level of R&D makes market competition more intense by lowering firms' marginal costs of production, then the resulting decline in profits will reduce their incentive to conduct R&D, implying that research consortia can result in less R&D. These results are echoed in later work by d'Aspremont and Jacquemin (1988) and others. This implies that research consortia which consist of direct competitors are less effective in creating innovative outcomes. In this literature it is typically assumed that firms are symmetrical in terms of their capabilities or knowledge, and firms seek to achieve a single R&D outcome in the most cost-efficient way.

In the management literature the motives for cooperation among firms are examined more extensively, and under quite different assumptions. Firms in research consortia are often recognized to possess heterogeneous capabilities, and they may or may not be direct competitors in the product market. The resource-based view suggests that a firm can be conceived as a portfolio of core competencies. Research consortia can be viewed as opportunities for one partner to internalize the skills or competencies of the other(s) to create next-generation competencies (Hamel 1991). Firms consist of a knowledge base, and this knowledge – particularly technological knowledge – is often 'tacit' (Polanyi 1958) and not easily diffused across the firm's boundaries. Organizational vehicles, such as research consortia, are required to affect this transfer (Kogut 1988). These learning-based arguments imply that a key objective of research consortia is complementary knowledge or skill-sharing among participants (Sakakibara 1997a).

This learning function of research consortia implies that the benefits firms can get from research consortia can differ by their ▶ absorptive capacity (Lane et al. 2006). Cohen and Levinthal (1989) showed that a high spillover rate in R&D among competitors can provide a positive incentive to conduct R&D when a company's own R&D increases its learning capability. Since research consortia are a 'forced' spillover scheme, participation gives firms an incentive to conduct more R&D, to increase their absorptive capacity and benefit from consortia.

There are costs of participating in research consortia (Jorde and Teece 1990). These include the potential leakage of proprietary information, the costs to monitor opportunistic behaviour of participants and to align interests among participants. Thus, a primary challenge in research consortia is to design the consortia to minimize such costs.

Empirical Studies on Research Consortia

Empirical studies which analyse research consortia have proliferated. The issues studied include the attributes of collaborating firms, their alliance experience and network position, consortia structure, governance mechanisms employed, and how they relate to the motives of participation, the

formation of consortia, and R&D spending and performance of consortia and participating firms. The success of empirical studies critically depends on the availability of detailed data. Participation in research consortia is an important strategic decision for firms, so firms do not necessarily announce their participation in private research consortia. Therefore, government-sponsored research consortia have been frequently used to obtain comprehensive data.

The most important research issue of research consortia is the determinant of their performance. Since firms participate in research consortia for their own benefits, the success of research consortia should be measured at the participant level, not at the consortium level. One can approach this issue on multiple levels. The first is the overall impact of the participation in research consortia on research productivity of participating firms. Branstetter and Sakakibara (1998) examine the data on Japanese government-sponsored research consortia. They found that if a firm participates in an additional project per year, it raises its patenting per R&D dollar (i.e., its research productivity) by between 4% and 8%.

A more disaggregated approach is to identify the characteristics of consortia that are associated with the increase of research productivity of participating firms. Branstetter and Sakakibara (2002) examine the same data. They focus on two major characteristics of research consortia: spillover potential and *ex post* product market competition among participating firms. They measure spillover potential as technological proximity among member firms in the technological space, and the level of *ex post* product market competition as the product market proximity of member firms. Their outcome measure is the number of patents taken by consortia participants in technological areas targeted by consortia. They find positive association between technological proximity and consortium outcomes, and a negative relationship between product-market proximity and consortium outcomes. In addition, they employ qualitative characteristics of consortia, and find that these consortia are most effective when they focus on basic research.

Selection Problem

Selection problem is a fundamental issue that researchers have to cope with when they conduct evaluation studies of research consortia, especially those sponsored by government. Since governments or consortia organizers seek to encourage firms with strong R&D capabilities to participate in order to maximize benefits, if we observe good outcomes from certain types of research consortia we cannot distinguish whether these consortia are effective or if only good firms participate in these consortia. This selection problem is the single greatest limitation of past research to measure the impact of public technology programmes (Klette et al. 2000).

The analyses of Branstetter and Sakakibara (2002) demonstrate a way to address the selection problem by utilizing detailed panel data. By employing the data of patenting in the targeted technologies before, during and after participation in a consortium by individual firms, they control for the pre-existing technological strength of a firm in the targeted technologies. Also, using observations on firms that did not participate in consortia as a control, they can extract the pure-participation effect. Finally, because they observe the same firms participating in multiple consortia, they are able to measure the marginal impact of different consortium characteristics and firm characteristics on research outcomes.

See Also

- ▶ Absorptive Capacity
- ▶ Alliances
- ▶ Collaborative Innovation
- ▶ Innovation
- ▶ Innovation Policy
- ▶ Inter-organizational Learning
- ▶ Joint Venture
- ▶ Research and Development (R&D) Alliances
- ▶ Research and Development (R&D) Organization

References

Branstetter, L.G., and M. Sakakibara. 1998. Japanese research consortia: A microeconometric analysis of industrial policy. *Journal of Industrial Economics* 46: 207–233.

Branstetter, L.G., and M. Sakakibara. 2002. When do research consortia work well and why? Evidence from Japanese panel data. *American Economic Review* 92: 143–159.

Cohen, W.M., and D.A. Levinthal. 1989. Innovation and learning: The two faces of R&D. *Economic Journal* 99: 569–610.

D'Aspremont, C., and A. Jacquemin. 1988. Cooperative and noncooperative R&D in duopoly with spillovers. *American Economic Review* 78: 1133–1137.

Hagedoorn, J. 1993. Understanding the rationale of strategic technology partnering: Interorganizational modes of cooperation and sectoral differences. *Strategic Management Journal* 14: 371–385.

Hamel, G. 1991. Competition for competence and interpartner learning within international strategic alliances. *Strategic Management Journal* 12: 83–103.

Jorde, T.M., and D.J. Teece. 1990. Innovation and cooperation: Implications for competition and antitrust. *Journal of Economic Perspectives* 4: 75–96.

Katz, M.L. 1986. An analysis of cooperative research and development. *RAND Journal of Economics* 17: 527–543.

Klette, T.J., J. Moen, and Z. Griliches. 2000. Do subsidies to commercial R&D reduce market failure? Microeconometic evaluation studies. *Research Policy* 29: 471–495.

Kogut, B. 1988. Joint ventures: Theoretical and empirical perspectives. *Strategic Management Journal* 9: 319–332.

Lane, P.J., B.R. Koka, and B. Pathak. 2006. The reification of absorptive capacity: A critical review and rejuvenation of the construct. *Academy of Management Review* 31: 833–863.

Polanyi, M. 1958. *Personal knowledge: Towards a post-critical philosophy*. Chicago: University of Chicago Press.

Sakakibara, M. 1997a. Heterogeneity of firm capabilities and cooperative research and development: an empirical examination of motives. *Strategic Management Journal* 18: 143–164.

Sakakibara, M. 1997b. Evaluating government-sponsored R&D consortia in Japan: Who benefits and how? *Research Policy* 26: 447–473.

Sakakibara, M., and M. Dodgson. 2003. Strategic research partnerships: Empirical evidence from Asia. *Technology Analysis and Strategic Management* 15: 228–245.

Spence, M. 1984. Cost reduction, competition, and industry performance. *Econometrica* 52: 101–121.

Vonortas, N.S. 1997. *Cooperation in research and development*. Boston: Kluwer.

Resilience

J. P. Macintosh
University College London, Institute for Security and Resilience Studies, London, UK

Abstract

Resilience is a word that is gaining increasing currency in the field of strategic management (Cascio, Foreign Policy 172:82–95, 2009) although not without some criticism (Rose, Environmental Hazards: Human and Policy Dimensions 6:1–16, 2007). Use of the word is evolving from its classical etymology and narrow engineering definition as bounce-back to the status quo ante. In life sciences resilience is taken to be evolutionary in nature. This understanding accords with the reality of living in dynamic networks, where our 'bounded rationality' (Simon, H.A. [1956] 1982. Reply: Surrogates for uncertain decision problems. In *Models of bounded rationality*, *vol. 1: Economic analysis and public policy*. Cambridge, MA: The MIT Press.) is increasingly dangerous to ignore. On these terms resilience is a realist concept for enabling bodies to bounce forward, innovating appropriately through learning from a past overtaken by events and exploration of the uncertainties ahead.

Definition Resilience is the enduring power of a body or bodies for transformation, renewal and recovery through the flux of interactions and flow of events. Decisive moments – crises proper – continually test for resilience; this process is evolutionary in nature. Resilience grows and decays as the innovative capacity to select activities from a variety of capabilities changes, enriched or not by learning.

The Evolutionary Meaning of Resilience

In this article, resilience is defined as the enduring power of a body or bodies for transformation, renewal and recovery through the flux of

interactions and flow of events. The measure of resilience is how bodies bounce forward to thrive on change in dynamic networks rather than bounce back to a status quo ante overtaken by events. Achieving the former evinces strategic ▶ leadership; the latter may be expedient but can invite mismanagement. There are many hypotheses that flow from the definition offered here; all must be left open to falsification.

Resilience is not a new word but its meaning continues to evolve through time. Those changes, particularly in the last century, are of great importance to strategic management, present and future. The word resilience works at the nexus of natural, life and social sciences (Rutter 2007). In the strategic management of security – whether national, international, transnational or human – resilience is now a much-cited keyword (Cascio 2009). Its richest strategic significance is perhaps at the interface of economics and ecosystems (Simon 2005; Hodgson and Knudson 2010), where the risks and uncertainties of evolving networks abound (Dorogovtsev and Mendes 2003).

Frequent use of any word can multiply confusions (Rose 2007). For narrow specialists it may indicate interdisciplinary overstretch. So it is necessary to be precise about underpinning assumptions and the association of meanings invoked by the word 'resilience'. Put in evolutionary rather than teleological terms the word can at least sidestep the trap of appearing to posit a grand unifying theory. In defining resilience it is also important to say what it is not. The strategic management definition is here described under three abridged subheadings:

- Decision takers' learning competencies;
- Combinations of real capability options; and,
- Transformational capacity for multi-innovation.

These lead to the provisional strategic conclusion that resilience is almost synonymous with ▶ competitiveness.

Decision Takers' Learning Competencies

Dynamic networks can make the decisions and actions of us all – strategic. This is why, in the wake of the September 2000 fuel protests and in anticipation of the financial crises triggered in 2007 (Jenkin 2010; IEO 2011), the UK government endorsed the concept of 'resilience to crises' (MacIntosh and Granatt 2001). At the direction of then Prime Minister Tony Blair, resilience began to be sown wide and deep. Crises – as decisive moments – and network contagion had found the competencies of elites and the multitude wanting.

Orthodox approaches to risk work in only limited ways. Misusing these methods exacerbates the harm done, yet learning continues to lag behind strategic requirements. The reality of radical uncertainty has been well characterized (Keynes 1921; Knight 1921). Decision takers' competencies can be enhanced by learning from scientific advances on several fronts, not least maths (Dorogovtsev and Goltsev 2008) and the uptake of these advances by popular social science (Watts 2004). Therefore, wilful ignorance perhaps underscores the agency problems that degrade strategic managers rather than any infestation of black swans.

Our ambivalence towards risk and uncertainty leaves us prone to irresilience. Ambivalence is easily tipped towards fear of losing the status quo. Rather than appropriate learning and unlearning, the uncertainty of self-organized criticality poised for cascading failures (Lewis 2011) adds to a sense of helplessness. This is unhealthy.

Contagion from super-spreading hubs (Haldane and May 2011) need not always mark a turn for the unendurable worse. Crises can be anticipated (Sornette 2009), their genealogy understood (Gorton 2012), and learning how to act decisively in such circumstances is feasible. It stems from our earliest experiences (Rutter 2012) and education (Tough 2013). The value of such learning becomes a mainstay to wealth creation (Beinhocker 2006) because healthy appetites for risk and uncertainty are fostered widely.

Dynamic networks require competencies for distributed decision-taking among diverse bodies not clones as hubs for irresilience. Leadership on these terms is less about leaders per se and more

about an enabling ethos for continuous learning within and between organizations (Augier and MacIntosh 2010). Yet as Rutter (1993) warns, understanding the value of resilience does not necessarily make it readily producible.

Combinations of Real Capability Options

The resilience challenge for strategic management is to continuously produce capabilities fit for evolving environments. Capabilities can be understood as evolving ecologies of competencies and technology (MacIntosh et al. 2012). This links the definition of capabilities as 'routines' that 'confer decision options' (Winter 2000) with meta-capabilities – that is, 'capabilities to acquire capabilities' (Teece et al. 1997). In so doing, evolutionary capabilities can combine the healthy exploitation of cognitive energy savings gained through selecting and performing well-honed drills with the greater rewards of exploring for breakthroughs that lead to disruptive ► innovation. This in no way underestimates the perils of learning myopia (Levinthal and March 1993) and the bias against exploration (March 2006).

Allowing 'exploitation to squeeze out exploration' is a failure of strategic management but it does occur. To avoid such failure aspirations must be kept high (Winter 2000) with strategy and leadership achieving two objectives:

- Integrating a wide variety of evolving capability options; and,
- Deepening the capacity for innovation into more composable options.

Neither of these can be achieved by organizational introspection (Dosi et al. 2000). All enterprises are enmeshed in dynamic risky relationships. Crises just make that obvious as the inadequacies of technology readiness levels and training needs analysis emerge. 'Transformation in contact' (Dannatt 2009) with adversity is made harder if the mesh of competencies and technology is thin and patchy. The agility and versatility for surge and mutual aid break down well before the exhaustion of combinations of real capability options. The resilience to overcome these organizational pathologies is not just an issue for integrated emergency management or improved project and programme management. It is strategic.

Transformative Capacity for Multi-Innovation

Confusing resilience with irresilience is unhealthy. Being overcome by the strategic challenges ahead is all too easy, particularly as the perils of learning myopia resonate with the great stagnation (Cowan 2011) and the uncertain prospects for a great rebalancing (Pettis 2013). A revolution in strategic management may seem warranted but such reformation tends to be undone by counter-reformation (Augier and March 2011). It is vital to have good diagnostics for resilience in terms of both depth and breadth.

Some idealists have called for the encouragement of 'de-growth' (Daly 1996), some urge the greater uptake of leisure (Skidelsky and Skidelsky 2012), whilst others advocate mass open innovation (Jeroen et al. 2010). Strategic management will find insurmountable problems with these prescriptions. An evolutionary process (Whitehead 1978) of learning that affirms resilience as transformational (Allen 2010), because it is grounded in empiricist combinations of real capability options, is more pragmatic. It has taken leading ecologists in the adaptive rut of 'panarchy' 30 years to realize this (Holling 1973, 2001).

Some ecologists are now acknowledging the economics of innovation (Moore and Westley 2011). Inasmuch as the move from seeing innovation as a (fractured) pipeline to an ecosystem is important, there are impediments that need to be acknowledged. First, academia's contribution to that ecosystem remains fragmented and largely entrenched at the invention end of a pipeline (Etzkowitz and Leydesdorf 2001). Secondly, irrespective of the unrelenting push of technological change at consumers, transformative research has been in deep deficit for decades (Braben 2004, 2008). Overcoming these impediments is vital to resilience and strategic management.

Composable design principles and diagnostics for enabling transformative research to enhance our capacity for multi-innovation is the ultimate measure of resilience.

The Value of Resilience

To endure involves becoming different (Prigogine 1997) rather than being the same. This quality of resilience cannot be effortless or risk-free, but it is vital. Selecting and enabling people with the cognitive competencies to deliver 'strategic resilience' (Marshalla and Ojiakoa 2010) helps. Resilience is not about buying protective measures to stop risk or actuarial insurance to displace it. Investing in resilience enhances the competitive fitness of bodies.

Strategic management increasingly recognizes that it must combine 'profitability' with 'growth' (Chakravarthy and Lorange 2010) to thrive in an uncertain world. Accounting standards that incorporate better measures of value (Pitelis and Vasilaros 2010) will be encouraging. Recent work on 'shared-value' (Porter and Kramer 2011) not only hints at healthier ways to reconcile competitiveness and cooperativeness but also the importance of more basic research into the value of resilience, for example, to be found in the advances of evolutionary economics (Foster and Metcalfe 2012). Strategic management worth the name will need these measures if, in the wake of the economic crises triggered in 2007, resilience is to grow rather than decay in the face of the even greater strategic challenges that lie ahead.

Enterprises or bodies that do not discover healthy combinations of capabilities through exploring and exploiting tend to fail sooner rather than later. Their irresilience becomes evident. Resilience lives with and seeks to learn from our bounded rationality (Simon 1955). Bodies that transform build resilience through the healthy uptake of innovation, both in response to and by shaping ever-changing environments.

See Also

- ▶ Competitiveness
- ▶ Innovation
- ▶ Leadership
- ▶ Organizational Design
- ▶ Organizational Learning

References

Allen, P.M. 2010. Complexity and management. In *Handbook of the philosophy of science*, Philosophy of Complex Systems, vol. 10, ed. C. Hooker. Amsterdam: Elsevier.

Augier, M., and J.P. MacIntosh. 2010. Beyond good and evil or a genealogy of ethics: Challenges for professional, unprofessional and amateur users of force. In *Military ethics: International perspectives*, ed. J. Stouffer and S. Seiler. Kingston: Canadian Defence Academy Press.

Augier, M.S., and J. March. 2011. *The roots, rituals, and rhetoric of change: North American Business Schools after the second World War*. Stanford: Stanford University Press.

Beinhocker, E.D. 2006. *The origin of wealth: Evolution, complexity and the radical remaking of economics*. London: Random House.

Braben, D.W. 2004. *Pioneering research: A risk worth taking*. Hoboken: Wiley.

Braben, D.W. 2008. *Scientific freedom: The elixir of civilisation*. Hoboken: Wiley.

Cascio, J. 2009. The next big thing: Resilience. *Foreign Policy* 172: 82–95.

Chakravarthy, B., and P. Lorange. 2010. *Profit or growth? Why you don't have to choose*. Switzerland: IMD.

Cowan, T. 2011. *The great stagnation*. London: Dutton.

Daly, H.E. 1996. *Beyond growth: The economics of sustainable development*. Boston: Beacon Press.

Dannatt, R. 2009. A perspective on the nature of future conflict. RIIA transcript of speech delivered at Chatham House, London, 15 May.

Dorogovtsev, S.N., and A.V. Goltsev. 2008. Critical phenomena in complex networks. *Review of Modern Physics* 80: 1275–1335.

Dorogovtsev, S.N., and J.F. Mendes. 2003. *Evolution of networks: From biological nets to the internet and WWW*. Oxford: Oxford University Press.

Dosi, G., R.R. Nelson, and S.G. Winter. 2000. *The nature and dynamics of organisational capabilities*. Oxford: Oxford University Press.

Etzkowitz, H., and L. Leydesdorf (eds.). 2001. *Universities and the global knowledge economy: A triple helix of University-Industry–Government Relations*. London: Continuum.

Foster, J., and J.S. Metcalfe. 2012. Economic emergence: An evolutionary economic perspective. *Journal of Economic Behavior & Organization* 82: 420–432.

Gorton, G.B. 2012. *Misunderstanding financial crises: Why we don't see them coming*. Oxford: Oxford University Press.

Haldane, A., and R. May. 2011. Systemic risk in banking ecosystems. *Nature* 469: 151–155.

Hodgson, G.M., and T. Knudsen. 2010. *Darwin's conjecture: The search for general principles of social and economic evolution*. London: University of Chicago Press.

Holling, C.S. 1973. Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics* 4: 1–23.

Holling, C.S. 2001. Understanding the complexity of economic, ecological, and social systems. *Ecosystems* 4: 390–405.
IEO (Independent Evaluation Office). 2011. IMF performance in the run-up to the financial and economic crisis: IMF surveillance in 2004–07. Available at http://www.ieo-imf.org/ieo/files/completedevaluations/crisis-%20main%20report%20(without%20moises%20signature).pdf. Accessed 9 May 2013.
Jenkin, B. 2010. Strategic defence and security review. Paper presented at House of Commons debate, 4 November, column 1117.
Jeroen, P.J., A.B.. de Jong, T. Kalvetc, and W. Vanhaverbekedef. 2010. Exploring a theoretical framework to structure the public policy implications of open innovation. *Technology Analysis & Strategic Management* 22: 877–896.
Keynes, J.M. 1921. *A treatise on probability*. London: Macmillan.
Knight, F.H. 1921. *Risk, uncertainty and profit*. Boston: Houghton Mifflin.
Levinthal, D.A., and J.G. March. 1993. The myopia of learning. *Strategic Management Journal* 14: 95–112.
Lewis, T. 2011. *Bak's sand pile: Strategies for a catastrophic world*. Williams: Agile Press.
MacIntosh, J.P., and M.S.D. Granatt. 2001. Building resilience: A new approach to crises. Cabinet Office.
MacIntosh, J.P., J.K. Mahrra, and L.R. Tyler. 2012. Transformation: DIG for realism. Working paper, Institute for Security & Resilience Studies at University College London. Available at http://www.ucl.ac.uk/isrs/publications/TransformationDIG4R. Accessed 6 June 2013.
March, J.G. 2006. Rationality, foolishness, and adaptive intelligence. *Strategic Management Journal* 27: 201–214.
Marshall, A., and U. Ojiako. 2010. From the myth of Prometheus to strategic resilience: Two cognitive paradigms linking risk and innovation. *Prometheus* 28: 343–360.
Moore, M.-L., and F. Westley. 2011. Surmountable chasms: Networks and social innovation for resilient systems. *Ecology and Society* 16: 5. Available at http://www.ecologyandsociety.org/vol16/iss1/art5/. Accessed 9 June 2013.
Pettis, M. 2013. *The great rebalancing: Trade, conflict, and the perilous road ahead of the world economy*. Oxford: Princeton University Press.
Pitelis, C.N., and V. Vasilaros. 2010. The determinants of value and wealth creation at the firm, industry, and national levels: A conceptual framework and evidence. *Contributions to Political Economy* 29: 33–58.
Porter, M.E., and M.R. Kramer. 2011. Creating shared value: How to reinvent capitalism – and unleash a wave of innovation and growth. *Harvard Business Review*, January–February, 1–17.
Prigogine, I. 1997. *The end of certainty: Time, chaos and the new laws of nature*. New York: Free Press.
Rose, A. 2007. Economic resilience to natural and man-made disasters: Multidisciplinary origins and contextual dimensions. *Environmental Hazards: Human and Policy Dimensions* 6: 1–16.
Rutter, M. 1993. Resilience: Some conceptual considerations. *Journal of Adolescent Health* 14: 626–631.
Rutter, M. 2007. Implications of resilience concepts for scientific understanding. *Annals of the New York Academy of Science* 1094: 1–12.
Rutter, M. 2012. Resilience as a dynamic concept. *Development and Psychopathology* 24: 335–344.
Simon, H.A. 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* 69: 99–188.
Simon, H.A. [1956] 1982. Reply: Surrogates for uncertain decision problems. In *Models of bounded rationality, vol. 1: Economic analysis and public policy*. Cambridge, MA: The MIT Press.
Simon, H.A. 2005. Darwinism, altruism and economics. In *The evolutionary foundation of economics*, ed. K. Dopfer. Cambridge: Cambridge University Press.
Skidelsky, R., and E. Skidelsky. 2012. *How much is enough? The love of money and the case for a good life*. London: Allen Lane.
Sornette, D. 2009. Dragon-kings, black swans and the prediction of crises. *International Journal of Terraspace Science and Engineering* 2: 1–18.
Teece, D.J., G. Pisano, and A. Shuen. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18: 509–533.
Tough, P. 2013. *How children succeed: Grit, curiosity and the hidden power of character*. London: Random House.
Watts, D.J. 2004. *Six degrees: The science of a connected age*. London: Vintage Books.
Whitehead, A.N. 1978. In *Process and reality*, corrected ed, ed. D.R. Griffen and D.W. Sherburne. New York: Free Press.
Winter, S.G. 2000. The satisficing principle in capability learning. *Strategic Management Journal* 21(special issue): 981–996.

Resource Allocation Theory

Joseph Bower
Harvard University, Business School, Boston, MA, USA

Abstract

This article considers the process of resource allocation, whereby an organization determines how best to apportion its factors of production between the various productive activities in which it wishes to engage. It is suggested that none of the academic approaches to date has provided an entirely coherent picture of the process, in part because of the contradictory models of the process that they generate.

The article goes on to consider the planning processes that are involved in assessing future projects and the way in which past outcomes feed into the assessment of future projects.

Definition Resource allocation is the process whereby an organization determines how to apportion its production factors among the various productive activities in which it aims to engage. The process has a number of different aspects, involving economic, social, political and technical considerations.

In the field of business management, there are few tasks for leadership as important as the allocation of the firm's resources. The field of economics conceives of firms as resource-allocating entities engaged in the production of goods and services. Sociology sees firms as providing social roles and income for their members. And management theory is devoted in large part to the planning and budgeting for the use of resources, while financial theory focuses in part on the use of capital resources.

Research that examines the resource allocation process reveals that none of the disciplines provides an adequate description of the process, or an adequate basis for theory-based prescription to guide the choices of managements. The reason is that resource allocation is an economic, technical, social and political process carried out by many different members of an organization occupying operating, management and leadership roles.

One theory focuses on the allocation of financial resources among options in the form of different projects presented to a top management conceived as a unitary actor. Another theory focuses on the preparation and approval of operating budgets. Still another is concerned with the way individuals are assigned to different roles in the organization, how their performance is incented and measured. And yet another theory is concerned with how the purpose of an organization should be conceived, and how a strategy should be devised to use resources so that the purpose may be achieved.

The challenge is that all of these theories are pertinent, but none is useful when applied independently of the others. And their prescriptions conflict, in part because of their contradictory assumptions about the process. Briefly, in finance, allocation is conceived as a problem of selection, by top management, from among projects that are adequately described by financial measures that are intrinsically comparable. All projects whose returns exceed the cost of capital should, in principle, be funded. Planning theory sees top management as providing aggregate goals that are factored into subgoals for each part of the business. The divisions then develop detailed business plans whose projected results will hopefully add up to the aggregate corporate goal. Where there is a gap, business unit goals are increased or the corporation designs an acquisition. The field of strategy conceives of a firm as a portfolio of business unit strategies that are more valuable as a package because of the leverage provided by the corporate strategy. Allocation then proceeds in two phases: a first in which choices are made among business units according to their prospects over time and their need for funds; and a second in which choices are made among alternative uses of funds to implement selected strategies.

The managers responsible for resource allocation know that the choice among projects they select (operating and capital budgeting) is important, but that a more important part of the process may well be the generation of possible projects, and then, among all the possible projects generated, the selection by lower levels of management of which projects to bring forward. Research has shown that these choices are influenced in important ways by the measurement and information systems of the firm, the budgeting and capital budgeting process, the organization, the measurement and reward of management performance and, embracing all, the degree to which the purpose of the firm has been articulated as an effective strategy for competing in changing markets.

In some of today's best-managed firms, the allocation of operating and investment capital is managed as part of a cycle that includes the planning and measurement of businesses, and the development, measurement and allocation of management talent. Whereas the allocation

Resource Allocation Theory, Table 1 Discounted actual results compared with discounted forecasts

Type of project	Mean[a] of PV actual results/PV forecasts
Cost reduction	1.1
Sales expansion	0.6
New products	0.1

[a]The variance around the means move from tight to wide as one goes from cost reduction to new products

problem in economics and finance is seen ultimately to be one of allocating scarce capital, modern approaches in management recognize that strategies are crafted over time as learning from implementation (feedback from operations, customers and competitive actions), enabling the refinement of strategies and the reallocation of human, physical and financial capital.

The importance of this observation is illustrated by simple results from early research showing the difference in the *ex post* results of projects of different types compared with the *ex ante* projections (see Table 1).

The numbers reveal a phenomenon remarked on by both Yogi Berra and Albert Einstein: the problem with forecasts is that they're about the future. The estimates of performance and returns in requests for funds may look concrete and precise, but the further out the results are projected and the more dependent they are on unknown factors (such as customer responses to new products), the less certain they may be. In this sense, a 14.2% projected return on an investment in cost reduction may be fairly assumed to be better than one projected at 12%. But returns on new products projected at 14% and 16% are clearly more or less the same thing since the ability to project is so poor.

Since managers understand this phenomenon their immediate question on seeing a request for capital is 'Who made the forecasts?' Where the source of a proposal is a credible general manager, attention shifts quickly to qualitative aspects of the plan, in order to see what strategic issues it raises. Indeed, in a kind of triage process, the attention of top management in resource allocation tends to be given to those plans coming from untested managers. Plans submitted by managers with good reputations tend to be approved quickly, and those from managers with poor reputations are deferred for further study; it is those in the middle that get a thorough review.

