P

Patent Citations

Michelle Gittelman
Rutgers University, Business School, Newark, NJ, USA

Abstract

This article discusses the importance of patent citations in outlining a 'paper trail' for the development of knowledge in a time of increasingly rapid technological development. It notes that the increased presence of patent data on the Internet has made it easier to achieve this process through the use of datasets such as those published by the National Bureau of Economic Research (NBER). It shows that the importance or originality of an invention can be determined to a large extent by the number of patent citations that it has received, although there are also problems in using patent citations to make such decisions. Finally, there is some discussion of the way in which patent citations can be used for research purposes.

Definition Patent citations are a means of providing a paper trail of the prior developments which any new patent draws upon. They have been widely adopted as a key measure of knowledge and a reflection of the relative importance of particular patents, and their dissemination has been greatly increased by the use of online patenting processes.

Patent citations have been widely adopted as a key measure of knowledge, contributing to a large body of empirical literature in strategic management, technological ▶ innovation and economic growth. In the United States, patent applicants and patent examiners are legally required to disclose all existing sources ('prior art'), patented and unpatented, written and otherwise, that might invalidate the patentability of an invention. Citations to the prior art are made for a number of reasons, among them to demonstrate (or refute) the novelty of claims, and to acknowledge antecedent sources that were important in developing a new invention. Citations contain remarkably rich and detailed information that is useful in studying many aspects of technological innovation: the identities and precise locations of individual inventors and the organizations to whom ▶ patents are assigned; detailed technological classes that correspond to each of a patent's claims; dates of patent application and granting; and bibliographic information about non-patent references, such as scientific publications. By encoding these aspects of patented inventions, citations provide a solution to a key problem in empirical research, noted by Paul Krugman, that '[k]nowledge flows are invisible; they leave no paper trail by which they may be measured and tracked' (Krugman 1991: 53, cited in Jaffe et al. 1993: 578).

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

Before the late 1990s these data were laborious to collect; however, falling costs, increased availability of patent data on the Internet, increases in computing power and advances in statistical methodologies have broadened the size of datasets and the scope of research. A major impetus to this trend has been the publication by the National Bureau of Economic Research (NBER) of their dataset, freely available to users, of US patents and their associated citations (Hall et al. 2001).

Uses of Patent Citations to Measure Knowledge

Citations are a central measure in empirical research, spanning topics broadly concerned with innovation, economic growth, knowledge diffusion, and the structure and evolution of technological fields. Pioneering research using patent citations found that citing patent pairs were more likely to be localized in space than a control group of non-cited patents, providing empirical support for the hypothesis that ▶ knowledge spillovers, despite being intangible and easily transmitted across space, are geographically localized (Jaffe et al. 1993). Evidence that spillovers are localized has important implications for the understanding of modern economic growth, and the geography of knowledge spillovers continues as a major theme of citation-based research (Gittelman 2007; Agrawal et al. 2008; Breschi and Lissoni 2009). In addition, citations have been applied to study a wide range of topics in management of technology, including (but not limited to) innovation in multinational corporations (Almeida 1996; Zhao 2006); competition in technological fields (Stuart and Podolny 1996; Ziedonis 2004); firm-level learning from external and internal sources (Mowery et al. 1996; Almeida et al. 2002); the impact of labour mobility and social networks on knowledge diffusion (Almeida and Kogut 1999; Singh 2005; Agarwal et al. 2009); and the relationship between publicly funded science and industrial innovation (Gittelman and Kogut 2003; Sorenson and Fleming 2004; Murray and Stern 2007). Citation-based indices have been developed that enable researchers to identify the technological breadth of individual inventions and the scope of their subsequent impact: the originality index measures the diversity of fields contained in backward citations, and the generality index is a forward-looking measure that captures the number of different fields citing back to a patent after it has been issued (Trajtenberg et al. 1997).

While nearly all patents contain at least one prior art citation, the majority of patents receive none or very few citations by other patents. This means that citations are disproportionately skewed towards a small number of patented inventions. Since so many patents – particularly in fields where patenting is widespread, such as electronics and information technology – have little individual value, it is useful to have a measure that identifies those patents that are important. Forward patent citations provide a means to identify such patents. Patent counts weighted by forward citations provide a more accurate measure of a firm's technological performance than simple patent counts, and cumulative forward citation counts offer a means of identifying particularly valuable inventions. Forward citations have been shown to be correlated with non-patent measures of technological, economic and social value, and have been used in a wide variety of studies concerned with the determinants of R&D productivity and high-impact innovations (e.g., Gittelman and Kogut 2003; Gittelman 2006; Singh and Fleming 2010).

Limits of Patent Citations as Measures of Knowledge

It is reasonable to assume that most citations are not made randomly but indicate some relationship between a citing and cited patent; a much stronger assumption, implicit in research using citations to measure knowledge flows, is that citations are a 'noisy signal' of the knowledge used in the inventive process. This latter assumption has been called into question in recent years. Patents are legal documents, and, as such, citations are shaped by complex rules and practices. Citations may reflect the strategic choices of patent

applicants concerned with protecting valuable proprietary knowledge or maximizing the value of their patent portfolios. Applicants may choose not to disclose important sources of their learning, particularly if they might invalidate important claims, or citations may be added that are important to supporting claims but were not actually used in the inventive process. Other citations may have little direct bearing on the patented invention, for instance to support a general point in a patent's description. Furthermore, prior art search is frequently conducted by professionals (attorneys, patent searchers and patent examiners) with no direct input into the inventive process. Indeed, a survey of patent inventors revealed that they were familiar with less than one third of the prior art listed on their own patents (Jaffe et al. 2000). Finally, applicants vary in the effort they expend on prior art search, which is costly and time-consuming, such that there is significant variation in the amount and quality of prior art contained in patents and the corresponding degree to which citations trace out the full scope of prior art applied to an invention. Therefore, care needs to be taken in the construction of variables and interpretation of statistical results using patent citations as a measure of inventors' prior knowledge.

One of the more salient problems in interpreting citations as inventor knowledge is that patent examiners, who represent the patent authority and have no direct input into the inventive process, add (and remove) citations to the list provided by applicants. While these citations reveal relevant prior art, they do not necessarily correspond to knowledge used by inventors in developing the citing invention. In 2001, the US patent office began publishing information that distinguished citations added by applicants from those added by patent examiners, which allowed for an analysis of their impact on the degree to which they add citations, and possibly influence or skew inferences made from pooled applicant and examiner citations. The share of examiner patents is very high: examiner-added citations accounted for 63 % of all citations in the 2000–2003 period, and about 40 % of patents contained no applicant-added citations at all (Alcácer et al. 2009). Many of the largest corporate owners of patents, particularly in fields such as electronics and computing, submit very little of their own prior art on their patents, an indication that they do not invest significant resources in prior art searches for the majority of their patents. Furthermore, the widespread assumption that examiners add random 'noise' to citations made by applicants is not borne out by the data: the two citations streams track each other quite closely. Self-citations (citations to an applicant's own prior patents) are more likely to be added by examiners than applicants themselves (Alcácer and Gittelman 2006). Individual examiners' characteristics influence their citation practices, and examiners tend to ignore applicant-added citations in evaluating claims. These patterns raise questions about the assumption that citations are a 'noisy signal' of applicant knowledge, and are suggestive that institutionally mediated factors are strong in shaping citation practices.

Moreover, the degree to which citations track real economic and technological activity has been questioned in recent empirical research. For instance, the system by which patents are classified into technology categories is designed to assist in patent retrieval and prior art search, and is not intended to correspond to industrial and economic activity. Therefore, it can be problematic to utilize classification codes to match similar patents, a core method in studies that rely on matched control group samples. Thompson and Fox Kean show that seminal findings about the localization of knowledge spillovers based on the matched-case control method are sensitive to the level at which patents are matched (three, six or nine-digit technology classifications). In a study of the widely licensed patent on recombinant DNA, citations to the patent omitted close to 90 % of organizations licensing the patent, an indication that forward citations correlate weakly with actual use of an invention. The validity of forward citations as a measure of the economic value of patents has also come under scrutiny: while there is ample evidence showing a correlation between forward citations and various measures of economic value, a recent study revealed that the magnitude of the effect is

relatively small, with forward citations explaining less than 5 % of the variation in economic value of patents as reported by inventors (Gambardella et al. 2008).

Policy-Oriented Research Using Citations

The patent system was designed to provide protection of intellectual property and thereby encourage innovation and investment in R&D. In addition to creating private incentives to innovate, patents publish detailed technical information, thereby diffusing knowledge that can further stimulate R&D in society at large. However, the patent system, as any institutional framework, can become misaligned with its intended goals and yield behaviours that are counter-productive to societal welfare. For instance, patents may be used as 'bargaining chips' in cross-licensing arrangements; firms may patent inventions that they never intend to commercialize in order to prevent others from doing so or to reap financial gains from litigation; and the patent office may grant patents on inventions that do not meet the requirements of patentability. Disclosure of prior art is a central component of a well-functioning patent system, and researchers have employed patent citations to analyse the operation of the patent system itself, with the aim of identifying weaknesses and developing policies to help improve the effectiveness and quality of patents and their governing institutional framework. This research represents a parallel stream to studies that use citations to measure knowledge; however, to the extent that more is known about the actual practice by which citations are generated and used, more accurate citation-based measures can be developed and applied to the study of knowledge-based economic activity.

See Also

- ▶ Innovation
- ▶ Innovation Networks
- ▶ Knowledge Spillovers
- ▶ Patents

References

Agarwal, R., M. Ganco, and R.H. Ziedonis. 2009. Reputations for toughness in patent enforcement: Implications for knowledge spillovers via inventor mobility. *Strategic Management Journal* 30: 1349–1374.

Agrawal, A., D. Kapur, and J. McHale. 2008. How do spatial and social proximity influence knowledge flows? Evidence from patent data. *Journal of Urban Economics* 64: 258–269.

Alcácer, J., and M. Gittelman. 2006. Patent citations as a measure of knowledge flows: The influence of examiner citations. *Review of Economics and Statistics* 88: 774–779.

Alcácer, J., M. Gittelman, and B. Sampat. 2009. Applicant and examiner citations in U.S. patents: An overview and analysis. *Research Policy* 38: 415–427.

Almeida, P. 1996. Knowledge sourcing by foreign multinationals: Patent citation analysis in the U.S. semiconductor industry. *Strategic Management Journal* 17: 155–165.

Almeida, P., and B. Kogut. 1999. Localization of knowledge and the mobility of engineers in regional networks. *Management Science* 45: 905–917.

Almeida, P., J.Y. Song, and R.M. Grant. 2002. Are firms superior to alliances and markets? An empirical test of cross-border knowledge building. *Organization Science* 13: 147–161.

Breschi, S., and F. Lissoni. 2009. Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography* 9: 439–468.

Gambardella, A., D. Harhoff, and B. Verspagen. 2008. The value of European patents. *European Management Review* 5: 85–89.

Gittelman, M. 2006. National institutions, public–private knowledge flows, and innovation performance: A comparative study of the biotechnology industry in the US and France. *Research Policy* 35: 1052–1068.

Gittelman, M. 2007. Does geography matter for science-based firms? Epistemic communities and the geography of research and patenting in biotechnology. *Organization Science* 18: 724–741.

Gittelman, M., and B. Kogut. 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science* 49: 366–382.

Hall, B., Jaffe, A. and Trajtenberg, M. 2001. *The NBER patent citations data file: Lessons, insights and methodological tools*. NBER Working paper No. 8498. Cambridge, MA: National Bureau of Economic Research. Available at <http://www.nber.org/papers/w8498> (Accessed 12 May 2013).

Jaffe, A.B.., M. Trajtenberg, and R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108: 577–598.

Jaffe, A.B.., M. Trajtenberg, and M.S. Fogarty. 2000. Knowledge spillovers and patent citations: Evidence

from a survey of inventors. *American Economic Review* 92: 215–218.
Krugman, P. 1991. *Geography and trade*. Cambridge: The MIT Press.
Mowery, D.C., J.E. Oxley, and B.S. Silverman. 1996. Strategic alliances and interfirm knowledge transfer. *Strategic Management Journal* 1: 777–791.
Murray, F., and S. Stern. 2007. Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. *Journal of Economic Behavior & Organization* 63: 648–687.
Singh, J. 2005. Collaborative networks as determinants of knowledge diffusion patterns. *Management Science* 51: 756–770.
Singh, J., and L. Fleming. 2010. Lone inventors as sources of breakthroughs: Myth or reality? *Management Science* 56: 41–56.
Sorenson, O., and L. Fleming. 2004. Science and the diffusion of knowledge. *Research Policy* 33: 1615–1634.
Stuart, T.E., and J.M. Podolny. 1996. Local search and the evolution of technological capabilities. *Strategic Management Journal* 17: 21–38.
Trajtenberg, M., R. Henderson, and A. Jaffe. 1997. University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology* 5: 19–50.
Zhao, M. 2006. Conducting R&D in countries with weak intellectual property rights protection. *Management Science* 52: 1185–1199.
Ziedonis, R.H. 2004. Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms. *Management Science* 50: 804–820.

Patent Exhaustion

C. Paul Wazzan
Berkeley Research Group, LLC, Los Angeles, CA, USA

Abstract

The classic interpretation of patent exhaustion, also known as the doctrine of first sale, holds that the first sale or licence of a patented product 'exhausts' the patent owner's rights or control relative to that product; subsequent users are free from patent claims. The product is de facto 'licensed'. This basic interpretation is eventually expanded to address method claims and self-replicating technologies.

Definition Patent exhaustion, also known as the doctrine of first sale, holds that the first sale or licence of a patented product 'exhausts' the patent owner's rights or control relative to that product.

Classic Patent Exhaustion

Patent exhaustion, also known as the doctrine of first sale, holds that the first sale or licence of a patented product 'exhausts' the patent owner's rights or control relative to that product. Simplistically, a licensed producer of patented widgets sells a widget to a retailer. The retailer is then free to use the widget however it sees fit. When the widget is ultimately packaged by the retailer with some other products and resold to the end user, that end user is not subject to a patent infringement claim by the producer of the widget. In economic terms, the patent holder's ability to require a licence from multiple parties along a production chain is limited (Osborne 2004).

Historically, the doctrine dates to the mid- to late nineteenth century, with cases decided in courts in the USA and UK. For example, see Bloomer v. McQuewan (1852), Mitchell v. Hawley (1872), Adams v. Burke (1873), Keeler v. Standard Folding Bed Co (1895), and Betts v. Willmott (1870–1871). Most European jurisdictions have employed relatively close variations of this classical interpretation (Dietz 1978), as does Japan (Kuroda and Katayama 2012). By the mid-twentieth century, the US Supreme Court had expanded the doctrine such that it applied to the sale of an item that did not completely practise the patent but did embody the patented invention (United States v. Univis Lens Co 1942).

Patent Exhaustion and Method Claims

Plaintiff LG Electronics licensed ▶ patents to Intel Corp. for use in its microprocessors. Quanta, a computer manufacturer, purchased Intel microprocessors and combined them with non-Intel components to make computers that practised LG's patents. LG sued Quanta for patent

infringement, and the district court granted Quanta's motion for summary judgement on the grounds that the licence from LG to Intel exhausted LG's rights to sue Intel's customers. On reconsideration, the district court denied summary judgement, holding that the patent exhaustion doctrine did not apply to method claims. The Federal Circuit Court of Appeals affirmed. The US Supreme Court (Quanta Computer, Inc. v. LG Electronics, Inc 2008) reversed and ruled in favour of Quanta, stating that 'methods may be embodied in a product, the sale of which exhausts patent rights'.

Patent Exhaustion and Self-Replicating Technologies

Monsanto produces patented transgenic soybean seeds that are resistant to glyphosate-based herbicides, such that weeds are destroyed by application of the herbicide while the soya plants themselves are immune. Seeds harvested from plants grown from Monsanto's transgenic seeds are like wise herbicide resistant – in other words they 'self-replicate' (Chin 2012). Monsanto claimed that use of harvested seeds (instead of repurchasing seeds from Monsanto for each new planting) was patent infringement. The district court granted summary judgement of infringement for Monsanto (Monsanto Co. v. Bowman 2009). Citing the doctrine of exhaustion, Mr Bowman appealed to the Federal Circuit, which upheld (Monsanto Co. v. Bowman 2011). The US Supreme Court granted certiorari on 5 September 2012, with respect to whether the Federal Circuit erred by (1) refusing to find patent exhaustion in patented seeds even after an authorized sale, and (2) creating an exception to the doctrine of patent exhaustion for self-replicating technologies. In May 2013, the US Supreme Court upheld the decision of the Court of Appeals for the Federal Circuit that Monsanto's patent rights had not been exhausted by an authorized sale of seeds intended for consumption (Bowman v. Monsanto 2013). The Court ruled that the exhaustion doctrine did not permit Bowman to use his purchased seeds to make additional seeds without Monsanto's permission. It should be noted that the Court did not extend its ruling to any other self-replicating technologies.

Conclusions

This article presents the classic definition of patent exhaustion and demonstrates how that definition was expanded by the US Supreme Court to include method claims under certain conditions (e.g., when the methods are embodied within the product). We conclude with a discussion of a 2011 Federal Circuit decision (now under review by the US Supreme Court) that created an exception for 'self-replicating' technologies.

See Also

- ▶ Licensing
- ▶ Patents

References

Adams v. Burke. 1873. 84 U.S. (17 Wall.) 453, 455–456.
Betts v. Willmott. 1870–1. L.R. 6 Ch App 239.
Bloomer v. McQuewan. 1852. 55 U.S. (14 How.) 539, 553–554.
Bowman v. Monsanto Co. 2013. U.S. 569, 11–796.
Chin, Y.W. 2012. Licensing, patent exhaustion, and self-replicating technologies: A case study. *The Licensing Journal* 32: 22–26.
Dietz, A. 1978. *Copyright law in the European community*. Alphen aan den Rijn: Sijthoff & Noordhoff.
Keeler v. Standard Folding Bed Co. 1895. 157 U.S. 659, 666–667.
Kuroda, K., and E. Katayama. 2012. Efforts to establish clear standards for exhaustion in Japan. *Washington Journal of Law, Technology & Arts* 7: 515–535.
Mitchell v. Hawley. 1872. 83 U.S. (16 Wall.) 544, 547.
Monsanto Company LLC v. Bowman. 2009. 686 F. Supp. 2d 834 (S.D. Ind.).
Monsanto Company LLC. v. Bowman. 2011. 657 F. 3d 1348 (Fed. Cir.).
Osborne, J.W. 2004. A coherent view of patent exhaustion: A standard based on patentable distinctiveness. *Santa Clara Computer & High Technology Law Journal* 20: 643–693.
Quanta Computer, Inc. v. LG Electronics, Inc. 2008. 170 L. Ed. 2d 996.
United States v. Univis Lens Co. 1942. 86 L. Ed. 1408.

Patents

Peter Grindley
Berkeley Research Group, London, UK

Abstract

Patents grant inventors rights to exclude others from using their inventions for a limited time, in return for public disclosure of the inventions. They encourage innovation by providing inventors with greater ability to earn a return on their inventions. Patents also underpin the technology market, stimulating diffusion. Patents may be used in many ways: to protect an innovation from being copied, to gain freedom-to-design via ▶ cross-licensing, to preclude others from getting patents on the firm's developments and to earn an optimal return on innovation by ▶ licensing.

Definition A patent is a government grant that provides the owner with exclusive rights to exclude others from using an invention for a limited time in return for public disclosure. Patents protect and reward inventions and provide a basis for the commercialization and diffusion of technology.

Patents are a critical component of protecting and stimulating invention, particularly by individuals. A patent is a government grant that awards an inventor exclusive rights to exclude others from using an invention for a limited period of time in return for its public disclosure. After the patent period, typically 20 years from grant, the invention may be freely used by others. For a US patent an invention must be novel, non-obvious and useful. The specification must disclose enough information so that a person of 'ordinary skill in the art' should be able to reproduce it. Patents in most other countries are similar and are increasingly harmonized internationally (OECD 1994; USPTO 2012a).

Purpose

A patent is not a right to use an invention, as this may also require the use of other proprietary inventions. Rather, it allows the inventor to exclude others from its use and to commercialize the invention more widely without fear of copying. The potential earnings act as a stimulus to innovation. In return the patent is published, so that the knowledge becomes known, may stimulate other inventions and will be available without restrictions after patent expiration. Patents also enable inventions to be licensed more easily, further assisting diffusion.

The owner may use the patent right in many ways. It may put the invention into practice itself by making and selling products, license its use to others, bar others from using it and not use it itself, or allow unrestricted use. All these are permitted to the patent owner. Any firm wishing to compete with the new invention must either invent an equivalent product that does not use the patented technology, or license the technology from the inventor.

There is another strategy too: infringement. A competitor or user might simply decide to infringe, believing such behaviour may pass unnoticed. There may also be genuine uncertainty about the validity of the patent and this could motivate some firms to ignore potential infringement.

Patent Value

Not all inventions have the same value – most have little value and many are never implemented (Moore 2005). While only a small percentage of inventions are valuable, for those that are the returns can be large. Rarest are the breakthrough inventions, which can create new market segments, such as the Xerox copier or Pilkington float glass. There is also considerable variation in the quality of the patents protecting innovation. Patent validity is often challenged and a patent may only be considered confirmed once it has been successfully litigated for validity, usually a complex and expensive process. Until such time its value is probabilistic (Lemley and Shapiro 2005). The protection from imitation is also limited. Many patents are 'invented around' within a few years, perhaps as many as 60 % within 4 years (Mansfield et al. 1981).

As a result, the distribution of patent values is highly skewed, with very few extremely valuable patents, a middle range of patents of various values and a large number of low-value or worthless patents (Schankerman 1998; Barney 2002). It is estimated that the top 5 % of patents may account for about 70 % of total patent value, and the top 1 % for 40 % of total value. More than half of US patents are allowed to lapse before the full patent period.

Patent rights in the USA became stronger in 1982 following the setting up of the US Court of Appeals for the Federal Circuit (CAFC) to hear patent litigation appeals. This gave greater consistency over patent litigation and increased the likelihood that a patent would be ruled valid and enforceable. The use of patenting has grown from around 60,000 US utility patents granted per year in 1982 to 166,000 in 2001 and about the same number per year since then (USPTO 2012b).

The most prolific use of patenting, and the greatest increase in patent numbers, is in information and communication technologies (ICT) industries as well as pharmaceuticals and biotechnology, speciality chemicals and medical and scientific instruments. Over 40 % of US patents granted in 2008 were in ICT.

Patents are valid only in the issuing country. A major invention may need to be patented in several countries to obtain sufficient worldwide coverage to make copying in the remaining countries uneconomic. There are usually maintenance fees to keep the patent valid, and the patent owner may allow low-value patents to lapse before full term.

Patent Strategies

Patents can be an important component of innovation strategy, providing legal protection to enable the inventor to commercialize its invention more effectively. The owner may use patent protection in several possible ways: to protect its own products, to sell or license the rights to use it to others, to block a competitor but not to exercise it itself, or simply to ensure that a competitor does not patent its technology. In some cases it may not assert the patent or allow its free use without royalties. These are all legitimate ways of using patents (Fox 1998).

Commercialization options depend on market circumstances and on the legal and economic strength of the patents. Strong IP rights expand the ways in which patents may be used. They also facilitate trade in knowledge and undergird the technology ▶ licensing market. In extracting value some specific considerations are:

- A patent itself is not enough to earn value until it is combined with complementary assets in manufacture and marketing. These may be accessed within the firm or by out-licensing the technology (Teece 1986).
- Although the patent must explain how the invention works in sufficient detail to be reproducible there may also be significant know-how required before the innovation can be used practically. The need for essential (tacit) know-how may strengthen protection.
- The skewed distribution of patent values suggests that firms may need to follow up many unprofitable ideas in search of the few that pay off.
- Innovation is a continuous process and an innovator is unlikely to be able to rest on patent protection for long. Improvements and 'invent arounds' deplete the value of patents and an innovating firm will need to stay ahead of the technology. Dynamic capabilities in innovation are a key to success in technology-related industries (Teece 2007).

The use of patents also differs according to the type of industry.

- In 'discrete' technology industries, such as pharmaceuticals, biotechnology and speciality chemicals, individual technologies operate within boundaries and do not overlap very much. Patents may be used mainly for traditional exclusion reasons, and also for out-licensing and countering litigation. Strategies focus on 'portfolio optimization' with

active patenting to protect a new technology, targeting high-value innovations or innovations that threaten competitors.
- In 'complex' technology industries, technologies are interdependent, with cumulative innovation and overlapping patent portfolios from multiple firms, such as in ICT. Patents are more likely to be used for ▶ cross-licensing, trading and preventing litigation. Strategies focus on 'portfolio maximization' to acquire patents for defensive use, as well as royalty earnings (Grindley and Teece 1997; Hall 2009).

Alternatives to Patents

Patents are only one means of protection, albeit an important one. Several surveys indicate that in many industries lead time and secrecy may be equally or more important. The most effective protection depends on the circumstances. Patents enable firms to protect innovations that are not amenable to secrecy. For discrete technologies secrecy and know-how may be used in combination with patents. In complex technologies, in which the underlying technology may be widely known, patents and lead time (rapid incremental innovation) may be more effective (Levin et al. 1987; Cohen et al. 2000).

Patents have an advantage in that they are more easily tradable than other forms of IP. Transfer of secret know-how is difficult and costly, with transactional problems of information transfer and opportunism. By contrast, patents are well defined with clear legal rights. This lends them to the technology market, expanding the ways in which an invention may be commercialized.

Patent Portfolios and Thickets

The role of patents is often explained in terms of a single invention and single patent. In practice the situation is more intricate. Even in discrete technologies most innovations build on previous innovations and are likely to be part of a stream of improvements. In complex technologies such as ICT there may be 'thickets' of patents from different owners, all of which may need to be accessed to make and sell products. The contribution of an individual patent to product value may be hard to identify within a portfolio when there are many relevant patents. Fragmented patent holdings in complex industries lead to extensive cross-licensing. Firms' patent strategies may aim to build portfolios, by research activities or purchasing patents, to use as bargaining chips in cross-licensing negotiations, as well as to protect their technology. However, a focus on patent numbers at the expense of quality may be self-defeating. A firm may target patents that block or are complementary to competitors' portfolios as those of most value in cross-licensing (Hall and Ziedonis 2001; Wagner and Parchomovsky 2005).

Antitrust Issues

Traditionally, there has been a 'tension' between antitrust and patent law – a concern that patent protection might be 'too strong' or misused to exclude rivals anti-competitively. Yet both approaches have the same aims of promoting innovation and competition, seen from different angles. A patent grants the inventor temporary monopoly rights over the use of its invention, which it may exercise in many ways, promoting dynamic competition. The patent owner must not stray beyond the 'natural scope' of the patent, such as by requiring royalty payments after patent expiration or using a patent as a pretext for other restrictions, otherwise there may be antitrust concerns (Lemley 2007).

See Also

- ▶ Appropriability
- ▶ Cross-Licensing
- ▶ Innovation Strategies
- ▶ Licensing

References

Barney, J. 2002. A study of patent mortality rates: Using statistical survival analysis to rate and value patent assets. *AIPLA Quarterly Journal* 30: 317–352.

Cohen, W., R. Nelson, and J. Walsh. 2000. *Protecting their intellectual assets: Appropriability conditions and Why U.S. Manufacturing Firms Patent (or not)*, NBER Working paper No. 7552. Cambridge, MA: NBER.

Fox, S. 1998. Intellectual property management: From theory to practice. In *Profiting from intellectual capital*, ed. P. Sullivan. New York: Wiley.

Grindley, P., and D. Teece. 1997. Managing intellectual capital: Licensing and cross-licensing in electronics. *California Management Review* 39: 8–41.

Hall, B. 2009. The use and value of patent rights. Paper presented at UK IP Ministerial Forum on the Economic Value of Intellectual Property, 10 June.

Hall, B., and R. Ziedonis. 2001. The patent paradox revisited: An empirical study of patenting in the U.S. semiconductor industry, 1979–95. *RAND Journal of Economics* 32: 101–128.

Lemley, M. 2007. A new balance between IP and antitrust. *Southwestern Journal of Law and Trade in the Americas* 13: 1–20.

Lemley, M., and C. Shapiro. 2005. Probabilistic patents. *Journal of Economic Perspectives* 19: 75–98.

Levin, R., A. Klevorick, R. Nelson, and S. Winter. 1987. Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity* 18: 783–831.

Mansfield, E., M. Schwartz, and S. Wagner. 1981. Imitation costs and patents: An empirical study. *Economic Journal* 91: 907–918.

Moore, K. 2005. Worthless patents. *Berkeley Technology Law Journal* 20: 1521–1552.

OECD (Organisation for Economic Co-operation and Development). 1994. *The measurement of scientific and technological activities: Using patent data as science and technology indicators*. Paris: OECD.

Schankerman, M. 1998. How valuable is patent protection? Estimates by technology field. *RAND Journal of Economics* 29: 77–107.

Teece, D. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* 15: 285–305.

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

USPTO (United States Patent and Trademark Office). 2012a. http://www.uspto.gov/web/offices/pac/doc/general/#patent. Accessed 13 Feb 2012.

USPTO (United States Patent and Trademark Office). 2012b. http://www.uspto.gov/about/stats/index.jsp. Accessed 13 Feb 2012.

Wagner, P., and G. Parchomovsky. 2005. Patent portfolios, research paper no. 05–25. *University of Pennsylvania Law Review* 154: 1–77.