All of this is understood by the general managers responsible for business units – be they divisions or groups. They know that when they approve a strategic plan, budget or capital request, they are investing – putting on the line with their signature – their reputation for good judgement. They are close enough to the business to make informed judgements about the forecasts and the reasoning that underlies the numbers. And they make their judgements based upon their understanding of how they will be measured and rewarded. If their career will be wrecked by a project that goes awry, they will tend to be very risk averse – regardless of what the corporation says about their approach to growth and risk. In effect, whereas the corporation may face a portfolio of opportunity with a distribution of returns, they believe they are facing a win–lose proposition. Win and they may get a bonus. Lose and they lose their job. It is in this way that profitable corporations may turn conservative as their executives favour more certain returns – the cost reductions as opposed to the new products.

In turn, the operating managers of the divisions – the functional and first-level general managers – face similar choices. A factory manager, for example, is often assessed closely with metrics directly or closely related to capacity utilization. It should not be surprising, then, that factory managers are often reluctant to propose the building of new capacity until they are sold out. It is very unusual to find firms that build production capacity enough ahead of demand to protect share in a rapidly growing market. Strategically, it makes absolute sense for the corporation, but the resource allocation process must be managed so that the downside risk of that strategic decision is not borne solely by the factory manager.

A related problem in large organizations is that the resource allocation process is captured by the finance function. At that point, projects that cannot readily fit into the format used for managing the process often fall by the

wayside. New businesses that should be funded as if they were entrepreneurial ventures – phased commitment of funds with progress milestones triggering subsequent phases – are treated the same way as cost reductions with estimates met precisely, including annual returns. It is no wonder that in companies managed by heavy finance functions truly new businesses are overfunded and fail, because no room has been allowed for learning, or underfunded and fail, because inadequate resources have been invested in market development.

Whatever the specifics of the process in a given corporation, the realized strategy of the organization is manifest in the pattern of past resource allocation and operations. While intent can give direction to that pattern, it is how resources are committed and then used that determines what the organization has actually done.

See Also

- ▶ Decision-Making
- ▶ Organizational Culture
- ▶ Organizational Learning
- ▶ Resource-Based View
- ▶ Resource Dependence
- ▶ Strategic Decision-Making

Resource Dependence

Jeffrey Pfeffer
Stanford University, Graduate School of Business, Stanford, CA, USA

Abstract

Resource dependence is a macro-organizational theory that posits the following. Because organizations are open systems transacting with their environments, there are inter-organizational power-dependence relations that constrain organizational behaviour. To mitigate these constraints, firms do things such as appoint boards of directors to co-opt external dependence or engage in mergers and joint ventures to absorb interdependence. Moreover, the ability to deal with critical external constituencies influences the distribution of power *within* firms, so that inter-organizational power dynamics permeate organizational boundaries.

Definition Resource dependence is a macro-level theory of organizations that emphasizes the consequences of power-dependence relations between focal organizations and others in their environment for organizational behaviours such as mergers and board composition as well as for internal organizational power dynamics.

In the late 1960s, an open systems view of organizations (Scott 1998) took hold. Because organizations necessarily transact with other entities in their environments to acquire inputs such as financing or material, and to sell products and services to customers in exchange for money, organizations depend on their environments to survive. Survival is a primary goal of virtually all organizations. As such, neither profit nor non-profit nor governmental organizations are closed, self-contained entities, and therefore cannot be fully analysed or understood as such. Following Emerson's (1962) theory of power-dependence relations, in which power is the obverse of dependence, interdependence with entities in the environment leads to the possibility of these external actors influencing an organization, and vice versa. And proceeding from Thompson's (1967) insight that organizations will try to buffer their technical core from perturbations from the environment so they can operate as efficiently and effectively as possible, it logically follows that companies will attempt to manage relationships of external interdependence so as to preserve as much autonomy as possible.

Pfeffer and Salancik (1978) presented the results of empirical tests of resource dependence theory and elaborated its logic. There was evidence that US companies with more dependence on government contracts were more responsive to

affirmative action pressures (Salancik 1979) and that Israeli managers were more likely to accede to government requests to invest in development areas to the extent that they sold more to the government (Pfeffer 1972c). Burt (1983) showed that constraints emanating from resource dependence were economically significant, as the degree of constraint predicted profit margins, with less constrained sectors enjoying higher levels of profit.

Organizations were not only constrained by their dependence on other organizations in their environment but they also took actions to manage those constraints. Pfeffer (1972b) found that the percentage of sales and, to a lesser extent, purchases with other industries explained variations in inter-industry merger patterns, and did so even after controlling for industry profitability, a result subsequently replicated by Finkelstein (1997). Pfeffer (1972a) reported that the composition of boards of directors could be understood in part by examining patterns of transactional – buying and selling – interdependence, a result also demonstrated using more refined analytical methods by Burt et al. (1980). Patterns of joint ventures (Pfeffer and Nowak 1976) and alliances (Gulati and Gargiulo 1999) also mirrored the structure of transactional interdependence.

The third premise of resource dependence theory – that external contingencies would be reflected in internal power dynamics – has also received some empirical support. Thornton and Ocasio (1999), in their study of the book publishing industry, traced the change in who controlled publishing companies to the shift in industry emphasis from books and their content to financial considerations. Pfeffer (1992) noted that power inside electric utilities evolved from engineers to lawyers and those with business backgrounds, as the critical dependencies shifted from those with a more technical orientation to dealing with regulatory bodies and the financial markets. Pfeffer's (1973) study of the composition of hospital boards also showed that governing bodies reflected the power of various industry groups to the extent that such groups were powerful in the organizations' environments.

Critique

There have been three major critiques of resource dependence theory and its empirical foundations. First, scholars from the population ecology tradition (e.g., Hannan and Freeman 1989) argued that resource dependence, with its emphasis on adaptive strategic action in response to environmental contingencies, overlooked population dynamics – differential founding and survival rates – as important determinants of the distribution of organizational characteristics observed in organizational populations. Population ecology also argued that resource dependence overestimated the likelihood of organizational adaptation and downplayed how much inertia existed inside companies (Hannan and Freeman 1984).

Second, institutional theory (e.g., Scott 1995), which originally de-emphasized power dynamics in its focus on rules, roles, norms and institutional environments, more recently expanded its theoretical focus to acknowledge that laws and normative constraints themselves sometimes evolve from power dynamics among organizations. In that sense, institutional theorists like Scott argue that resource dependence can be seen as a special case of the more general institutional theory, which also takes an open-systems view.

Third, Davis (2009) has argued that the world (accurately) described by resource dependence at the time of its development in the late 1960s has changed so dramatically as to call into question the current relevance of the theory. Specifically, Davis maintains that the rise in power of the capital markets has made organizational attempts to manage interdependence through mergers and similar strategies moot, and has so severely restricted managerial attempts to negotiate autonomy as to make any theory with a managerialist orientation not very useful.

See Also

- ▶ Interlocking Directorates
- ▶ Organization Theory
- ▶ Organizational Ecology

References

Burt, R.S. 1983. *Corporate profits and co-optation*. New York: Academic.

Burt, R.S., K.P. Christman, and H.C. Kilburn. 1980. Testing a structural theory of corporate cooptation: Interorganizational directorate ties as a strategy for avoiding market constraints on profits. *American Sociological Review* 45: 821–841.

Davis, G.F. 2009. The rise and fall of finance and the end of the society of organizations. *Academy of Management Perspectives* 23: 27–44.

Emerson, R.M. 1962. Power-dependence relations. *American Sociological Review* 27: 31–41.

Finkelstein, S. 1997. Interindustry merger patterns and resource dependence: A replication and extension of Pfeffer (1972). *Strategic Management Journal* 18: 787–810.

Gulati, R., and M. Gargiulo. 1999. Where do interorganizational networks come from? *American Journal of Sociology* 104: 1439–1493.

Hannan, M.T., and J.H. Freeman. 1984. Structural inertia and organizational change. *American Sociological Review* 49: 149–164.

Hannan, M.T., and J.H. Freeman. 1989. *Organizational ecology*. Cambridge, MA: Harvard University Press.

Pfeffer, J. 1972a. Size and composition of corporate boards of directors: The organization and its environment. *Administrative Science Quarterly* 17: 218–228.

Pfeffer, J. 1972b. Merger as a response to organizational interdependence. *Administrative Science Quarterly* 17: 382–394.

Pfeffer, J. 1972c. Interorganizational influence and managerial attitudes. *Academy of Management Journal* 15: 317–330.

Pfeffer, J. 1973. Size, composition, and function of hospital boards of directors: A study of organizational-environment linkage. *Administrative Science Quarterly* 17: 218–228.

Pfeffer, J. 1992. *Managing with power: Politics and influence in organizations*. Boston: Harvard Business School Press.

Pfeffer, J., and P. Nowak. 1976. Joint ventures and interorganizational interdependence. *Administrative Science Quarterly* 21: 398–418.

Pfeffer, J., and G.R. Salancik. 1978. *The external control of organizations: A resource dependence perspective*. New York: Harper & Row.

Salancik, G.R. 1979. Interorganizational interdependence and responsiveness to affirmative action: The case of women and defense contractors. *Academy of Management Journal* 22: 375–394.

Scott, W.R. 1995. *Institutions and organizations*. Thousand Oaks: Sage.

Scott, W.R. 1998. *Organizations: Rational, natural, and open systems*. Englewood Cliffs: Prentice Hall.

Thompson, J.D. 1967. *Organizations in action*. New York: McGraw-Hill.

Thornton, P.H., and W. Ocasio. 1999. Institutional logics and the historical contingency of power in organizations: Executive succession in the higher education publishing industry, 1958–1990. *American Journal of Sociology* 105: 801–843.

Resource Redeployment

Gwendolyn Lee
Hough Graduate School of Business and Warrington College of Business Administration, Gainesville, FL, USA

Abstract

The article considers the issue of resource redeployment, which involves changes in the management of a firm's resources in different organizational and market settings. It considers the extent to which resources can be readily redeployed, which will be reflected in the degrees of similarity between the original setting and the new setting to which resources are redeployed. The article reviews developments in the literature about the subject of resource redeployment and explicates resource redeployment as a dynamic process of how resources are reconfigured as firms shift their boundaries.

Definition Resource redeployment is a process of resource management involving the decision and implementation of how to apply the firm's resources to different organizational and market settings.

Resource redeployment is a process of resource management involving the decision and implementation of how to apply the firm's resources to different organizational and market settings. Understanding how to deploy firm resources to compete in product markets has been a central concern in strategy research (Selznick 1957; Penrose 1959; Rubin 1973; Wernerfelt 1984; Barney 1991; Peteraf 1993;

Helfat et al. 2007). However, the issue of resource redeployment has received comparatively less attention. This entry reviews the emerging stream of literature that seeks to explicate resource redeployment as a dynamic process of how resources are reconfigured as firms shift their boundaries.

To What Extent Can Resources Be Redeployed?

For resources to be redeployed, they need to be fungible, where fungibility refers to an attribute of a resource that facilitates its application to different organizational and market settings (Teece 1980; Anand and Singh 1997; Anand 2004). However, resources vary in their degree of fungibility. Resources are less fungible when they are more 'sticky', in that their development is a result of the firm's path-dependent actions accumulated over time (Dierickx and Cool 1989). As illustrated in Sirmon et al. (2008: 924), 'managers will find it more difficult to redeploy a resource such as a large, specialized, and complex manufacturing facility than human capital'.

Besides fungibility, an inherent attribute, managers may be limited by bounded rationality in that the applications they find are restricted to closely related settings, even for otherwise fungible resources. Industry- and business-specific routines (Nelson and Winter 1982; Levinthal 1991) as well as context-specific ones (Galunic and Rodan 1998) reduce the probability that firms will recognize how resources might be applied in a different setting. In addition, resource redeployment is constrained by a firm's historical, political and cultural context, along with psychological costs associated with change (Oliver 1997). These constraints on resource redeployment suggest that resources are more likely to be redeployed when the initial setting in which the resources were developed and the subsequent setting to which the resources will be applied are more similar. Anand (2004) finds that the more similar the target and the acquirer are in technological profiles, the higher the number of ▶ acquisitions the acquirer undertakes for resource redeployment.

How Common Is Resource Redeployment?

The phenomenon of resource redeployment is more common than generally recognized. Resource redeployment may manifest as a firm's entry into a new business coupled with (complete or partial) exit from the firm's other business. In response to changing market conditions, firms shift from one business to another over time while utilizing the same (or an expanded) resource base (Helfat and Eisenhardt 2004). For instance, when diminishing demand causes industries to decline, as occurred, for example, in the tobacco industry, firms entered other consumer product businesses by leveraging cash generated plus the skills in marketing and distribution that they developed in their core business (Miles 1982). Firms like Du Pont, in response to diminishing government demand for military products, entered related industries by utilizing the managerial and skill base that it had built up in meeting demand before industry decline (Chandler 1990). Lockheed responded by diversifying into commercial satellite launching and applying the technological expertise of its defense business systems unit to commercial settings for collecting traffic tickets and child support payments. Other firms, such as Raytheon, diversified by acquiring complementary assets from firms outside their industry (Anand 2004).

Resource redeployment is also commonly observed after horizontal acquisitions – that is, the acquisition of one firm by another in the same industry. Targets and acquirers frequently redeploy resources, especially resources that face market failure (Capron et al. 1998). Specifically, acquirers redeploy R&D, manufacturing and marketing resources to and from targets, and redeploy managerial and financial resources to targets. For instance, the target may use the acquirer's sales network; the acquirer's managers may spend part of their time helping the target create new managerial capabilities.

What Is the Empirical Evidence on the Antecedents of Resource Redeployment?

Similarity in Settings of Resource Application

The similarity between the initial setting in which the resources were developed and the subsequent setting to which the resources will be applied may affect the potential for absorbing and exploiting the redeployed resources. Similarity in terms of overlap in knowledge stocks and knowledge flows enhances firms' absorptive capacity (Cohen and Levinthal 1990), and thus the ability to use the redeployed resource. Similarity in terms of key success factors makes the competitive environment more familiar; therefore, value-creating opportunities are more likely to be abundant for the redeployed resource. Capron et al. (2001) found empirical evidence that shows that the more similarities there are between the acquiring and target businesses, the more the acquirers are found to redeploy resources to and from targets.

The Extent of Resource Asymmetry Between Target and Acquirer

The extent of resource asymmetry between target and acquirer is found empirically to be an important antecedent of post-acquisition resource redeployment (Capron et al. 1998). The relative pre-acquisition strength of the merging businesses along five resource dimensions affects the direction of redeployment. The empirical evidence suggests that firms frequently seek targets with strengths that the acquiring firm can use, or seek targets with weaknesses that the acquiring firm can overcome.

Geographic Scope

Target geographic scope is empirically shown to affect resource redeployment (Anand et al. 2005). Global scope of the target has a strong and significant influence on post-acquisition resource redeployment from the target to the acquirer. This finding shed light on the literature in international business. Resource redeployment serves a mechanism through which multinational enterprises enable intra-corporate knowledge transfer, highlighting the potential to gain access to skills and recombine firms' geographically distributed resources.

Regulatory Macro Institutions

Capron and Guillen (2009) show that a country's underlying ideology as to how the corporation should be governed has a strong effect on the extent of post-acquisition redeployment of resources to and from the target. When the rights of the target's employees in the target country are better protected, there is less resource redeployment of technology, marketing or management to and from the target. This finding adds the nature of regulatory macro institutions as an antecedent of resource redeployment.

What Is the Empirical Evidence on the Performance Implications of Resource Redeployment?

Capabilities of the Merged Firm

Capron and Mitchell (1998) find significant performance implications of bilateral redeployment following horizontal acquisitions, where firms redeploy resources both from an acquirer to a target and from the target to the acquirer. A high degree of bilateral redeployment is found to be associated with improvements in the acquiring and target businesses' combined capabilities. Specifically, R&D skills, product quality, product cost and output flexibility are found to improve with a high degree of bilateral redeployment for technical (product innovation, manufacturing), commercial (sales networks, brand names, marketing expertise), administrative (supplier relationships, logistic expertise, managerial capabilities, staff personnel) and financial resources. Shorter time-to-market is found to be associated with a high degree of bilateral redeployment for all but commercial resources. The findings suggest which types of resources are best suited to bilateral redeployment.

Performance of the Merged Firm

Capron and Hulland (1999) find redeployment of three key marketing resources (brands, sales

forces and general marketing expertise) following horizontal acquisitions to affect the performance of the merged firm. First, product quality was found to increase with redeployment of the brand resource from acquirer to target. Second, product line breadth was found to improve significantly both when the sales force resource was redeployed from acquirer to target and when the general marketing expertise resource was redeployed from target to acquirer. Third, geographic coverage expanded with redeployment of the brand resource from acquirer to target and the general marketing expertise resources from target to acquirer. The findings suggest that resource redeployment can contribute to post-acquisition performance by enhancing the merging firms' revenues.

Post-Acquisition Asset Divestiture

Capron et al. (2001) show that post-acquisition resource redeployment leads to asset divestiture from the business that receives the redeployed resources, but not from the business that contributes the new resources. As explained in Capron et al. (2001), acquisitions generate excess resources in the process of redeploying resources across the merging firms. The process of redeployment tends to create redundancies and conflicts with existing resources. The firm will then tend to divest excess physical assets, shut surplus facilities and lay off surplus employees.

As such, divestiture will occur from the business that receives the redeployed resources, whether the recipient is the acquirer or the acquisition target. The finding in Capron et al. (2001) raises an alternative perspective that contrasts with the prevailing view, where divestiture is argued to occur at the business that provides the resources. Divestiture occurs at the target after providing the resources because the acquirer has already captured what it intended to obtain from the acquisition (Duhaime and Grant 1984; Hitt et al. 1990).

Market Exit

Lee et al. (2010) find that the relatedness between a firm's new business and the firm's other businesses increases the firm's speed of exit from the new business, because of the potential for resource redeployment. If the new business fails, the firm with other businesses closely related to the new business is likely to have more opportunities for resource redeployment inside the firm. These opportunities represent a vibrant internal market, allowing more of the investment in the new business to be recouped. As such, exit would occur sooner if the performance of the new business falls below expectations. This logic is consistent with Lieberman (1990)'s finding that ▶ multiproduct companies, as compared with single-product firms, were more likely to exit from declining chemical product markets. However, this logic contradicts the classic resource-based argument where business relatedness is expected to decrease exit because of the prospects for better performance due to synergy. In reconciling the contradicting views, Lee et al. (2010) develop a theoretical model and an empirical test for diagnosing the divergent effects on exit emanating from contemporaneous sharing of resources versus resource redeployment between businesses over time.

See Also

- ▶ Acquisition Strategy
- ▶ Market Structure
- ▶ Multiproduct Companies
- ▶ Organizational Restructuring
- ▶ Resource-Based Theories
- ▶ Resource-Based View
- ▶ Theory of the Firm

References

Anand, J. 2004. Redeployment of corporate resources: A study of acquisition strategies in the US defense industries, 1978–1996. *Managerial and Decision Economics* 25: 383–400.

Anand, J., and H. Singh. 1997. Asset redeployment, acquisitions and corporate strategy in declining industries. *Strategic Management Journal* 18(summer special issue): 99–118.

Anand, J., L. Capron, and W. Mitchell. 2005. Using acquisitions to access multinational diversity: Thinking beyond the domestic versus cross-border M&A

comparison. *Industrial and Corporate Change* 14: 191–224.
Barney, J.B. 1991. Firm resources and sustained competitive advantage. *Journal of Management* 17: 99–120.
Capron, L., and M. Guillen. 2009. National corporate governance institutions and post-acquisition target reorganization. *Strategic Management Journal* 30: 803–833.
Capron, L., and J. Hulland. 1999. Redeployment of brands, sales forces, and general marketing management expertise following horizontal acquisitions: A resource-based view. *Journal of Marketing* 63: 41–54.
Capron, L., and W. Mitchell. 1998. Bilateral resource redeployment and capabilities improvement following horizontal acquisitions. *Industrial and Corporate Change* 7: 453–484.
Capron, L., P. Dussauge, and W. Mitchell. 1998. Resource redeployment following horizontal acquisitions in Europe and North America, 1988–1992. *Strategic Management Journal* 19: 631–661.
Capron, L., W. Mitchell, and A. Swaminathan. 2001. Asset divestiture following horizontal acquisitions: A dynamic view. *Strategic Management Journal* 22: 817–844.
Chandler Jr., A.D. 1990. *Scale and scope: The dynamics of industrial capitalism*. Cambridge, MA: Belknap Press of Harvard University Press.
Cohen, W.M., and D. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* 35: 128–152.
Dierickx, I., and K. Cool. 1989. Asset stock accumulation and sustainability of competitive advantage. *Management Science* 35: 1504–1514.
Duhaime, I.M., and J.H. Grant. 1984. Factors influencing divestment decision-making: Evidence from a field study. *Strategic Management Journal* 5: 301–318.
Galunic, D.C., and S. Rodan. 1998. Resource recombinations in the firm: Knowledge structures and the potential for Schumpetarian recombination. *Strategic Management Journal* 19: 1193–1201.
Helfat, C.E., and K.M. Eisenhardt. 2004. Inter-temporal economies of scope, organizational modularity, and the dynamics of diversification. *Strategic Management Journal* 25: 1217–1232.
Helfat, C.E., S. Finkelstein, W. Mitchell, M.A. Peteraf, H. Singh, D.J. Teece, and S.G. Winter. 2007. *Dynamic capabilities: Understanding strategic change in organizations*. Malden: Blackwell.
Hitt, M.A., R.E. Hoskisson, and R.D. Ireland. 1990. Mergers and acquisitions and managerial commitment to innovation in M-form firms. *Strategic Management Journal* 11(summer special issue): 29–47.
Lee, G. K., T. B. Folta, and M. B. Lieberman. 2010. Relatedness and market exit. *Best Paper Proceedings of the Sixty-ninth Annual Meeting of the Academy of Management*.
Levinthal, D.A. 1991. Organizational adaptation and environmental selection-interrelated processes of change. *Organization Science* 2: 140–145.
Lieberman, M.B. 1990. Exit from declining industries: 'Shakeout' or 'stakeout'? *RAND Journal of Economics* 21: 538–554.
Miles, R.H. 1982. *Coffin nails and corporate strategies*. Englewood Cliffs: Prentice Hall.
Nelson, R.R., and S.G. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Belknap Press of Harvard University Press.
Oliver, C. 1997. Sustainable competitive advantage: Combining institutional and resource-based views. *Strategic Management Journal* 18: 697–713.
Penrose, E. 1959. *The theory of the growth of the firm*. White Plains: M. E. Sharpe.
Peteraf, M.A. 1993. The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal* 14: 171–191.
Rubin, P.H. 1973. The expansion of firms. *Journal of Political Economy* 81: 936–949.
Selznick, P. 1957. *Leadership in administration: A sociological interpretation*. New York: Harper & Row.
Sirmon, D.G., S. Gove, and M.A. Hitt. 2008. Resource management in dyadic competitive rivalry: The effects of resource bundling and deployment. *Academy of Management Journal* 51: 919–935.
Teece, D.J. 1980. Economics of scope and the scope of the enterprise. *Journal of Economic Behavior & Organization* 1: 223–247.
Wernerfelt, B. 1984. A resource-based view of the firm. *Strategic Management Journal* 5: 171–180.

Resource-Based Theories

Margaret Peteraf
Dartmouth College, Department of Management, Tuck School of Business, Hanover, NH, USA

Abstract

Resource-based theories, which include theories of resource acquisition and accumulation (such as strategic factor market theory and the competitive lifecycle), theories of the firm (such as the knowledge-based view) and theories of ▶ sustainable competitive advantage (such as dynamic capabilities and the relational view), all share a set of basic assumptions. These begin with the view that the firm comprises a bundle of productive resources and capabilities, and that heterogeneity in performance across firms stems from an underlying heterogeneity in their resources and capabilities.

Definition Resource-based theories are theories of the nature, behaviour and/or performance of firms, in which the unit of analysis is a resource or capability (or a bundled set of resources and capabilities) that a firm possesses, controls or accesses preferentially. Resource-based theories may concern either business-level or corporate-level strategy issues (or both). Some are equilibrium-based, while others employ dynamic, process-oriented explanations.

Resource-based theories are theories of the nature, behaviour and/or performance of firms, in which the unit of analysis is a resource or capability (or a bundled set of resources and capabilities) that a firm possesses, controls or accesses preferentially. Resource-based theories may concern either business-level or corporate-level strategy issues (or both). Some are equilibrium-based, while others employ dynamic, process-oriented explanations.

In general, resource-based theories view the firm in terms of its productive resources and capabilities, and attribute heterogeneity in performance across firms to an underlying heterogeneity in their resources and capabilities (Penrose 1959; Rumelt 1984; Wernerfelt 1984). Resource-based theories are concerned not with generic inputs or with resources and capabilities that are freely available to all firms. Rather, they are concerned with those that enable a firm to perform activities that create more value than the activities of rival firms, giving the firm a competitive advantage (Peteraf and Barney 2003). Moreover, they are particularly concerned with resources and capabilities whose characteristics support the attainment of a sustainable competitive advantage and facilitate the generation and capture of rents. That is to say, their focus is on resources that are relatively scarce, valuable, resistant to both imitation and substitution, durable and imperfectly mobile (Dierickx and Cool 1989; Barney 1991; Amit and Schoemaker 1993; Peteraf 1993; Collis and Montgomery 1995). Resources and capabilities that have such characteristics are said to constitute a firm's strategic assets (Amit and Schoemaker 1993) or comprise its core competence (Prahalad and Hamel 1990). While resource-based theories may concern tangible resources, attention has been concentrated on intangible and socially complex resources such as reputation, organizational culture and knowledge (e.g., Barney 1988; Dierickx and Cool 1989; Conner and Prahalad 1996).

Resource-based theories range widely in their thrust, which is not surprising since the ▶ resource-based view has proven to be a highly versatile tool. An important set of these is concerned with explaining sustainable competitive advantage and persistent profitability differences among firms. This includes much of the classic work on the resource-based view, such as Rumelt (1984), Barney (1991), Amit and Schoemaker (1993) and Peteraf (1993). But it also includes more dynamic approaches to sustainable advantage – most notably, work on dynamic capabilities (Teece et al. 1997; Eisenhardt and Martin 2000; Zollo and Winter 2002; Helfat et al. 2007; Teece 2007). Moreover, it includes work such as the 'relational view' (Dyer and Singh 1998), which extends the resource-based logic to situations that expand beyond the boundaries of a single firm. Some of the work in this general category focuses on specific aspects of the puzzle, such as how the value created is captured (e.g., Coff 1999; Lippman and Rumelt 2003), and how taking a resource-based view of substitutes can help firms mitigate their threat (Peteraf and Bergen 2003).

A second group of resource-based theories is directed principally towards explaining the origin of heterogeneous resource and capability positions (Maritan and Peteraf 2011). One strand of this literature concerns the way in which firms acquire resources in strategic factor markets (Barney 1986; Makadok 2001; Maritan and Florence 2008). Some of it extends the strategic factor market logic into the realm of mergers and acquisitions initiated for the purpose of acquiring complex combinations of resources and capabilities (e.g., Barney 1988; Capron and Shen 2007). A second strand concerns the processes by which firms develop and accumulate resources and capabilities internally (Dierickx and Cool 1989). Resource-based

theories in this tradition span a much broader set of topical domains. For example, they include work on the management of resource stocks and flows, including resource allocation processes (e.g., Maritan 2001) and the 'systems perspective' on resource and capability management (e.g., Gary et al. 2008). They also include work concerned primarily with the properties of the resource accumulation process (e.g., Pacheco-de-Almeida et al. 2008). A third strand concerns evolutionary mechanisms as an explanatory factor for heterogeneity in resources and capabilities across firms. This literature includes theories on the role of initial firm endowments or prehistories in the evolution and deployment of capabilities (e.g. Helfat and Lieberman 2002; Ahuja and Katila 2004). It also includes work on capability evolution and the capability lifecycle (Helfat and Raubitschek 2000; Helfat and Peteraf 2003). Other work, such as that taking a 'routine-based' perspective, connects elements of evolutionary economics to elements of the resource-based view (e.g., Winter 1995; Karim and Mitchell 2000).

Other resource-based theories aim to contribute a new theory of the firm and address questions regarding the nature, boundaries and growth of firms. Much of the work subsumed under the rubric of the 'knowledge-based view' is essentially resource-based in its approach. Examples of this include Kogut and Zander (1992), Grant (1996), and Nickerson and Zenger (2004). Other work concerns, more directly, issues regarding the scope of the firm and its boundaries. This includes resource-based work on corporate strategy, such as Montgomery and Wernerfelt (1988), Montgomery and Hariharan (1991), and Helfat and Eisenhardt (2004).

New resource-based theories continue to emerge. For example, resource-based reasoning is now being applied to theories of entrepreneurship, where competitive advantage may be temporary, offering Schumpeterian rather than ▶ Ricardian rents (e.g., Kor et al. 2007; Foss et al. 2008). Even in this setting, however, the basic resource-based reasoning concerning value creation and competitive advantage still applies (Peteraf and Barney 2003).

See Also

- ▶ Competitive Advantage
- ▶ Competitive Heterogeneity
- ▶ Firm Resources
- ▶ Imitability
- ▶ Imperfect Resource Mobility
- ▶ Resource-Based View
- ▶ Ricardian Rents

References

Ahuja, G., and R. Katila. 2004. Where do resources come from? The role of idiosyncratic situations. *Strategic Management Journal* 25: 887–907.