Path Dependence in Technologies and Organizations

Carolina Castaldi[1], Giovanni Dosi[2] and Evita Paraskevopoulou[3]

[1]Technical University Eindhoven, Eindhoven, The Netherlands

[2]Scuola Superiore Sant'Anna, Pisa, Italy

[3]Universidad Carlos III Madrid, Getafe, Madrid, Spain

Abstract

Path dependence may be defined over a spectrum of phenomena ranging from mere dependence upon initial conditions to strong dependence upon a specific unfolding of events. It is observable at various layers of the economic system, ranging from the individual up to the aggregate system level. At the technology level, path dependence shows up when there is a persistence and lock-in to particular technological choices. Path dependence is also ubiquitous in the evolution and patterns of decision-making of organizations. The structure and rigidity of organizational memory, as well as the processes of interpretation, information retrieval and action formation of organizations, are fundamental sources of path dependence.

Definition Path dependence captures the idea that *history matters*. The notion is a key one within evolutionary economics and has powerful application to the understanding of irreversibilities in technological and organizational change.

The concept of path dependence captures the idea that *history matters*. Analytical approaches entailing path dependence stand against the mainstream development of economics as an 'ahistorical system of thought' (David 2001: 32). The notion is a key one within evolutionary economics (Nelson and Winter 1982), and has found powerful applications to the understanding of irreversibilities in technological and organizational change. Path dependence may be defined over

a spectrum of phenomena ranging from mere dependence upon ▶ initial conditions all the way to strong dependence upon a specific unfolding of events (see also the degrees of 'historicity' defined in Bassanini and Dosi 2001; David 2001; Castaldi and Dosi 2006).

Levels of Observation and Sources

Path dependence is observable at various layers of the economic system, ranging from the individual up to the aggregate system level.

Individual decision-making and learning tend to be path-dependent as soon as decisions are taken sequentially over time, reflect uncertainty or imperfect information, or depend on local interactions, and even more so if one accepts that preferences are endogenous in the first place (Aversi et al. 1999).

At the *technology* level, path dependence shows up when there is a persistence and lock-in to particular technological choices, reinforced by increasing returns in the production or adoption of technologies and products, and positive feedbacks and network externalities. Technological innovation and diffusion in fact often display *dynamic* increasing returns unravelling over time (Castaldi and Dosi 2006; Dosi and Nelson 2010; and more specifically on industrial dynamics, Antonelli 1999). A famous example, out of many, of lock-in to a suboptimal technology is the QWERTY keyboard supported by the path-dependent reproduction of users' skills (see David 1985; note, however, that precisely this example turned out to be controversial as an 'inferior lock-in': cf. Key 2013, and the discussion which follows in that issue of *Research Policy*). Another quite general source of path dependence entailing positive feedback is grounded in agglomeration economies, plausibly an important driver of the emergence of industrial districts such as Silicon Valley (Krugman 1991; Kenney and Von Burg 1999).

Path dependence is ubiquitous also in the evolution and patterns of decision-making of *organizations*. Organizational path dependence has been linked to various factors that explain persistence of organizational choices and that emphasize the importance that past events bear for the future orientation of organizations (see Sydow et al. 2009, for an extensive discussion). Imprinting, idiosyncratic learning and structural inertia (Stinchcombe 1965; Hannan and Freeman 1984; Argote 1999; Beckman and Burton 2008), to mention the most obvious ones, are the usual suspect mechanisms leading to path-dependent reproduction of organizational knowledge and behaviours. This is linked to the ways organizations elicit stored information, that is, their ability to remember. The structure and rigidity of organizational memory, as well as the processes of interpretation, information retrieval and action formation of organizations, are fundamental sources of path dependence (Dosi et al. 2011).

The features of selection processes are an important source of path dependence whenever evolutionary fitness (i.e., competitiveness of firms, technologies, etc.) depends in non-trivial ways upon multiple traits. In such cases, selection happens on a fitness landscape with multiple local maxima that are determined by (possibly random) initial conditions (Levinthal 1997; Castaldi and Dosi 2006). Organizations typically compete on such complex landscapes, and interrelated technological and behavioural traits are responsible for the path-dependent reproduction of organizational arrangements (Marengo 1996; Levinthal 1997, 2000). Moreover, the link between what firms do and the way they are selectively rewarded in the market is utterly opaque for at least three reasons: (i) the complexity of the environments where they operate; (ii) the mentioned multiple 'epistatic correlations' among behavioural and technological traits; and (iii) significant lags between organizational actions and performance-revealing feedbacks. In such circumstances, path dependence is also likely to be fuelled by behavioural/procedural and 'cognitive' forms of inertia (Tripsas and Gavetti 2000). At the organizational level, failure to account for the changes of the environment where an entity operates, and persistent reproduction of interpretative frameworks and actions, lead essentially to cognitive and operational lock-ins (i.e., *competence traps*).

In fact, these latter properties apply to many other formal organizations in addition to business firms (such as public agencies, trade unions etc.)

and to many institutional arrangements, for example, ethical codes and habits of thought (Dosi 1995). As argued by David (1994), institutions are a fundamental carrier of history. The attractiveness of 'doing things the way we know' can often act as an obstacle to change and lock individuals, organizations and whole economic systems into suboptimal behaviours and problem-solving heuristics. A famous example of the consequences of path-dependent individual decision-making relates to the segregation phenomena (Schelling 1971).

As countries can be characterized by combinations of complementary institutions, path dependence also strongly affects national dynamics (e.g., see the discussion on national systems of innovation: Lundvall 1992; Nelson 1993; Kogut 1993, and the evidence of persistence of national specializations).

Note, in any case, that evolution does not need to equate to progress, as one can identify many examples of path-dependent dynamics going *from better to worse* (see the story of Easter Island in Diamond 1995).

Escape Routes

Tackling 'bad path dependencies' involves different sorts of remedies with different degrees of intentionality. First, de-locking may rely on environmental shocks, on the arrival of new knowledge bases and, consequently, new paradigms. Relatedly, deviant behaviours may 'autocatalyse' and aggregately account for shifts in the system orientation (Castaldi and Dosi 2006). Within organizations, path-breaking routes include the purposeful loss of memory, changes in the organizational structure, increasing 'cognitive dissonance' between organizational cognitive frames and action repertoire, and management and labour turnover (Garud and Karnøe 2001; Dosi et al. 2011).

Formal Representations

Path-dependent phenomena have been modelled using mathematical tools such as non-linear dynamics and chaos (Brock and Malliaris 1989; Brock 1993), stochastic processes such as generalized Pólya urns (Arthur 1994; Dosi and Kaniovski 1994), and have borrowed models and concepts from (evolutionary) biology (e.g., on the dynamics of the evolution of fitness landscapes, see Kauffman 1989). Moreover, the broad field of complexity has been the fertile ground for multidisciplinary research on path dependence (see Frenken 2006).

Open Questions

Understanding path dependence is seriously hampered from an empirical point of view by the fact that in social sciences one generally observes only one of the many possible histories. Nevertheless, Gould (1977) has suggested the power of trying to imagine what would remain unchanged if the tape of evolution were run twice. The risk is the one of *ex post* evolutionary rationalizations, but plenty of opportunities are offered by available mathematical and conceptual models. A major challenge is the one of conceptualizing hierarchically nested evolutionary processes, allowing for slowly changing macro institutions, which in turn structure faster microdynamics of adaptation.

See Also

- ▶ Initial Conditions
- ▶ Learning and Adaptation
- ▶ Lock-in Effects
- ▶ Technology Adoption

References

Antonelli, C. 1999. The economics of path-dependence in industrial organization. *International Journal of Industrial Organization* 15: 643–675.

Argote, L. 1999. *Organizational learning: Creating, retaining and transferring knowledge*. Boston: Kluwer Academic.

Arthur, W.B. 1994. *Increasing returns and path dependence in the economy*. Ann Arbor: University of Michigan Press.

Aversi, R., G. Dosi, G. Fagiolo, M. Meacci, and C. Olivetti. 1999. Demand dynamics with socially evolving preferences. *Industrial and Corporate Change* 8: 353–408.
Bassanini, A.P., and G. Dosi. 2001. When and how chance and human will can twist the arms of Clio. In *Path dependence and creation*, ed. R. Garud and P. Karnøe. Mahwah: Lawrence Erlbaum Associates.
Beckman, C.M., and M.D. Burton. 2008. Founding the future: Path dependence in the evolution of top management teams from founding to IPO. *Organization Science* 19: 3–24.
Brock, W.A. 1993. Pathways to randomness in the economy: Emergent nonlinearity and chaos in economics and finance. *Estudios Economicos* 8: 3–54.
Brock, W.A., and A.G. Malliaris. 1989. *Differential equations, stability and chaos in dynamic economics*. Amsterdam: North-Holland.
Castaldi, C., and G. Dosi. 2006. The grip of history and the scope for novelty: Some results and open questions on path dependence. In *Understanding change*, ed. A. Wimmer and R. Koessler. Basingstoke: Palgrave Macmillan.
David, P.A. 1985. Clio and the economics of QWERTY. *American Economic Review* 75: 332–337.
David, P.A. 1994. Why are institutions the 'carriers of history'? Path dependence and the evolution of conventions, organizations and institutions. *Structural Change and Economic Dynamics* 5: 205–220.
David, P.A. 2001. Path dependence, its critics and the quest for 'historical economics'. In *Evolution and path dependence in economic ideas: Past and present*, ed. P. Garrouste and S. Ioannides. Cheltenham: Edward Elgar.
Diamond, J. 1995. Easter's end. *Discover* 16: 63–69.
Dosi, G. 1995. Hierarchies, market and power: Some foundational issues on the nature of contemporary economic organization. *Industrial and Corporate Change* 4: 1–19.
Dosi, G., and Y. Kaniovski. 1994. On 'badly behaved' dynamics: Some applications of generalized urn schemes to technological and economic change. *Journal of Evolutionary Economics* 4: 93–123.
Dosi, G., and R.R. Nelson. 2010. Technical change and industrial dynamics as evolutionary processes. In *Economics of innovation*, ed. B.H. Hall and N. Rosenberg. Amsterdam: Elsevier.
Dosi, G., L. Marengo, E. Paraskevopoulou, and M. Valente. 2011. The value and dangers of remembrance in changing worlds: A model of cognitive and operational memory of organizations. LEM Working paper, in progress.
Frenken, K. 2006. *Innovation, evolution and complexity theory*. Cheltenham: Edward Elgar.
Garud, R., and P. Karnøe (eds.). 2001. *Path dependence and creation*. Mahwah: Lawrence Erlbaum Associates.
Gould, S.J. 1977. *Ever since Darwin*. New York: Norton.
Hannan, M.T., and J. Freeman. 1984. Structural inertia and organizational change. *American Sociological Review* 49: 149–164.
Kauffman, S.A. 1989. Adaptation on rugged fitness landscapes. In *Lectures in the sciences of complexity*, vol. 1, ed. D.L. Stein. New York: Addison Wesley.
Kenney, M., and U. Von Burg. 1999. Technology, entrepreneurship and path dependence: Industrial clustering in Silicon Valley and Route 128. *Industrial and Corporate Change* 8: 67–103.
Key, N.M. 2013. The QWERTY problem. *Research Policy*, forthcoming.
Kogut, B.M. (ed.). 1993. *Country competitiveness: Technology and the organizing of work*. New York: Oxford University Press.
Krugman, P. 1991. Increasing returns and economic geography. *Journal of Political Economy* 99: 483–499.
Levinthal, D. 1997. Adaptation on rugged landscapes. *Management Science* 43: 934–950.
Levinthal, D. 2000. Organizational capabilities in complex worlds. In *The nature and dynamics of organizational capabilities*, ed. G. Dosi, R.R. Nelson, and S.G. Winter. Oxford: Oxford University Press.
Lundvall, B.-Å. 1992. *National systems of innovation: Toward a theory of innovation and interactive learning*. London: Pinter Publishers.
Marengo, L. 1996. Structure, competence and learning in an adaptive model of the firm. In *Organization and strategy in the evolution of the enterprise*, ed. G. Dosi and F. Malerba. London: Macmillan.
Nelson, R.R. (ed.). 1993. *National innovation systems: A comparative analysis*. New York: Oxford University Press.
Nelson, R.R., and S.G. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Belknap Press of Harvard University Press.
Schelling, T.C. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* 1: 143–186.
Stinchcombe, A. 1965. Social structures and organizations. In *Handbook of organizations*, ed. J.G. March. Chicago: Rand McNally.
Sydow, J., G. Schreyong, and J. Koch. 2009. Organizational path dependence: Opening the black box. *Academy of Management Review* 34: 689–709.
Tripsas, M., and G. Gavetti. 2000. Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic Management Journal* 21: 1147–1161.

Penrose Effect, the

Danchi Tan
National Chengchi University, Department of International Business, Taipei, Taiwan, China

Abstract

► Penrose, Edith T. (**1914–1996**) book (1959) provides a theory of the growth of the firm, maintaining that the binding constraint on the rate of the growth of the firm arises from the

limited capacities of its existing management. This managerial constraint on the rate of growth of the firm is known as *the Penrose effect*. The article discusses the reasons why the Penrose effect arises, reviews how this effect has been empirically examined, and suggests some future research directions.

Definition A computational simulation is a dynamic, process-oriented model instantiated on a computer. These can range from traditional economic models (expressed as equations) to more abstract constructs and processes (expressed as objects, agents, operators and algorithms).

The Penrose effect refers to the managerial constraint on the rate of growth of the firm. The effect predicts negative intertemporal correlations in the growth rates of firms: a fast-growing firm will encounter managerial problems and thus slow down its growth in the subsequent time period.

► Penrose, Edith T. (**1914–1996**) seminal book (1959) provides a theory of the growth of the firm, maintaining that the binding constraint on the rate of the growth of the firm arises from the limited capacities of its existing management. This managerial constraint on the rate of growth of the firm is known as *the Penrose effect* (Hay and Morris 1991: 347). According to Penrose, planning and managing growth requires inputs from managers who have experiences that are internal to the firm (Mahoney and Pandian 1992; Kor and Mahoney 2000). 'If a firm deliberately or inadvertently expands its organization more rapidly than the individuals in the expanding organization can obtain the experience with each other and with the firm that is necessary for the effective operation of the group, the efficiency of the firm will suffer . . . and a period of "stagnation" may follow' (Penrose 1959: 47). The Penrose effect therefore predicts negative intertemporal correlations in the growth rates of firms: a fast-growing firm will encounter managerial problems and thus slow down its growth in the subsequent time period.

Penrose's theory suggests that managers with experiences internal to the firm, or *internally experienced managers*, provide managerial and entrepreneurial services that are crucial for the growth of the firm (Penrose 1959: 31, n. 1). First, given that a firm is essentially an administrative organization, it relies on managerial knowledge to direct and coordinate productive resources. The process of decision-making and coordination requires internally experienced managers because it is too complex to be codified as a management 'blueprint' that newly hired managers could implement (p. 46).