Amit, R., and P.J.H. Schoemaker. 1993. Strategic assets and organizational rent. *Strategic Management Journal* 14: 33–46.

Barney, J.B. 1986. Strategic factor markets: Expectations, luck, and business strategy. *Management Science* 32: 1230–1241.

Barney, J.B. 1988. Returns to bidding firms in mergers and acquisitions: Reconsidering the relatedness hypothesis. *Strategic Management Journal* 9: 71–78.

Barney, J. 1991. Firm resources and sustained competitive advantage. *Journal of Management* 17: 99–120.

Capron, L., and J.C. Shen. 2007. Acquisitions of private vs. public firms: Private information, target selection, and acquirer returns. *Strategic Management Journal* 28: 891–911.

Coff, R.W. 1999. When competitive advantage doesn't lead to performance: The resource-based view and stakeholder bargaining power. *Organization Science* 10: 119–133.

Collis, D.J., and C.A. Montgomery. 1995. Competing on resources. *Harvard Business Review* 73: 118–128.

Conner, K.R., and C.K. Prahalad. 1996. A resource-based theory of the firm: Knowledge versus opportunism. *Organization Science* 7: 477–501.

Dierickx, I., and K. Cool. 1989. Asset stock accumulation and sustainability of competitive advantage. *Management Science* 35: 1504–1511.

Dyer, J.H., and H. Singh. 1998. The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review* 23: 660–679.

Eisenhardt, K.M., and J.A. Martin. 2000. Dynamic capabilities: What are they? *Strategic Management Journal* 21: 1105–1121.

Foss, N.J., P.G. Klein, Y.Y. Kor, and J.T. Mahoney. 2008. *Strategic Entrepreneurship Journal* 2: 73–94.

Gary, M.S., M. Kunc, J.D.W. Morecroft, and S.F. Rockart. 2008. System dynamics and strategy. *System Dynamics Review* 24: 407–429.

Grant, R.M. 1996. Toward a knowledge-based theory of the firm. *Strategic Management Journal* 17: 109–122.

Helfat, C.E., and K.M. Eisenhardt. 2004. Inter-temporal economies of scope, organizational modularity, and the dynamics of diversification. *Strategic Management Journal* 25: 1217–1232.

Helfat, C.E., and M.B. Lieberman. 2002. The birth of capabilities: Market entry and the importance of pre-history. *Industrial and Corporate Change* 11: 725–760.

Helfat, C.E., and M.A. Peteraf. 2003. The dynamic resource-based view: Capability lifecycles. *Strategic Management Journal* 24: 997–1010.

Helfat, C.E., and R.S. Raubitschek. 2000. Product sequencing: Co-evolution of knowledge, capabilities and products. *Strategic Management Journal* 21: 961–979.

Helfat, C.E., S. Finkelstein, W. Mitchell, M. Peteraf, H. Singh, D. Teece, and S. Winter. 2007. *Dynamic capabilities: Understanding strategic change in organizations*. Malden: Blackwell.

Karim, S., and W. Mitchell. 2000. Path-dependent and path-breaking change: Reconfiguring business resources following acquisitions in the US medical sector, 1978–1995. *Strategic Management Journal* 21: 1061–1081.

Kogut, B., and U. Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* 3: 383–397.

Kor, Y.Y., J.T. Mahoney, and S.C. Michael. 2007. Resources, capabilities and entrepreneurial perceptions. *Journal of Management Studies* 44: 1187–1212.

Lippman, S.A., and R.P. Rumelt. 2003. A bargaining perspective on resource advantage. *Strategic Management Journal* 24: 1069–1086.

Makadok, R. 2001. Toward a synthesis of the resource-based and dynamic-capability views of rent creation. *Strategic Management Journal* 22: 387–401.

Maritan, C.A. 2001. Capital investment as investing in organizational capabilities: An empirically grounded process model. *Academy of Management Journal* 44: 513–531.

Maritan, C.A., and R.E. Florence. 2008. Investing in capabilities: Bidding in strategic factor markets with costly information. *Managerial and Decision Economics* 29: 227–239.

Maritan, C.A., and M.A. Peteraf. 2011. Building a bridge between resource acquisition and resource accumulation. *Journal of Management* 37: 1374–1389.

Montgomery, C.A., and S. Hariharan. 1991. Diversified expansion by large established firms. *Journal of Economic Behavior & Organization* 15: 71–89.

Montgomery, C.A., and B. Wernerfelt. 1988. Diversification, Ricardian rents, and Tobin's q. *RAND Journal of Economics* 19: 623–632.

Nickerson, J.A., and T.R. Zenger. 2004. A knowledge-based theory of the firm: The problem-solving perspective. *Organization Science* 15: 617–632.

Pacheco-de-Almeida, G., J.E. Henderson, and K.O. Cool. 2008. Resolving the commitment versus flexibility trade-off: The role of resource accumulation lags. *Academy of Management Journal* 51: 517–536.

Penrose, E.T. 1959. *The theory of the growth of the firm*. New York: Wiley.

Peteraf, M.A. 1993. The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal* 14: 179–191.

Peteraf, M., and J.B. Barney. 2003. Unraveling the resource-based tangle. *Managerial and Decision Economics* 24: 309–323.

Peteraf, M.A., and M.E. Bergen. 2003. Scanning dynamic competitive landscapes: A market-based and resource-based framework. *Strategic Management Journal* 24: 1027–1041.

Prahalad, C.K., and G. Hamel. 1990. The core competence of the corporation. *Harvard Business Review* 68: 79–91.

Rumelt, R.P. 1984. Toward a strategic theory of the firm. In *Competitive strategic management*, ed. R. Lamb. Englewood Cliffs: Prentice Hall.

Teece, D.J. 2007. Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal* 28: 1319–1350.

Teece, D.J., G. Pisano, and A. Shuen. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18: 509–533.

Wernerfelt, B. 1984. A resource-based view of the firm. *Strategic Management Journal* 5: 171–180.

Winter, S.G. 1995. Four Rs of profitability: Rents, resources, routines and replication. In *Evolutionary and resource based views: A synthesis*, ed. C.A. Montgomery. Amsterdam: Kluwer Academic Publishers.

Zollo, M., and S.G. Winter. 2002. Deliberate learning and the evolution of dynamic capabilities. *Organization Science* 13: 339–351.

Resource-Based View

Ting Xiao[2], Asli M. Arikan[1,2] and Jay B. Barney[2]
[1]Georgia State University, Robinson College of Business, Atlanta, GA, USA
[2]Ohio State University, Fisher College of Business, Columbus, OH, USA

Abstract

The resource-based view (RBV) has emerged as one of several important explanations of persistent firm performance differences in the field of strategic management. After passing through an intense period of theoretical development and proliferation in the early 1980s and early 1990s, basic RBV logic was established and began to have an impact on empirical research in the field. At the same

time, resource-based logic began to influence theoretical and empirical work in other non-strategic management disciplines including human resource management, marketing, management information systems, and operations research, among others.

Definition The resource-based view is a theory that explains variance in firm performance as a function of differences in the resources and capabilities of competing firms.

The Theoretical History of the RBV

The resource-based view (RBV), like any theory, draws on prior theoretical work in developing its predictions and prescriptions. In the case of the RBV, important prior theoretical work comes from at least four sources: (1) the traditional study of distinctive competencies, (2) Ricardian economics, (3) Penrosian economics and (4) the study of the anti-trust implications of economics. Each of these prior theories will be briefly discussed in turn.

Traditional Work on Distinctive Competencies

Since at least 1911, scholars have tried to answer the question, 'Why do some firms persistently outperform others?' Before economic approaches to answering this question began to dominate this discussion (beginning with Porter 1979 and continuing with Porter 1980, 1981, 1985), this effort focused on what were known as a firm's distinctive competencies. Distinctive competencies are those attributes of a firm that enable it to pursue a strategy more efficiently and effectively than other firms (Selznick 1957; Learned et al. 1969; Snow and Hrebiniak 1982; Hitt and Ireland 1985, 1986).

Ricardian Economics

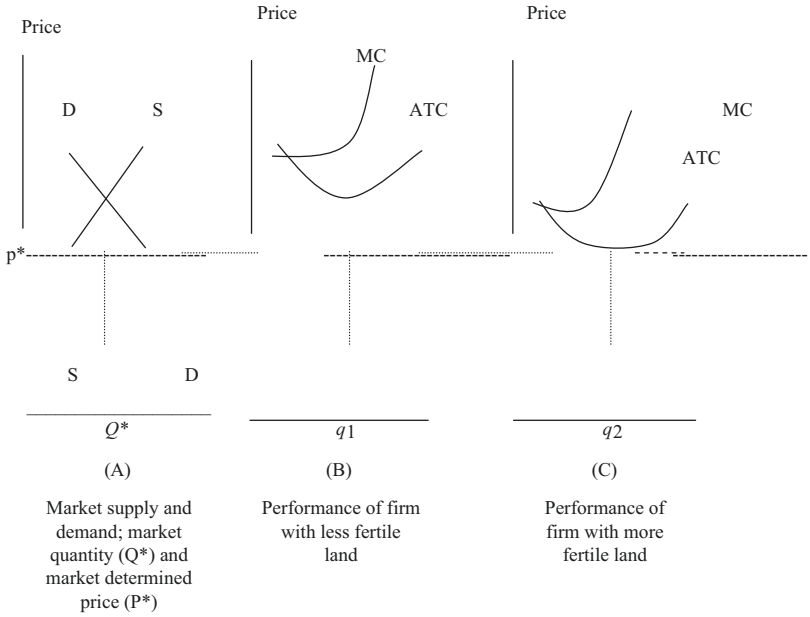
The next major influence on the evolution of the RBV, Ricardian Economic, focused on the economic consequences of the 'original, unaugmentable, and indestructible gifts of Nature' (Ricardo 1817). Much of this early work focused on the economic consequences of owning land.

Unlike many factors of production, the total supply of land is relatively inelastic in supply. In these settings, it is possible for those that own higherquality factors of production with inelastic supply to earn an economic rent. An economic rent is a payment to an owner of a factor of production in excess of the minimum required to induce that factor into employment (Hirshleifer 1980).

Ricardo's argument concerning land as a factor of production is summarized in Fig. 1. Imagine that there are many parcels of land suitable for growing wheat. Further, suppose that the fertility of these different parcels of land varies from high fertility (low costs of production) to low fertility (high costs of production). The long-run supply curve for wheat in this market can be derived as follows: at low prices, only the most fertile land will be cultivated; as prices rise, production continues on the very fertile land and additional crops are planted on less fertile land; at still higher prices, even less fertile land will be cultivated. This analysis leads to the simple market supply curve presented in panel A of Fig. 1. Given market demand, P^* is the market-determined price of wheat in this market.

Now consider the situation facing two different kinds of firms. Both of these firms follow traditional profit-maximizing logic by producing a quantity (q) such that marginal cost equals marginal revenue. However, this profit-maximizing decision for the firm with less fertile land (in panel B of Fig. 1) generates zero economic profit. On the other hand, the firm with more fertile land (in panel C of Fig. 1) has average total costs less than the market-determined price and thus is able to earn an economic rent.

In traditional economic analysis, the economic rent earned by the firm with more fertile land should lead other firms to enter into this market, to obtain some land and begin production of wheat. However, all the land that can be used to produce wheat in a way that generates at least zero economic profits given the market price P^* is already in production. In particular, there is no



Resource-Based View, Fig. 1 Ricardian rents and the economics of land with different levels of fertility

more very fertile land left, and fertile land (by assumption) cannot be created. This is what is meant by land being inelastic in supply. Thus the firm with more fertile land and lower production costs has a higher level of performance than farms with less fertile land, and this performance difference will persist, since fertile land is inelastic in supply.

Traditionally, most economists have implicitly assumed that relatively few factors of production have inelastic supply (Hirshleifer 1980). Most economic models presume that, if prices for a factor rise, more of that factor will be produced, increasing supply and ensuring that suppliers will earn only normal ▶ economic rents. However, the RBV suggests that numerous resources used by firms are inelastic in supply and are possible sources of economic rents. Thus although labour per se is probably not inelastic in supply, highly skilled and creative laborers may be. Similarly, although individual managers are probably not inelastic in supply, managers who can work effectively in teams may be. And although top managers may not be inelastic in supply, top managers who are also institutional leaders (as suggested by Selznick and others) may be. Firms that own (or control) these kinds of resources may be able to earn economic rents by exploiting them.

Penrosian Economics

In 1959 ▶ Edith Penrose published *The Theory of the Growth of the Firm*. Penrose's objective was to understand the process through which firms grow and the limits of growth. Traditional economic models had analysed firm growth using the assumptions and tools of neoclassical microeconomics (Penrose 1959). Most important of these, for Penrose, was the assumption that firms could be appropriately modelled as if they were relatively simple production functions.

This abstract notion of what a firm is had and continues to have utility in some circumstances. However, in attempting to understand constraints on the growth of firms, Penrose (1959) concluded that this abstraction was not helpful. Instead, she argued that firms should be understood, first, as an administrative framework that links and coordinates activities of numerous individuals and groups, and second, as a bundle of productive resources. The task facing managers was to exploit the bundle of productive resources controlled by a firm through the use of the administrative framework that had been created in a firm. According to Penrose, the growth of a firm is limited by (1) the productive opportunities that exist as a function of the bundle of productive

resources controlled by a firm, and (2) the administrative framework used to coordinate the use of these resources.

Besides looking inside a firm to analyse the ability of firms to grow, Penrose made several other contributions to what became the RBV. First, she observed that the bundles of productive resources controlled by firms can vary significantly by firm – that firms, in this sense, are fundamentally heterogeneous even if they are in the same industry. Second, Penrose adopted a very broad definition of what might be considered a productive resource. Where traditional economists (including Ricardo) focused on just a few resources that might be inelastic in supply (such as land), Penrose began to study the competitive implications of such inelastic productive resources as managerial teams, top management groups and entrepreneurial skills. Finally, Penrose recognized that, even within this extended typology of productive resources, there might still be additional sources of firm heterogeneity. Thus in her analysis of entrepreneurial skills as a possible productive resource, Penrose observed that some entrepreneurs are more versatile than others, that some are more ingenious in fundraising, that some are more ambitious, and that some exercise better judgement.

The Anti-Trust Implications of Economics

Economists have always been interested in the social policy implications of the theories they develop. One of the most important ways that economics has been used to guide social policy is in the area of anti-trust regulation. Based on the conclusion that social welfare is maximized when markets are perfectly competitive, economists have developed various techniques for describing when an industry is less than perfectly competitive, what the social welfare implications of this imperfect competition are, and what remedies, if any, are available to enhance competitiveness and restore social welfare (Scherer 1980).

One of the most obvious ways that an industry may be less than perfectly competitive is when that industry is dominated by only a single firm (the condition of monopoly) or by a small number of cooperating firms (the condition of oligopoly). In both these settings, according to traditional economic analyses, prices will be higher than in a competitive market, and thus social welfare will be diminished.

In the early 1970s, a small group of anti-trust scholars began to question traditional approaches to anti-trust regulation. For example, in 1973, Demsetz published an article in the *Journal of Law and Economics* that argued that a firm earning persistent superior performance cannot be taken as prima facie evidence that that firm was engaging in anti-competitive activities. Indeed, anticipating the RBV, Demsetz argued that some firms may enjoy persistent performance advantages either because they are lucky or because they are more competent in addressing customer needs than other firms. Demsetz (1973: 3) argues:

> Superior performance can be attributed to the combination of great uncertainty plus luck or atypical insight by the management of a firm... Even though the profits that arise from a firm's activities may be eroded by competitive imitation, since information is costly to obtain and techniques are difficult to duplicate, the firm may enjoy growth and a superior rate of return for some time...
>
> Superior ability also may be interpreted as a competitive basis for acquiring a measure of monopoly power. In a world in which information is costly and the future is uncertain, a firm that seizes an opportunity to better serve customers does so because it expects to enjoy some protection from its rivals because of their ignorance of this opportunity or because of their inability to imitate quickly.

While developed in the context of discussions of anti-trust regulation, Demsetz clearly anticipates some important tenets of resource-based logic.

The Development of Resource-Based Theory

Early Resource-Based Contributions

Perhaps the first resource-based publication in the field of strategic management identified as such was by Wernerfelt (1984). Ironically, Wernerfelt's resource-based arguments did not grow out of any

of the four theoretical traditions identified above. Rather, Wernerfelt's argument is an example of the dualistic reasoning that is common in economics. Such reasoning suggests that it is possible to restate a theory originally developed from one perspective with concepts and ideas developed in a complementary (or dual) perspective. For example, in microeconomics, it is possible to develop economic theories of decision-making using either utility theory, revealed preference theory, or state preference theory; in finance, it is possible to estimate the value of an investment using the Capital Asset Pricing Model (CAPM) or Arbitrage Pricing Theory. Wernerfelt (1984) attempted to develop a theory of competitive advantage based on the resources a firm develops or acquires to implement product market strategy as a complement or dual of Porter's (1980) theory of competitive advantage based on a firm's product market position.

One of Wernerfelt's (1984) primary contributions was recognizing that competition for resources and among firms based on their resource profiles can have important implications for the ability of firms to gain advantages in implementing product market strategies. In this way, Wernerfelt anticipated some of the critical elements of the RBV as it developed in the 1990s.

In the same year that Wernerfelt (1984) published his paper, Rumelt (1984) published a second resource-based paper in a book of readings that emerged from a conference on strategic management. While these papers addressed similar kinds of issues, they did not refer to each other. Where Wernerfelt (1984) focused on establishing the possibility that a theory of firm performance differences could be developed in terms of the resources that a firm controls, Rumelt began describing a strategic theory of the firm, that is, a theory explaining why firms exist and that focused on the ability of firms to generate economic rents. At its most general level, such a theory would suggest the conditions under which firms, as an example of hierarchical governance (Williamson 1975, 1985), would be a more efficient way to create and appropriate economic rents than other forms of governance, including markets. Rather than firms existing as efficient ways to minimize the threat of opportunism in transactions – as suggested by the transactions cost theorists (Williamson 1975) – Rumelt (1984) was exploring the rent generating and appropriating characteristics of firms.

The third resource-based article published in the field of strategic management is Barney (1986). Like Wernerfelt (1984), Barney (1986) suggests that it is possible to develop a theory of persistent superior firm performance based on the attributes of the resources a firm controls. However, Barney (1986) moves beyond Wernerfelt (1984) by arguing that such a theory can have very different implications from theories of competitive advantage based on the product market positions of firms.

Barney (1986) introduces the concept of strategic factor markets as the market where firms acquire or develop the resources they need to implement their product market strategies. He shows that if strategic factor markets are perfectly competitive, the acquisition of resources in those markets will anticipate the performance those resources will create when used to implement product market strategies. This suggests that, if strategic factor markets are perfectly competitive, even if firms are successful in implementing strategies that create imperfectly competitive product markets, those strategies will not be a source of economic rents. Put differently, the fact that strategic factor markets can be perfectly competitive implies that theories of imperfect product market competition are not sufficient for the development of a theory of economic rents.

Of course, strategic factor markets are not always perfectly competitive. Barney (1986) suggests two ways that such markets can be imperfectly competitive and thus two ways that firms can acquire or develop the resources they need to implement product market strategies in ways that generate economic rents. First, following Demsetz (1973), in the face of uncertainty, firms can be lucky. That is, if all the firms competing in a particular strategic factor market expect that resources acquired there will generate v levels of value in product markets, the price for those resources will quickly rise to v. However, if the actual value these resources can generate

is $v + x$, where x is some positive number, then firms that acquire this resource for v will earn an economic rent.

Second, also following Demsetz (1973), it may be the case that a particular firm has unusual insights about the future value of the resources it is acquiring or developing in a strategic factor market. Firms with these special insights will generally not overpay for a resource (when the market determined price for that resource is greater than its actual value in implementing a product market strategy) and will generally be able to acquire or develop undervalued resources (when the market-determined price for that resource is less than its actual value in implementing a product market strategy). By avoiding errors and taking advantage of opportunities, firms with special insights can earn economic rents. Barney then shows that many other apparent competitive imperfections in strategic factor markets are actually special cases of these other two competitive imperfections.

Barney (1986) concludes his paper by suggesting that the resources a firm already controls are more likely to be sources of economic rents than resources it acquires from external sources. This is because the resources a firm already controls were acquired or developed in a previous strategic factor market where their price was a function of the expected value of those resources in that market. However, if a firm can find new ways to use a resource to implement product market strategies, this new resource use would not have been anticipated in the original factor market and thus can be a source of economic rents.

Dierickx and Cool (1989) extended Barney's (1986) argument by describing what it is about the resources a firm already controls that may make it possible for that resource to generate economic rents. Following Rumelt's (1984) discussion of ▶ isolating mechanisms, Dierickx and Cool (1989) suggest that resources that are subject to time compression diseconomies, that are causally ambiguous, that are characterized by interconnected asset stocks, or that are characterized by asset mass efficiencies are less likely to be subject of strategic factor market competition than other kinds of resources. Many of the attributes of a firm's resources that prevent them from being subject to strategic factor market competition identified by Dierickx and Cool (1989) are later discussed and applied by Barney (1991a).

Together, these three papers – Wernerfelt (1984), Rumelt (1984), and Barney (1986) as extended by Dierickx and Cool (1989) – outline some of the basic principles of resource-based logic. These papers suggest that it is possible to develop a theory of persistent superior firm performance using a firm's resources as a unit of analysis.

These three papers also suggest some of the attributes that resources must possess if they are to be a source of sustained superior firm performance – Rumelt's (1984) concepts of value and 'isolating mechanisms' and Barney's (1986) notion that resources already controlled by a firm are more likely to be a source of economic rents than other kinds of resources. They further suggest that it is the bundle of special resources possessed by a firm that may enable a firm to gain and sustain superior performance.

Resource-Based Theory

Beginning in the 1980s and continuing through the 1990s, resource-based theory has been developed through the publication of numerous papers in a wide variety of journals. Some of the key definitions, assumptions, assertions, and predictions of this body of literature are presented here.

Definitions

Because resource-based theory is a theory, it is important to begin by defining some of its critical terms. First among these is the term *resources*. While this term has been defined elsewhere (e.g., Rumelt 1984; Wernerfelt 1984; Barney 1991b, 2001) current use of the term suggests the following definition:

> *Resources* are the tangible and intangible assets firms use to conceive and implement their strategies.

As was suggested earlier in this article, firms develop or acquire resources in *strategic factor*

markets. These markets may or may not be perfectly competitive.

The economic and strategic value of these tangible and intangible resources also varies. In general, resources are valuable when they enable a firm to develop and implement strategies that have the effect of lowering a firm's net costs and/or increasing a firm's net revenues beyond what would have been the case if these resources had not been used to develop and implement these strategies.

Of course, the tangibility of firm resources is a matter of degree. Resources that are typically more tangible include, but are not limited to, a firm's financial capital (e.g., equity capital, debt capital, retained earnings, leverage potential) and physical capital (e.g., the machines and buildings it owns). Resources that are typically less tangible include, but are not limited to, a firm's human capital (e.g., the training, experience, judgement, intelligence, relationships, and insights of individual managers and workers in a firm) and organizational capital (e.g., attributes of collections of individuals associated with a firm, including a firm's culture, its formal reporting structure, its reputation in the market place and so forth).

Assumptions

Resource-based theory, like all theories, adopts several assumptions. Many of these assumptions are consistent with other theories of persistent superior firm performance and thus will not receive particular attention here. For example, resource-based logic adopts the assumption that firms are profit-maximizing entities and that managers in firms are boundedly rational. Over and above these basic assumptions, resource-based logic makes two additional assumptions that distinguish it from other strategic management theories: the assumption of *resource heterogeneity* and the assumption of *resource immobility* (Barney 1991a). These assumptions are:

> *Resource heterogeneity*: competing firms may possess different bundles of resources.
>
> *Resource immobility*: these resource differences may persist.

Note that these two assumptions suggest that resource heterogeneity and immobility *may* exist. These assumptions do not suggest that all firms will always be unique in ways that are strategically relevant. Rather, these assumptions suggest that some firms, some of the time, may possess resources that enable them to more effectively develop and implement strategies than other firms and that these resource differences can last.

The concept of heterogeneity incorporates two attributes of firm resources: scarcity and non-substitutability (Barney 1991a). A firm resource is *scarce* when the demand for that resource is greater than its supply. A resource is *non-substitutable* when no other resources can enable a firm to conceive and implement the same strategies as efficiently or effectively as the original resource. The concept of immobility suggests that some resources, some of the time, may be *inelastic in supply*, that is, more of a particular resource is not forthcoming even though demand for that resource is greater than its supply. Firm resources may vary in the extent to which they are scarce, non-substitutable, and inelastic in supply.

Empirical Tests of Resource-Based Logic

A wide variety of hypotheses derived from resource-based theory have been examined in the strategic management and other literatures. Some of this work is briefly described below.

Industry Versus Firm Effects on Firm Performance

Initial work done by Schmalensee (1985) and Wernerfelt and Montgomery (1988) on industry versus firm effects in explaining variance in firm performance was inconsistent with resource-based expectations. In particular, this work suggested that industry effects were more important than firm effects. However, in 1991, Rumelt published an article that contradicted these earlier findings: Rumelt argued that previous work had applied the wrong methods or had used inadequate data to evaluate the relative impact of industry and firm effects on firm performance. After

solving these problems, Rumelt's results were consistent with resource-based expectations. Several authors have replicated Rumelt's results (e.g., Brush and Bromiley 1997; McGahan and Porter 1997; Mauri and Michaels 1998). Some of these are critical of Rumelt's findings, but primarily in terms of the small corporate effect that Rumelt (1991) identified (Brush and Bromiley 1997). However, all these replications continue to document that firm effects are a more important determinant of firm performance than industry effects, although the relative size of these effects can vary by industry.

Resources and Firm Performance

The bulk of empirical resource-based work in the field of strategic management has focused on identifying resources that have the attributes that resource-based theory predicts will be important for firm performance and then examining whether or not the predicted performance effects exist. The performance effects of a wide variety of different types of firm resources have been examined, including a firm's history (e.g., Collis 1991; Barnett et al. 1994; Rao 1994), employee know-how (e.g., Hall 1992, 1993; Glunk and Wilderom 1998), its integrative capability (e.g., Henderson and Cockburn 1994), its innovativeness (e.g., Bates and Flynn 1995; McGrath et al. 1996), its culture (e.g., Moingeon et al. 1998), and its network position (e.g., Baum and Berta 1999; McEvily and Zaheer 1999), to name just a few. Overall, results are consistent with resource-based expectations.

There are, however, a few studies that generate results that are inconsistent with resource-based expectations. For example, Poppo and Zenger's (1995) analysis of vertical integration is more consistent with transactions cost economics than resource-based theory.

Resources and Corporate Strategy

The impact of resources on corporate strategies has also been examined empirically. One of the most important findings in this area is that SIC-code based measures of strategic relatedness must be augmented by resource-based measures to capture the full performance effects of diversification strategies (e.g., Robins and Wiersema 1995; Farjoun 1998). Moreover, only when the basis of a diversification strategy is valuable, rare, and costly to imitate can firms expect such a strategy to generate superior firm performance (Markides and Williamson 1996).

Applications of Resource-Based Theory Beyond Strategic Management

Resource-based theory has also been used to generate testable hypotheses beyond the field of strategic management. This includes research in human resource management (e.g., Wright and McMahan 1992; Huselid 1995), marketing (e.g., Capron and Hulland 1999), entrepreneurship (e.g., Alvarez and Busenitz 2001), information systems (e.g., Ray et al. 2005), operations management (e.g., Klassen and Whybark 1999) and technology and innovation management (e.g., Helfat 1997).

Critique

The resource-based view has been criticized along several dimensions. Some suggest that it is only a view and not a theory. Others have argued that the analytical frameworks associated with the resource-based view, in particular, the VRIO framework of Barney (1991a), are tautological (Priem and Butler 2001). Despite these criticisms, the resource-based view continues to inform a great deal of research in strategy and related fields.

See Also

- ▶ Alliance Capability
- ▶ Business Strategy
- ▶ Causal Ambiguity
- ▶ Economic Rent
- ▶ Isolating Mechanisms
- ▶ Penrose, Edith T. (1914–1996)
- ▶ Uncertain Imitability
- ▶ Variance Decomposition

References

Alvarez, S.A., and L.W. Busenitz. 2001. The entrepreneurship of resource-based theory. *Journal of Management* 27: 755–775.

Barnett, W.P., H.R. Greve, and D.Y. Park. 1994. An evolutionary model of organizational performance. *Strategic Management Journal* 15: 11–28.

Barney, J.B. 1986. Strategic factor markets: Expectations, luck and business strategy. *Management Science* 32: 1512–1514.

Barney, J.B. 1991a. Firm resources and sustained competitive advantage. *Journal of Management* 17: 99–120.

Barney, J.B. 1991b. The resource-based view of strategy: Origins, implications, and prospects. *Journal of Management* 17: 97–211.

Bates, K.A., and E.J. Flynn. 1995. Innovation history and competitive advantage: A resource-based view analysis of manufacturing technology innovations. *Academy of Management Journal* 38: 235–239.

Baum, J., and Berta. 1999. Sources, dynamics, and speed: A longitudinal behavioral simulation of interorganizational and population-level learning. *Advances in Strategic Management* 16: 155–184.

Brush, T.H., and P. Bromiley. 1997. What does a small corporate effect mean? A variance components simulation of corporate and business effects. *Strategic Management Journal* 18: 325–835.

Capron, L., and J. Hulland. 1999. Redeployment of brands, sales forces, and general marketing management expertise following horizontal acquisitions: A resource-based view. *Journal of Marketing* 63: 41–54.

Collis, D.J. 1991. A resource-based analysis of global competition: The case of the bearings industry. *Strategic Management Journal* 12: 49–68.

Demsetz, H. 1973. Industry structure, market rivalry, and public policy. *Journal of Law and Economics* 16: 1–9.

Dierickx, I., and K. Cool. 1989. Asset stock accumulation and sustainability of competitive advantage. *Management Science* 35: 1504–1511.

Farjoun, M. 1998. The independent and joint effects of the skill and physical bases of relatedness indiversification. *Strategic Management Journal* 19: 11–630.

Glunk, U., and C.P.M. Wilderom. 1998. Predictors of organizational performance in small and medium-sized professional service firms. *International Journal of Technology Management* 16: 23–36.

Hall, R. 1992. The strategic analysis of intangible resources. *Strategic Management Journal* 13: 35–144.

Hall, R. 1993. A framework linking intangible resources and capabilities to sustainable competitive advantage. *Strategic Management Journal* 14: 607–618.

Helfat, C.E. 1997. Know-how and asset complementarity and dynamic capability accumulation: The case of R&D. *Strategic Management Journal* 18: 339–360.

Henderson, R., and I. Cockburn. 1994. Measuring competence? Exploring firm effects in pharmaceutical research. *Strategic Management Journal* 15: 63–84.

Hirshleifer, J. 1980. *Price theory and applications*. Englewood Cliffs: Prentice Hall.

Hitt, M.A., and R.D. Ireland. 1985. Corporate distinctive competence, strategy, industry and performance. *Strategic Management Journal* 6: 273–293.

Hitt, M.A., and R.D. Ireland. 1986. Relationships among corporate level distinctive competencies, diversification strategy, corporate structure, and performance. *Journal of Management Studies* 23: 401–416.

Huselid, M. 1995. The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal* 38: 635–672.

Klassen, R.D., and D.C. Whybark. 1999. The impact of environmental technologies on manufacturing performance. *Academy of Management Journal* 42: 599–615.

Learned, E.P., C.R. Christensen, K.R. Andrew, and W. Guth. 1969. *Business policy*. Homewood: Irwin.

Markides, C.C., and P.J. Williamson. 1996. Corporate diversification and organizational structure: A resource-based view. *Academy of Management Journal* 39: 340–367.

Mauri, A.J., and M.P. Michaels. 1998. Firm and industry effects within strategic management: An empirical examination. *Strategic Management Journal* 19: 211–219.

McEvily, B., and A. Zaheer. 1999. Bridging ties: A source of firm heterogeneity in competitive capabilities. *Strategic Management Journal* 20: 1133–1156.

McGahan, A.M., and M.E. Porter. 1997. How much does industry matter, really? *Strategic Management Journal* 18: 15–30.

McGrath, R.G., M.-H. Tsui, S. Venkataraman, and I.C. MacMillan. 1996. Innovation, competitive advantage and rent: A model and test. *Management Science* 42: 389–403.

Moingeon, B., B. Ramanantsoa, E. Metais, and J.D. Orton. 1998. Another look at strategy-structure relationships: The resource-based view. *European Management Journal* 16: 297–305.

Penrose, E.T. 1959. *The theory of the growth of the firm*. New York: Wiley.

Poppo, L., and T. Zenger. 1995. Opportunism, routines, and boundary choices: A comparative test of transaction cost and resource-based explanations for make-or-buy decisions. *Academy of Management Journal* 38: 42–46.

Porter, M.E. 1979. How competitive forces shape strategy. *Harvard Business Review* 57: 137–156.

Porter, M.E. 1980. *Competitive strategy*. New York: Free Press.

Porter, M.E. 1981. The contribution of industrial organization to strategic management. *Academy of Management Review* 6: 609–620.

Porter, M.E. 1985. *Competitive advantage*. New York: Free Press.

Priem, R.L., and J.E. Butler. 2001. Is the resource-based 'view' a useful perspective for strategic management research? *Academy of Management Review* 26: 22–40.

Rao, H. 1994. The social construction of reputation – certification contests, legitimation, and the survival of organizations in the American automobile-

industry – 1895–1912. *Strategic Management Journal* 15: 29–44.
Ray, G., W.A. Muhanna, and J.B. Barney. 2005. Information technology and the performance of the customer service process in North American insurance companies: A resource-based analysis. *Management Information Systems Quarterly* 29: 625–652.
Ricardo, D. 1817. *Principles of political economy and taxation*. London: J. Murray.
Robins, J., and M.F. Wiersema. 1995. A resource-based approach to the multibusiness firm: Empirical analysis of portfolio interrelationships and corporate financial performance. *Strategic Management Journal* 16: 277–299.
Rumelt, R. 1984. Toward a strategic theory of the firm. In *Competitive strategic management*, ed. R. Lamb. Englewood Cliffs: Prentice Hall.
Rumelt, R. 1991. How much does industry matter? *Strategic Management Journal* 12: 167–185.
Scherer, F.M. 1980. *Industrial market structure and economic performance*. Boston: Houghton Mifflin.
Schmalensee, R. 1985. Do markets differ much? *American Economic Review* 75: 341–351.
Selznick, P. 1957. *Leadership in administration*. New York: Harper and Row.
Snow, C.C., and L.G. Hrebiniak. 1982. Top-management agreement and organizational performance. *Human Relations* 35: 1139–1157.
Wernerfelt, B. 1984. A resource-based view of the firm. *Strategic Management Journal* 5: 171–180.
Wernerfelt, B., and C.A. Montgomery. 1988. Tobin's *q* and the importance of focus in firm performance. *American Economic Review* 78: 246–250.
Williamson, O.E. 1975. *Markets and hierarchies: Analysis and antitrust implication*. New York: Free Press.
Williamson, O.E. 1985. *The economic institutions of capitalism*. New York: Free Press.
Wright, P.M., and G.C. McMahan. 1992. Theoretical perspectives for strategic human resources management. *Journal of Management* 18: 295–320.

Return on Invested Capital (ROIC)

Carliss Y. Baldwin
Harvard University, Business School, Boston, MA, USA

Abstract

Return on invested capital (ROIC) is a financial measure of the profitability of a firm or business unit. If it is greater than the business's cost of capital, then reinvestment of earnings increases shareholder ▶ value. The ROIC also determines a maximum self-sustaining growth rate for the business in the absence of outside funding. Finally, for businesses engaged in Schumpeterian competition, innovators with an ROIC advantage can drive out their predecessors by making them unprofitable. In this fashion, relative ROIC determines an innovation's potential for 'creative destruction'.

Definition Return on invested capital (ROIC) is usually defined as the ratio of net operating profit after tax (NOPAT) to the sum of net fixed assets (NFA) and net working capital (NWC). The denominator equals the capital needed to operate the business. For a stand-alone firm, this is the same as the sum of the firm's debt (D) and equity (E).

$$ROIC = \frac{NOPAT}{NFA + NWC} = \frac{NOPAT}{D + E}$$

Return on invested capital (ROIC) is a financial ratio that measures the profitability of a business. The business can be a stand-alone firm or a unit in a larger corporation, but it must have a separate income statement and identifiable assets for the measure to be applicable.

Calculation

The analyst's objective in calculating an ROIC is to gauge the earning power of a business in relation to the capital needed to run it. ROIC is typically computed as net operating profit after tax (NOPAT) divided by the sum of net fixed assets (NFA) and net working capital (NWC). Fixed assets are net of accumulated depreciation, while net working capital is defined as current asset balances (operating cash, accounts receivable, inventories, prepaid expenses) less 'spontaneous' liabilities such as accounts payable, accrued expenses and accrued taxes. The overriding principle with respect to liabilities is to separate the 'money that you pay for', namely, debt and equity, from 'free' financing that comes from trade credit and lags in payments to employees and the

government. Reserve accounts should be classified according to this principle: if they carry an interest rate or share in the profits of the business, they are part of the invested capital base; if they do not, they should be subtracted from total assets in order to arrive at the invested capital figure.

On the asset side of the balance sheet, it is important to separate the assets needed to run the business from non-operating assets such as excess cash or securities and property held for investment purposes. Here the overriding principle is whether or not the business needs this asset in order to operate. Similarly, interest on excess cash and non-recurring profits (or losses) on sales of investments should be omitted in the calculation of net operating profit.

Dupont Identity

ROIC can be decomposed according to the following identity:

$$ROIC = \frac{NOPAT}{Sales} \cdot \frac{Sales}{Invested\ Capital}$$

The first term measures the business's profit margin while the second measures the 'turnover' (or efficiency) of its capital. This formula was invented by F. Donaldson Brown at Du Pont Company in the early part of the 20th century. It was used at Du Pont to evaluate the performance of businesses and to allocate capital to new investment opportunities (Chandler 1977: 446–448; Hounshell 1998).

ROIC as a Measure of Business Attractiveness

The ROIC measure can be used to determine whether an existing business is attractive in the sense of being a worthwhile target of investment and a candidate for growth. The basic test is to compare the business's ROIC with a weighted average cost of capital (WACC) based on the opportunity cost of debt and equity. For public companies in advanced capital markets, the WACC can be estimated using an equilibrium asset pricing model such as the capital asset pricing model (CAPM). For privately held firms and those in emerging markets, one can use target rates of return obtained from each class of investor.

A business whose ROIC is consistently above its weighted average cost of capital is an attractive opportunity. Accordingly, reinvesting earnings in the business (to fund its growth) is consistent with the goal of shareholder value maximization: shareholders will earn more from reinvestment than from comparably risky opportunities in the external capital market.

In a so-called perfect market with no taxes or transaction costs, shareholders are indifferent to the choice between reinvesting earnings versus raising external funds to finance the growth of a business (Modigliani and Miller 1958). But in actual capital markets, dividends are subject to taxation and new capital issuance is subject to transaction costs. Thus, Myers (1984) hypothesized that reinvested earnings should be the preferred method of funding business growth. (This is known as the 'pecking order' hypothesis.)

However, *reinvesting earnings in a business whose ROIC is less than its cost of capital destroys shareholder value*. If ROIC is less than the cost of capital, shareholders' alternative opportunities are more attractive than the business; hence shareholders are better served if earnings are paid out in the form of dividends and/or share repurchases.

Continuing to invest in unattractive businesses is a form of managerial agency cost: managers who behave this way are sometimes called empire builders. The problem can be addressed via a corporate takeover: the value destroyed by suboptimal reinvestment allows a takeover specialist to pay a premium to obtain control of the firm. Once in control, the new owner can institute more efficient investment policies and return the cash to the investors. Jensen (1986, 1993) argued that such agency costs were common in the 1970s and 1980s. In his view, hostile corporate takeovers and the leveraged buyout (LBO) form of corporate organization had the effect of reducing

managers' incentives to reinvest earnings in unattractive businesses.

ROIC and 'Sustainable Growth'

A business with constant operating ratios and no debt, which pays no dividends and does not issue new shares, can grow at the rate of its ROIC. A quick proof is as follows: let $I(t)$ denote the firm's invested capital at the beginning of time period t; and $\Delta I(t)$ the increase in invested capital during the period. Assuming a constant turnover ratio (see above), the business's rate of growth, $g(t)$ is $\Delta I(t)/I(t)$. Let $\Pi(t)$ denote the business's profit (NOPAT) during the period. If the firm pays no dividends, then all of its profit is available for reinvestment, and if it obtains no external capital in the form of new debt or equity, then this is the *only* money available:

$$\max \Delta I(t) = \Pi(t).$$

Substituting, we obtain:

$$g(t) \equiv \frac{\Delta I(t)}{I(t)} = \frac{\prod(t)}{I(t)} \equiv ROIC\ (t)$$

The growth rate in this formula is known as the firm's 'sustainable growth rate'. Note: for a firm with debt, which pays a constant percentage of earnings as dividends, $g(t) = (1 - d) \cdot ROE$, where d is the dividend payout ratio and ROE is the firm's return on equity.

Donaldson (1984) reported that 'self-sustaining growth' formulas were used as a planning tool at major US corporations whose goal was to maintain 'self-sufficiency', that is, they did not wish to rely on the capital markets to fund their growth. He also derived the two forms of the equation given above.

For some time, the formula was regarded as a mere curiosity: following Modigliani and Miller's (1958) argument, most scholars believed that the capital markets were virtually frictionless and thus a firm could always raise new capital at a fair price. Today, however, capital market frictions based on asymmetric information, agency and transaction costs are well-established facts. For equity issuance by large public corporations, investment banking fees alone create a wedge of 2–5% between the cost of internally generated funds and external funds. The difference is higher for small firms, family-owned firms and firms in emerging markets. Because of such frictions, sustainable growth is often an explicit or implicit constraint on business strategy, hence a first-order concern for business strategists and strategy scholars.

ROIC and Innovation

A business's ROIC determines how it will fare in competition against rivals in the product market. ▶ Joseph Schumpeter famously argued that a process of change and 'creative destruction' is the very essence of capitalism. 'New combinations', even though they temporarily disrupt the status quo, bring about material progress and economic development in the long run (Schumpeter 1934, 1942).

The ROIC measure lies at the very heart of the process of creative destruction. In the first place, for products and markets where there is no pre-existing competition, ROIC will be used by financiers to determine whether the new product or market is worth funding. In expectation, the profits from the new venture divided by the capital needed to create it must exceed the opportunity cost of capital. Completely new products or markets are relatively rare, however. In the more common case, new entrants compete against incumbents in an existing market. In such cases, the firm with the highest ROIC is most likely to succeed.

Baldwin and Clark (2006) shows that in head-to-head competition, subject to conditions discussed below, a firm with an 'ROIC advantage' vis-à-vis rivals, that is, a higher ROIC for all feasible product prices, can grow faster and more profitably in every time period (because of the sustainable growth formula). Furthermore, over time, it can drive its rival's ROIC below the cost of capital, while maintaining its own ROIC above that mark. In such circumstances, a value-maximizing rival will be forced to exit. A non-value-maximizing rival might hang on,

but would find itself cut off from the capital markets and possibly the target of a takeover (Jensen 1986). This argument can be applied to any of Schumpeter's 'new combinations': products, production processes, sources of supply, and organizational forms and practices.

Of course, there are boundary conditions on this analysis. Government regulation protecting incumbents can overturn the result. Large firms with deep pockets can drive out nascent challengers with superior ROICs, although, as Schumpeter realized, the large firms may do better to acquire or pre-empt such rivals. Finally, diversified companies can use profitable businesses to subsidize failing ones, but the wisdom of this practice, especially in the presence of an active takeover market, is questionable.

Thus, in head-to-head competition, absent subsidies to the weaker rival, the firm with the higher ROIC will prevail and the one with the lower ROIC will be forced to withdraw. An ROIC advantage is thus the acid test of an innovation's potential for 'creative destruction'.

Summary

ROIC is an operational measure of the profitability of a business relative to the capital employed. It measures the attractiveness of the business relative to the opportunity cost of capital, and reveals whether reinvestment of earnings is worthwhile. Further, it determines the maximum growth rate the business can sustain without recourse to outside funding. Finally, for businesses engaged in Schumpeterian competition, innovators with an ROIC advantage will, over time, drive their predecessors to the edges of their market and even out of existence; thus, ROIC is *the* criterion of innovative success and survival.

See Also

- ▶ Agency Problems
- ▶ Agency Theory
- ▶ Capital Asset Pricing Model (CAPM)
- ▶ Corporate Strategy
- ▶ Profit
- ▶ Resource Allocation Theory
- ▶ Schumpeter, Joseph (1883–1950)
- ▶ Value

References

Baldwin, C. Y. and Clark, K. B. 2006. *Architectural innovation and dynamic competition: The smaller "Footprint" strategy*. Harvard Business School Working paper No. 07-014. Available at: http://www.hbs.edu/faculty/Pages/item.aspx?num=22503. Accessed 28 Jan 2013.

Chandler, A.D. 1977. *The visible hand: The managerial revolution in American business*. Cambridge, MA: Harvard University Press.

Donaldson, G. 1984. *Managing corporate wealth*. New York: Praeger.

Hounshell, D. A. 1998. Measuring the return on investment in R&D: Voices from the past, visions of the future. In *Assessing the value of research in the chemical sciences: Report of a workshop*. Washington, DC: National Academy Press. Available at: http://www.nap.edu/catalog/6200.html. Accessed 28 Jan 2013.

Jensen, M.C. 1986. Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review* 76: 323–329.

Jensen, M.C. 1993. The modern industrial revolution, exit, and the failure of internal control systems. *Journal of Finance* 48: 831–880.

Modigliani, F., and M. Miller. 1958. The cost of capital, corporation finance and the theory of investment. *American Economic Review* 48: 261–297.

Myers, S.C. 1984. The capital structure puzzle. *Journal of Finance* 39: 575–592.

Schumpeter, J.A. 1934. *The theory of economic development*. Cambridge, MA: Harvard University Press.

Schumpeter, J.A. 1942. *Capitalism, socialism, and democracy*. New York: Harper & Brothers.

Reverse Engineering

Bhavna Hariharan
Kozmetsky Global Collaboratory, Stanford, CA, USA

Abstract

Reverse engineering is employed to understand the physical and functional details of a product with the intentions of replicating or

redesigning the original. Used extensively with physical products and software, this entry provides an overview of the method, technology, and legal issues associated with it.

Definition Reverse engineering is a method of working backwards from a finished product to understand how it works, how it was created with the intention of replication and redesign.

Introduction

Reverse engineering (RE) is a method employed to understand the physical and functional details of a product with the intention of replication, re-creation, or redesign by discovering both the dimensions and the design decisions that were made to arrive at the final product. Considered a subset of systems engineering, RE is undertaken to improve "one's own product as well as to analyze a competitor's product" (Cross et al. 1992). It has been described as "fundamentally directed to discovery and learning" (Samuelson and Scotchmer 2002) "albeit of a path already taken" (Pooley 1997).

RE is applied to complex systems which can be broken down into a hierarchy of subsystems and elements. It has been widely used with hardware and software but has also been applied to product design pedagogy (e.g., Otto and Wood; Becker 2007), studying biological diversity (e.g., Csete and Doyle 2002), complex cultural concepts (Taves 2015), and understanding the brain (e.g., O'Connor et al. 2009; Cauwenberghs 2013) among others.

Overview of the RE Method

RE is "starting with the known product and working backwards to divine the process which aided in its development or manufacture" (Kewanee Oil Corp v. Bicorn Corp. 1974). It is undertaken without access to original designer, drawings, or documentation about the product and the aim is to ascertain both functional and dimensional specification. Dimensional specifications include dimensions of the part, tolerances, materials used, and any processing done to it during manufacture and method of assembly. Functional specifications refer to the functions performed by the product, the expected performance and interaction of the parts within the product. In addition, RE also includes design intent, design requirements, constraints, and an understanding of the design alternatives that were available and the parameters considered to arrive at the design decisions that led to the final product.

RE has two subareas – Redocumentation and design recovery (Chikofsky and Cross II 1990). Redocumentation is using detailed dimensional and functional specifications to construct a blueprint of the product that would enable its manufacture and function without any additional information. Design recovery is the reclamation of the contextual information relevant to the part or product being reverse engineered to "fully understand what a program [or product] does, how it does, why it does it, and so forth" (Biggerstaff 1989).

RE of Physical Products

Rekoff describes RE as "the process of developing a set of specifications for a complex physical system by an orderly examination of specimens of that system" (Rekoff 1985). A generic five step process includes (1) assimilating existing information about the part as can be gleaned from existing documentation and contextual information, (2) postulate the elements that the product is made of to enable subsequent disassembly, (3) disassembly of the product into subsystems and elements and postulate their role in the overall system, (4) analyze, test, and dimension the elements to verify whether the postulates previously formulated are correct or need to be updated, and finally (5) complete documentation of the RE for future use and communication. (For a practicing RE, see Ingle 1994).

Dimensional information can be collected manually or using three-dimensional digitizers and scanners to generate a model of the part by collecting the (x, y, z) coordinates of all points that

make up the surface(s) of the part (point cloud). The widespread use of CAD technologies has resulted in updating the RE definition leaving its aims unchanged. RE is a "process in which designers acquire a design concept of a product from digitization of a physical model, and create the CAD model to realize approximation to the physical model: the model created can be reused, modified and optimized" (Zhang 2003).

New technologies for scanning increasingly complex parts are being invented and tested, Computer tomography for example (Marinsek and Paolasini 1999). Simultaneously advancements are being made in software that capture dimensional, functional, and design intent. (See Bradley and Currie 2005). As a case in point, knowledge based reverse engineering allows for expert knowledge about features and constraints to be added to point clouds. This enables the creation and maintenance of a model that is better suited for reuse and redesign (Durupt et al. 2011).

RE of Software

RE of software is undertaken "to learn about the structure and organization of the product or to learn its algorithm" (Association of the Bar of the City of New York 1989). It is done for commercial and noncommercial purposes and can result in competing or noncompeting alternative software. Another application of RE is maintenance of software systems. A common reason for RE of software is that it needs to work with updated operating systems or new microchips. "All too frequently, to interface adequately two pieces of software, some level of RE needs to be undertaken" (Hall 1992). RE techniques also "expedite new development by examining how similar systems are constructed. By examining the internals of another system, development team designers can make more informed design decisions for their situation" (Cross et al. 1992).

Typically, contextual information about design alternatives, constraints, and decisions made is not captured explicitly and remains as tacit knowledge of software developers. In cases where documentation is absent or people who developed the code have left or the code has been acquired from or maintained externally, the source code is the only reliable information source. RE can also be applied to databases. The difference is "while the main focus of code RE is on improving human understanding about how this information is processed, database RE tackles the question of what information is stored and how this information can be used in a different context" (Müller et al. 2000).

RE tools allow higher-level information to be extracted from source code by managing the complexities of understanding the software system. These tools generate alternate views, recover lost information, synthesize higher abstraction, and facilitate reuse (Chikofsky and Cross II 1990). RE tools are being developed and improved to understand ever more complex systems and interface with legacy systems.

Legal Issues

Software is protected by copyright law. This law protects the original expression of an idea but not any "idea, procedure, process, system, method of operation, concept, principle or discovery" (The Copyright Act of 1976). Software is different from other literary works in that it is a set of instructions that cause some action to be performed. Source code is written in higher-level languages that resemble English. These are converted into binary (sequence of 1 s and 0 s) during compilation and it is this object code that the computer understands. It is very difficult to differentiate between the code and the functions it performs. Copyright law protects the original source code but not the tasks that it performs.

Furthermore, the Uniform Trade Secrets Act offers protection for "any formula, pattern, device or compilation of information which gives him an opportunity to obtain an advantage over those who do not know or use it" (Restatement of Torts), but it "does not forbid the discovery of the trade secret by fair and honest means, e. g., independent creation or RE" (416 U.S. at 491). (See US Trade Secrets Act 2016). In many cases, it becomes necessary to arrive at the source code by working backwards from nonreadable machine language and careful

consideration is required before using RE on a competitor's product (Beherens and Levary 2008). There are arguments, both in support of RE as a legal way to reduce monopoly and enable innovation and also on RE as infringing on the rights of innovators and companies that come up with novel software (See Raskind 1985; Samuelson 1990; Gilbert-Macmillan 1993).

Conclusion

The roots of RE can be traced back to World War II, the Industrial Revolution or even the building of the pyramids in 2560 BC (Messler 2013). It has found application in different fields, and the technology for accomplishing RE has evolved to enable understanding increasingly complex systems. Research continues on RE related hardware, software, and legal disputes.

References

Becker, Katrin. 2007. Instructional ethology: Reverse engineering for serious design of educational games. In *Proceedings of the 2007 conference on future play*, 121–128. ACM.

Beherens, C. Brian, and Reuven R. Levary. 2008. Legal aspects of reverse engineering. In *Reverse engineering*, 195–206. London: Springer.

Biggerstaff. 1989. Look up from Cross2, chikofsky and May.

Bradley, Colin, and Bernadette Currie. 2005. Advances in the field of reverse engineering. *Computer-Aided Design and Applications* 2(5): 697–706.

Cauwenberghs, Gert. 2013. Reverse engineering the cognitive brain. *Proceedings of the National Academy of Sciences* 110(39): 15512–15513.

Chikofsky, Elliot Jcross II. 1990. Reverse engineering and design recovery: A taxonomy. *ABI/Inform Global*.

Cross, James H., Elliot J. Chikofsky, and Charles H. May. 1992. Reverse engineering. *Advances in Computers* 35: 199–254.

Csete, Marie E., and John C. Doyle. 2002. Reverse engineering of biological complexity. *Science* 295(5560): 1664–1669.

Durupt, Alexandre, Sebastien Remy, and Guillaume Ducellier. 2011. Reverse engineering of a piston using knowledge based reverse engineering approach. In *Global product development*, 683–690. Berlin/Heidelberg: Springer.

Gilbert-Macmillan, Kathleen. 1993. Intellectual property law for reverse engineering computer programs in the European community. *Index to Legal Periodicals and Books*: 247–264.

Hall, Pav. 1992. Overview of reverse engineering and reuse research. *Web of Science Core Collection*: 239–249.

Ingle, Kathryn A. 1994. *Reverse engineering*. New York: McGraw-Hill Professional Publishing.

Kewanee oil co v Bicron corp. 1974. 416 US 470, 476.

Marinsek, G., and S. Paolasini. 1999. Reverse engineering. In *AMST'99*, 557–564. Vienna: Springer.

Messler, Robert. 2013. *Reverse engineering: Mechanisms, structures, systems & materials*. McGraw Hill Professional.

Müller, Hausi A., Jens H. Jahnke, Dennis B. Smith, Margaret-Anne Storey, Scott R. Tilley, and Kenny Wong. 2000. Reverse engineering: A roadmap. In *Proceedings of the conference on the future of software engineering*, 47–60. ACM.

O'Connor, Daniel H., Daniel Huber, and Karel Svoboda. 2009. Reverse engineering the mouse brain. *Nature* 461(7266): 923–929. https://doi.org/10.1038/nature08539.

Ordem, de Advogados de Nova Iorque. 1989. Reverse engineering and intellectual property law. *The Record of the Association of the Bar of the City of New York*.

Raskind, Leo J. 1985. Reverse engineering, unfair competition, and fair use. *Periodicals Index Online*: 385.

Rekoff, M.G. 1985. On reverse engineering. *INSPEC:* 244–252.

Restatement of Torts § 757, Comment (b). 1939.

Samuelson, Pamela. 1990. Reverse-engineering someone else's software: Is it legal? *IEEE Software* 7(1): 90–96.

Samuelson, Pamela, and Suzanne Scotchmer. 2002. The law and economics of reverse engineering. *The Yale Law Journal* 111(7): 1575–1663.

Taves, Ann. 2015. Reverse engineering complex cultural concepts: Identifying building blocks of 'religion'. *Journal of Cognition and Culture* 15(1–2): 191–216. https://doi.org/10.1163/15685373-12342146.

The Copyright Act of 1976. 17 USC §§ 101–810

Reverse Internalization

Stephen Tallman
University of Richmond, Robins School of Business, Richmond, VA, USA

Abstract

Reverse internalization suggests that foreign locations are no longer just markets to be internalized by multinational enterprises (MNEs), but are becoming important sources of global competitive advantage as well. This suggests

three reversals in the modern global marketplace. First, source reversal, or the sourcing of assets from host, rather than home, markets; second, content reversal, a new focus on profiting from organizing capabilities rather than from unique technologies; and third, process reversal, by which the internalization of transactions for key assets is replaced by increasing use of external, more market-like, means of exchange.

Definition The process by which multinational enterprises source assets from host markets, rather than home; profit from skills in creating global customer value rather than exploitation of specific technologies; and employ market and hybrid means of governance as well as internal control to access the economic value of these assets.

Reverse internalization expresses the sense that modern multinational enterprises (MNEs) source their resources and capabilities from foreign host markets as well as their domestic or home markets. It also suggests that traditional perceptions of strategic assets as unique and superior technologies or brands with global application may ignore local and organizational capabilities that also generate competitive advantage. Finally, it suggests that the traditional model of internal or hierarchical control of the strategic processes of the MNE is giving way to more external or market-based means of governance.

Internalization: The Conventional Model

The traditional model of the MNE has at its core the concept of 'internalization', that firms in possession of unique assets derived from their home markets will exploit these assets in foreign markets through foreign direct investment (FDI) that allows their extension and application across country borders while retaining control within the boundaries of the firm (Buckley and Casson 1976). The fullest expression of this model is the Eclectic Model of FDI proposed by ▶ Dunning, John **(1927–2009)** (1988). In this model, Dunning proposes that three conditions must be met in order for firms to expand beyond their home country borders: firms must own a set of unique assets, these assets must be augmented by foreign location-bound factors of production, and the most efficient means of controlling the process of exploiting these assets must be through internal means. The creation and expansion of the MNE is the outcome of a series of decisions to internalize previous or potential external market transactions.