Second, internally experienced managers also influence the development of newly recruited personnel. To be able to provide managerial services that are economically valuable to the firm, newly recruited personnel need to learn 'the best way of doing things in the particular set of circumstances in which they are working' (p. 52). Internally experienced managers help to develop the new recruits by providing them with tacit knowledge of the ways things work within the firm, and by laying out plans to aid the newly hired personnel to learn on the job and gain requisite experience. As a result, 'the amount of activity that can be planned [by internally experienced managers] at a given time limits the amount of new personnel that can profitably be absorbed in the "next period"' (p. 49).

Third, internally experienced managers provide entrepreneurial services in defining the product opportunities of their firm. Penrose suggests that a firm is not only an administrative organization but a collection of product resources, where the choice of different uses of these resources over time is determined by administrative decisions (Penrose 1959: 24). She maintains that external conditions such as product or factor markets are never a serious barrier to growth, because 'there is not an effective limit to the amount of any kind of productive resources that the firm can obtain at a price' (p. 43), and, more importantly, 'there are opportunities for profitable investment open somewhere in the economy' (p. 43). Thus, the product opportunity of a firm is restricted to the extent to which its entrepreneur fails to recognize opportunities for expansion, is unwilling to act or is unable to respond to the opportunity (p. 32).

Internally experienced managers therefore provide indispensible services for the growth of the firm. Because such managers must be developed

within the firm and cannot be hired from outside, firms face an inelastic supply of managerial services in the short run. Rapid growth of a firm in one time period is likely to be followed by a temporary period of stagnation because the firm cannot adjust its managerial inputs to the desired level in a timely enough manner to manage increased complexity after the growth (Penrose 1955).

The Penrose effect is incorporated within mainstream macroeconomic (Uzawa 1969) and microeconomic theories of investment as a source of dynamic adjustment costs of the firm (Mortensen 1973; Rubin 1973; Slater 1980; Treadway 1970). While case studies support the Penrose effect (Penrose 1960; Richardson 1964), Shen (1970) presents the first large-sample empirical evidence for this effect. He suggests that the Penrose effect accounted for the negative correlation coefficients between the growth rates of 4,000 Massachusetts manufacturing plants for the periods 1948–1953 and 1953–1957. Shen maintains that if not for managerial constraints, growth rates of these plants in successive time periods would be positively correlated because larger plants enjoy increasing returns to scale for labour (Shen 1970: 706) and therefore can grow faster.

Shen's study shows empirical evidence of the Penrose effect at the plant level. Subsequent studies provide supportive, but not robust, empirical evidence of the effect at the firm level. Gander (1991) suggests that, due to the Penrose effect, there are decreasing returns to managerial resources (i.e., managerial diseconomies). He thus predicts that the growth rate of firm size is matched by a greater growth rate of managers. Gander tests this prediction using aggregate two-digit SIC US industry data and finds support for the 1977–1980 period but not for the 1983–1986 period. Orser et al. (2000) examine the correlation between two consecutive years of revenue on a sample of small and medium-sized Canadian firms. They report that fewer than one-quarter of these firms had two consecutive years of revenue increases.

More recent empirical studies indicate that organizational forms may influence the need for the services of internally experienced managers, which affects the magnitude of the Penrose effect. Thompson (1994) and Shane (1996) maintain that the use of contractual forms of expansion reduces the need for managers to be directly involved in daily operations and thus mitigates the managerial constraints. They show that US firms following a franchise strategy grew faster than those firms that expanded by establishing hierarchical outlets. Tan (2003) investigates whether the Penrose effect exists for foreign direct investments. She did not find that a fast-growing Japanese firm in a US industry grew more slowly in the subsequent time period. Tan suggests that the reduced need for coordinating foreign subsidiaries, and the use of multidivisional organizational structures within a multinational firm, may relieve some managerial diseconomies that one would expect from domestic firm expansion. In a follow-up study, Tan and Mahoney (2005) examine and corroborate the impact of the need for cross-border coordination on the Penrose effect. They find that Japanese investors in the United States were more vulnerable to the Penrose effect in industries where close coordination within multinational firms is required. Tan (2009) further corroborates the importance of coordination in the Penrose effect by comparing acquisitive and organic (greenfield) growth. She shows that acquisitive entry relieves managerial constraints and enables faster growth than organic (greenfield) entry when the headquarters and subsidiaries are weakly interdependent; however, when there is a need for strong interdependence between headquarters and subsidiaries, acquisitive entry results in slower growth. Hutzschereuter et al. (2011) draw from a sample of German multinational firms and show that cross-border coordination results in even greater managerial constraints when a multinational firm expands into culturally distant and diverse foreign markets. In general, extant theoretical modelling and empirical works on the Penrose effect have focused on (the limit of) the growth of a firm in its *existing* business activities. The focus has led these studies to emphasize the *managerial* services, as opposed to the *entrepreneurial* services, of internally experienced managers, in that their roles mainly consist

of managing daily operations, coordinating (complementary) activities and developing recently hired recruits.

Future research could fruitfully explore the Penrose effect on the growth of a firm *beyond* its existing business activities through the processes of product diversification and international expansion. This suggested research direction would require greater attention being paid to the *entrepreneurial* services of internally experienced managers, since it is the entrepreneur's view of productive opportunities that leads the firm into new product and geographical areas (Augier and Teece 2007).

Examining the Penrose effect in the growth of firms in new product and geographical markets would also require researchers to address managerial learning more fully. Extant works on the Penrose effect mainly focus on the learning of newly recruited personnel and emphasize the acquisition of internal experience. However, moving into new markets also requires internally experienced managers to learn, because they need to develop new technological knowledge (Chandler 1990) and location-specific knowledge (Pitelis and Verbeke 2007). Several recent works began to explore this issue. Tan and Mahoney (2007) find that Japanese firms were more capable of achieving growth in consecutive time periods in the entered US industries when the environment is more conducive to the learning of local market knowledge and when the firms send more expatriates to develop the local personnel. Goerzen and Beamish (2007) show that the use of expatriates by Japanese firms contributes more to their US subsidiary performance when these firms have greater experience of the local market. Drawing on a large sample of Swedish firms, Lockett et al. (2011) find that while fast ▶ organic growth leads to low subsequent organic growth, fast acquisitive growth generates fast organic growth in successive periods because acquisitions expand a firm's productive opportunity by bringing in new resources and enabling learning new knowledge. Future research exploring this issue may prove useful in the increasingly important development processes within firms.

See Also

- ▶ Managerial Resources and Capabilities
- ▶ Organic Growth
- ▶ Penrose, Edith T. (1914–1996)
- ▶ Resource-Based View

References

Augier, M., and D.J. Teece. 2007. Dynamic capabilities and multinational enterprise: Penrosean insights and omissions. *Management International Review* 47: 175–192.

Chandler, A.D. 1990. *Scale and scope: The dynamics of industrial capitalism*. Cambridge, MA: Harvard University Press.

Gander, J.P. 1991. Managerial intensity, firm size and growth. *Managerial and Decision Economics* 12: 261–266.

Goerzen, A., and P.W. Beamish. 2007. The Penrose effect: 'Excess' expatriates in multinational enterprises. *Management International Review* 47: 221–239.

Hay, D.A., and D.J. Morris. 1991. *Industrial economics*. New York: Oxford University Press.

Hutzschereuter, T., J.C. Voll, and A. Verbeke. 2011. The impact of added cultural distance and cultural diversity on international expansion patterns: A Penrosean perspective. *Journal of Management Studies* 48: 305–329.

Kor, Y.Y., and J.T. Mahoney. 2000. Penrose's resource-based approach: The process and product of research creativity. *Journal of Management Studies* 37: 109–139.

Lockett, A., J. Wiklund, P. Davidsson, and S. Girma. 2011. Organic and acquisitive growth: Re-examining, testing, and extending Penrose's growth theory. *Journal of Management Studies* 48: 48–74.

Mahoney, J.T., and J.R. Pandian. 1992. The resource-based view within the conversation of strategic management. *Strategic Management Journal* 13: 363–380.

Mortensen, D.T. 1973. Generalized costs of adjustment and dynamic factor demand theory. *Econometrica* 41: 657–665.

Orser, B.J., S. Hogarth-Scott, and A.L. Riding. 2000. Performance, firm size, and management problem solving. *Journal of Small Business Management* 38: 42–58.

Penrose, E.T. 1955. Limits to the growth and size of firms. *American Economic Review* 45: 531–543.

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

Penrose, E.T. 1960. The growth of the firm: A case study of the Hercules Powder Company. *Business History Review* 34: 1–23.

Pitelis, C., and A. Verbeke. 2007. Edith Penrose and the future of the multinational enterprise: New research directions. *Management and International Review* 47: 139–149.

Richardson, G.B. 1964. The limits to a firm's rate of growth. *Oxford Economic Papers* 16: 9–23.
Rubin, P.H. 1973. The expansion of firms. *Journal of Political Economy* 81: 936–949.
Shane, S.A. 1996. Hybrid organizational arrangements and their implications for firm growth and survival: A study of new franchisors. *Academy of Management Journal* 39: 216–234.
Shen, T.Y. 1970. Economies of scale, Penrose-effect, growth of plants and their size distribution. *Journal of Political Economy* 78: 702–716.
Slater, M. 1980. The managerial limitations to a firm's rate of growth. *Economic Journal* 90: 520–528.
Tan, D. 2003. The limits to the growth of multinational firms in a foreign market. *Managerial and Decision Economics* 24: 569–582.
Tan, D. 2009. Foreign market entry strategies and post-entry growth: Acquisitions vs. greenfield investments. *Journal of International Business Studies* 40: 1046–1063.
Tan, D., and J.T. Mahoney. 2005. Examining the Penrose effect in an international business context: The dynamics of Japanese firm growth in US industries. *Managerial and Decision Economics* 26: 113–127.
Tan, D., and J.T. Mahoney. 2007. The dynamics of Japanese firm growth in US industries: The Penrose effect. *Management International Review* 47: 259–279.
Thompson, R.S. 1994. The franchise life cycle and the Penrose effect. *Journal of Economic Behavior & Organization* 24: 207–218.
Treadway, A.B.. 1970. Adjustment costs and variable inputs in the theory of the competitive firm. *Journal of Economic Theory* 2: 329–347.
Uzawa, H. 1969. Time preference and the Penrose effect in a two-class model of economic growth. *Journal of Political Economy* 77: 628–652.

Penrose, Edith T. (1914–1996)

Christos N. Pitelis[1] and Liudmyla Svystunova[2]
[1]Brunel University London, Uxbridge, UK
[2]University of Bath, Bath, UK

Abstract

Edith Penrose is one of the founding figures in the economic theory of the (multinational) firm and (international) strategic management. Her now famous book, *The Theory of the Growth of the Firm*, has become a canonical reference to the currently dominant resource-, knowledge-, and (dynamic) capabilities-based approaches to business strategy and has branched out to numerous related fields. In this entry we briefly discuss Penrose's background, her seminal book, and her wider contribution to and influence on organizational economics, strategic management, international business, economic development, and public policy.

Life and Work

Edith Elura Tilton was born on 29 November 1914 in Los Angeles. In 1936 she graduated from the University of California at Berkeley with a BA in economics.

At Berkeley, she was taught by and assisted the economist E. F. Penrose ('Pen'). She followed Pen as a researcher when he became the Economic Adviser to the US Ambassador in Britain. That brought her into contact with prominent economists, including Schumpeter, Keynes, and Meade.

In 1947, Penrose began her master's and doctoral studies at Johns Hopkins University, supervised by Fritz Machlup, co-director of a project on the growth of firms. Her fieldwork at the Hercules Powder Company began the research that eventually led to her book *The Theory of the Growth of the Firm* (TGF). On completion of her doctorate in 1951, she became Lecturer and Research Associate at Johns Hopkins.

Penrose continued work on TGF at the Australian National University in Canberra (1955–1957) and then at the University College of Arts and Sciences in Baghdad (1957–1959), where she developed an interest in the Arab world, the oil industry, the international firm, and developing countries. She later co-authored *Iraq*: *International Relations and National Development* (1978) with Pen.

In 1959, Penrose obtained a joint Readership in Economics with reference to the Middle East at the London School of Economics and the School of Oriental and African Studies, London, where in 1964 she took the first Chair of Economics with Special Reference to Asia. From 1978 and until her retirement in 1984, she served as Professor of Political Economy and Associate Dean for Research and Development at INSEAD.

Following retirement, Penrose led an active life, serving on several governing bodies and committees and advising businesses and countries. Increasing recognition of her early work set her thinking again about theories of the firm. She became interested in how firms were changing and had considered the idea of a theory of the death of the firm, which metamorphosed into the metamorphosis of the firm (see Penrose 2008). Edith Penrose passed away on 11 October 1996.

Penrose's Contribution

Penrose's magnum opus is undoubtedly TGF, first published in 1959. "The Growth of the Firm, A Case Study: The Hercules Powder Company," completed in 1956 and published in 1960, was originally intended to be part of TGF. Hence it is safe to view the two pieces as part of an integrated whole and consider its main arguments.

For Penrose, mainstream economic theory viewed "the 'firm' as primarily a set of supply and demand functions" (Penrose 1985: 6). It was not equipped for "the analysis of the expansion of the innovating, multi-product, 'flesh and blood' organization that businessmen call firms" (p. 12) and in particular in explaining firm growth. However, Penrose chose not to challenge the extant theory of the "firm" as part of the theory of price and production, so long as it cultivates its own garden and we cultivate ours' ([1959] 2009: 9).

The Penrosean Firm and the Market

Penrose viewed firms as bundles of resources, under internal direction, for the production of goods and services, sold in markets for a profit. Their boundaries were defined by the area of coordination and "authoritative communication." Firms differed from markets in that transactions in markets did not take place within administrative coordination. Entrepreneurs were in search of profits, and firms did not maximize profits in the conventional neoclassical economics sense of equating marginal costs to marginal revenues. However, firms desired and pursued an increase in total long-term profits "for the sake of the firm itself and in order to make more profit through expansion" (Penrose [1959] 2009: 29).

For Penrose, resources rendered multiple services. The heterogeneity of services from resources gave each firm its unique character. Effective use of resources and innovation took place when resources were combined with each other. The external environment was seen as an "image" in the mind of the entrepreneur. Firms' activities were governed by their "productive opportunity," which involved a dynamic interaction between the internal and the external environment and included all of the productive possibilities that its entrepreneurs could see and leverage. The long-term profitability, growth, and survival of firms depended on their establishing "relatively impregnable bases" (RIBs) (Penrose [1959] 2009: 137) from which to adapt and extend their operations in an uncertain, changing, and competitive world. A new technological base required the firm to achieve a "competence" in some significantly different area of technology (Pitelis 2009b).

There are two major categories of "causes" of growth for Penrose: those external to the firm and those internal. Penrose suggested that external causes (raising capital, demand conditions, etc.), while of interest, "cannot be fully understood without an examination of the nature of the firm itself" (1955: 532). The problem as she saw it was "the internal incentives to and limits on growth – a theory of the growth of the firm that does not relate to fortuitous external events" (p. 532).