The MNE is defined by its internalization of international market (external) transactions for intermediate goods, services and knowledge derived from its home market but sold in foreign host markets either directly (via licensing) or by incorporation into final products (via exporting). This process is covered in detail in the entries on ▶ internalization theory. A critical assumption of this classic model is that firms generate their unique assets, whether hard assets, patents, trade secrets or complex processes in their home markets. Another assumption is that internalization of control is a response to conditions of market failure for both the final product and the underlying technologies that can be both exogenous (shipping costs, trade barriers) and endogenous (information asymmetries, opportunistic partners).

Reverse Internalization in the Global Marketplace

Contrary to the conventional model, both casual observation and empirical research have made clear that not all profitable resources of MNEs originate in the home country market, and indeed they show that MNEs often sell products developed in foreign markets using technologies developed, at least in part, in those same markets. Further, MNEs are creating significant value not by creating unique product and process technologies that require the protection of internalization, but by managing global value chains through which a variety of unique technologies are accessed via market, hybrid (alliance) and internal transactions and bundled for global markets. The

MNE profits from the sort of organizational, logistical and relational capabilities that have long been seen as complementary, not profit-generating, assets. The trend among global firms is movement towards flexibility in the sourcing of, the content basis for and the means of control over their strategic assets – all reversals of the classic internalization model.

Sourcing Reversals

MNEs are finding valuable assets in what were originally seen as foreign host markets. No longer just sources of new demand for existing products, these locations are providing inputs of resources to MNEs and forcing the development of capabilities that can be applied profitably in both global and multiple local markets. In addition, MNEs are building valuable organization capabilities through their global operations that allow these firms to present unique value to customers in many widespread markets, gaining economic value from coordinating international activities rather than from competing in host national markets (Tallman and Fladmoe-Lindquist 2002). An example is GE's use of 'reverse innovation' to develop a portable ultrasound machine in China which is now being rolled out in the USA and other developing markets (Immelt et al. 2009). Another example is Apple's iPhone, which is assembled in China from outsourced parts gathered in many locations and then sold in worldwide markets. Apple designs and markets the iPhone and orchestrates its production, but provides none of the core electronic technology or traditional production in the USA. We see that, rather than sourcing assets with economic value from the home country, firms are gaining value by sourcing assets and activities in foreign markets or by operating across markets. Offshore production is not limited only to low-cost manufacturing, and certainly not only to local markets. Indeed, foreign subsidiaries are now given global strategic mandates to not just produce, but to innovate and develop new products for world markets – to include sales back into the home market (Birkinshaw 1996).

Content Reversals

The traditional internalization model of the MNE focuses on ownership of key technologies and brands, typically sourced from the home market, and the drivers of real value in the goods and services offered in final markets – home or host. Reverse internalization suggests that, while access to the latest technologies may be vital, the real economic benefits accrue to the MNE with the best capabilities for identifying and assembling the most innovative bundles of assets and distributing the resulting unique final goods and/or services. Conventional models have assigned the greatest economic value to unique and defensible knowledge resources, and have treated the complex organizing capabilities that surround these resources as complementary – important, but not unique – and therefore not typically or justifiably the basis for competitive advantage (Teece 1986). The modern MNE, however, gains its competitive advantage from two non-traditional sources. First, it identifies host market assets that are not technological innovations, but rather reflect overlooked market niches in host countries that can be exploited worldwide, as in the GE example above. Second, the MNE prospers less from internal development and ownership of unique technologies than from its ability to deliver these technologies efficiently to worldwide customers through its skills at combination, organization, networking, distribution, sales and support. Thus, what had been considered complementary capabilities have become organizational competences that are the keys to profitability for information-age MNEs.

Process Reversals

As we see above, the traditional model of the MNE anticipates that firms will want to use internal, hierarchically controlled, transactions to move their most important rent-yielding assets across geographical borders and organizational boundaries in order to protect their value from opportunistic or incompetent partners. In the

rapidly evolving dynamic industries of the information age, however, asset protection has become less important than innovation and learning (Tallman and Fladmoe-Lindquist 2002). The essence of value is creation, not protection, so that the value of internalization has dropped. Internal processes are good for ensuring that valuable assets are not exposed to potential loss, but restrict innovation and creativity to wholly owned units, company employees and internal procedures. Strong networks of suppliers, partners and subsidiaries are needed to create the new ideas, technologies, products and business models for future success. By providing widely differing ideas for combination into novel approaches to products and markets, open networks using external and hybrid transactions offer multiple opportunities for the creation of value in the new global economy (Kogut and Zander 1992). Mutual needs for innovation, the importance of global logistics and the rapid development and destruction of reputation in the information technology age provide value incentives for cooperation without impeding innovation – access has replaced control in valuing transactions.

In conclusion, the term 'reverse internalization' distinguishes the developing tendency for MNEs to look at foreign locations not as opportunities to internalize markets but as possible sources of competitive advantage. This implies that key assets do not always arise from home market activities, that advanced technologies are not the only sources of global competitive advantage and that internal governance of the transmission, combination and application of resources is no longer the optimum means of control.

See Also

- ▶ Buckley and Casson
- ▶ Capability and Multinational Enterprises (MNEs)
- ▶ Dunning, John H. (1927–2009)
- ▶ Internalization Theory
- ▶ Reverse Knowledge Transfer

References

Birkinshaw, J.M. 1996. How multinational subsidiary mandates are gained and lost. *Journal of International Business Studies* 27: 467–496.

Buckley, P.J., and M. Casson. 1976. *The future of the multinational enterprise*. London: Macmillan.

Dunning, J.H. 1988. The eclectic paradigm of international production: A restatement. *Journal of International Business Studies* 19: 1–32.

Immelt, J., V. Govindarajan, and C. Trimble. 2009. How GE is disrupting itself. *Harvard Business Review* 87: 56–65.

Kogut, B., and U. Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* 3: 383–397.

Tallman, S., and K. Fladmoe-Lindquist. 2002. Internationalization, globalization, and capability-based strategy. *California Management Review* 45: 116–135.

Teece, D.J. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* 15: 285–305.

Reverse Knowledge Transfer

Tina C. Ambos
University of Geneva, School of Economics and Management, Geneva, Switzerland

Abstract

This article describes the meaning, origins and relevance of "reverse knowledge transfers" and gives a brief overview of the recent academic debate in the area of intra-multinational corporation knowledge (information, technology and know-how) transfers.

Definition Reverse knowledge transfers are intra-organizational exchanges of information, technology or know-how from international subsidiaries (host countries) to corporate headquarters (home countries).

Reverse knowledge transfers are intra-organizational exchanges of information, technology or know-how from international subsidiaries (host countries) to corporate headquarters (home countries). The term "reverse" is used to contrast these transfers with the more conventional form of

"forward" transfers from headquarters to subsidiaries, and "lateral" transfers between subsidiaries.

The issue of reverse knowledge transfer has gained significance in international business research since the late 1990s and parallels the recognition that headquarters also act as receivers of knowledge from their internationally dispersed subsidiaries. This debate contributed to shaping the contemporary conceptualizations of the multinational corporation (Hedlund 1986; Bartlett and Ghoshal 1989; Nohria and Ghoshal 1997; Doz et al. 2001). It is recognized that subsidiaries are able to create or source valuable knowledge that may provide benefits for other organizational entities, and headquarters may absorb and/or disseminate this knowledge further. The term was inspired by Lars Hakanson and Robert Nobel's studies on "reverse technology transfers" (2000, 2001) as well as Mo Yamin's (1995) research.

In the international context, the importance of "forward" transfers of knowledge from headquarters to overseas subsidiaries has long been emphasized (Dunning 1958; Vernon 1966), but the acceptance of subsidiaries as valuable sources of knowledge has only become relevant with the subsidiary-focused stream of research (see Paterson and Brock (2002) for a review), which was led by research on the internationalization of innovative activities and other types of influential units in the MNC (cf. Birkinshaw 1998; Frost 1998; Holm and Pedersen 2000). The distinction between the different knowledge sources and the transfer direction of organizational knowledge exchange has been provoked by the debates on the roles organizational units play as entities that are embedded in the multinational organization as well as in their host country (e.g., Gupta and Govindarajan 2000; Andersson and Forsgren 2002; Foss and Pedersen 2002; Mudambi 2002).

Since the 1990s, several studies have investigated the antecedents, amount and success factors of reverse knowledge transfers in different functional settings in the multinational corporation. Central constructs in these studies in the domain of the subsidiary are the host country context (as a determinant for ▶ knowledge sourcing and creation) and subsidiary roles (Buckley et al. 2003; Yang et al. 2008; Criscuolo 2009). In the headquarters' domain the coordination mechanisms (Björkman et al. 2004; Frost and Zhou 2005; Ambos et al. 2006; Asmussen et al. 2013; Rabbiosi 2011), and headquarters' ▶ absorptive capacity have been found to play a crucial role. But relational variables, such as the geographic and cultural distance between headquarters and subsidiaries, have also been subjects of investigation. Reverse knowledge transfers have often been associated with activity in "peripheral" organizational entities (e.g., Ambos and Ambos 2009).

While a number of studies, including those mentioned above, have focused exclusively on the phenomenon of "reverse" knowledge transfer, the recognition of the importance of reverse transfers in general has led to a consciousness that the transfer direction is a key variable in the study of cross-border knowledge flows in the multinational corporation.

See Also

- ▶ Absorptive Capacity
- ▶ Knowledge Sourcing
- ▶ Technology Transfer

References

Ambos, T.C., and B. Ambos. 2009. Location choice, management and performance of international R&D investments in peripheral economies. *International Journal of Technology Management* 49: 24–41.

Ambos, T.C., B. Ambos, and B.B. Schlegelmilch. 2006. Learning from foreign subsidiaries: An empirical investigation of headquarters' benefits from reverse knowledge transfers. *International Business Review* 15: 294–312.

Andersson, A., and M. Forsgren. 2002. The strategic impact of external networks: Subsidiary performance and competence development in the multinational corporation. *Strategic Management Journal* 23: 979–996.

Asmussen, C.G., N.J. Foss, and T. Pedersen. 2013. Knowledge transfer and accommodation effects in multinational corporations: Evidence from European subsidiaries. *Journal of Management* 39: 1397–1429.

Bartlett, C.A., and S. Ghoshal. 1989. *Managing across borders: The transnational solution*. Boston: Harvard Business School Press.

Birkinshaw, J. 1998. Multinational subsidiary evolution: Capability and charter change in foreign-owned

subsidiary companies. *Academy of Management Review* 23: 773–796.
Björkman, I., W. Barner-Rasmussen, and L. Li. 2004. Managing knowledge transfer in MNCs: The impact of headquarters control mechanisms. *Journal of International Business Studies* 35: 443–455.
Buckley, P.J., J. Clegg, and H. Tan. 2003. The art of knowledge transfer: Secondary and reverse transfer in China's telecommunications manufacturing industry. *Management International Review* 43: 67–93.
Criscuolo, P. 2009. Inter-firm reverse technology transfer: The home country effect of R&D. *Industrial and Corporate Change* 18: 869–899.
Doz, Y., J.F.P. Santos, and P.J. Williamson. 2001. *From global to metanational*. Boston: Harvard Business School Press.
Dunning, J.H. 1958. *American investment in British industry.* London: George Allen & Unwin.
Foss, N.J., and T. Pedersen. 2002. Transferring knowledge in MNCs: The role of sources of subsidiary knowledge in organizational context. *Journal of International Management* 8: 49–67.
Frost, T.S. 1998. *The geographic sources of innovation in the multinational enterprise: US subsidiaries and host country spillovers, 1980–1990*. Dissertation, Sloan School of Management, Massachusetts Institute of Technology, Boston.
Frost, T.S., and C. Zhou. 2005. R&D co-practice and "reverse" knowledge integration in multinational firms. *Journal of International Business Studies* 36: 676–687.
Gupta, A.K., and V. Govindarajan. 2000. Knowledge flows within multinational corporations. *Strategic Management Journal* 21: 473–496.
Hakanson, L., and R. Nobel. 2000. Technology characteristics and reverse knowledge transfer. *Management International Review* 40(Special issue): 29–48.
Hakanson, L., and R. Nobel. 2001. Organizational characteristics and reverse knowledge transfer. *Management International Review* 41: 395–420.
Hedlund, G. 1986. The hypermodern MNC: A heterarchy? *Human Resource Management* 25: 9–25.
Holm, U., and T. Pedersen. 2000. *The emergence and impact of MNC centres of excellence: A subsidiary perspective*. London: Macmillan.
Mudambi, R. 2002. Knowledge management within multinational firms. *Journal of International Management* 8: 1–7.
Nohria, N., and S. Ghoshal. 1997. *The differentiated network: Organizing multinational corporations for value creation*. San Francisco: Jossey-Bass.
Paterson, S.L., and D.M. Brock. 2002. The development of subsidiary-management research: Review and theoretical analysis. *International Business Review* 11: 139–163.
Rabbiosi, L. 2011. Subsidiary roles and reverse knowledge transfer: An investigation of the effects of coordination mechanisms. *Journal of International Management* 17: 97–113.
Vernon, R. 1966. The product cycle hypothesis in a new international environment. *Oxford Bulletin of Economics and Statistics* 41: 255–266.
Yamin, M. 1995. Determinants of reverse transfer: The experience of UK multinationals. In *New challenges for European and International Business, proceedings of the 21st annual EIBA conference, Urbino, Italy*, ed. R. Schiattarella. European International Business Associations.
Yang, Q., R. Mudambi, and K.E. Meyer. 2008. Conventional and reverse knowledge flows in multinational corporations. *Journal of Management* 43: 882–902.

Ricardian Rents

Frederic Sautet
George Mason University, Mercatus Center, Arlington, VA, USA

Definition Originally associated with land, a Ricardian rent is the result of the possession of a natural or man-made idiosyncratic, scarce factor. Like profit, a Ricardian rent is a surplus earning above the costs necessary to deploy and use a resource. Unlike profit, however, it would continue exist in a hypothetical state of equilibrium as long the resource remained scarce.

In the eighteenth century, the Physiocrats gave land a special status in the economy. It was, according to François Quesnay, the only productive input. All wealth came from the land. Adam Smith and A. R. Turgot did not entirely side with the Physiocrats on that point and saw industry (e.g., Smith's pin factory) also as a source of wealth. But land kept a special status in the writings of many classical economists, including David Ricardo. He was concerned, among other things, with explaining the earnings that accrue to different groups in society and understanding the impact of land appropriation on commodity prices, independently of the influence of labour and capital.

To that effect, Ricardo developed and presented the law of rent in his famous treatise *On the Principles of Political Economy and Taxation* (1821). Therein he stated: 'rent is that portion of the produce of the earth which is paid to the

landlord for the use of the original and indestructible powers of the soil' (Ricardo 1821: 33). In this approach, it is the scarcity of fertile land that determines the amount of rent paid to the landlord. The more fertile the land, the more rent it will earn relative to the marginal land, that is, rent-free land or land that is not fertile enough to generate rent, given the same inputs of labour and capital. Rent is thus a surplus above the costs necessary to till a scarce and fertile land.

In Ricardo's view, rent is not part of the cost of production because land is a different category of input. Land is different not so much because it is the only source of wealth, as the Physiocrats saw it, but because it is in fixed (or quasi-fixed) supply. Rents are earned because land cannot be expanded. Since the most fertile land is always cultivated first, it will earn rent above the marginal rent-free land.

In the neoclassical model of perfect competition, factors are paid their marginal discounted value product, which leaves no surplus of any kind. In this standard, there are no rents. Rents indicate inefficiencies. These can be, for instance, the result of heterogeneous resources, as is the case with pieces of land of different fertility. Ricardian rents are an abnormality in the neoclassical model, as they should always be competed away but are not, because of the particularities of some factors. In this sense, the notion of Ricardian rent has come to refer, in modern economics, to the surplus earning of land, or of any another resource, natural or man-made, which is generally in fixed supply (e.g., a unique human capital, some particular knowledge embedded in an extraordinary team). Contrary to entrepreneurial ▶ profit (and ▶ schumpeterian rents), which are transient, Ricardian rents would continue to exist in a hypothetical state of equilibrium as long as the resource remained scarce.

In Ricardo's model, when land of inferior fertility is abundant, rent is low or absent. A rent is only a payment for fertility of a scarce land; it is a scarce resource's surplus earning. It can thus be seen as some sort of 'super normal profit' or 'above normal earning'. This has led to the idea that rents exist because of a differential in earnings with a 'normal' situation. But rents are not due to a differential. If there were only two pieces of land of homogeneous fertility, rents could still exist (on both pieces of land) as long as land as a whole were scarce relative to its demand (and given the costs of labour and capital). In the context of industrial organization theory and management, Ricardian rents may reflect the difficulty for competitors to expand competencies to match an idiosyncratic resource (Teece and Coleman 1998: 820).

Ricardo's law of rent is an important insight into the determination of land (and other factors) prices. But instead of taking a pure Ricardian approach, it may be preferable to view rent as the unit price of the services of any good (not just land). In such a case, the value of any good is a function of its expected future contributions in the productive process (i.e., the present value of its rents). All resources earn rents, and all rents reflect the value of the resource to the production process. This way of looking at the notion of rent avoids the difficulty of having to define 'normal' returns, and simply rests on the idea that since resources are all heterogeneous they do not equally contribute to production, and thus do not command the same prices. In equilibrium, every resource will earn a rent according to its contribution to the production process. In the disequilibrium context, resources may be mispriced and ▶ entrepreneurial rents (i.e., pure profit) may emerge as the *ex ante–ex post* differences in the earnings of resources.

See Also

- ▶ Economic Rent
- ▶ Entrepreneurial Rents
- ▶ Managerial Rents
- ▶ Profit
- ▶ Quasi-Rent
- ▶ Schumpeterian Rents

References

Ricardo, D. 1821. *On the principles of political economy and taxation*, 3rd ed. London: John Murray.

Teece, D.J., and M. Coleman. 1998. The meaning of monopoly: Antitrust analysis in high-technology industries. *The Antitrust Bulletin* 43: 801–857.

Risk and Uncertainty

K. Francis Park[1] and Zur Shapira[2]
[1]New York University, Stern School of Business, New York, NY, USA
[2]New York University, Leonard N Stern School of Business, New York, NY, USA

Abstract

Risk is the situation under which the decision outcomes and their probabilities of occurrences are known to the decision-maker, and uncertainty is the situation under which such information is not available to the decision-maker. Research on decision-making under risk and uncertainty has two broad streams: normative and descriptive. Normative research models how decision *should* be made under risk and uncertainty, whereas descriptive research studies how decisions under risk and uncertainty are *actually* made. Descriptive studies have exposed weaknesses of some normative models in describing people's judgment and decision-making and have compelled the creation of more intricate models that better reflect people's decision under risk and uncertainty.

Definition

Risk refers to decision-making situations under which all potential outcomes and their likelihood of occurrences are known to the decision-maker, and uncertainty refers to situations under which either the outcomes and/or their probabilities of occurrences are unknown to the decision-maker. How decision-makers perceive risk and uncertainty depends on the context of the decision and the characteristic of the decision-maker.

Body of Text

Risk and uncertainty are commonly used terms both in decision-making research and in life. The lay usages of the words, however, significantly diverge from the researcher's definitions. Oxford English Dictionary defines risk as "a situation involving exposure to danger," and it defines uncertainty as "the state of being not able to be relied on, not known or definite." The definitions that decision-making researchers assign to these terms are very different. Conceptually, risk defined as variance (or standard deviation) does not only involve exposure to danger, but it can also include potential for gain. Researcher's definition of uncertainty is not too different from that of Oxford English Dictionary, but there are different types of uncertainty that require more detailed descriptions. In this entry, we first discuss the definitions of risk and uncertainty and then outline few broad streams of research on this topic. Generally speaking, two different research perspectives exist in this domain: normative and descriptive. The two perspectives study different aspects of decision-making under risk and uncertainty, and they complement each other.

Early Conception of Risk and Uncertainty

Knight (1921), in defining risk and uncertainty, made an explicit distinction between the two terms. Knight defined risk as decision-making situations in which all potential outcomes are identified and their respective probabilities of occurrences are known. He characterized uncertainty as situations in which either the possible decision outcomes or the specific probabilities associated with such outcomes are unknown to the decision-maker. Luce and Raiffa (1957, p. 13) echoed Knight's distinction between risk and uncertainty. They made distinctions among certainty, risk, and uncertainty. Whereby decision-makers are in the realm of certainty, if they know the outcome that their decision will lead to. Decision-makers are in the domain of risk, if their decision is known to lead to a certain outcome with certain probability, and they are in the realm of uncertainty, if the probabilities associated with the outcomes are either unknown or meaningless.

The key distinction between risk and uncertainty lies in the quantifiability of the number of possible decision outcomes and the probabilities of their occurrences. Uncertainty often refers to aspects of decision-making which are not easily quantified. As Knight argued, there can be uncertainty regarding both the decision outcomes and the probabilities of the outcome occurrences. There can also be uncertainty about decision-maker's preferences (March 1978). Quantifiability is the key characteristic of risk, but the definition of risk and the measurement of risk are not straightforward. In fact, the process of defining risk in real life often reflects a political process that depends on the decision-maker, the technologies being considered, and the characteristics of decision problem (Fischhoff et al. 1984). These factors affect the two-step process of defining risk; the first step is to determine which consequences or outcome dimensions to include and the next step is to construct risk indices based on the consequences selected in the previous step.

The St. Petersburg Paradox: Expected Value and Expected Utility

Once the definition and the measurement of risk are available, we then need a rule for making decision under risk. According to statistical decision theory, the expected value rule is considered as the best rule to follow when making risky decisions (Raiffa 1968). The expected value rule entails multiplying the value of the outcomes by their respective probabilities of occurrences and then summing them up. Decisions that have higher values after this calculation are considered better than decisions with lower values. While simplistic, this rule fails to accurately depict people's actual decision-making. Consider this coin-flipping game. In this game, a player is asked to guess the face of a coin toss. If she guesses wrong on the first $(n - 1)$ times but guesses correctly on the nth trial, she wins $\$2^n$. Thus, the expected value of this game is the summation of \$1 infinite number of times or simply infinity. In other words, under the expected value rule, people should be willing to pay an infinite amount of money to play this game. In reality, however, people are willing to pay only about \$2.50 to play this game. This finding is known as the St. Petersburg Paradox, and Bernoulli (1738) offered a different decision rule that can account for this paradox: the *expected utility* rule. The term utility refers to the psychological value that people assign to certain decision outcomes (Savage 1954; von Neumann and Morgenstern 1944). Utility differs from value in that utility may differ from one context to another for the same decision-maker and from one decision-maker to another, that is, the utility that people assign to certain outcomes are subjective and context dependent. Under the expected utility rule, people's utility from playing the aforementioned coin-flipping game may differ from the game's expected value. People's marginal utility decreases as outcome value increases, and people's expected utility function is concave. As a consequence, players may be unwilling to pay an infinite amount of money to play the game. People, however, not only assign subjective utility to outcome values, but they also assign subjective probabilities to achieving those outcomes. The theory of subjective expected utility accounts for such tendency as well (Savage 1954).

Two Additional Paradoxes and Different Theories

The Allais Paradox

The theory of expected utility, however, also falls short in accurately depicting people's decision-making in many cases. The Allais paradox is a classic example that illustrates the violation of expected utility rule in people's actual decision-making (Allais 1953). Allais presented two different sets of gambles to people and found that people's shift in preference between the two presentations of gambles violated the expected utility rule's notion of consistency. The Allais paradox reflects the expected utility rule's failure to account for people's preference for certainty, and alternative theories of risk have been offered to better capture such tendency.

Rank-Dependent Utility Theory and Cumulative Prospect Theory

In reviewing a large number of studies, Weber (1994) pointed out that the assumption of independence between the probability and utility of an outcome, which is essential to expected utility, is often violated in practice. She discussed a class of non-expected utility models for choice under risk and uncertainty, called rank-dependent utility (RDU) models. These models originally proposed by Quiggin (1982) and Yaari (1987) hold that people evaluate the probabilities of outcome by ranking those outcomes initially and then looking at the cumulative probabilities of obtaining these outcomes. Quiggin's (1982) theory of anticipated utility theory extended the expected utility theory partly by using a weaker form of the independence of irrelevant alternatives axiom. Quiggin's theory is essentially a generalization of the expected utility theory set forth by von Neumann and Morgenstern (1944). Quiggin's theory overweighs unlikely extreme outcomes, which is analogous to people's tendency to overweight low probability events such as winning a lottery on the one hand and getting in a plane accident on the other. Schmeidler's (1989) extension of the expected utility model covers situations involving uncertainty to account for the Ellsberg paradox (that will be discussed below).

As Weber (1994) noted, RDU models hold that decision weights can be nonlinear and as such they can describe phenomena that were characterized earlier as optimistic and pessimistic behavior. Indeed, Kahneman and Tversky revised their notion of nonlinear decision weights in prospect theory (1979) into cumulative probabilities (Tversky and Kahneman 1992) in a way that appear to avoid the problem of violation of stochastic dominance.

Extreme Events

People's tendency to overweight unlikely extreme events has many implications. For instance, in finance, a premium for risk (i.e., fat tail) exists due to investors' fear of disasters (Bollerslev and Todorov 2011a, b). Approximately 5% of equity premium is thought to be attributable to the compensation for rare disaster events. People, however, only overweight the probability of rare events when they are provided with summary description of possible outcomes; when people make decisions based on their experience, they instead underweight the probability of rare events (Hertwig et al. 2004; Ungemach et al. 2009).

Lampel et al. (2009) argued that rare events have two elements, a probability estimate on the one hand and a non-probability element that is based on the enacted salience of such events. The latter implies that in addition to the probability component of rare events, people enact such event by focusing on the unique and unusual features of such events in a constructionist manner that differs from the scientific way of assigning probabilities to such events.

Recent research by Ülkumen et al. (2016) suggests that in conceiving of uncertain events, people distinguished between events that are knowable (like the length of the river Nile, which the decision-maker does not to know) and those that are completely random (a toss of a fair coin) and that natural language provides us with cues about the way they conceive of these two elements. Extreme events pose a big problem for calculated probabilities of such events (i.e., fat tails), and focusing on the different elements in people's conceptualization of such events may lead to a better understanding of the creation of subjective estimates of probabilistic events.

The Ellsberg Paradox

So far we have addressed people's decision-making under risk and have not considered their decision-making under uncertainty. Savage (1954) argued that all uncertainties can be reduced to risk. More specifically, under Savage's theory of subjective expected utility, decision-makers are thought to behave similarly under uncertainty as they do under risk, if their *subjective* assessment of the probabilities of the outcomes is the same in both cases. The Ellsberg (1961) paradox, however, illustrates that this may not be the case. Ellsberg showed that people's decision-making under risk and uncertainty violates the assumptions of the expected utility theory, and he asserted that such violation occurs because people are averse to ambiguity, which refers to a condition

between complete ignorance and risk. Such aversion to ambiguity has been shown empirically as well in the context of decisions of insurance companies (Kunreuther et al. 1993) and has been shown to be moderated by decision-maker's confidence in his or her judgment and skill (Heath and Tversky 1991). Ambiguity aversion, however, is only present when a person is presented with both clear and vague prospects, and it disappears in a noncomparative context in which comparison between prospects is absent (Fox and Tversky 1995).

Descriptive Approaches

It is difficult to model people's decision-making under risk and uncertainty. Another stream of research, instead, *describes* how people make decisions. These researchers focus on how people perceive and interpret risk and uncertainty and how dispositional characteristics and contextual factors influence their choices under risk and uncertainty. Some researchers attempt to establish attitude toward risk as a stable individual trait that is linked to personality or culture (Douglas and Wildavsky 1982), but there is lack of consensus and consistency among studies that address the potential link between risk taking and dispositional characteristics (Slovic 1964). Factors other than dispositional characteristics have also been studied for their effect on risk taking. People's risk preference has been found to hinge on the framing of the decision problem (Tversky and Kahneman 1981) and on people's mood and feelings at the time of decision-making (Hastorf and Isen 1982; Loewenstein et al. 2001; Rottenstreich and Hsee 2001). Thaler and Johnson (1990) illustrated that prior outcomes affect people's risk preference. More specifically, people become relatively more risk seeking after a prior gain, which they term as the house-money effect. The reason is that immediately after a gain, people do not assimilate the gain into their own assets and treat it as the "house money" that they can be more risk seeking with. People become more risk seeking, if they get a chance to break even, like gamblers on horse races at the end of the day. They term this as the break-even effect. Those effects can be explained by shifts of the reference point or by introducing another reference point such as survival (March and Shapira 1992). Moreover, learning from experience is thought to make people more risk averse over time (March 1996).

Risk Measurement and Absolute Risk

To better assess people's risk preferences, we need a measurement of risk. Descriptive research on risk measurement has compared and contrasted verbal and numerical measures of risk. Erev and Cohen (1990) found that people are more comfortable expressing risk verbally, but they prefer to receive numerical information about risk. There is consistency in the interpretation of verbally represented outcome probabilities within a decision-maker but not between different decision-makers (Budescu and Wallsten 1985). Furthermore, when provided with numerical estimates of probabilities, people tend to invoke rule-based decision-making process, whereas when given verbal measures of risk, people become more associative, and their decisions become more intuitive (Windschitl and Wells 1996). More closely related to the issue of risk measurement, Zimmer (1983) found that people use about five to six expressions to describe an entire range of probabilities. Beyth-Marom (1982) found that seven category scales to be sufficient to capture risk. In sum, people utilize a small number of expressions to conceptualize risk; thus, risk measurement may require a small number of category scales to capture people's risk preferences.

Normative perspective on decision-making under risk and uncertainty attempts to avoid all the complications that arise from individual characteristics of decision-makers. These researchers have created a riskiness measure that is independent of the decision-maker. Aumann and Serrano (2008) defined the riskiness of a gamble based on an individual who is indifferent between taking and not taking the gamble in question. More specifically, they measured risk by taking the reciprocal of the absolute risk aversion of that individual. Foster and Hart (2009), on the other hand, defined risk by identifying the critical

wealth level of decision-maker. When the decision-maker's wealth falls below that critical level, a gamble is defined to be risky.

Neural Correlates

Some researchers have taken the descriptive approach to studying people's decision-making under risk and uncertainty deeper into people's brains. Because it is our brain where all the decision-making process takes place, these researchers argue that it is important to study brain activities to understand decision-making (Camerer et al. 2005). These researchers investigate where risk attitude is located in our brains and how certain brain activity is related to people's decision-making under risk and uncertainty. Risk-seeking deviations from optimal financial decisions activate a certain region in the brain that differs from brain regions that get activated when risk-averse deviations from optimal choices are made (Kuhnen and Knutson 2005). Hsu et al. (2005) have also shown that high level of ambiguity is associated with higher activation of amygdala and orbitofrontal cortex of our brains. Furthermore, the amygdala is also responsible for people's general tendency to be averse to losses (De Martino et al. 2010).

Risk and Uncertainty in Strategy Research

Strategy research often borrows measurement of risk from other fields such as finance and decision theory. Studies that borrow risk measurement from finance literature use β from the capital asset pricing model, while studies that use the measurements of decision theory literature create certain variance measurements of accounting returns such as ROA (Reuer and Leiblein 2000; Ruefli et al. 1999). Such measurements of risk are used to study the relationship between risk and firm strategy or firm performance. On the other hand, uncertainty in strategy literature has its distinct definition. Strategy literature defines a firm's uncertainty to be arising from conflict and interdependence with its environmental elements, and strategy research studies how firms reduce this type of uncertainty by managing their environmental conflict and interdependence (e.g., Hoffmann 2007; Martin et al. 2015).

Conclusion

There are many different approaches to studying people's decision-making under risk and uncertainty. Normative and descriptive approaches are the two main perspectives, which complement each other. Modeling people's choices under risk and uncertainty is becoming more intricate as we know more about how people actually make decisions under risk and uncertainty. Simply put, risk can be considered as a subset of uncertainty that is quantifiable or measurable, whereas uncertainty refers to ignorance about potential outcomes or their respective likelihood of occurrences.

References

Allais, M. 1953. Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine. *Econometrica* 21 (4): 503–546.

Aumann, R.J., and R. Serrano. 2008. An economic index of riskiness. *Journal of Political Economy* 116 (5): 810–836.

Bernoulli, D. 1738. Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperiales Petropolitanae* 5: 175–192.

Beyth-Marom, R. 1982. How probable is probable? A numerical translation of verbal probability expressions. *Journal of Forecasting* 1 (3): 257–269. https://doi.org/10.1002/for.3980010305.

Bollerslev, T., and V. Todorov. 2011a. Tails, fears, and risk premia. *Journal of Finance* 66 (6): 2165–2211. https://doi.org/10.1111/j.1540-6261.2011.01695.x.

Bollerslev, T., and V. Todorov. 2011b. Estimation of jump tails. *Econometrica* 79 (6): 1727–1783.

Budescu, D.V., and T.S. Wallsten. 1985. Consistency in interpretation of probabilistic phrases. *Organizational Behavior and Human Decision Processes* 36 (3): 391–405. https://doi.org/10.1016/0749-5978(85)90007-X.

Camerer, C.F., G.F. Loewenstein, and D. Prelec. 2005. Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature* 43 (1): 9–64.

De Martino, B., C.F. Camerer, and R. Adolphs. 2010. Amygdala damage eliminates monetary loss aversion. *Proceedings of the National Academy of Sciences of the United States of America* 107 (8): 3788–3792. https://doi.org/10.1073/pnas.0910230107.

Douglas, M., and A. Wildavsky. 1982. How can we know the risks we face? Why risk selection is a social process.

Risk Analysis 2 (2): 49–58. https://doi.org/10.1111/j.1539-6924.1982.tb01365.x.

Ellsberg, D. 1961. Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics* 75 (4): 643–669.

Erev, I., and B.L. Cohen. 1990. Verbal versus numerical probabilities: Efficiency, biases, and the preference paradox. *Organizational Behavior and Human Decision Processes* 45 (1): 1–18. https://doi.org/10.1016/0749-5978(90)90002-Q.

Fischhoff, B., S.R. Watson, and C. Hope. 1984. Defining risk. *Policy Sciences* 17 (2): 123–139. https://doi.org/10.1007/BF00146924.

Foster, D.P., and S. Hart. 2009. An operational measure of riskiness. *Journal of Political Economy* 117 (5): 785–814.

Fox, C.R., and A. Tversky. 1995. Ambiguity aversion and comparative ignorance. *Quarterly Journal of Economics* 110 (3): 585–603.

Hastorf, A., and A.M. Isen, eds. 1982. *Cognitive social psychology*. New York: Elsevier.

Heath, C., and A. Tversky. 1991. Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty* 4 (1): 5–28. https://doi.org/10.1007/BF00057884.

Hertwig, R., G. Barron, E.U. Weber, and I. Erev. 2004. Decisions from experience and the effect of rare events. *Psychological Science* 15 (8): 534–539.

Hoffmann, W.H. 2007. Strategies for managing a portfolio of alliances. *Strategic Management Journal* 28 (8): 827–856.

Hsu, M., M. Bhatt, R. Adolphs, D. Tranel, and C.F. Camerer. 2005. Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310 (5754): 1680–1683. https://doi.org/10.1126/science.1115327.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (2): 263–292.

Knight, F.H. 1921. *Risk, uncertainty, and profit*. New York: Hart, Schaffer & Marx.

Kuhnen, C.M., and B. Knutson. 2005. The neural basis of financial risk taking. *Neuron* 47 (5): 763–770. https://doi.org/10.1016/j.neuron.2005.08.008.

Kunreuther, H., R.M. Hogarth, and J. Meszaros. 1993. Insurer ambiguity and market failure. *Journal of Risk and Uncertainty* 7 (1): 71–87. https://doi.org/10.1007/BF01065315.

Lampel, J., J. Shamsie, and Z. Shapira. 2009. Experiencing the improbable: Rare events and organizational learning. *Organization Science* 20 (5): 835–845. https://doi.org/10.1287/orsc.1090.0479.

Loewenstein, G.F., C.K. Hsee, E.U. Weber, and N. Welch. 2001. Risk as feelings. *Psychological Bulletin* 127 (2): 267–286.

Luce, R.D., and H. Raiffa. 1957. *Games and decisions: Introduction and critical survey*. New York: Wiley.

March, J.G. 1978. Bounded rationality, ambiguity, and the engineering of choice. *Bell Journal of Economics* 9 (2): 587–608.

March, J.G. 1996. Learning to be risk averse. *Psychological Review* 103 (2): 309–319. https://doi.org/10.1037/0033-295X.103.2.309.

March, J.G., and Z. Shapira. 1992. Variable risk preferences and the focus of attention. *Psychological Review* 99 (1): 172–183.

Martin, G., R. Gözübüyük, and M. Becerra. 2015. Interlocks and firm performance: The role of uncertainty in the directorate interlock-performance relationship. *Strategic Management Journal* 36 (2): 235–253. https://doi.org/10.1002/smj.2216.

Quiggin, J. 1982. A theory of anticipated utility. *Journal of Economic Behavior and Organization* 3 (4): 323–343. https://doi.org/10.1016/0167-2681(82)90008-7.

Raiffa, H. 1968. *Decision analysis: Introductory lectures on choices under uncertainty*. Oxford: Addison-Wesley.

Reuer, J.J., and M.J. Leiblein. 2000. Downside risk implications of multinationality and international joint ventures. *Academy of Management Journal* 43 (2): 203–214. https://doi.org/10.2307/1556377.

Rottenstreich, Y., and C.K. Hsee. 2001. Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science* 12 (3): 185–190. https://doi.org/10.1111/1467-9280.00334.

Ruefli, T.W., J.M. Collins, and J.R. Lacugna. 1999. Risk measures in strategic management research: Auld lang syne? *Strategic Management Journal* 20 (2): 167–194.

Savage, L.J. 1954. *The foundations of statistics*. New York: Wiley.

Schmeidler, D. 1989. Subjective probability and expected utility without additivity. *Econometrica* 57 (3): 571–587.

Slovic, P. 1964. Assessment of risk taking behavior. *Psychological Bulletin* 61 (3): 220–233.

Thaler, R.H., and E.J. Johnson. 1990. Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science* 36 (6): 643–660. https://doi.org/10.1287/mnsc.36.6.643.

Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211 (4481): 453–458. https://doi.org/10.1126/science.7455683.

Tversky, A., and D. Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5 (4): 297–323. https://doi.org/10.1007/BF00122574.

Ülkümen, G., C.R. Fox, and B.F. Malle. 2016. Two dimensions of subjective uncertainty: Clues from natural language. *Journal of Experimental Psychology: General* 145 (10): 1280–1297. https://doi.org/10.1037/xge0000202.

Ungemach, C., N. Chater, and N. Stewart. 2009. Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychological Science* 20 (4): 473–479. https://doi.org/10.1111/j.1467-9280.2009.02319.x.

von Neumann, J., and O. Morgenstern. 1944. *Theory of games and economic behavior*. Princeton: Princeton University Press.

Weber, E.U. 1994. From subjective probabilities to decision weights: The effect of asymmetric loss functions on the evaluation of uncertain outcomes and events. *Psychological Bulletin* 115 (2): 228–242. https://doi.org/10.1037/0033-2909.115.2.228.

Windschitl, P.D., and G.L. Wells. 1996. Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology: Applied* 2 (4): 343–364. https://doi.org/10.1037/1076-898X.2.4.343.

Yaari, M.E. 1987. The dual theory of choice under risk. *Econometrica* 55 (1): 95–115. https://doi.org/10.2307/1911158.

Zimmer, A.C. 1983. Verbal versus numerical processing of subjective probabilities. In *Decision making under uncertainty*, ed. R.W. Scholz, 159–182. Amsterdam: North-Holland.

Risk Aversion

Dan Lovallo and Elif Ketencioglu
The University of Sydney, Faculty of Economics and Business, Sydney, NSW, Australia

Abstract

The traditional view of risk aversion is based on expected utility theory, which states that decision makers maximize the expected utility of outcomes. In some studies, expected utility theory has been argued to be an inadequate representation of ▶ decision-making under ▶ risk and uncertainty, and its ability to explain risk attitude has been challenged. The most valuable contribution to understanding risk aversion has been prospect theory, where gains and losses are perceived differently, and the overall utility is defined in terms of changes in values relative to a reference point.

Definition Risk aversion is a preference for certainty over uncertainty. Based on expected values, a risk averse person may prefer a certain outcome with a lower pay-off over an uncertain outcome with a higher pay-off.

A person is said to be risk averse if s/he accepts a sure gain over a gamble; risk neutral if s/he is indifferent between the sure gain and the gamble; and risk seeking if the value of the sure gain is higher than the gamble of choosing the guaranteed option. Earlier attempts to understand ▶ decision-making under uncertainty explain risk-taking behaviour by expected values where the monetary outcomes are weighed by their probabilities.

$E(x_1, p_1; \ldots x_n, p_n) = p_1x_1 + \ldots + p_nx_n$; where x_i is the monetary value (price) and p_i is the expected probability $(p_1 + \ldots + p_n = 1)$.

In this context, given a choice between a gamble and a sure payment, the decision maker may choose the gamble due to the higher expected value. The St Petersburg paradox by Bernoulli (1954) highlights the major inadequacy of expected values in evaluating risk by showing that identical risks may have different values for individuals and that a small amount of money could have infinite expected value, yielding irrational decisions. The paradox asks the fair price of entry to a game for a single player where a coin is tossed at each stage. The game starts at one dollar and the amount is doubled for each 'head', and the game ends at the first 'tail'. Bernoulli proposed using utilities rather than monetary values to evaluate risk, which initiated the emergence of expected utility theory and provided a solution for the paradox.

Expected Utility Theory

Expected utility theory is based on Von Neumann and Morgenstern's (1944) utility hypothesis, in which rational individuals maximize their expected utility function. Utilities of outcomes are weighted by their probabilities in the expected utility function. In its simplest form, the overall utility function is as follows: $U(x_1, p_1; \ldots x_n, p_n) = p_1u(x_1) + \ldots + p_nu(x_n)$; where p_i is the probability of outcome x_i and $u(x_i)$ is regarded as a function of total assets $(p_1 + \ldots + p_n = 1)$.

This theory assumes that risk-averse individuals have a concave utility function for wealth, meaning that the marginal utility of wealth diminishes as wealth increases. Thus, the degree of risk aversion is observed in the concavity of the utility function, where risk-neutral individuals have

linear utility functions and risk-seeking individuals have convex utility functions.

The focus of expected utility theory has been criticized for failing to accurately explain risk-taking behaviour. In an early challenge, Allais (1953) showed that risk preferences may change among individuals and that the psychology of risk could result in subjective probabilities. This implies that probability weights will not be linear, as stated in expected utility theory; thus, the difference between a change of probabilities from 0.99 to 1.00 has more impact on preferences than a change from 0.10 to 0.11. Savage (1954) further argued that not only the probabilities but also the utility functions will differ across individuals, resulting in subjective expected utility functions. This notion is supported by Kahneman and Tversky's (1972) writing on subjective probabilities, which argues that people will not always follow the principles of probability theory in their risk attitudes. Tversky and Kahneman further showed that the decision weight of a probability of 0.95 could be in fact 0.80 (1972). Another example that violated the consistent risk attitude assumption is given by Samuelson's paradox (Samuelson 1963), where an individual declines a bet in which he could gain $200 or lose $100 with 0.5 probabilities for each, but is willing to accept a string of 100 such bets. This paradox clearly shows how misleading expected utility theory can be in interpreting risk preferences.

Other arguments show that when the stakes are small enough, decision makers will tend to be risk neutral (Arrow 1971). A further extension of this argument is given by Rabin and Thaler (2001), who argue that, independent of the context, if a utility maximizer always turns down moderate stakes, he will turn down large stakes as well. In the context of insurance, expected utility theory predicts that people will buy insurance with high premiums and coverage. Yet most insurance policies have low premiums and limited coverage (Rabin and Thaler 2001). Further violations are observed in the context of the framing of options in terms of gains or losses (Tversky and Kahneman 1986). Furthermore, the source of uncertainty as a known or unknown proportion of risk (Ellsberg 1961) are shown to yield systematically different preferences. March has argued that, although ambiguity changes the expected outcome of normative behaviour in choice and preferences, it is not necessarily a fault to be corrected. Instead, he argues that inconsistent and ambiguous preferences promote a contextual change in risk taking that adjusts for future consequences.

Prospect Theory

Kahneman and Tversky (1979) demonstrated several problems that violate the axioms of expected utility theory and proposed a descriptive model of risk taking, namely prospect theory. In this model, utility is defined over gains and losses relative to a reference point, contrary to attributed utilities in expected utility theory. The theory also distinguishes the decision weights and probabilities of an outcome. Decision weights are attributed probabilities of outcomes by the decision maker for a set of prospects. Risk in prospect theory is captured by changes in value relative to the reference, and the value function is given as follows: $V(x, p; y, q) = \pi(p)v(x) + \pi(q)v(y)$; where $(x, p; y, q)$ is a prospect with outcomes x and y, and probabilities of p and q respectively $(p + q = 1)$. π reflects the decision weight of probabilities and the impact of probability on the overall value of prospect, and v reflects the subjective value of an outcome relative to a reference point; hence, it measures the changes of value from the reference point.

The three important aspects of prospect theory are certainty, reflection and isolation effects (Kahneman and Tversky 1979). The certainty effect is defined as overweighting outcomes that are certain relative to those that are merely probable. This results in a value function that is concave for gains, convex for losses and steeper for losses than for gains. The reflection effect shows that people will display risk-averse behaviour for gains and risk-seeking behaviour for losses. For instance, people would be willing to enter a gamble with a small probability of a large prize or pay for insurance against a small probability of large loss. It further implies that certainty will increase risk-averse behaviour for losses and desirability of gains. The isolation effect refers to the discarding

of common components of a prospect, which results in inconsistent preferences when the same choice is presented in different forms. Each of these clearly shows that the axioms of expected utility theory are violated.

Loss Aversion

Rabin (2000) argues that expected utility theory will incorrectly predict the relationship between risk aversion in modest and large stakes, and further states that the risk-averse behaviour in modest-to-small-stake risk choices in experiments cannot be explained by the utility-maximizing behaviour. Hence, loss aversion is provided as an adequate explanation for modest scale risk aversion. Loss aversion, also a part of prospect theory, states that people are more willing to avoid losses than to acquire gains (Kahneman et al. 1991). This explains why people may reject a 50-50 gamble with equal stakes when they view it as a loss rather than a gain.

Loss aversion is argued to cause the status quo bias, which biases individuals to stick with the status quo due to the perceived higher disadvantages of making a change (Samuelson and Zeckhauser 1988). When participants were faced with a choice between making risky investments, staying neutral or retaining the status quo, most chose the status quo. The results suggest that the bias increases with the number of alternatives, meaning that the risk-averse behaviour will increase when the decision maker is faced with a large set of outcomes. In another study, Kahneman and Lovallo (1993) found the same results for risk-seeking behaviour for single, low-probability gambles with risk aversion for repeated gambles. Moreover, loss aversion is argued to be reinforced by the status quo bias in the managerial context, thus creating a more risk-averse environment for managerial decision-making.

A general conclusion for risk preferences has been that individual choices are best explained by changes in utility relative to a reference point rather than a fixed amount of utility attained for a certain outcome (Kahneman et al. 1991).

See Also

- ▶ Decision-Making
- ▶ Prospect Theory and Strategic Decision-Making
- ▶ Risk and Uncertainty
- ▶ Risk-Taking

References

Allais, P.M. 1953. Le comportement de l'homme rationnel devant le risque, critique des postulats et axiomes de l' Ecole Americaine. *Econometrica* 21: 503–546.

Arrow, K. 1971. *Essays in the theory of risk-bearing*. Chicago: Markham.

Bernoulli, D. [1738] 1954. Exposition of a new theory on the measurement of risk, trans. L. Sommer. *Econometrica* 22: 22–36.

Ellsberg, D. 1961. Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics* 75: 643–669.

Kahneman, D., and D. Lovallo. 1993. Timid choices and bold forecasts. *Management Science* 39: 17–31.

Kahneman, D., and A. Tversky. 1972. Subjective probability: A judgment of representativeness. In *Judgment under uncertainty: Heuristics and biases*, ed. D. Kahneman, P. Slovic, and A. Tversky. New York: Cambridge University Press.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47: 236–291.

Kahneman, D., and A. Tversky. 1984. Choices, values and frames. *American Psychologist* 39: 341–350.

Kahneman, D., J.L. Knetsch, and R.H. Thanler. 1991. Anomalies: The endowment effect, loss aversion and status quo bias. *Journal of Economic Perspectives* 5: 193–206.

Rabin, M. 2000. Risk aversion and expected utility theory: A calibration theorem. *Econometrica* 68: 1281–1292.

Rabin, M., and R.H. Thaler. 2001. Anomalies: Risk aversion. *Journal of Economic Perspectives* 15: 219–232.

Samuelson, P. 1963. Risk and uncertainty: A fallacy of large numbers. *Scientia* 98: 10–13.

Samuelson, P., and W. Zeckhauser. 1988. Status quo bias in decision making. *Journal of Risk and Uncertainty* 1: 7–59.

Savage, L.J. 1954. *The foundations of statistics*. New York: Wiley.

Tversky, A., and D. Kahneman. 1986. Rational choice and the framing of decisions. *Journal of Business* 59: S251–S278.

Tversky, A., and D. Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5: 297–323.

von Neumann, J., and O. Morgenstern. 1944. *Theory of games and economic behavior*. Princeton: Princeton University Press.

Risk-Taking

K. Francis Park[1] and Zur Shapira[2]
[1]New York University, Stern School of Business, New York, NY, USA
[2]New York University, Leonard N Stern School of Business, New York, NY, USA

Abstract

Risk taking is the willingness to accept the level of risk associated with a certain decision. In other words, it refers to making decision that entails risk. In this article, we focus on managerial risk-taking. We discuss the variable risk preference model in depth, summarize the effect of performance feedback on risk-taking, and highlight some of the strategy research on risk-taking that takes the upper echelon perspective of organizations. The variable risk preference model assumes that managers have two reference points in making risky decisions – aspiration point and survival point – and which reference point they pay attention to affects their risk-taking behavior. Performance feedback has an influence on risk-taking that often makes decision-makers take a more longitudinal perspective and change their risk attitudes. Some strategy research on managerial risk-taking that adopts the upper echelon perspective examines the linkages between trait and risk-taking behavior and between incentive structure and risk-taking.

Definition Risk-taking refers to decision-making under risk or the willingness to accept the risk involved with a decision. Managerial risk-taking focuses on how managers in organizations perceive risk and make risky decisions. Manager's willingness to take risk depends on her firm's wealth position relative to certain reference points – aspiration point, success point, and survival points – as well as her characteristic as an individual decision-maker and the incentive structure in which she is in.

Body of Text

Risk-taking refers to making decision under risk. In this article, we discuss how strategic decisions under risk are made. We will specifically address how managers make decisions involving risk and how they perceive and interpret risk. We will then review some of strategy research on risk-taking, mainly focusing on how CEOs and top management teams make risk-taking decisions. Strategy research employs multiple methodologies, which range from interviews to formal modeling, in addressing such research question.

Managerial Risk-Taking

Manager's perception or definition of risk is different from the assumptions made in normative models of risk-taking. Managers are familiar with the expected utility rule, but they do not use it (MacCrimmon and Wehrung 1986). More specifically, managers do not have a mean and variance approach to risk in that they only consider the downside risk of potential outcomes of their choices (March and Shapira 1987; Shapira 1995). Moreover, Shapira showed that managers care more about the magnitude of possible outcomes than the probabilities associated with the outcomes. Managers are most concerned about large magnitude of potential losses, and their definition of risk is whether there is a possibility of incurring huge losses.

In addition, taking risk when investing in financial markets is done in a passive manner since investors cannot influence what goes on in such market. However, Shapira also found that managers do not passively accept risk (in settings other than pure financial markets) in that they do not view risk to be inherent in the situation that they find themselves in. Instead, managers believe that risk can be managed. Managing risk, managers argue, is what differentiates risk-taking from gambling. In fact, the many successful executives are the biggest risk takers, but the more mature executives are more risk averse (MacCrimmon and Wehrung 1990). Managers

also tend to be overly optimistic in making forecasts but, at the same time, overly cautious in making choices (Kahneman and Lovallo 1993).

Performance Feedback and Subsequent Risk Taking

Decision-maker's risk-taking propensity, however, is not constant over time. In fact, decision-makers learn through making decisions and receiving feedback based on the decisions, and this feedback affects their risk-taking propensity. Investors suffer from "myopic loss aversion," which refers to the combination of loss aversion and their tendency to evaluate equity portfolio frequently (Benartzi and Thaler 1995, p. 75). Myopically loss averse investors, however, take more risk when they evaluate their equity portfolio less frequently (Thaler et al. 1997). As the frequency of performance evaluation increases, decision-makers become more risk averse (Gneezy and Potters 1997). Decision makers are more risk averse for gains than for losses, thus, if they conform to standard learning models, they learn to favor less risky alternatives in the positive domain (March 1996) and to favor a certain outcome over an uncertain alternative (Denrell 2007). When such adaptive learning process is slow and imprecise, however, the likelihood of engaging in risky activities increases (Denrell and March 2001).

Performance feedback has a strong impact on firms' risk-taking. In analyzing a large and longitudinal data set, Ref and Shapira (2017) show how performance feedback, when a firm is near its aspiration point as well as far away from it, can lead firms to change their risk-taking in a significant manner.

Decision Targets and Risk Preference

Although managers appear not to follow the mean and variance approach in risk-taking, researchers have investigated the empirical relationship between mean and variance of firm performance. Bowman's (1980, 1982) paradox refers to the negative relationship between mean and variance of firm performance. More specifically, Bowman found that firms with higher average profit have lower variability in their profits as well. This means that low-performing firms are more risk-taking than high-performing firms. Low performance and lack of slack increase risk-taking, but the risks that are taken may also lead to poor returns (Bromiley 1991). Fiegenbaum and Thomas (1988) offered a prospect theory (Kahneman and Tversky 1979) type explanation to the paradox. They showed that a firm's target returns play an important role in its risk-taking behavior. They found that firms that are below their return target are risk seeking, whereas firms that are above their return target are risk averse.

Similarly, risk preference depends on the decision-maker's position relative to her standard (Fishburn 1977). Laughhunn et al. (1980) reported similar findings to what Fiegenbaum and Thomas (1988) did. More specifically, Laughhunn, Payne, and Crum found that managers of firms below their target returns were more risk seeking than those above their target returns. What is more interesting is that this relationship is present only when the potential for what they term as ruinous losses was not present. In other words, managers of firms below their target returns are risk seeking only if there is no possibility of incurring an unacceptable magnitude of loss. When the potential for ruinous loss is present, managers do not become risk seeking even if the firms are positioned below their target returns. This finding highlights that managers not only care about their firms' position relative to their target returns but that they also care about other factors, such as the possibility of incurring an unacceptable magnitude of loss, that affect their risk preferences. These risk tendencies are explained by the model to be presented next.

The Variable Risk Preference Model

Another reference point that needs to be considered is the survival point (March 1988; March and Shapira 1992). March and Shapira argued that decision-makers have two reference points – aspiration point and survival point – to which they attend to one at a time. According to the

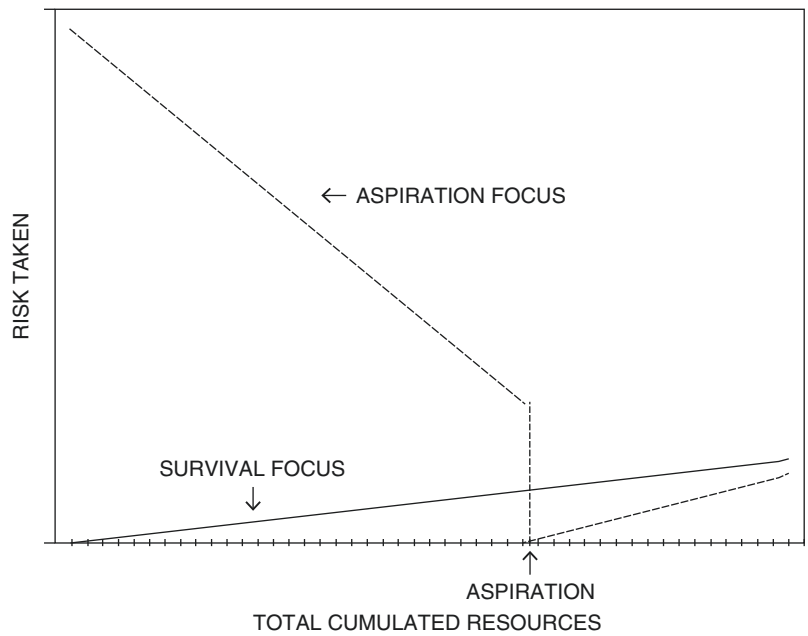
behavioral theory of the firm, as performance level falls below the aspiration point, firms engage in more search and more risky behaviors (Cyert and March 1963). Building on this point, managers are thought to attend to one of the aforementioned reference points and which reference point they pay attention to affects their risk preferences. Survival point is the point at which all cumulative resources are exhausted, and the aspired level of return is the target performance level, which is thought to adapt and change through experience (Lewin et al. 1944; Shapira 1995). When focusing on survival, a manager's risk preference has a monotonically increasing relationship with the total cumulated resources. In other words, as the magnitude of the total cumulated resources increase, managers take more risk. Under aspiration focus, managers are extremely risk seeking at very low levels of the total cumulated resources, but they become less risk seeking as they approach their aspiration level. Once they position above their aspiration level, they become risk averse, but then they start to become more risk seeking again as they move further above their aspiration level (Fig. 1).