There were two basic reasons why there are incentives for growth endogenous to the firm, which were self-reinforcing, leading to opportunities for further expansion. First is the claim that the execution of any plan requires resources, which are in excess of those strictly necessary for its execution. Second, on completion of a plan, managerial resources will be released. Crucially, "the services that the firm's management is capable of rendering will tend to increase between the time when the plan is made and the time when the execution is completed" (Penrose 1955: 533).

Penrose attributed the ubiquitous presence of unused resources to arguments by Charles Babbage, Austin Robinson, and Sargent Florence such as the "balance of processes" or "the principle of multiples," which implied that

> a firm would have to produce on a vast scale if it were to use fully the services of all the resources required for much smaller levels of output." (Penrose 1955: 533)

Managerial services were of particular importance in this context, in part because they were available to the firm only in limited amounts. In addition, the completion of expansion plans created resources as personnel gained additional experience and released resources.

An increase in knowledge caused the productive opportunity of the firm to change in ways not directly related to changes in the environment and contributed to the "uniqueness" of such individual firms (Penrose [1959] 2009: 48). This was particularly true because while some knowledge is "objective" (transmittable), some took the form of "experience" (what today we call 'tacit'), which was hard to transmit. Experience rendered managerial services firm-specific.

For Penrose, unused productive services for the enterprising firm simultaneously represented a challenge to innovate, an incentive to expand, and a source of competitive advantage. They facilitated innovation within the firm and were "a selective force in determining the direction of expansion" (Penrose [1959] 2009: 77).

Once it was recognized that firms are to be defined in terms of resources, not products, and given the resources' potential versatility, demand conditions could not limit a firm's expansion, and "diversification" became the normal state of affairs. In this sense, and in the absence of traditional managerial diseconomies, the existence of which Penrose questioned, there were limits to growth but not to the size of firms and that was determined by the rate at which experienced managerial staff could plan and implement plans. The services of "inherited" managerial resources controlled the amount of new managerial resources that could be absorbed, thus limiting the rate of growth of firms.

Firm Growth, Business Strategy, and Industry Organization

The above account is well known (see Penrose and Pitelis 1999; Pitelis 2009b). Less known is Penrose's use of these ideas in explaining vertical integration, mergers and acquisitions, industrial concentration, the scope for small firms, and competition policy.

For Penrose, firms integrated vertically in part because they might be able to produce more cheaply for their own requirements (Penrose 1956, [1959] 2009). However, they had to set this against the diversion of resources from potentially more profitable activities. Mergers and acquisitions could be motivated in part by the need to acquire productive services. Targets were likely to complement or supplement the acquiring firm's existing activities.

Concentration in a growing economy emerged when larger firms as a group grew faster than the smaller firms and therefore the economy as whole (Penrose 1956, [1959] 2009). Larger and older firms had a "competitive advantage" over smaller firms not only in terms of efficient advantages (size, experience, access to funds, etc.) but also because of "monopolistic power" (Penrose 1956: 64). In a growing economy, and given limits to firm growth, large firms were unlikely to take advantage of all opportunities open to them, allowing potentially profitable opportunities for smaller firms. These were the "interstices" of the economy. Limits to the large firms' growth rate and big business competition could lead to a decline in concentration, albeit not the absolute size of large firms.

On competition, Penrose saw a strong case for the big firm and for "big business competition," especially "with respect to the rate of development of new technology and new and improved products" (Penrose [1959] 2009: 229). The "basic dilemma" was that competition induces innovation but while the growth of firms may be efficient, the consequent size might lead to industry structures which impeded growth.

The Multinational Enterprise and the Political Economy of International Trade and Relations

Penrose had a long and enduring interest in the multinational enterprise (MNE), foreign direct investment (FDI), and their political economy. Many of her major publications were on this topic, some preceding TGF (and even Hymer's

[1960] 1976 pioneering contribution; see Penrose 1956, 1985, 1987, 1995, 1996).

Penrose considered the MNE as the natural outcome of the pressures for growth and was initially reluctant to recognize any fundamental differences between MNEs and domestic firms. However, in her 1996 paper she acknowledged that cross-national differences in culture, language, and similar considerations suffice to "justify separate treatment of international firms" (Penrose 1996: 1720).

While Penrose came close to, she did not ultimately develop a fully fledged theory of FDI and the MNE, nor did she claim to have done so. She attributed FDI to efficiency-related advantages of the parent company, which included resources and experience and what she termed the "indefinable advantage" of the internal operations of an ongoing concern (Penrose 1956, 1971). She saw subsidiaries as independent entities able to develop their own advantages, but in doing so, she arguably failed to recognize the possibility of leveraging subsidiary skills and to relate this to the theory of the MNE and FDI. Ironically, Penrose was also one of the first scholars to employ the concept of transfer prices as a major MNE advantage (1968a, 1968b), which left her only one step away from a theory of the FDI based on the internalization of transfer price advantages (Pitelis 2011).

Still, Penrose's knowledge and learning perspective had added cognitive and entrepreneurial elements that are missing from extant theory of the MNE (Pitelis 2007a). These could be highlighted in terms of Dunning's (1988) ownership, location, internationalization (OLI) approach, which is now widely seen as a general framework for the MNE and FDI. For Dunning, the choice of FDI versus alternatives such as licensing is predicated upon the co-existence of ownership (O) advantages, internalization or integration (I) advantages, and locational (L) advantages.

In TGF, O advantages were mostly efficiency advantages because they were the result of an endogenous knowledge/innovation process and only became monopolistic when firms erected barriers to entry or used other restrictive practices.

While TGF did not address L-related advantages, it is consistent with the Penrosean perspective to suggest that firms would locate where existing resources/knowledge can add value to their own resources and operations. This is in line with Dunning's discussion of asset and institution-seeking locational advantages, with more recent attempts to build a theory of the metanational and with work on institutional advantages (Jones and Pitelis 2015).

Penrose did not deal with I advantages in the specific context of the MNE, but she had dealt extensively with integration, which she considered to be an earlier and more accurate term for "internalization." Her argument for horizontal integration was the acquisition of valuable managerial resources "as a means of obtaining the productive services and knowledge that are necessary for a firm to establish itself in a new field," which is "often far more important than the elimination of competition and the reduction of the costs of entry" (Penrose [1959] 2009: 112). Penrose had stated that opportunities arising from the nature of the productive resources of the firm can provide it with an advantage in the production of some of its own requirements. Alongside market opportunities in the case of forward integration, competitive pressures, and special problems arising from uncertainty, all played a role in helping to explain vertical integration (Penrose [1959] 2009: 128).

Applying such ideas to the MNE would suggest FDI which is induced by superior capabilities and is resource/knowledge and institution seeking. This is a more entrepreneurial, forward-looking theory of the MNE than the more static "internalization" and OLI perspectives (Pitelis and Teece 2010; Jones and Pitelis 2015).

Penrose also had a longstanding interest in the relation between MNEs and developing countries. An early advocate of joint ventures as means of reducing local antagonisms and facilitating MNE growth, she also recognized their potential problems, such as the threat to the MNE role and structure (Penrose 1971). She recognized that despite the potential benefits of FDI, the needs of developing states and their governments' agendas differed from those of MNEs. She saw appropriate government policies and institutions, and firm yet

amicable negotiations, as essential for enabling the governments of developing countries to receive the maximum possible net benefits from FDI (Penrose 1968a, 1968b), but also attributed at least some of the MNE costs to developing countries to governmental policies that gave MNEs "excessive protection" (Penrose 1973). Overall, she advocated greater independence for MNEs from specific countries to ease fears that foreign control from more powerful countries could undermine the independence of others (see Pitelis 2011).

Penrose also highlighted the complexity of development, in that the importation of technological skills required the development of institutions and social attitudes to successfully adapt these technologies to a particular culture. She argued that the "third world's" economic development requires qualified leaders and officials, alongside "political leadership," or even "political entrepreneurship" (Klein et al. 2010), and emphasized the role of institutions and policies, now a critical element of the modern theory of development.

Penrose's analyses of transfer pricing, dumping, and protectionism challenged mainstream neoclassical views and were very modern in the context of "new international trade" and related theories (see Penrose 1968a, 1968b, 1990; Pitelis 2009b). She suggested that "restriction on the repatriation of profits under some circumstances may be a useful means of ensuring, for a while, continued foreign investment" (Penrose 1962: 138), and argued that "dumping is endemic in the system, an integral part of the competition among large, diversified, research-based, integrated companies" (Penrose 1990: 185).

Influence

Penrose and Neoclassical Economic Theory

An important focus of "managerial theories of the firm" at around the time of Penrose's writing was on the extent to which managerially run firms could pursue objectives different from short-term profit maximization and on the implications of such behavior for "managerial capitalism." Penrose's role in this context was seen in terms of providing justification for the motivation for growth and the "Penrose effect." Concerning the former, Penrose (1985) admitted that profits and growth could not be treated as "equivalent criteria for the selection of investment programs" (p. 8). Nevertheless, she found:

> the assumption that managers of firms try in general to make as much money as seems practical to be not only the most useful, but in fact the only general assumption from which reasonably general conclusions can be drawn. (p. 12)

The remarkably similar and independent work of Alfred Chandler (1962) supported her theory of growth, while Oliver Williamson's (1981) analysis of the M-form organization supported her view that growing firms "expand their ability to manage growth efficiently, with minimum interference with on-going operation" (Penrose 1985: 11).

Penrose quoted, not disapprovingly, various applications of the "Penrose effect" in various contexts (Pitelis 2009b). While in its evolutionary context, the "Penrose effect" simultaneously described and determined firms' limits to endogenous growth and the receding boundaries of the firm, out of context it could be seen as just another reason why there can be constraints to "optimal growth." That was formalized in models of firm growth, optimal investment, and "optimal growth" (for a summary, see Pitelis 2009b). Such applications of Penrose's work mostly overlooked her key insight: the endogenous, production-side growth advantages associated with the knowledge-creation process through specialization and division of labor, in an evolving, cohesive shell called a firm.

In contrast to the neoclassical theory, Penrose's approach did not involve rational optimizing agents, did not focus on the efficient allocation of scarce resources alone, and it did not look for an equilibrium. In her theory, knowledge was not just "tacit" and hard to transmit but was also not known in advance because of uncertainty and because of being created in the context of an evolutionary process and through the purposeful actions of economic agents, not least firms. In addition, knowledge was not scarce in the conventional sense, in that its use by someone need not necessarily always exclude somebody else from using it. If anything, the exchange of

knowledge can actually help to enhance it. This contrasted with the neoclassical assumption of perfect (even if asymmetric) information, and it may help us to appreciate Penrose's apparent failure to make significant inroads into neoclassical economics (Pitelis 2009b). This, however, contrasted with her influence on the emerging field of business strategy.

Penrose and the Resource, Knowledge, and (Dynamic) Capabilities-Based Perspective

In the 1980s, Penrose became aware of the burgeoning literature drawing on her work, including the then emerging resource-based, competence-based (dynamic), capabilities-based, and/or knowledge-based approaches to strategy. She had also come across some of the early literature, notably Teece's (1982) seminal article, which combined Penrose-inspired resource-based ideas and transaction cost ideas in order to explain the multiproduct firm. In her last (1996) paper on the topic, Penrose recognized the importance of transaction cost issues, which she considered to be one of the "two major types of explanation for the growth of firms in a market economy" (1996: 1717), the second being the Resource-Based View (RBV) she had helped to found. The Penrose-inspired resource, knowledge, and, more recently, dynamic capabilities-based perspective now arguably dominate strategic management and organization science (Pitelis 2009b; Teece 2014).

More recently, the emergent literature on strategic human resource management (SHRM) and (strategic) entrepreneurship employed TGF and the RBV as one of their main pillars, and so did the Dynamic Capabilities (DCs) perspective (for entrepreneurship, see Pe'er et al. 2014; for international entrepreneurship, see Naldi and Davidsson 2013; for SHRM, see Ployhart and Moliterno 2011; for the DCs perspective, see Augier and Teece 2008; Teece 2014). Building on Penrosean ideas, Pitelis and Teece (2009, 2010) revisited the nature and essence of the firm and the MNE in terms of market and business ecosystem creation and co-creation, as opposed to in terms of market failure alone. The (strategic) marketing literature also drew on RBV ideas (Hunt 2011). Resource and capability-based arguments are now central to the explanation of firm-level sustainable advantage (Teece 2007; Pitelis 2009a; Jacobides et al. 2012).

At the more macroeconomic level, independently developed endogenous growth and capabilities-related ideas are now influential in macroeconomic theories of endogenous growth, in economic development and in the sustainable advantage of firms and nations (Pitelis and Teece 2016). Penrose's ideas have also been synthesized with those of the behavioral school (Pitelis 2007b).

Today, TGF remains one of the leading books on the theory of the firm with circa 27,000 citations in Google Scholar.

An Overall Assessment

Penrose's Main ("Single" in her Words) Argument and Some Recent Critiques and Debates

Penrose's major contribution and influence pertain to both her "single" argument and her method-epistemology. The argument was not simply about the theory of the growth of the firm, it was about the theory of (growth of) knowledge. Her argument and method-epistemology involved a dynamic interplay between induction and deduction, structure and agency, in the context of a history-informed path-dependent evolutionary dynamic, shaped by actors' conscious, yet path-dependent and structure-molded actions, in the context of a purposive organization, the firm, operating under conditions of uncertainty (Jones and Pitelis 2015), and it was more suited to the concerns of nonneoclassical economics and strategy scholars.

For Loasby (1999), Penrose's work reinvented the classical resource-creation tradition, while Richardson (1999) claimed that Penrose's contribution arguably provided a tool that could help to deal with both coordination and growth. Penrose's endogenous knowledge perspective went beyond the classical economics contributions both in terms of the learning-induced endogenous growth and in terms of explaining why and how the size

of the market is itself determined by specialization and the division of labor (and vice versa).

Penrose's ideas can arguably make a significant contribution to the issue of the "nature of the firm" too, namely, the question of why firms exist (Coase 1937). Penrose took this for granted. Coase equated the "nature" of the firm with the "employment relation," attributing this to the efficiency benefits derived from reductions in market transaction costs. Penrose's approach can complement this argument. The "employment relation" can be explained in terms of efficiency gains, through productivity enhancements, through endogenous innovation/equals knowledge-growth (see Pitelis and Wahl 1998).

Some early disputes questioned the extent to which Penrose's book was influential in the early resource-based contributions (Barney 1991; Peteraf 1993) and asked whether the RBV was tautological and/or nonoperationalizable. The Barney/Peteraf variant of modern RBV, which drew on Chicago school economists such as Demsetz (1972) and emphasized valuable, rare, inimitable and nonsubstitutable-type resources as intraorganizational barriers to competition, was more about rents in equilibrium. The last part of TGF that addressed the issue of firm size, artificial barriers to entry and the like, and her references to RIBs could also be seen easily in "rents in equilibrium" terms (Pitelis 2004). However, her main emphasis was on the capture of created value through learning and innovation, Hence, for Penrose, firms created but also tried to appropriate value both at quasi-equilibrium (monopoly) and at disequilibrium through building RIBs.