The model therefore has two main drivers, position and focus, which may lead firms to take different risks form the same asset position if they focus on different reference points. For example, firms that find themselves way below their aspiration level but still focus on it will take much higher risk than firms who find themselves in the same asset position but focus on the survival point.

Shapira interviewed practicing managers and showed that managerial risk taking, in fact, closely adhere to the descriptive model, and through analyzing organizational data, Miller and Chen (2004) found further evidence for this model. The variable risk preference model has been utilized to complement the transaction cost economics theory's inadequate behavioral assumption of risk neutrality (Chiles and McMackin 1996). It was used to explain betting behavior of contestants in the Jeopardy! television game (Boyle and Shapira 2012) and has been further extended to account for three reference points: survival point, aspiration point, and success point (Hu et al. 2011).

Signal Detection Theory and the Project Selection Model

Signal detection theory examines how decision-makers handle signal and noise. Shapira (1995)



Risk-Taking, Fig. 1 Risk taken as a function of cumulated resources for fixed focus models of variable risk (March and Shapira 1992, p. 175)

used this framework to examine the values and utilities of outcomes that were prone to erroneous judgments and analyzed such decisions in terms of the type I and type II errors. The variable risk preference model depicts how manager's risk preference changes as her attention shifts between the two reference points. While the variable risk preference model describes manager's risk preference prior to making any strategic decision, the project selection model reflects how a manager's different risk attitude affects project selection and its outcome evaluation. The project selection model describes how manager's *ex ante* evaluation of a project is linked with the project's *ex post* outcome evaluation and when strategic surprises occur (Lampel and Shapira 2001; Shapira 1995).

Managers accept any project with the *ex ante* evaluation that is above a certain threshold x_c, and they reject projects if the *ex ante* evaluation is below that critical value. Similarly, projects with outcomes above a certain threshold y_c are considered as successes, while projects with outcomes below that critical value are considered as failures. The values of x_c and y_c determine the likelihood of project success, project failure, and strategic surprise. The project selection model has its roots in the signal detection theory (Green and Swets 1966). The signal detection theory, however, focuses on the probability side of signal detection, while the project selection model highlights the utility aspect of project selection and evaluation. The project selection model has been used to study project selection, technological foresights and oversights (Garud et al. 1997), risk-sharing incentive contracts (Shapira 1993), and hurricane evacuation decision (Dye et al. 2014).

Risk-Taking and the Upper Echelon Perspective

So far we have referred to several strategy research studies that address the topic of risk-taking in organizations as we explicated the details of the two models: the variable risk preference model and the project selection model. Other strategy research that studies risk-taking behavior often adopt the upper echelons perspective (Hambrick and Mason 1984), which posits that strategic decisions and performance levels of organizations depend on the characteristics of top managers. Since this article focused mainly on managerial risk-taking, we briefly review some of the studies that adopt the upper echelons perspective in studying managerial risk taking. These studies can be broadly classified as either linking CEO traits to her risk-taking decisions or investigating the role of agency (Eisenhardt 1989) and incentive in CEO risk-taking. Researchers that study CEO traits generally assume risk attitude to be a stable personality characteristic, whereas researchers that look at the role of agency and incentive on CEO risk-taking view risk attitude to be context dependent. The proxy behaviors that are interpreted as high level of risk-taking are often R&D expenditure, mergers and divestitures, and the adoption of a new technology or innovation.

Trait

As mentioned above, research focusing on the link between manager trait and risk-taking propensity assumes risk attitude to be a stable characteristic of an individual. The two constructs that received the most attention are hubris (i.e., overconfidence) and narcissism. Manager hubris is related with risk-seeking behavior but only when managerial discretion is high (Li and Tang 2010). Narcissistic CEOs' risk attitude, on the other hand, is contingent on the social feedback that they receive. Narcissistic CEOs become more risk seeking after receiving social praise (Chatterjee and Hambrick 2007) and more risk seeking in the domain in which there is higher audience engagement (Gerstner et al. 2013).

Agency and Incentives

Another stream of research looks at the effect of agency and incentives on managerial risk taking. These researchers assume that manager's risk attitude is context dependent,and, more importantly, assume that managers serve their interest by varying their risk propensity in making strategic decisions. As CEOs wealth is more closely linked to stock volatility, CEOs become more risk-taking (Coles et al. 2006). Similarly, CEOs that receive stock options are more risk-taking than those that do not (Deutsch et al. 2011; Sanders and

Hambrick 2007). Fixed incentive scheme, on the other hand, decreases risk-taking (Wright et al. 2007). More specifically, CEO restricted stock value relative to her reference point is linked to R&D intensity; negative deviation increases R&D, while positive deviation decreases it (Lim 2015). Relatedly, the risk preference of CEOs that are given stock options depend on the prior performance of the firm (Lim and McCann 2014). Wright et al. (1996) found that institutional ownership has a positive relationship with corporate risk-taking.

Conclusion

Managerial risk attitudes and decision-making under risk deviate from what is assumed in the expected utility theory. Managers consider downside risk and are particularly sensitive to the possibility of incurring a large magnitude losses. They pay less attention to probability estimates but, instead, pay more attention to the magnitude of possible outcomes. Furthermore, managers believe they can control the degree of risk involved in their decisions. The variable risk preference model and the project selection model depict some of these tendencies of managerial decision-making under risk. Other streams of strategy research report that CEO traits, agency, and incentive structure affect manager's risk-taking propensity.

References

Benartzi, S., and R.H. Thaler. 1995. Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics* 110 (1): 73–92.

Bowman, E.H. 1980. A risk-return paradox for strategic management. *Sloan Management Review* 21 (3): 17–31.

Bowman, E.H. 1982. Risk seeking by troubled firms. *Sloan Management Review* 23 (4): 33–42.

Boyle, E., and Z. Shapira. 2012. The liability of leading: Battling aspiration and survival goals in the *Jeopardy!* Tournament of champions. *Organization Science* 23 (4): 1100–1113. https://doi.org/10.1287/orsc.1110.0690.

Bromiley, P. 1991. Testing a causal model of corporate risk taking and performance. *Academy of Management Journal* 34 (1): 37–59. https://doi.org/10.2307/256301.

Chatterjee, A., and D.C. Hambrick. 2007. It's all about me: Narcissistic chief executive officers and their effects on company strategy and performance. *Administrative Science Quarterly* 52 (3): 351–386.

Chiles, T.H., and J.F. McMackin. 1996. Integrating variable risk preferences, trust, and transaction cost economics. *Academy of Management Review* 21 (1): 73–99.

Coles, J.L., N.D. Daniel, and L. Naveen. 2006. Managerial incentives and risk-taking. *Journal of Financial Economics* 79 (2): 431–468. https://doi.org/10.1016/j.jfineco.2004.09.004.

Cyert, R.M., and J.G. March. 1963. *A behavioral theory of the firm*. Englewood Cliffs: Prentice Hall.

Denrell, J. 2007. Adaptive learning and risk taking. *Psychological Review* 114 (1): 177–187. https://doi.org/10.1037/0033-295X.114.1.177.

Denrell, J., and J.G. March. 2001. Adaption as information restriction: The hot stove effect. *Organization Science* 12 (5): 523–538. https://doi.org/10.1287/Orsc.12.5.523.10092.

Deutsch, Y., T. Keil, and T. Laamanen. 2011. A dual agency view of board compensation: The joint effects of outside director and CEO stock options on firm risk. *Strategic Management Journal* 32 (2): 212–227. https://doi.org/10.1002/smj.876.

Dye, K.C., J.P. Eggers, and Z. Shapira. 2014. Trade-offs in a tempest: Stakeholder influence on hurricane evacuation decisions. *Organization Science* 25 (4): 1009–1025. https://doi.org/10.1287/orsc.2013.0890.

Eisenhardt, K.M. 1989. Agency theory: An assessment and review. *Academy of Management Review* 14 (1): 57–74.

Fiegenbaum, A., and H. Thomas. 1988. Attitudes toward risk and the risk-return paradox: Prospect theory explanations. *Academy of Management Journal* 31 (1): 85–106. https://doi.org/10.2307/256499.

Fishburn, P.C. 1977. Mean-risk analysis with risk associated with below-target returns. *American Economic Review* 67 (2): 116–126.

Garud, R., P.R. Nayyar, and Z. Shapira. 1997. Technological innovation, oversights, and foresights: An overview. In *Technological innovation: oversights, and foresights*, ed. R. Garud, P. Nayyar, and Z. Shapira, 3–12. New York: Cambridge University Press.

Gerstner, W.-C., A. König, A. Enders, and D.C. Hambrick. 2013. CEO narcissism, audience engagement, and organizational adoption of technological discontinuities. *Administrative Science Quarterly* 58 (2): 257–291. https://doi.org/10.1177/0001839213488773.

Gneezy, U., and J. Potters. 1997. An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics* 112 (2): 631–645.

Green, D.M., and J.A. Swets. 1966. *Signal detection theory and psychophysics*. New York: Wiley.

Hambrick, D.C., and P.A. Mason. 1984. Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review* 9 (2): 193–206. https://doi.org/10.5465/AMR.1984.4277628.

Hu, S., D. Blettner, and R.A. Bettis. 2011. Adaptive aspirations: Performance consequences of risk preferences

at extremes and alternative reference groups. *Strategic Management Journal* 32 (13): 1426–1436. https://doi.org/10.1002/smj.960.

Kahneman, D., and D. Lovallo. 1993. Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science* 39 (1): 17–31. https://doi.org/10.1287/mnsc.39.1.17.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (2): 263–292.

Lampel, J., and Z. Shapira. 2001. Judgmental errors, interactive norms, and the difficulty of detecting strategic surprises. *Organization Science* 12 (5): 599–611. https://doi.org/10.1287/orsc.12.5.599.10094.

Laughhunn, D.J., J.W. Payne, and R. Crum. 1980. Managerial risk preferences for below-target returns. *Management Science* 26 (12): 1238–1249. https://doi.org/10.1287/mnsc.26.12.1238.

Lewin, K., T. Dembo, L. Festinger, and P.S. Sears. 1944. Level of aspiration. In *Personality and the behavior disorder*, ed. J.M. Hunt, 333–378. New York: Ronald Press.

Li, J., and Y. Tang. 2010. CEO hubris and firm risk taking in China: The moderating role of managerial discretion. *Academy of Management Journal* 53 (1): 45–68. https://doi.org/10.5465/AMJ.2010.48036912.

Lim, E.N.K. 2015. The role of reference point in CEO restricted stock and its impact on R&D intensity in high-technology firms. *Strategic Management Journal* 36 (6): 872–889. https://doi.org/10.1002/smj.2252.

Lim, E.N.K., and B.T. McCann. 2014. Performance feedback and firm risk taking: The moderating effects of CEO and outside director stock options. *Organization Science* 25 (1): 262–282. https://doi.org/10.1287/orsc.2013.0830.

MacCrimmon, K.R., and D.A. Wehrung. 1986. *Taking risks: The management of uncertainty*. New York: Free Press.

MacCrimmon, K.R., and D.A. Wehrung. 1990. Characteristics of risk taking executives. *Management Science* 36 (4): 422–435.

March, J.G. 1988. Variable risk preferences and adaptive aspirations. *Journal of Economic Behavior and Organization* 9 (1): 5–24.

March, J.G. 1996. Learning to be risk averse. *Psychological Review* 103 (2): 309–319. https://doi.org/10.1037/0033-295X.103.2.309.

March, J.G., and Z. Shapira. 1987. Managerial perspectives on risk and risk taking. *Management Science* 33 (11): 1404–1418. https://doi.org/10.1287/mnsc.33.11.1404.

March, J.G., and Z. Shapira. 1992. Variable risk preferences and the focus of attention. *Psychological Review* 99 (1): 172–183.

Miller, K.D., and W.-R. Chen. 2004. Variable organizational risk preferences: Tests of the March-Shapira model. *Academy of Management Journal* 47 (1): 105–115. https://doi.org/10.2307/20159563.

Ref, O., and Z. Shapira. 2017. Entering new markets: The effect of performance feedback near aspiration and well below and above it. *Strategic Management Journal* 38 (7): 1416–1434. https://doi.org/10.1002/smj.2561.

Sanders, W.M.G., and D.C. Hambrick. 2007. Swinging for the fences: The effects of CEO stock options on company risk taking and performance. *Academy of Management Journal* 50 (5): 1055–1078. https://doi.org/10.5465/AMJ.2007.27156438.

Shapira, Z. 1993. Ambiguity and risk taking in organizations. *Journal of Risk and Uncertainty* 7 (1): 89–94.

Shapira, Z. 1995. *Risk taking: A managerial perspective*. New York: Russell Sage Foundation.

Thaler, R.H., A. Tversky, D. Kahneman, and A. Schwartz. 1997. The effect of myopia and loss aversion on risk taking: An experimental test. *Quarterly Journal of Economics* 112 (2): 647–661.

Wright, P., S.P. Ferris, A. Sarin, and V. Awasthi. 1996. Impact of corporate insider, blockholder, and institutional equity ownership on firm risk taking. *Academy of Management Journal* 39 (2): 441–463. https://doi.org/10.2307/256788.

Wright, P., M. Kroll, J.A. Krug, and M. Pettus. 2007. Influences of top management team incentives on firm risk taking. *Strategic Management Journal* 28 (1): 81–89. https://doi.org/10.1002/smj.548.

Rivalry and Collusion

Christopher Pleatsikas
Berkeley Research Group, LLC, Emeryville, CA, USA

Abstract

In economics, rivalry is the process by which firms compete to gain customers. Rivalry can be expressed through changes to the price and non-price (e.g., features and non-price terms of sale) characteristics of a good or service or through competition to produce an innovation (e.g., to add new features to an existing product or to develop new products). Collusion indicates the reduction or elimination of rivalry, as firms act collectively to suppress competition through coordination of prices, features and/or direct manipulation of the market (e.g., by deciding which firms will supply specific customers). With few exceptions, explicit collusion among rivals is treated as a violation of antitrust laws.

Definition Rivalry is the process by which firms compete to gain customers as expressed through changes to the price and non-price characteristics of a good or service. Collusion is collective action by firms to reduce or eliminate rivalry through coordination of prices, features and/or direct manipulation of the market.

Competition and Collusion

The concept of competition is fundamentally rooted in the notion of rivalry between or among economic entities in their efforts to obtain and retain consumers. Rivalry occurs within a market, which, in economics, is defined as the competitive space that includes all close substitutes for the product of interest. According to economic principles, rivalry implies continuous vying for consumers among suppliers and is driven by the profit motive and entrepreneurial incentives. In competitive markets, firms have strong incentives to offer products and services that match consumers' preferences in order to stay competitive and ensure efficient production. Rivalry also provides a catalyst for innovation by encouraging firms to find new, different and more advanced means to satisfy consumer needs. Rivalry is not confined to competing merely on price, but encompasses all aspects of the provision of a good or service, such as the sale terms, appearance, quality and performance of the product in question.

Collusion is defined as an agreement among firms to cooperate in order to avoid rivalry (Stigler 1964). Such an agreement can provide benefits to the parties through higher prices, lower costs and/or lower risks. However, these benefits are achieved at the expense of other economic entities (customers or suppliers). It is generally associated with horizontal competitors (i.e., competitors that supply substitutes for each other's products). Collusion can also occur in vertical chains (e.g., between distributors and retailers or manufacturers and distributors), although the ultimate objective is to reduce or eliminate rivalry between horizontal competitors or sets of horizontal competitors (e.g., at multiple levels of a distribution chain simultaneously). There are varied means of achieving an agreement among firms to cooperate, and cooperation can be tacit or explicit, although, as discussed below, tacit cooperation is often treated differently from explicit cooperation under antitrust laws.

The Effects of Collusion

Collusion, by increasing prices (either directly or through reducing product quality or reducing production volumes), lowers equilibrium output in a market. This in turn decreases consumer welfare and economic efficiency in a manner similar to the losses that are associated with monopoly compared with other, more competitive, market outcomes. While the price increases and output decreases may not be equivalent to those that would prevail under a single-firm monopoly, the objective is to raise prices above and/or limit output below those that would prevail in a competitive market (i.e., supracompetitive prices and/or subcompetitive output). Because collusion among competitors is illegal in many jurisdictions, collusion is, by its very nature, almost always carried out in secret.

In contrast to the secretive nature of most collusive agreements, the term 'cartel' has been traditionally used (both in its lay sense and by economists) to refer to formal – and often public – associations of otherwise independent firms. Cartels, like more secretive collusive agreements, have the objective of limiting competition to increase profits. Particularly prior to the Second World War, cartels operated openly and often with the participation or overt support of governmental authorities. Since the demise of most state-supported cartels, the term 'cartel' has been used to refer to any collusive agreement (whether legal or illegal).

There are numerous specific means by which collusive agreements attempt to increase profits: they may explicitly set prices; they may divide markets (also referred to as market allocation) geographically, by customer type or in other ways; they may require that specific customers must deal only with designated suppliers; they may be aimed at rigging a bidding process

(e.g., for government contracts or for supply to other firms); and/or they may restrict non-price terms, such as sales terms (Carton and Perfloff 2005).

The Stability of Collusive Activity

Research has demonstrated that most cartels and collusive agreements are unstable in the sense that they do not generally persist for long periods of time (although there are exceptions). For example, Levenstein and Suslow (2006) found that the average duration was 5–7 years, although there was considerable dispersion in the results. Collusive agreements face four main challenges to their durability (Motta 2005; Marshall and Marx 2012):

- *Coordination*. Coordination is required in order for competitors to align prices and/or outputs. However, precisely because collusion requires competitors to agree on market outcomes, such coordination can be difficult (e.g., different firms may have different expectations as to the share of the market they should achieve). Coordination becomes even more difficult as market conditions change and firms differ in their views as to how these changes should be reflected in any agreements among them. Successful cartels must develop a structure that allows the agreement to adapt to changing market conditions.
- *Cheating*. The incentive to cheat on any collusive agreement may be strong, as a party to the agreement may profit by undercutting prices or increasing production. Durable cartels must develop means for detecting and punishing cheating. However, the very act of punishment – which often involves cutting prices to punish the cheater – may undermine the agreement.
- *Entry*. Because collusive agreements aim to restrict output and/or increase price they increase incentives for other firms to enter a market. Consequently, participants must either deter/discourage entry or co-opt significant entrants. Successful cartels attempt to create barriers to entry (e.g., through regulations or restricted access to vital inputs).
- *Detection*. Antitrust authorities have devised both quantitative (e.g., analyses of pricing patterns) and legal methods for detecting collusion. An example of the latter is the leniency programmes many antitrust authorities have implemented that give limited or full immunity to the first party to cooperate in uncovering and prosecuting cartels (Aubert et al. 2006). These programmes have become even more significant as an increasing number of jurisdictions have followed the United States' example in criminalizing some types of collusive conduct.

Factors That Can Facilitate Collusion

Neither economists nor policymakers have been able to develop tools to predict when, under what circumstances and in which industries collusion will occur. At best, economists have only been able to identify a number of factors that make collusion more likely, although the absence of any or even a number of these factors does not indicate that collusion could not occur. These factors can be divided into two broad classes – structural factors that describe features of the market in question and behavioural factors that describe how firms in the market may operate.

Some of the most important structural factors include (Carton and Perloff 2005; Motta 2005):

- Market concentration. Coordination may be easier when there are fewer firms in the market.
- Symmetry. Coordination may be easier if the firms in the market have similar market shares and/or similar levels of vertical integration. It also may be easier if firms make products that are relatively homogeneous (i.e., are closer substitutes for the products made by competitors).
- Entry conditions. Sustaining coordination is easier the more difficult it is for new competitors to enter.
- Buyer power. Coordination is more difficult to the extent that buyers' can exercise significant market power.

- Frequency/regularity of orders. To the extent that orders occur infrequently and/or are for large volumes (relative to market size), the incentives to cheat may be increased, thereby undermining coordination.

In addition to structural factors, certain behavioural/market operation features may facilitate coordination. Price transparency is one example – when all customers tend to be charged the same price and/or prices tend to be well known in the marketplace (e.g., as with posted prices). Also, certain rivalry and collusion pricing rules, such as resale price maintenance policies (i.e., manufacturers specifying the prices that downstream sellers can charge) or the prevalence of sales terms that allow incumbent suppliers to match rivals' prices to existing customers may facilitate coordination because they may make it easier to detect cheating.

Tacit Coordination and Oligopoly

While explicit collusion between or among rivals is almost always illegal today in advanced economies, tacit coordination among rivals occupies a more ambiguous legal position. The existence of coordination among rivals that is tacit implies that no explicit agreement exists. Instead, the participants in a market may recognize their mutual interdependence in terms of pricing and output decisions and act accordingly (Stigler 1964).

Most economic theories of oligopoly – that is, markets characterized by firms that have a considerable degree of market power, but insufficient market power to dictate price without regard to competition – are based on theories relating to tacit coordination. Based initially on principles propounded in the nineteenth century (by Cournot and Bertrand), economists have developed increasingly sophisticated 'game theories' to explain how firms in oligopolistic markets will coordinate prices and/or output without the need for explicit agreements among them. These theories explain the observation that, in the real world, under many market circumstances, particularly when firms repeatedly interact over time, competitive firms can achieve stable market outcomes in which prices exceed (often substantially) the prices that would prevail in a perfectly competitive market. As a result, it is often both difficult to discern when observed prices are the result of explicit collusion as opposed to tacit coordination and, further, even when explicit collusion has been found to occur, the degree to which prices have exceeded the levels that would have occurred in the absence of explicit collusion.

Joint Ventures

Finally, there is one type of explicitly cooperative endeavour that may be specifically exempted from adverse antitrust enforcement. These are certain types of ▶ joint venture between competitors. In some cases joint ventures enable competitors to combine certain manufacturing assets when those assets, as standalone entities, would not be financially viable over the long run. More commonly, research and development joint ventures allow separate firms to combine their efforts so as to develop future products more effectively than if they had pursued these efforts individually. In such cases commercialization is pursued separately by the firms. Finally, in the last two or three decades, firms have increasingly relied on a specialized type of joint venture known as a standard setting organization (SSO). SSOs bring together intellectual property owners, device suppliers and customers to develop standards for such products as mobile phones and information storage devices to facilitate the adoption and dissemination of new technology. All these types of joint ventures may be formally approved or vetted by antitrust authorities or receive informal approval (for the limited purposes for which they were formed) in order that their participants may avoid antitrust penalties.

See Also

▶ Game Theory
▶ Joint Venture

References

Aubert, C., W. Kovacic, and P. Rey. 2006. The impact of leniency and whistleblowing programs on cartels. *International Journal of Industrial Organization* 24: 1241–1266.

Carton, D.W., and J.M. Perloff. 2005. *Modern industrial organization*, 4th ed. Boston: Pearson Addison-Wesley.

Levenstein, M.C., and V.Y. Suslow. 2006. What determines cartel success? *Journal of Economic Literature* 64: 43–95.

Marshall, R.C., and L.M. Marx. 2012. *The economics of collusion: Cartels and bidding rings*. Cambridge, MA: MIT Press.

Motta, M. 2005. *Competition policy: Theory and practice*. Cambridge: Cambridge University Press.

Stigler, G. 1964. A theory of oligopoly. *Journal of Political Economy* 72: 44–61.c.

Role of Lead Users in Innovation, the

Eric Von Hippel
Massachusetts Institute of Technology, Sloan School of Management, Boston, MA, USA

Abstract

Due to rapidly falling innovation design and communication costs, users are increasingly able to innovate for themselves without the assistance of producers. Innovation by users and 'lead users' may be larger in aggregate than innovation by producers in aggregate, and innovation research, policies and practices are adapting accordingly.

Definition 'Users' are firms or individual consumers that expect to benefit from *using* a product or a service. Lead users are a subset of all users, defined by two characteristics: (1) They are ahead of the majority of users in their populations with respect to an important market trend, and (2) they expect to gain relatively high benefits from a solution to the needs they have encountered at the market's leading edge.

'Users' are firms or individual consumers that expect to benefit from *using* a product or a service. In contrast, producers are those that expect to benefit from *selling* a product or a service. Innovation user and innovation producer are the two general 'functional' relationships between innovator and innovation. Users are unique in that they alone benefit *directly* from innovations. All others (here lumped under the term 'producers') must sell innovation-related products or services to users, indirectly or directly, in order to profit from innovations. Thus, in order to profit, inventors must sell or license knowledge related to innovations, and manufacturers must sell products or services incorporating innovations.

In this overview we first discuss evidence for the scale and scope of product and service innovations by users and lead users, and then note some implications of this increasingly important phenomenon.

Innovation by Users

Both qualitative observations and quantitative research in a number of fields clearly document the important role users play as first developers of products and services later sold by manufacturing firms. Quantitative studies of user innovation document that many of the most important and novel products and processes in a range of fields have been developed by user firms and by individual users. Thus, Enos (1962) reported that nearly all the most important innovations in oil refining were developed by user firms. Von Hippel (1988) found that users were the developers of about 80% of the most important scientific instrument innovations, and also the developers of most of the major innovations in semi-conductor processing. Pavitt (1984) found that a considerable fraction of invention by British firms was for in-house use. Shah (2000) found that the most commercially important equipment innovations in four sporting fields tended to be developed by individual users.

Empirical studies of specific fields, from medical equipment to sports equipment, show that *many* users – from 10% to nearly 40% – engage in developing or modifying products (von Hippel 2005). The first representative study of consumer product development by consumers in the UK finds that users in aggregate invest 1.4 times

more into consumer product development than do all producers of consumer products in the UK combined (von Hippel et al. 2011). Clearly, innovation by users is a major phenomenon.

Lead Users

Studies of innovating users (both individuals and firms) show them to have the characteristics of 'lead users' (Urban and von Hippel 1988; Herstatt and von Hippel 1992; Olson and Bakke 2001; Lilien et al. 2002); that is, they are ahead of the majority of users in their populations with respect to an important market trend, and they expect to gain relatively high benefits from a solution to the needs they have encountered there. The correlations found between innovation by users and lead user status are highly significant, and the effects are very large (Morrison et al. 2000; Franke and Shah 2003).

Since lead users are at the leading edge of the market with respect to important market trends, one can guess that many of the novel products they develop for their own use will appeal to other users too, and so might provide the basis for products manufacturers would wish to commercialize. This turns out to be the case. A number of studies have shown that many of the innovations reported by lead users are judged to be commercially attractive and/or have actually been commercialized by manufacturers.

Research provides a firm grounding for these empirical findings. The two defining characteristics of lead users and the likelihood that they will develop new or modified products have been found to be highly correlated (Morrison et al. 2004). In addition, it has been found that the higher the intensity of lead user characteristics displayed by an innovator, the greater the commercial attractiveness of the innovation that lead user develops (Franke and von Hippel 2003a).

Corporate Strategy and Lead User Innovation

Baldwin and von Hippel (2011) model and compare the economics of user design of new products with the economics of producer design. They find that, due to steady reductions in innovation design costs as a result of improvements in computerized design tools, and due to steady reductions in communication costs as a result of the Internet, innovation by single users and by users working openly and collaboratively increasingly compete with producers' internal innovation development processes. Producers must learn to adapt to this emerging reality by adapting their business models to acquisition of innovation designs from users.

Modification of producers' innovation processes to *systematically* search for and further develop innovations created by lead users can provide producers with better innovation performance. A natural experiment conducted at 3M illustrates this possibility. Annual sales of lead user product ideas generated by the average lead user project at 3M were conservatively forecast by management to be more than eight times the sales forecast for new products developed in the traditional manner – $146 million as against $18 million per year (Lilien et al. 2002).

Innovations developed by users can also achieve widespread diffusion when those users themselves become producers, setting up a firm to produce their innovative product(s) for sale. Shah (2000) showed this pattern in sporting goods fields. In the medical field, Lettl and Gemünden (2005) document a pattern in which innovating users take on many of the entrepreneurial functions needed to commercialize the new medical products they have developed, but do not themselves abandon their user roles. In juvenile products, Shah and Tripsas (2007) document a pattern in which user–innovators themselves found new firms. Finally, in a Kaufmann Foundation survey of US entrepreneurship, Shah et al. (2011) find that, across all fields of entrepreneurship, '46.6% of innovative startups [those founded to commercialize an innovation] founded in the United States that survive to age 5 years are founded by users'.

See Also

► Collaborative Innovation
► Open Innovation

- ▶ Open Source
- ▶ Product Innovation
- ▶ User Innovation

References

Baldwin, C.Y., and E. von Hippel. 2011. Modeling a paradigm shift: From producer innovation to user and open collaborative innovation. *Organization Science* 22: 1399–1417.

Enos, J.L. 1962. *Petroleum progress and profits: A history of process innovation*. Cambridge MA: The MIT Press.

Franke, N. and von Hippel, E. 2003a. *Finding commercially attractive user innovations*. Working paper No. 4402–03. Cambridge, MA: The MIT Sloan School of Management.

Franke, N., and S. Shah. 2003. How communities support innovative activities: An exploration of assistance and sharing among end-users. *Research Policy* 32: 157–178.

Herstatt, C., and E. von Hippel. 1992. From experience: Developing new product concepts via the lead user method: A case study in a 'low tech' field. *Journal of Product Innovation Management* 9: 213–222.