Other criticisms of resource-based ideas come from Michael Porter and Oliver Williamson, whose views still inform major "competing" perspectives to Penrose's on firm (strategy). Both have dealt with resources and competences (e.g., Porter 1999; Williamson 1999). While Porter was largely dismissive of what he saw as vague concepts, Williamson had wider-ranging critiques, the major one being the apparently tautological nature of the perspective and the lack of operationalizability and supporting evidence. Yet Penrose's contribution has not been criticized as tautological (Pitelis 2009b). The predictive and prescriptive aspects of her theory are both operationalizable and testable, as evidenced by an extensive list of empirical studies that support Penrosean views (see Nason and Wiklund 2015).

Further Limitations and Generalizations

A major limitation of TGF, shared by both Porterian and Williamsonian perspectives on strategy, relates to intraorganizational conflict (Cyert and March 1963), in particular, whenever there exists a "principal-agent" relationship (see also Klein et al. 2012). An integration of behavioral and Penrose-inspired RBV (Pitelis 2007b) would suggest that intrafirm knowledge generation could engender endogenous innovation and growth through the generation and leverage of "excess resources" and "slack." "Slack" may alleviate conflict but also supply motivational and psychological reasons for obedience. Given resource availability, it may enable endogenous growth and innovation. Intrafirm knowledge generation can inform management as to why, whether, and how to leverage excess resources to alleviate conflict, breed success, and engender a virtuous cycle of endogenous growth and innovation.

Concluding Remarks and Implications for Managerial Practice and Public Policy

By weaving endogenous knowledge, innovation and growth, human resources, the role of "image" and "productive opportunity," and the dynamic interaction between internal and external, agency and structure, Penrose went beyond the existing economic theories to provide what is considered to be the first economics-based, yet interdisciplinary, organizational theory of the firm that helps to bridge economic and organizational theories of firms, organizations, institutions, and strategy.

While Penrose advocated no managerial practice per se, she felt that rents, of both the monopoly type and through the building of RIBs, were important for the long-term successful expansion of firms and their sustainable competitive advantage (Pitelis 2004, 2009a). The building of RIBs, however, is itself predicated on the successful redeployment of resources, competences, and other

advantages in a dynamic, changing, uncertain environment where the key to long-term success is to build technology bases through perennial innovations, knowledge creation, and by internalizing the Schumpeterian process of "creative destruction," which for Penrose forced large firms to become more and more "creative" ([1959] 2009: 94). The message for management practice is to focus on firm and market creation and co-creation through appropriable and perennial innovation (Pitelis and Teece 2010; Pitelis 2012).

However, Penrose felt that good managerial practice might not suffice and advocated suitable anti-trust policies by government, aimed at marrying firm-level sustainable advantage with nationwide sustainable advantage, including the support of innovation and value creation through (big business) competition, small firm creation, and growth (latterly in networks). She also saw an important role for government in developing countries. Such aspects of Penrose's work are becoming increasingly relevant and could become especially influential in the field of sustainable development.

See Also

- ▶ Dynamic Capabilities
- ▶ International Business
- ▶ Multinational Corporations
- ▶ Resource-Based Theories
- ▶ Resource-Based View
- ▶ Teece, David J. (Born 1948)

Selected Works

1955. Research on the business firms: limits to growth and size of firms. *American Economic Review* 45(2): 531–543.

1956. Foreign investment and the growth of the firm. *Economic Journal* 66: 220–235.

1960. The growth of the firm: A case study: The Hercules Powder Company. *Business History Review* 34: 1–23.

1962. Some problems of policy towards direct private foreign investment in developing countries. Middle East Economic Papers Lebanon: American Research Bureau/American University of Beirut.

1968a. *The large international firm in developing countries*: *The international petroleum industry*. London: Allen & Unwin.

1968b. Problems associated with the growth of international firms. *Tijdschriftvoor Vennootschappen*, *Vereinigingen en Stichtingen* 9. In (1971) *The growth of firms*, *Middle East oil and other essays*, ed. E. T. Penrose. London: Frank Cass & Co., Ltd.

1971. *The growth of firms*: *Middle East oil*, *and other essays*. London: Cass.

1973. The changing role of multinational corporations in developing countries. Paper submitted to the United Nations Groups of Eminent Persons to Study the Impact of Multinational Corporations on Development and on International Relations, Geneva.

1985. The theory of the growth of the firm twenty-five years later. *Acta Universitatis Upsaliensis*: *Studia Oeconomicae Negotiorum* (*Uppsala Lectures in Business*) 20: 1–16.

1987. Multinational corporations. In *The New Palgrave*: *A Dictionary of Economics*. London: Macmillan.

1990. Dumping, 'unfair' competition and multinational corporations. *Japan and the World Economy* 1:181–187.

1995. *The theory of the growth of the firm,* 3rd ed. Oxford: Oxford University Press.

1996. Growth of the firm and networking. In *International encyclopaedia of business and management*. London: Routledge.

2008. Strategy/organization and the metamorphosis of the large firm. *Organization Studies* 29: 1117–1124.

[1959] 2009. *The theory of the growth of the firm,* 4th ed. Oxford: Oxford University Press.

References

Augier, M., and D.J. Teece. 2008. Strategy as evolution with design: Dynamic capabilities and the design and evolution of the business enterprise. *Organization Studies* 29: 1187–1208.

Barney, J. 1991. Firm resources and sustained competitive advantage. *Journal of Management* 17: 99–120.
Chandler, A.D. 1962. *Strategy and structure: Chapters in the history of the American industrial enterprise*. Cambridge, MA: The MIT Press.
Coase, R.H. 1937. The nature of the firm. *Economica* 4: 386–405.
Cyert, R.M., and J.G. March. 1963. *A behavioral theory of the firm*. Englewood Cliffs: Prentice Hall.
Demsetz, H. 1972. When does the rule of liability matter? *Journal of Legal Studies* 1: 13–28.
Dunning, J.H. 1988. The eclectic paradigm of international production: A restatement and some possible extensions. *Journal of International Business Studies* 19: 1–31.
Hunt, S.D. 2011. The theory of monopolistic competition, marketing's intellectual history, and the product differentiation versus market segmentation controversy. *Journal of Macromarketing* 31: 73–84.
Hymer, S. H. [1960] 1976. *The international operations of national firms: A study of foreign direct investment*. Cambridge, MA: The MIT Press.
Jacobides, M.G., S.G. Winter, and S.M. Kassberger. 2012. The dynamics of wealth, profit, and sustainable advantage. *Strategic Management Journal* 33: 1384–1410.
Jones, G., and C. Pitelis. 2015. Entrepreneurial imagination and a demand and supply-side perspective on the mne and cross-border organization. *Journal of International Management* 21: 309–321.
Klein, P.G., J.T. Mahoney, A.M. McGahan, and C.N. Pitelis. 2010. Toward a theory of public entrepreneurship. *European Management Review* 7: 1–15.
Klein, P.G., J.T. Mahoney, A.M. McGahan, and C.N. Pitelis. 2012. Who is in charge? A property rights perspective on stakeholder governance. *Strategic Organization* 10: 304–315.
Loasby, B.J. 1999. The significance of Penrose's theory for the development of economics. *Contributions to Political Economy* 19: 31–46.
Naldi, L., and P. Davidsson. 2013. Entrepreneurial growth: The role of international knowledge acquisition as moderated by firm age. *Journal of Business Venturing* 29: 687–703.
Nason, R.S., and J. Wiklund. 2015. An assessment of resource-based theorizing on firm growth and suggestions for the future. *Journal of Management*, forthcoming, https://doi.org/10.1177/0149206315610635.
Pe'er, A., I. Vertinsky, and T. Keil. 2014. Growth and survival: The moderating effects of local agglomeration and local market structure. *Strategic Management Journal* 37: 541–564.
Penrose, P., and C.N. Pitelis. 1999. Edith Elura Tilton Penrose: Life, contribution and influence. *Contributions to Political Economy* 18: 3–22.
Peteraf, M.A. 1993. The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal* 14: 179–191.
Pitelis, C.N. 2004. Edith Penrose and the resource-based view of (international) business strategy. *International Business Review* 13: 523–532.
Pitelis, C.N. 2007a. Edith Penrose and a learning-based perspective on the MNE and OLI. *Management International Review* 47: 207–219.
Pitelis, C.N. 2007b. A behavioral resource-based view of the firm: The synergy of Cyert and March (1963) and Penrose (1959). *Organization Science* 18: 478–490.
Pitelis, C.N. 2009a. The co-evolution of organizational value capture, value creation, and sustainable advantage. *Organization Studies* 30: 1115–1139.
Pitelis, C.N. 2009b. Introduction: Edith Penrose's 'the theory of the growth of the firm' fifty years later. In *The theory of the growth of the firm*, 4th ed, ed. E.T. Penrose. Oxford: Oxford University Press.
Pitelis, C.N. 2011. Globalization, development, and history in the work of Edith Penrose. *Business History Review* 85: 65–84.
Pitelis, C. 2012. Clusters, entrepreneurial ecosystem co-creation, and appropriability: A conceptual framework. *Industrial and Corporate Change* 21: 1359–1388.
Pitelis, C.N., and D.J. Teece. 2009. The (new) nature and essence of the firm. *European Management Review* 6: 5–15.
Pitelis, C.N., and D.J. Teece. 2010. Cross-border market co-creation, dynamic capabilities and the entrepreneurial theory of the multinational enterprise. *Industrial and Corporate Change* 19: 1247–1270.
Pitelis, C.N., and D.J. Teece. 2016. The strategic SCA of nations: A capabilities-based industrial strategy for social value co-creation. Available at: http://ssrn.com/abstract=2749110. Accessed 10 Aug 2016.
Pitelis, C.N., and M. Wahl. 1998. Edith Penrose: Pioneer of stakeholder theory. *Long Range Planning* 31: 252–261.
Ployhart, R.E., and T.P. Moliterno. 2011. Emergence of the human capital resource: A multilevel model. *Academy of Management Review* 36: 127–150.
Porter, M.E. 1999. Michael Porter on competition. *Antitrust Bulletin* 44: 841–880.
Richardson, G.B. 1999. Mrs Penrose and neoclassical theory. *Contributions to Political Economy* 18: 23–30.
Teece, D.J. 1982. Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior & Organization* 3: 39–63.
Teece, D.J. 2007. Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal* 28: 1319–1350.
Teece, D.J. 2014. The foundations of enterprise performance: Dynamic and ordinary capabilities in an (economic) theory of firms. *Academy of Management Perspectives* 28: 353–367.
Williamson, O.E. 1981. The modern corporation: Origins, evolution, attributes. *Journal of Economic Literature* 19: 1537–1569.
Williamson, O.E. 1999. Strategy research: Governance and competence perspectives. *Strategic Management Journal* 20: 1087–1108.

Perfect Competition

Christopher Pleatsikas
Berkeley Research Group, LLC, Emeryville, CA, USA

Abstract

Perfect competition describes one of the two endpoints of the continuum used to categorize market conditions. As such, it is the antithesis of the other continuum endpoint – monopoly. Whereas monopoly refers to the circumstance where there is just one seller of a product in a market perfect competition refers to a circumstance where there are an infinite number of sellers in a market and ▶ competition is unrestrained except by market forces. The concept of perfect competition is often traced to the ideas articulated in the eighteenth century by Adam Smith of 'natural price' and the reduction in price that occurs as the number of sellers increases. The concept was refined, ultimately through mathematical formulation, resulting, by the 1950s, in the concept as we understand it today. Perfect competition can best be perceived as a benchmark used to illustrate other economic concepts. No real-world market can satisfy all its requirements.

Definition Perfect competition is a theoretical benchmark concept in economics that results in the achievement, in the long run, of maximum efficiency, and is used as the basis against which to measure market performance for other theoretical and real-world market structures and other economic concepts.

Perfect competition is a term used in economics to describe one of the two endpoints of the continuum used to categorize market conditions. As such, it is the antithesis of the other continuum endpoint – monopoly. Whereas monopoly, in the sense the term is used in economic theory (as opposed to its use in antitrust analysis), refers to the circumstance where there is just one seller of a product in a market (i.e., it describes a circumstance in which there is the absence of ▶ competition), perfect competition refers to a circumstance where there are an infinite number of sellers in a market, and competition is unrestrained except by market forces.

The concept of perfect competition is often traced to the twin ideas articulated in the eighteenth century by Adam Smith of 'natural price' and the reduction in price that occurs as the number of sellers increases. It was further developed through the 'rule of unlimited competition' set forth by Cournot in the nineteenth century (Stigler 1957). The concept was yet further refined and developed through logical analysis and, ultimately, through mathematical formulation, resulting, by the 1950s, in the concept as we understand it today (e.g., see Arrow 1959; Debreu 1959; McKenzie 2002; Weintraub 2002).

In mathematical terms, in a perfectly competitive market price is set at ▶ marginal cost. The marginal cost is just equivalent to the opportunity costs of making a good. The point at which demand – the salient characteristics of demand being captured by a 'demand curve' (often illustrated by a straight line in economic texts) which represents the marginal benefit to society of a product – and supply – again a curve, representing the marginal cost to society of producing a good – intersect defines both the price at which the product is sold and the quantity supplied at that price. In the long run, at equilibrium in a perfectly competitive market, marginal revenue equals average revenue, which is the market price. This equilibrium is stable (unless perturbed by some exogenous event), in the sense that no producer has any incentive to produce more or less and no buyer is willing to purchase more or less.

While some economists (and many non-economists) contend that few perfectly competitive markets exist in the real world – the wheat market (in terms of producers, i.e., farmers) is sometimes cited as one of the few examples of a perfectly competitive real-world market – strictly speaking, as the concept is understood in economic theory, there are no real-world markets that satisfy all the manifold, very restrictive conditions required for a market to be perfectly

competitive (e.g., Samuelson 1965; Pindyck and Rubinfeld 2004). These conditions include:

- Infinite number of sellers, each willing to supply the product at a certain price.
- Infinite number of buyers, each willing to buy the product at a certain price.
- Price-taking, in the sense that no buyer and no seller (and no feasible combination of buyers or sellers) is able to influence price. Each buyer and each seller takes the price as given. The implication of this fact is that any seller who attempts to raise price, even by a very small amount, above the 'competitive market' price will lose all sales, and any buyer who attempts to secure product at a price that undercuts, even by a small amount, the 'competitive market' price will find no seller willing to provide any product to her/him.
- No entry and exit barriers. This implies that firms incur no non-recoverable costs if they enter and none if they exit (while the definitions of entry barriers and exit barriers are controversial among economists, they generally agree that the lack of non-recoverable entry and entry costs are consistent with the lack of entry or exit barriers). Furthermore, entry and exit into perfectly competitive markets is assumed to be instantaneous.
- Homogenous, perfectly divisible outputs. All firms sell identical products and buyers perceive the products of any one producer to be perfect substitutes for those of any other producer. Perfect divisibility implies that output is continuously variable and that any output level is feasible.
- No transaction costs. Transaction costs are assumed to be zero both on the production side and the buyer side. Therefore all factors of production are perfectly mobile (and can thus be reallocated in response to changes in demand) instantaneously and without cost, and buyers incur no costs to purchase products.
- Perfect information. Both buyers and sellers possess all relevant information and perfect foresight. No one has any informational advantage.
- Constant returns to scale in production and no technological advantage. Any technological progress is immediately propagated throughout the market.
- Profit maximization. Firms are assumed to sell at the point at which marginal cost equals marginal revenue. In long-run equilibrium, marginal cost would equal average cost. So each firm would just cover its costs (a condition that economists refer to as 'zero economic profits', or profits just sufficient to cover all variable costs, and provide a return to capital just sufficient to cover the opportunity cost of capital) (Varian 2005).
- No ▶ externality. Each firm bears all the costs of its production and imposes no uncompensated costs on others.

No real-world market can satisfy all these requirements. Perfect competition can be analogized to a hypothesized frictionless surface used to illustrate certain physics concepts. As such, perfect competition can best be perceived as a pedagogical tool or benchmark used by economists to illustrate other economic concepts. For example, a perfectly competitive market, in contrast to most real-world markets, is in equilibrium in the long run. It is also both productively and allocatively efficient – that is, it results in production at least cost (productively efficient), and production occurs at the point where the marginal benefit to society is equal to the marginal cost of production (allocatively efficient).

More to the point, because a perfectly competitive market in equilibrium by definition maximizes social welfare, it is the means by which the performance of other, more realistic, ▶ market structure can be defined and measured. In other words, the performance of other market models can be defined in terms of the deviations from social welfare or prices or cost structures that characterize them in comparison to the perfectly competitive market model. These alternative market models are collectively termed 'imperfectly competitive'. They can violate any one (or several or all) of the assumptions that underpin the perfectly competitive market model, but generally (although not always), imperfectly competitive markets are characterized by relatively few sellers, non-standardized, differentiated products,

barriers to entry and imperfect information available to buyers, sellers or, more often, both.

The deviations of alternative market models from the perfectly competitive benchmark may be best represented by the concept of economic negligibility, which is central to the notion of perfect competition (Aumann 1964). Economic negligibility implies that no agent within the economic system – either on the selling or the buying side – can affect outcomes, that is, prices or quantities. Stated differently, economic negligibility implies that no participant in a perfectly competitive market has any degree of market power. In the other market models developed by economists, economic negligibility is discarded and agents can affect outcomes. Of course, in real-world markets, firms continuously vie for competitive advantage against their actual and potential rivals and strive to earn above-competitive rates of return on their investments. The prospect of above-competitive returns, which can often be achieved by at least some firms in the real world, motivates entrepreneurs and managers and energizes market competition.

The most obvious foil to perfect competition is the classical monopoly, whose most salient characteristic is the single seller (or, in the case of the monopsonist, the single buyer) that can extract positive economic profits (e.g., returns in excess of the opportunity cost of capital) because it faces no competition. That is, the monopolist can choose its price (subject only to the specific characteristics of the demand curve, but not to any competitive constraint), its output and its profit level. Unlike the economically negligible participant in a perfectly competitive market, the monopolist exercises substantial market power – the power to price without regard to competitive constraint.

Other imperfectly competitive economic models, such as oligopoly (few sellers) and monopolistic competition (multiple sellers of differentiated products), vary the amount and/or duration of market power available to agents. Consequently, agents in such markets are described as having some degree of market power – that is, some degree of control over price. Practically, this means that, unlike agents in perfectly competitive markets, agents in imperfectly competitive markets can increase price without necessarily losing all their customers. In what are described as 'oligopolistic markets', there are few sellers of identical or differentiated products. For example, in oligopolistic markets firms are generally aware of their influence over price, are cognizant that their pricing and output decisions are interdependent with the corresponding decisions of other firms, and can often earn rates of return that exceed opportunity costs. In 'monopolistically competitive' markets, competitive firms sell differentiated products that are viewed as only imperfectly substitutable for the products of other firms. Imperfect substitution gives each firm some degree of market power and allows them to charge prices exceeding marginal costs, at least in the short run (the degree of long-run power being related, at least in part, to entry conditions). In mathematical terms, firms have some discretion over price and/or product quality because they face downward-sloping demand curves.

Many other economic concepts, too, are usefully analysed as deviations from the conditions that are required in order for a market to be perfectly competitive and, viewed through that lens, illustrate the implications of such deviations on market performance and social welfare. Just some of the more prominent examples include:

- Externalities, which are uncompensated costs or benefits that economic entities impose or confer on other economic entities. The existence of externalities makes perfect competition impossible because prices no longer represent social costs. As a result, the existence of externalities, when they have significant effects, requires some type of market intervention, such as government regulation (e.g., environmental regulation to reduce pollution or internalize its costs to the entity that causes it).
- Public goods are products that, if supplied to one person, are available to others at no additional cost. An example is national defence. In contrast with private goods, for which consumption of a unit by one party precludes consumption of that same unit by another party, public goods cannot be supplied by private markets, even perfectly competitive ones,

because a supplier cannot confine consumption of the good to those who pay for it. As a consequence, no individual supplier would provide such a good (i.e., because it could not obtain adequate compensation or, in the extreme, any compensation for it through private transactions). Since collective action is required to supply a public good, supply violates one of the fundamental assumptions associated with perfectly competitive markets, which posits economic negligibility for any entity or group of entities (Pearce 1992).

- Entry and exit costs. Whatever one's definition of entry and exit barriers, non-recoverable costs associated with these actions introduce friction into the market and are both common in the real world and incompatible with perfectly competitive markets. They reduce the potential for arbitrage, the 'lubricant' that facilitates adjustment to equilibrium (and which is assumed to occur instantaneously in perfectly competitive markets).

For each and every assumption that underpins the economic model of perfect competition, a similar analysis of deviations from the model's requirements could be developed. As one deviates further from the idealized model, more realistic – in the sense of describing real-world market conditions – market characteristics emerge, and more complex and more nuanced market behaviour (and its price and non-price implications) can be analysed. This encapsulates perfect competition's real-world relevance – as a theoretical benchmark for assessing social welfare implications (Carlton and Perloff 2005).

Finally, it should be noted that, because perfect competition is a pedagogical economic tool and not a descriptor of real-world markets, it is not the benchmark used in antitrust analysis to determine whether conduct is anti-competitive or whether a merger would substantially lessen competition (e.g., see Pleatsikas and Teece 2001). The proper economic benchmark for gauging firm behaviour in an antitrust context is a workably competitive market. In a workably competitive market, some (or even all) market participants may have some market power (i.e., some discretion over price), but no market participant has a substantial degree of market power (which, as defined by economists, indicates an entity that has no competitive constraint on its ability to price, or for whom competitive constraints are relatively unimportant). In a workably competitive market, at any specific point in time, prices can deviate from underlying costs and the deployed technologies can deviate from the most efficient ones currently available. However, in such markets, economic forces drive the market, albeit not instantly, towards efficient prices, outputs and costs.

See Also

- ▶ Competition
- ▶ Contestability
- ▶ Externality
- ▶ Marginal Cost
- ▶ Market Structure

References

Arrow, K.J. 1959. Toward a theory of price adjustment. In *The allocation of economic resources*, ed. M. Abramovitz. Stanford: Stanford University Press.

Aumann, R.J. 1964. Markets with a continuum of traders. *Econometrica* 32: 39–50.

Carlton, D.W., and J.M. Perloff. 2005. *Modern industrial organization*, 4th ed. Boston: Pearson Addison-Wesley.

Debreu, G. 1959. *The theory of value*. New York: Wiley.

McKenzie, L.W. 2002. *Classical general equilibrium theory*. Cambridge, MA: The MIT Press.

Pearce, D.W. (ed.). 1992. *The MIT dictionary of economics*, 4th ed. Cambridge, MA: The MIT Press.

Pindyck, R.S., and D.L. Rubinfeld. 2004. *Microeconomics*, 6th ed. Upper Saddle River: Prentice Hall.

Pleatsikas, C., and D. Teece. 2001. Economic fallacies encountered in the law and economics of antitrust: Illustrations from Australia and New Zealand. *Trade Practices Law Journal* 9: 73–94.

Samuelson, P.A. 1965. *Foundations of economic analysis*, 2nd ed. Cambridge, MA: Harvard University Press.

Stigler, G. 1957. Perfect competition, historically contemplated. *Journal of Political Economy* 65: 1–17.

Varian, H.R. 2005. *Intermediate microeconomics: A modern approach*, 7th ed. New York: W.W. Norton.

Weintraub, R. 2002. *How economics became a mathematical science*. Durham: Duke University Press.

Performance Measures

Anita M. McGahan
University of Toronto, Rotman School of Management, Toronto, ON, Canada

Abstract

Organizational performance is at the heart of the field of strategy and ▶ innovation management. Indeed, the field of strategy is frequently defined as the study of organizational performance (Porter 1980, 1985; Barney 2011). This entry begins by describing the challenges, opportunities and trade-offs associated with various measures of corporate performance, and then turns to new measures of the innovative performance of organizations. The core of the article is a discussion of the increasing interest in the field in expanded measures of organizational performance along several dimensions. The entry concludes by pointing to several promising future directions for research on this topic.

Definition In the field of strategy and innovation management, "performance measures" are constructs that describe organizational output. Classical examples of performance measures are profitability, financial value and patent rates. While performance may also be measured at the individual level, only organizational performance is considered in this entry.

The Purpose of the Publicly Traded Firm: Profitability

The prevailing view in the field of management is that the purpose of the corporation is to generate a *return on invested capital* for equity investors as residual claimants on the firm's profitability (Jensen and Meckling 1976). This view, inherited mainly from the discipline of economics, derives from theories that emphasize the role of the firm as a vehicle for pooling the resources of investors under limited liability and is controversial (Stout 2012). *Profitability* under this conceptualization is often measured as contemporaneous *return on equity*. The operationalization of this measure was advanced through the widespread adoption, especially in the 1970s and 1980s, of *the DuPont equation*, which partitions return on equity into components associated with profit margins, asset turnover and financial leverage.

The classical approach typically addresses the dynamics of performance by considering that the firm's responsibility to its residual claimants – namely shareholders – is to maximize not only contemporaneous performance but future performance. This objective is put to practice through the maximization of the net present ▶ value of profitability over time. Indeed, many first-year finance courses in business schools continue to train students through Excel modelling on how to project profits into the future and then calculate their present value using discount rates that reflect the cost of capital, which itself is driven in part by assessments of operating risk.

The advantage of this approach is its emphasis on the intertemporal character of the components of profitability, but a disadvantage is the unitary nature of the discount rate used to calculate present values, since this rate does not change to reflect the actualized changes in risk that might arise in the future as various uncertainties are resolved and unanticipated events arise. A second disadvantage is that the approach does not account for the interrelationships between discount rates and profitability. These two problems have been addressed in part by the development of real-options theory and analysis, which are techniques for modelling the future resolution of uncertainty. Yet a third disadvantage of the classical approach, even accounting for the advances of real-options theory, is a fundamental critique regarding the absence of an accounting for the fact that a company might not survive certain types of outcomes. As a result, interest in the field has mounted since the early 1990s on the trade-offs between profit maximization and survival prospects (discussed in the next section). A fourth disadvantage is that the classical approach to the firm as a profit-maximizing entity

does not deal directly with the prospects of ► innovation, and of the challenges and opportunities tied to innovation (discussed in the subsequent section). Finally, the classical approach to the firm is at odds with analyses that demonstrate that the legal basis for the corporation awards equity investors residual claims on corporate profits only in bankruptcy (Davis 2009; Klein et al. 2012; Stout 2012).

The classical approach to profitability has retained its power, despite these disadvantages, in part because of the utility of stock prices as an indicator of investor expectations about the future. In theory, *stock prices* are driven by investors' fundamental assessments of a firm's prospects for future profitability discounted to reflect expectations about risks of all types (including discontinuities in performance). Tobin's Q is a measure of prospective return that deals with some of the problems of calculating a discount rate by taking the total valuation placed on the firm by both equity and bond claimants and then dividing the total by the salvage value of the firm's assets (often proxied by the book value of assets). *Economic value added*, a concept which became popular in the mid-1990s, takes this concept a step further by examining how cash flows are related to investments. However, each of these concepts rests on the abilities of residual claimants (i.e., stockholders and/or bondholders) to attach a value to their claims that reflects the firm's actual, fundamental prospects for profitability. The dot-com crash of 2001 called into question the accuracy of these models as the values of many corporate securities were priced speculatively rather than fundamentally in a bout of "irrational exuberance", to use a phrase coined by then Federal Reserve Chairman Alan Greenspan. The stock market crash of 2007 and 2008 further exacerbated the problems of these measures by highlighting how the complexities of derived instruments (i.e., "derivatives") pervasively shaped valuations even on simple securities such as corporate equities. These failures of accurate valuation have led to burgeoning interest in correlated risk, a topic explored at the end of this entry as a direction for future research.

Trade-Offs Between Profitability and Survival

During the 1980s and 1990s, the field of strategy and innovation management engaged scholars of organizational and population ecology (Hannan and Freeman 1989) in a new type of discussion regarding the purpose of the firms. Scholars from the ecology tradition examined firms as analogous to organisms, and sought to explain their long-term prospects for *survival*. Under this view, profitability was suppressed as a temporary artefact of short-term competitive dynamics rather than long-term prospects. Concepts such as the "Red Queen effect" (Barnett 2008) emerging from this tradition demonstrated that companies locked into dense competitive fields could become stuck with poor performance in efforts to survive the selection pressure created by the density. A critical element of this perspective was the idea that firms could not easily adjust their positions in the field to avoid competition.

At about the same time, a concept called "*sustainability*" was developing in the classical side of the field, informed by the traditions of economics, and especially by the subfield of industrial organization (Ghemawat 1986, 1991). This concept emphasized that the capacity of a firm to generate a stream of profits over time depended on a number of factors such as preferential access to important raw materials, scale economies and relationship-based lock-in. Because these constructs were difficult to tie systematically and operationally to projections of financial flows, scholars tended to model sustainability by examining whether and how profitability itself changed over time (Mueller 1986; Ghemawat 1991; McGahan and Porter 1999). During the mid-1990s, important advances from the field of accounting on the ► balanced scorecard (Kaplan and Norton 1992) were imported into strategy and innovation management. The idea of the balanced scorecard involved examining interrelated and complementary indicators of profitability to overcome the distortions created by high-powered incentives on singular performance metrics. This idea represented an early advancement of

behavioural reasoning into the strategy and innovation management literature.