Lettl, C., and H. Gemünden. 2005. The entrepreneurial role of innovative users. *Journal of Business and Industrial Marketing* 20: 339–346.

Lilien, G.L., P.D. Morrison, K. Searls, M. Sonnack, and E. von Hippel. 2002. Performance assessment of the lead user idea generation process. *Management Science* 48: 1042–1059.

Morrison, P.D., J.H. Roberts, and E. von Hippel. 2000. Determinants of user innovation and innovation sharing in a local market. *Management Science* 46: 1513–1527.

Morrison, P.D., J.H. Roberts, and D.F. Midgley. 2004. The nature of lead users and measurement of leading edge status. *Research Policy* 33: 351–362.

Olson, E.L., and G. Bakke. 2001. Implementing the lead user method in a high technology firm: A longitudinal study of intentions versus actions. *Journal of Product Innovation Management* 18: 388–395.

Pavitt, K. 1984. Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy* 13: 343–373.

Shah, S. 2000. *Sources and patterns of innovation in a consumer products field: Innovations in sporting equipment*. Working paper No. 4105. Cambridge, MA: The MIT Sloan School of Management.

Shah, S.K., and M. Tripsas. 2007. The accidental entrepreneur: The emergent and collective process of user entrepreneurship. *Strategic Entrepreneurship Journal* 1: 123–140.

Shah, S.K., S. Winston Smith, and E.J. Reedy. 2011. *Who are user entrepreneurs? Findings on innovation, founder characteristics and firm characteristics*. Kansas City: Kauffman Foundation Report.

Urban, G.L., and E. von Hippel. 1988. Lead user analyses for the development of new industrial products. *Management Science* 34: 569–582.

Von Hippel, E. 1988. *The sources of innovation*. New York: Oxford University Press.

Von Hippel, E. 2005. *Democratizing innovation*. Cambridge, MA: The MIT Press.

Von Hippel, E., de Jong, J. P. J. and Flowers, S. 2011. Comparing business and household sector innovation in consumer products: Findings from a representative survey in the UK. *Management Science*, forthcoming at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1683503

Rosenbloom, Richard S. (1933–2011)

Michael A. Cusumano
Massachusetts Institute of Technology, Sloan School of Management, Cambridge, MA, USA

Abstract

Richard S. 'Dick' Rosenbloom was the David Sarnoff Professor of Business Administration at the Harvard Business School in Boston, Massachusetts. According to the official obituary from the university (Harvard Business School, Harvard Business School Professor Richard Rosenbloom dies at 78, press release, 25 October, 2011), he was on the faculty from 1957 to 1998, and is best known for studying the management of research, development and ▶ innovation in large firms, and the strategic uses of technology for competitive advantage and economic development.

Born in Springfield, Massachusetts, in 1933, the eldest son of the manager of a furniture factory, Rosenbloom studied chemistry at Harvard College and then completed an MBA in 1956 and a DBA in 1960 at the Harvard Business School. He began teaching in the Production and Operations Management Unit of the school in 1959. His MBA courses included Advanced Production Problems, Competitive Strategy, Management of Technology,

Manufacturing Policy and Operations Management as well as Technology, Business, and the Modern Society. He was director of the school's doctoral programme from 1970 to 1974 and served as Associate Dean for Research and Course Development from 1976 to 1980. He was also a consultant to Xerox and other firms and industry organizations, helped found the Jerusalem Institute of Management, and was a director of nine public companies in the United States, Great Britain and Israel, including Arrow Electronics.

Rosenbloom was the author or editor of numerous articles, books, and case studies. *Technology and Information Transfer*: *A Survey of Practice in Industrial Organization* (1970, with F. Wolek) collected data from more than 2000 scientists and engineers and analysed how large firms communicated technical information and used this in R&D operations. From 1983 to 1997 he co-edited *Research on Technological Innovation*, *Management*, *and Policy*, which published current papers on leading topics in the field. He also co-edited (with W. Spencer) and contributed to *Engines of Innovation*: *U.S. Industrial Research at the End of an Era* (1997).

His most widely cited article was published in 2002, 'The role of the business model in capturing value from ▶ innovation: evidence from Xerox Corporation's technology spin-off companies' (with H. Chesbrough). This examines how Xerox pioneered plain-paper copying with proprietary technology and leasing machines but then found it difficult to commercialize innovations from Xerox PARC in computing and communications, which required different ways of competing, such as direct sales and partnerships. The second most cited article came in 1995, 'Explaining the attacker's advantage: technological paradigms, organizational dynamics, and its value network' (with Clayton Christensen). The authors argue that a firm's 'value network' – how it competes and solves customer problems – has a major impact on whether incumbents or new entrants introduce the most successful innovations. In the computer disk-drive industry, they found that successful innovations are those that fit within the firm's value network. New entrants led in innovations that addressed customer needs in different ways and had an advantage when the technology disrupted existing value networks and business models.

Two other highly cited articles analysed competition in the video-recorder industry and helped lay the foundation for the value network research. The first appeared in 1987: 'Technological pioneering and competitive advantage: the birth of the VCR industry' (with M. Cusumano). This traces the origins of video recording technology at Ampex Corporation in the United States during the 1950s and then the race to create a home video-cassette recorder among Sony, Japan Victor Corporation (JVC), Matsushita-Panasonic and Toshiba in Japan, Ampex and RCA in the United States, and Philips in Europe. Sony and JVC won this competition with the Betamax (introduced in 1974) and VHS (1975) machines, respectively. The authors argue that Sony and JVC prevailed by continually experimenting with failed designs for 15 years. They gradually accumulated manufacturing and design capabilities that other firms were unable to duplicate for mass production. The second article appeared in 1992: 'Strategic maneuvering and mass-market dynamics: the triumph of VHS over Beta' (with M. Cusumano and Y. Mylonadis). This completes the story by analysing how JVC and its parent, Matsushita-Panasonic, won the competition with Sony by widely licensing the VHS standard and encouraging broad distribution of pre-recorded VHS tapes. The wider availability of VHS devices and tapes created a positive feedback loop or 'bandwagon effect' (today described as a 'network externality') that eventually led VHS to a nearly 100% market share. By 1980, VCRs had become the largest selling consumer electronics product in history and was an early example of how important it is for an innovator firm to understand strategic concepts such as platforms, complementary products and services, as well as network effects. The fields of strategic management as well as technological innovation are much indebted to Rosenbloom for his pioneering work on these topics.

See Also

- ▶ Innovation
- ▶ Research and Development (R&D) Investment

References

Harvard Business School. 2011. Harvard Business School Professor Richard Rosenbloom dies at 78, press release, 25 October.

Selected Works

Rosenbloom, R.S., and H. Chesbrough. 2002. The role of the business model in capturing value from innovation: Evidence from Xerox corporation's technology spin-off companies. *Industrial and Corporate Change* 11:529–555.

Rosenbloom, R.S., and C. Christensen. 1995. Explaining the attacker's advantage: Technological paradigms, organizational dynamics, and its value network. *Research Policy* 4:233–257.

Rosenbloom, R.S., and M. Cusumano. 1987. Technological pioneering and competitive advantage: The birth of the VCR industry. *California Management Review* 29:51–76.

Rosenbloom, R.S., and F. Wolek. 1970. *Technology and information transfer: A survey of practice in industrial organization*. Boston: Harvard Business School.

Rosenbloom, R.S., M. Cusumano, and Y. Mylonadis. 1992. Strategic maneuvering and mass-market dynamics: The triumph of VHS over Beta. *Business History Review* 66:51–94.

Rugged Landscapes

Sendil Ethiraj
University of Michigan, Stephen M. Ross School of Business, Ann Arbor, MI, USA

Abstract

A fitness landscape plots the performance of an organization as a function of its many choices. Such a landscape becomes rugged when the choices are interdependent. Rugged landscapes direct attention to the trade-offs that are inevitable when decisions are interdependent. This entry first defines the concept of rugged landscapes and then briefly explicates its utility.

Definition Rugged landscapes are fitness landscapes with many peaks and valleys.

What Are Rugged Landscapes?

The notion of fitness landscapes originated in evolutionary biology in the work of Wright (1932) to describe the evolution of organisms and was subsequently formalized and further developed by Kauffman (1993). The fitness landscape constitutes a mapping between the attributes of organisms and their fitness. The metaphor of rugged fitness landscapes was designed to capture three essential features of evolution: (1) a rich diversity of organisms that emanates from a fixed genetic space, (2) the relationship between genetic neighbours on the fitness landscape, and (3) the consequences of selection on such a population of organisms. Rugged landscapes with these properties were simulated via a computer program to study their consequences for evolution. Levinthal (1997) built on a specific rugged landscape model – Kauffman's (1993) ▶ NK models – to study the interaction between selection and adaptation. Subsequently, a large number of papers have employed similar models to study a range of questions in strategic management (see, e.g., Rivkin 2000; Rivkin and Siggelkow 2003; Ethiraj and Levinthal 2004a, b; Lenox et al. 2006).

An NK landscape is defined by two parameters, N and K. The parameter N captures the number of decision choices (say within an organization), and K the interdependence among the N decisions. Most models assume that each of the N choices can take on two states, 0 or 1, resulting in 2^N possible configurations of choices. Each choice makes a discrete contribution to the overall performance of the organization, and the contribution of a choice depends on the state (0 or 1) of that choice and K other choices. For each possible state of a choice and its K-associated choices, the

contribution of the choice to overall performance is drawn from the uniform distribution U[0,1]. Overall performance of a configuration of choices is then an average over the contributions of the N choices. When $K = 0$, there are no interaction effects; each choice can be altered without affecting the contributions of other choices, and so the landscape is relatively smooth. When $K > 0$, the landscape becomes rugged. For instance, consider when $K = 2$. The contribution of each choice then depends on the value of two other choices. The resulting performance landscape becomes less correlated in the sense that neighbouring configurations of choices might not exhibit similar performances. When K increases, the landscape becomes more rugged and less correlated. Thus, the NK model affords the simulation of a tunable family of rugged landscapes.

A simple example helps illuminate the intuition behind the NK model. Consider a firm that is fully described by two strategic choices: (1) have a simple or complex product line, (2) train or do not train the sales force. In this stylized organization, if the choice of product line is independent of the sales force, then the firm does not face a rugged performance landscape. It can make independent choices about product lines and sales force training without considering potential interaction effects. In contrast, if the choice of product line is conditional on the requisite training of the sales force, the choices are interdependent. In this latter case, there are two locally optimal strategies. One combination unites a complex product line with a trained sales force, whereas the other pairs a simple product line with an untrained sales force. The other two combinations result in poor performance either because the sales force is overtrained for the simple product line or because the sales force is insufficiently trained to handle the complex product line. Thus, as the number of interdependent decisions increase, the trade-off among the choices also increases. Such trade-offs among the choices are what give rise to ruggedness in the performance landscape.

Why Should We Care About Rugged Landscapes?

The structure of the NK landscape constitutes several features that fit with how organizations make decisions. This, in turn, makes it a useful basis with which to study enduring and important problems in organization theory and strategy.

Creation and Persistence of Firm Heterogeneity

'Where does firm heterogeneity come from?' and 'why does it persist?' are enduring and important questions in strategic management (Nelson and Winter 1978; Lippman and Rumelt 1982; Nelson 1991). The NK model combined with a variety of myopic, ▶ local search processes (Cyert and March 1963) and some dispersion in initial conditions implies sustained heterogeneity among organizations that populate the fitness landscape. This is because firms cannot escape from local optima via a ▶ local search process on a rugged landscape. With even modest N, *many* configurations are possible (for instance, 4,096 for $N = 12$, a firm with just 12 choices). In most industries, we do not observe that degree of diversity of organizations, suggesting that for a value of N that might reasonably capture the space of decisions in actual organizations, the potential space of solutions is much greater than the search capacity of organizations. This naturally produces variation in choices (and performance) that, in turn, is persistent (see, e.g., Ethiraj et al. 2008; Yayavaram and Ahuja 2008).

Modelling a Variety of Boundedly Rational Search Processes

Studying strategic management involves modelling managerial behaviour as it occurs in the real world, with all the constraints pertaining to bounded rationality, uncertainty, individual/managerial agency, political processes and so on. A workhorse that provides the platform for the consideration of these myriad effects is the behavioural theory of the firm (Cyert and March 1963). The NK structure affords the modelling of several search processes outlined in Cyert and March (1963), including problemistic search,

local search, search via coalitions, sequential search, experiential search and so on. In recent years several papers have considered the effects of a variety of search processes using the NK structure (Gavetti and Levinthal 2000; Rivkin and Siggelkow 2003; Siggelkow and Levinthal 2003; Gavetti et al. 2005). The advantage of the NK structure lies in making possible the study of phenomena that cannot otherwise be studied due to data unavailability and/or the absence of the controlled conditions required to draw inferences even if data were available.

Flexibility to Accommodate Interdependence at Different Levels

While we understand that interdependence is pervasive and important in organizations, most studies are limited in their consideration of interdependence at just one or two levels in the organizations. Interdependence may span individuals, units within organizations, organizations within an industry and so on. Such nested interdependencies at various levels are often not observable in empirical work. How interdependencies across levels affect managerial behaviour is a topic that becomes tractable using the NK structure. Indeed, several papers have begun examining the effects of such nested interdependencies (Rivkin and Siggelkow 2003; Lenox et al. 2006).

In sum, the notion of rugged landscapes borrowed from evolutionary biology, in a short span of a decade and a half, has had a significant impact on our collective understanding of fundamental issues in strategic management, such as the sources and persistence of firm heterogeneity. The interested reader may find the articles cited here to be a useful starting point to obtaining an overview of the promise and potential (both future and realized) of this literature in tackling fundamental questions in strategic management.

See Also

- ▶ Computational Simulation
- ▶ Distant Search
- ▶ Local Search
- ▶ NK Models

References

Cyert, R.M., and J.G. March. 1963. *A behavioral theory of the firm*. Englewood Cliffs: Prentice Hall.

Ethiraj, S.K., and D.A. Levinthal. 2004a. Bounded rationality and the search for organizational architecture: An evolutionary perspective on the design of organizations and their evolvability. *Administrative Science Quarterly* 49: 404–437.

Ethiraj, S.K., and D.A. Levinthal. 2004b. Modularity and innovation in complex systems. *Management Science* 50: 159–173.

Ethiraj, S.K., D.A. Levinthal, and R.R. Roy. 2008. The dual role of modularity: Innovation and imitation. *Management Science* 54: 939–955.

Gavetti, G., and D. Levinthal. 2000. Looking forward and looking backward: Cognitive and experiential search. *Administrative Science Quarterly* 45: 113–137.

Gavetti, G., D.A. Levinthal, and J.W. Rivkin. 2005. Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal* 26: 691–712.

Kauffman, S.A. 1993. *The origins of order: Self-organization and selection in evolution*. New York: Oxford University Press.

Lenox, M.J., S.F. Rockart, and A.Y. Lewin. 2006. Interdependency, competition, and the distribution of firm and industry profits. *Management Science* 52: 757–773.

Levinthal, D.A. 1997. Adaptation on rugged landscapes. *Management Science* 43: 934–950.

Lippman, S., and R. Rumelt. 1982. Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *Bell Journal of Economics* 13: 418–453.

Nelson, R.R. 1991. Why do firms differ, and how does it matter? *Strategic Management Journal* 12: 61–74.

Nelson, R.R., and S.G. Winter. 1978. Forces generating and limiting concentration under Schumpeterian competition. *Bell Journal of Economics* 9: 524–548.

Rivkin, J. 2000. Imitation of complex strategies. *Management Science* 46: 824–844.

Rivkin, J., and N. Siggelkow. 2003. Balancing search and stability: Interdependencies among elements of organizational design. *Management Science* 49: 290–321.

Siggelkow, N., and D.A. Levinthal. 2003. Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation. *Organization Science* 14: 650–669.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In *Proceedings of the sixth international congress on genetics*, ed. D.F. Jones. Ithaca: Genetics Society of America.

Yayavaram, S., and G. Ahuja. 2008. Decomposability in knowledge structures and its impact on the usefulness of inventions and knowledge-base malleability. *Administrative Science Quarterly* 53: 333–362.

Rumelt, Richard (Born 1942)

Anne Marie Knott[1] and David G. Hoopes[2]
[1]Washington University in St Louis, Olin Business School, St Louis, MO, USA
[2]California State University Dominguez Hills, Department of Management, College of Business Administration and Public Policy, Carson, CA, USA

Abstract

Richard P. Rumelt has written a number of seminal papers and books. His first book, based on his dissertation, changed the language of corporate strategy and clarified important concepts like diversification and economic performance. His ideas on resources, firm and performance heterogeneity isolating mechanisms, causal ambiguity, uncertain imitability and corporate coherence have defined many of strategic management's central debates.

Background

Richard P. Rumelt is the Harry and Elsa Kunin Chair in Business and Society at the University of California-Los Angeles Anderson School of Management. After receiving a Bachelor's degree and while working on a Master's degree in electrical engineering from the University of California-Berkeley, he worked as a systems engineer at the Jet Propulsion Labs (JPL). Rumelt earned a DBA from Harvard, taught there, and helped open a Harvard extension in Iran. In 1976, he accepted a position at UCLA where he has worked ever since (except for a 3-year leave at INSEAD). He has consulted extensively with many of the most important companies in the world and the US government.

Rumelt has contributed seminal research to the study of ▶ diversification, the ▶ resource-based view (RBV) and two of the fundamental questions of strategy: (1) why firms differ, and (2) what value corporate headquarters add (Rumelt et al. 1994).

Rumelt's Scholarship

Rumelt's career has focused on issues central to the subject of business and corporate strategy. He has been concerned with the description of firm and competitive behaviour as well as prescription as to how firms should be managed.

Four of Rumelt's scholarly works have been widely recognized as pivotal in defining the direction of the field in both subject and method. These four pivotal contributions are as follows.

Strategy, Structure, and Economic Performance

This book was Rumelt's doctoral dissertation (Rumelt 1974). In *Strategy and Structure*, Alfred D. Chandler Jr. had studied the growth of large US corporations and explored whether their growth strategies (diversification versus vertical integration versus pure scale) led to differing organizational structures. Rumelt's study sets three tasks: establishing the empirical validity of Chandler's hypothesis, developing a more nuanced scheme of both strategy and structure, and relating both strategy and structure to economic performance (mainly profitability).

After studying the histories of 246 companies over 20 years, it became apparent to Rumelt that strategy and structure were correlated and causally linked. As firms moved to a strategy of diversification, they moved from functional (e.g., marketing, finance, production) departments to structures based on decentralized product divisions. In retrospect, this seems a simple, almost necessary observation. Yet it was novel when first proposed. At that time, research on structure was largely focused on the issue of whether organizations were 'organic' or 'mechanistic'. This theoretical focus blinded the research community to the fact that the modal form of structure in large firms had been drastically transformed over the past 20 years.

With regard to a more nuanced description of strategy, Rumelt's analysis of 20 years of annual reports, 10-Ks and articles on 246 companies over 20 years gave him an ecologist's view of their main lines of development from which he drew a

new taxonomy. For example, among related-diversified companies, some were 'related constrained', keeping the preponderance of their businesses related to a common core skill or resource, whereas a 'related-linked' company diversified via chains of relationships into businesses that were, at the ends of the chain, quite disparate. Similarly, 'unrelated' businesses were divided into passive holding companies and acquisitive conglomerates, then a new species on the scene.

With regard to economic performance, return-on-capital was highest for related-constrained firms and lowest for vertically integrated firms, with differences being statistically significant. (Strategy researchers tend to use internal investment efficiency measures because they are primarily interested in the fact that some firms are able to invest in positive net present value projects more often than others, whereas measures of shareholder return reflect the degree to which performance surprises rather than surpasses.)

Rumelt's study was instrumental in converting research in strategy from a case-study format to one which examined patterns of behaviour and performance across firms. In each of the years 1976–1981, there was at least one session at the Academy of Management purely devoted to presenting follow-on studies to his. In 2014, Google Scholar shows over 5,100 citations for the book and its follow-on journal article (Lippman and Rumelt 1982).

Uncertain Imitability: An Analysis of Interfirm Differences in Efficiency Under Competition

According to neoclassical economic theory, competition should equalize profit rates (Lippman and Rumelt 1982). According to industrial organization theory (circa 1980), higher than normal profits are due to firms colluding on price behind entry barriers. Michael Porter's influential book *Competitive Strategy* took this schema and flipped it on its head: good strategy, according to Porter, consisted of getting into industries where rivalry was diminished and where there were high entry barriers (Teece 1984).

By contrast, the field of strategy had been born in the early 1960s from case studies of industries which showed that firms within the same industry did things quite differently from one another and performed differently. Rumelt began to collaborate with Steven Lippman and together they fleshed out a model of firm-level performance. Production (or cost) functions were drawn from a known distribution and imitative attempts were drawn from the same distribution. Competing firms were perfect price-takers. Firms would drop out of the industry when unprofitable. And entry would stop when the surviving incumbents were sufficiently efficient to make the expectation from a new random draw unappealing. Lippman's expertise at modelling stopping rules in stochastic processes helped put the theory on a solid mathematical footing.

In one simple model Lippmann and Rumelt predicted (1) a strong association between concentration and profitability (as an association rather than causally linked), (2) a strong association between market share and profitability (again, not causally linked), and (3) the persistence of both observed firm and industry profits above the cost of capital despite perfect price-taking.

Published in the *RAND Journal of Economics*, the paper was pivotal in moving the discussion of competitive performance from the industrial organization paradigm to one emphasizing the barriers to imitability among firms. In a subsequent, less technical, article Rumelt dubbed these 'isolating mechanisms' (Rumelt 1984). The perspective morphed into the resource-based view, and continues to be an important concept in strategy research.

In simple terms, the resource-based view posits that the locus of profits in the economy is not the corporation, or even the business, but the special difficult-to-imitate resources that a firm owns or controls. It became a building block for the ▶ dynamic capabilities perspective (Teece et al. 1997; Teece 2004, 2007, forthcoming). Lippman and Rumelt (1982) and Rumelt (1984) also served as the foundation for discussions of

firm and ▶ competitive heterogeneity. This 'fundamental question' was addressed by Rumelt et al. (1994) when they asked, 'How do firms differ?'

How Much Does Industry Matter?

The issue of performance differences among firms within an industry is central to modern strategy thinking (Rumelt 1991). Rumelt (1982) made a study of Compustat data showing that firms within industries have rates of return that differ from one another more than industries differ from one another. A decade later, Richard Schmalensee published a lead article in the *American Economic Review* which reported on a variance components decomposition of rates of return in the newly collected Federal Trade Commission (FTC) line-of-business data. He concluded that industry effects predominated, justifying the industrial organization (IO) economists' focus on industry structure.

Rumelt took the article apart piece by piece and revealed Schmalensee's mistake. Schmalensee allowed his industry variance component to capture any and all sources of industry-to-industry differences, but restricted his business-unit differences to market share effects. Rumelt applied to the FTC and, eventually, got access to this data. Running the more complete variance components decomposition, Rumelt found the business unit effects strongly dominated industry effects. In fact, in US manufacturing activities, the variance among business unit effects was six times larger than was the variance among industry effects. Furthermore, there was almost no evidence of any significant corporate effect (that is, the variance in corporate performance was fully accounted for by industry and business-unit variances.)

The *Strategic Management Review* published the paper. The Strategic Management Society recognized it as a classic, naming it 'Best Paper' 1996, the first year the paper was eligible. (A paper must have been published five or more times prior to the award date to be eligible.)

Rumelt's paper puts empirical teeth in the strategic management perspective which sees the locus of profit in the economy is at the level of the firm (or below), not the industry. Other scholars have performed numerous replications of this study in other settings and with other data.

The basic results continue to hold.

Understanding Corporate Coherence: Theory and Evidence

This study returned to the subject of corporate diversification patterns, first studied by Rumelt 22 years earlier in his thesis (Teece et al. 1994). The paper made two important contributions. First, it created a metric for measuring corporate coherence that did not require one to classify firms. Second, it provided a much-improved economic footing for thinking about corporate diversification and coherence.

Rumelt's original work on diversification required him to study the history of a company and then make a judgement about which 'category' it fell within. Other researchers followed suit, but many complained of the heavy investment in time required. The popular alternative was to use an entropy measure of diversity based on counting SIC codes. The obvious problem with this is that SIC codes are not uniformly different from one another, so counting the number in which a firm participates is fairly meaningless.

The new idea in this paper was to use the survivor principle to measure coherence. That is, two SIC codes are related if they are often combined within the same firm and distant if they are not. More precisely, the paper counts the number J_{ik} of times that SIC code i was paired with code j within firms. It then compares this with the expected number of pairings were SIC codes are fixed in number but randomly assigned to firms with no concern for coherence (a hypergeometric random variable). The measure of pairwise coherence became the number of standard deviations that J_{ik} fell from its hypergeometric mean.

From this, two overall metrics were generated. WAR measured the average coherence among activities in the firm, the average taken across all

distinct pairs. WARN measured the average coherence between neighbouring activities within a firm, neighbourliness being taken from positions on the maximum spanning tree. The upshot was that as firms added activities, WAR unsurprisingly increased but WARN did not. Firms were not expanding by jumping farther and farther afield from existing activities. Rather, they followed the natural relatedness among activities. This was a nice empirical match to the concept of 'related linked' Rumelt had developed earlier.

Because it provided a new and reasonable way of measuring coherence (diversification), and because it presented a modern analysis of the rationale for multi-business corporations, this paper breathed new life into the study of diversification. The reasoning and methods were picked up by many other scholars.

There is much work Rumelt has done beyond these four studies. We would especially mention two papers with Steven Lippman that helped to establish the micro-foundations of the resource-based view (Lippman and Rumelt 2003a, b) and expanded upon it (Liggett et al. 2009).

Rumelt's paper with D. Bardolet and D. Lovallo titled 'The hand of corporate management in capital allocations: patterns of investment in multi- and single-business firms' (2010) has helped unite the fields of behavioural economics and corporate strategy. In particular, the paper shows that multi-business firms tend to invest more heavily in cash-needy low-profit businesses than do comparable single business firms. However, there is *no* evidence that this internal 'cross-subsidy' results in superior eventual performance.

Rumelt's most recent work has been *Good Strategy/Bad Strategy* (2011). The book departs from looking at strategy in purely economic terms and defines real-world strategy work as problem-solving – that is, a strategy is an approach to dealing with a high-stakes challenge. Given that definition, a 'bad strategy' is one that fails to identify the problem being solved or is simply goal-setting, or general affirmations of ambition and values, rather than problem-solving. This way of looking at strategy allows a unified framework for examining strategy in firms, non-profits and governments because, in each context, good strategy follows from identifying the key challenges and offering ways to deal with them. Too often this does not happen because leaders hold weak models of what strategy is or because the logic of financial budgeting drives out problem identification and problem-solving.

See Also

- ▶ Competitive Heterogeneity
- ▶ Diversification
- ▶ Dynamic Capabilities
- ▶ Resource-Based Theories
- ▶ Resource-Based View

Acknowledgements The authors would like to acknowledge the help provided by the Palgrave editorial team in completing this article.

References

Liggett, T.M., S.A. Lippman, and R.P. Rumelt. 2009. The asymptotic Shapley value for a simple market game. *Economic Theory* 40: 333–338.

Lippman, S.A., and R.P. Rumelt. 1982. Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *Bell Journal of Economics* 13: 418–438.

Lippman, S.A., and R.P. Rumelt. 2003a. The payments perspective: Micro-foundations of resource analysis. *Strategic Management Journal* 24: 903–927.

Lippman, S.A., and R.P. Rumelt. 2003b. A bargaining perspective on resource advantage. *Strategic Management Journal* 24: 1069–1086.

Rumelt, R.P. 1982. How important is industry in explaining firm profitability? Working paper, UCLA.

Teece, D.J. 1984. Economic analysis and strategic management. *California Management Review* 26: 87–110.

Teece, D.J. 2004. Knowledge and competence as strategic assets. In *Handbook on knowledge management 1*. Springer: Berlin.

Teece, D.J. 2007. Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal* 28: 1319–1350.

Teece, D.J. Forthcoming. *The foundations of enterprise performance: Dynamic capabilities in an (economic) theory of firms academy of management perspectives*.

Teece, D.J., G. Pisano, and A. Shuen. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18: 509–533.

Selected Works

1974. *Strategy, structure, and economic performance*. Boston: Harvard University Press.

1984. Towards a strategic theory of the firm. In *Competitive strategic management*, ed. R. Lamb. Englewood Cliffs: Prentice Hall.

1991. How much does industry matter? *Strategic Management Journal* 12: 167–185.

1994. In *Fundamental issues in strategy*, ed. D. Schendel and D.J. Teece. Boston: Harvard Business School Press.

1994 (with D.J. Teece, G. Dosi and S. Winter). Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior & Organization* 23: 1–30.

2010 (with D. Bardolet and D. Lovallo). The hand of corporate management in capital allocations: Patterns of investment in multi- and single-business firms. *Industrial and Corporate Change* 19: 591–612.

2011. *Good strategy/bad strategy: The difference and why it matters*. New York: Crown Business.