By the mid-1990s, scholars from the population ecology and economics traditions became interested in relationships between profitability and survival, and particularly the microstructure of the trade-offs between survival odds and contemporaneous performance. Insights developed initially by Abernathy and Utterback (1978) and others were further advanced by Bower and Christensen (1995), Myers and Rosenbloom (1996), Christensen (1997) and Tripsas (1997). S-curves, the "innovator's dilemma" and industry-innovation cycles became central to the discourse on performance. Today, this literature has been extended to generate new insights about spin-offs and entrepreneurial performance – topics discussed below.

Innovative Performance

By the 1990s, scholars at the core of the field turned in earnest to direct measurements of the innovation performance of organizations. Contingent theorizing on the connections between contemporaneous performance and survival odds raised a series of questions regarding mechanisms of *growth* made especially relevant by the dot-com boom. At the same time, a general decline in corporate research as well as newly available evidence on the challenges of commercialization stimulated interest in systems of innovation (Nelson 1993; Furman et al. 2002), markets for technologies (Arora et al. 2001; Gans and Stern 2003), adaptive organizational capabilities (Cohen and Levinthal 1990; Teece et al. 1997; Gittelman and Kogut 2003) and the appropriability of returns from innovation (Cockburn and Griliches 1989). The measures of innovation performance employed in each of these lines of research were diverse, and yet several central approaches were salient.

First, the NBER *patent* dataset developed by Hall et al. (2000) allowed scholars to score the importance of patented innovations by identifying their subsequent (i.e., "forward") *citations*. Several influential studies established a positive link between innovative performance and the financial value of corporations using citation-weighted patents (Cockburn and Griliches 1989; Jaffe et al. 1993). Over the course of the subsequent decade, a large literature emerged – with many studies authored by researchers associated with the NBER itself – using citation-weighted patents as a measure of innovative performance.

Second, a related and complementary literature on *learning* capabilities (sometimes called dynamic, adaptive and/or absorptive capabilities) emerged. One approach to measuring learning involved examining the pattern of *backward citations* on patents, although the relevance of such citations to the mechanism of learning was also called into question by researchers, who noted that citations were frequently added to patents by examiners rather than scientists (Alcacer and Gittelman 2006). Many other approaches to the measure of learning were adopted, including *survey* results and industry-specific idiosyncratic indicators such as scholarly ties among members of scientific teams. One of the most important surveys influencing the selection of such measures was conducted by a team of senior researchers in 1987 (Levin et al. 1987) and subsequently reissued 10 years later (Cohen et al. 1997).

Third, new datasets on *partnerships* (including, for example, joint ventures, alliances and contracts) were deployed to measure innovative performance as transaction cost theorizing suggested performance metrics associated with the efficiency of transactions as a singular and legitimate unit of analysis. Because partnering capabilities were identified and identifiable through empirical analysis, and because relationship-based metrics of performance were theoretically robust, a large literature associated effective partnering with innovative performance.

Emerging Areas of Enquiry

As ideas regarding contemporaneous and intertemporal organizational performance developed, scholars turned their sights to a number of additional fruitful areas for generating insights into performance. Three of these are outlined here.

First, interest has increased in expanding the role of performance measures to reflect processes related to the emergence of new mechanisms of value creation, and especially entrepreneurial companies, non-profits and governmental organizations. The performance of these organizations is difficult to assess because they generally do not generate profits. Thus, alternative measures of performance are required to compare the relative success of each type of organization. For entrepreneurial firms, relevant measures are typically identified as *time-to-IPO, value-at-IPO*, and the *background and network ties* of the founders, financiers and executives. For non-profits and governmental organizations, the challenge is even greater because of the absence of consistency in objectives, which makes *value creation* even more difficult to quantify. New research in this area points to *growth*, *legitimacy* and *coalitional support* as potentially valuable cross-sectional measures of performance.

Second, widespread critiques of management and of business schools have led to interest in expanding the role of performance metrics even of large, publicly traded corporations. A movement that originated to promote the idea of ▶ corporate social responsibility, which originally emphasized the importance of corporate charity, community engagement and environmental responsibility, yielded several metrics for performance, including the *triple bottom line* on environmental, social and corporate performance. These ideas have developed in a number of directions. One major area of current research involves examining *corporate social investment*, a concept which emphasizes that social practices generate returns for investors. A new metric called *shared value* has been introduced recently to the literature (Porter and Kramer 2011) for analysing the boundary conditions on which such a return is based. A second area of current research on this topic involves *stakeholder analysis*, which seeks to identify the conditions under which stakeholders other than investors may make a claim on the returns that corporations generate (Freeman 1984; Donaldson and Preston 1995). This area of work is currently becoming integrated with *property rights* theories in a synthesis that identifies how a corporation's investments are co-created in tandem with other actors in its ecosystem (Klein et al. 2012).

Third, a complementary area of enquiry examines the performance of the broader systems in which corporations are embedded. Bhutan's adoption of *happiness* as the national goal has incited new conversations in the field about the nature and purpose of commercial activity, including investments in innovation. Scholars with training in the fields of organization studies and sociology have advanced knowledge about corporate systems by examining *network ties*, *relational capital* (such as *trust*) and business *clusters* as important units of analysis for understanding performance.

Future Directions

Several directions in the field of organizational performance are evident as the research on this topic evolves. Five in particular stand out as particularly promising. First, a renewed and liberalized focus on *productivity* is emerging. While the concept has deep roots in the field, the construct has typically been applied to the analysis of corporate and country performance. Scholars from a range of backgrounds are now applying this idea to non-profit organizations, governmental bodies, networks and clusters.

Second, the field is now turning to *normative* studies of performance that consider the central role of *advocacy* and *rights*. At the heart of this critique is the claim that the purpose of the corporation is the generation of societal prosperity rather than wealth for equity holders (Davis 2009; Klein et al. 2012; Stout 2012). Some work in this area is emerging from the study of stakeholders and property rights, while more is emerging from organizational studies of *social movements* (Rao et al. 2000; Mair and Marti 2006).

Third, the financial market crisis of 2007 and 2008 has demonstrated the significance of *correlated risk* to the outcomes of organizational activity. While risk analysis has long been the province of scholars of finance and operations

management, it has not taken centre stage historically in strategy and innovation management. Scholars are now seeking to understand how organizational performance must be calibrated by an assessment of the risks in the investments undertaken to generate the performance.

Four, *behavioural analysis* is a descriptive technique that motivates and demands new theory. This type of analysis has been influential in finance and economics, and also carries promise for the field of strategy and innovation management. The fundamental approach involves examining how actors behave under various conditions and generate precise insights that reflect psychological biases – thus pointing to fruitful directions for further theorizing.

Finally, *multi-level analysis* of the contemporaneous performance of individuals, organizations, communities, networks and countries yields insights into the allocative and distributional consequences of performance improvements at any single level. This area of enquiry sheds light on whether and how organizations create value in a macroeconomic sense – and particularly whether corporate performance arises in tandem with improvements or at the detriment of the performance at other levels in society.

In sum, performance measurement is a robust and vibrant area of research in the field of strategy and innovation management.

See Also

- ▶ Balanced Scorecard
- ▶ Corporate Social Responsibility
- ▶ Innovation
- ▶ Sustainable Competitive Advantage
- ▶ Value

References

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

Alcacer, J., and M. Gittelman. 2006. Patent citations as a measure of knowledge flows: The influence of examiner citations. *Review of Economics and Statistics* 88: 774–779.

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

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

Barney, J.B. 2011. *Gaining and sustaining competitive advantage*. New York: Prentice Hall.

Bower, J.L., and C.M. Christensen. 1995. Disruptive technologies: Catching the wave. *Harvard Business Review* 73: 45–53.

Christensen, C.M. 1997. *The innovator's dilemma: When new technologies cause great firms to fail*. Boston: Harvard Business School Press.

Cockburn, I., and Z. Griliches. 1989. Industry effects and appropriability measures in the stock market's valuation of R&D and patents. *American Economic Review* 78: 419–423.

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

Cohen, W., R.R Nelson, and J. Walsh. 1997. *Appropriability conditions and why firms patent and why they do not in the U.S. manufacturing sector*. Working paper, Carnegie Mellon University.

Davis, G. 2009. *Managed by the markets: How finance reshaped America*. Oxford: Oxford University Press.

Donaldson, T., and L.E. Preston. 1995. The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review* 20: 65–91.

Freeman, R.E. 1984. *Strategic management: A stakeholder approach*. Boston: Pitman.

Furman, J.L., M.E. Porter, and S. Stern. 2002. The determinants of national innovative capacity. *Research Policy* 31: 899–933.

Gans, J.S., and S. Stern. 2003. The product market and the market for "ideas": Commercialization strategies for technology entrepreneurs. *Research Policy* 23: 333–350.

Ghemawat, P. 1986. Sustainable advantage. *Harvard Business Review* 64: 53–58.

Ghemawat, P. 1991. *Commitment: The dynamic of strategy*. New York: Free Press.

Gittelman, M., and B. Kogut. 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science* 49: 366–382.

Hall, B.H., A.B. Jaffe, and M. Trajtenberg. 2000. *Market value and patent citations: A first look. NBER working paper no. 7741*. Cambridge, MA: National Bureau of Economic Research.

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

Jaffe, A.B.., M. Trajtenberg, and R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108: 577–598.

Jensen, M.C., and W.H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3: 305–360.

Kaplan, R.S., and D.C. Norton. 1992. The balanced scorecard: Measures that drive performance. *Harvard Business Review* 70: 71–79.
Klein, P., J. Mahoney, A. McGahan, and C. Pitelis. 2012. A property rights approach for a stakeholder theory of the firm. *Strategic Organization* 10: 304–315.
Levin, R., A. Klevorick, R.R. Nelson, and S.G. Winter. 1987. Appropriating the returns from industrial R&D. *Brookings Papers on Economic Activity* 18: 783–820.
Mair, J., and I. Marti. 2006. Social entrepreneurship research. *Journal of World Business* 411: 36–44.
McGahan, A.M., and M.E. Porter. 1999. The persistence of shocks to profitability. *Review of Economics and Statistics* 81: 143–153.
Mueller, D.C. 1986. *Profits in the long run*. Cambridge: Cambridge University Press.
Myers, M.B., and R.S. Rosenbloom. 1996. Rethinking the role of industrial research. In *Engines of innovation: U.S. industrial research at the end of an era*, ed. R.-S. Rosenbloom and J.C. Spencer. Boston: Harvard Business School Press.
Nelson, R.R. (ed.). 1993. *National innovation systems: A comparative analysis*. Oxford: Oxford University Press.
Porter, M.E. 1980. *Competitive strategy*. New York: Free Press.
Porter, M.E. 1985. *Competitive advantage*. New York: Free Press.
Porter, M.E., and M. Kramer. 2011. Creating shared value. *Harvard Business Review* 89: 2–17.
Rao, H., C. Morrill, and M.B. Zald. 2000. Power plays: How social movements and collective action create new organizational forms. *Research in Organizational Behavior* 22: 237–281.
Stout, L.A. 2012. *The shareholder value myth: How putting shareholders first harms investors, corporations, and the public*. San Francisco: Berrett-Koehler Publishers.
Teece, D.J., G. Pisano, and A. Shuen. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18: 509–533.
Tripsas, M. 1997. Unraveling the process of creative destruction: Complementary assets and firm survival in the typesetter industry. *Strategic Management Journal* 18: 119–142.

Peters, Tom (Born 1942)

David Collins
University of Essex, Department of Management, Colchester, Essex, UK

Tom Peters was born in Baltimore in 1942. He holds engineering degrees from Cornell University (BCE, MCE) and business degrees (MBA, Ph.D.) from Stanford University. His reputation as a commentator on the business of management has also led to the award of a number of honorary degrees, including a doctorate from the State University of Management in Moscow.

Between 1966 and 1970, Tom Peters served in the US Navy and made two active service deployments to Vietnam with the construction battalion or *Seabees*. Following his naval service, Peters worked for the consulting firm Peat Marwick Mitchell, before taking up a post in the White House as a senior advisor on drug abuse. In 1974, he joined McKinsey and Company. Peters was made a partner in 1977, and remained with the firm until 1981, when he resigned in order to write up the results of a research project into 'organizational effectiveness' that he had conducted at the behest of his employer.

Peters' initial attempt to write up this project resulted in an unstructured manuscript running to some 1300 pages. In an attempt to tame this text, Peters' McKinsey colleague Bob Waterman was invited to assist in the writing process. In 1982, the fruit of their collaboration was published as *In Search of Excellence* (Peters and Waterman 1982). To the surprise of all concerned, this book quickly became a bestseller. Indeed, in 1983 it became the first non-biographical business book to top the *New York Times* bestseller list. It remained at the head of this list for two years, until it was toppled by Peters' next book (Peters and Austin 1985).

The early 1980s was a challenging period for the US. The economy was struggling under the burden of double-digit rates of interest, inflation and unemployment. It was, in addition, losing key markets to foreign competitors. Reacting to these developments, a range of commentators suggested that American managers would have to emulate the practices developed and employed by Japanese business. *In Search of Excellence* rejected this counsel, insofar as it set out to explore, and to reveal, the very best in American managerial practice.

The research into organizational effectiveness conducted by Peters employed a range of financial performance metrics to construct a sample of highly performing US organizations. On the

strength of interviews conducted in a subsample of these organizations, Peters and Waterman argued that excellence in business is the product of a commitment to eight organizational attributes:

1. A bias for action.
2. Close to the customer.
3. Autonomy and entrepreneurship.
4. Productivity through people.
5. Hand-on, value driven.
6. Stick to the knitting.
7. Simple form, lean staff.
8. Loose-tight properties.

Reaction to *In Search of Excellence* has been mixed. Indeed, mere mention of the name Tom Peters tends to excite and polarize opinion. Some, for example, protest that Tom Peters is nothing less than 'a guru': *The Economist* has dubbed Peters 'the uber-guru' of management, while *Fortune* – drawing upon slightly different cultural influences – has named Peters the 'Ur-guru' of management. Others, however, complain that Peters dispenses neither wisdom nor enlightenment. Carroll, for example, argues that *In Search of Excellence* offers an account of the complex business of management that is shallow and simplistic (Carroll 1983).

Watson (2001) takes an interesting position on the legacy of the excellence project. He acknowledges that *In Search of Excellence* has conceptual and methodological flaws. Yet he insists that this text offers an important and enduring contribution to the theory and practice of management because it recognizes the 'moral economy' that is the workplace. Even the most ardent of Peters' critics, however, would have to concede that *In Search of Excellence* effectively defined and created the 'popular management' segment of the publishing industry (Pagel and Westerfelhaus 2005).

The publisher Bloomsbury named *In Search of Excellence* 'The Greatest Business Book of All Time'. However, it would be a mistake to reduce Peters' philosophy to the contents of a now 30-year-old text. Thus we should note that Peters offers around one hundred seminars each year, and continues to rank among the highest paid performers on the international lecture circuit. Furthermore, we should acknowledge that since 1982 he has produced eight major books that have continued to develop and to refine his concern with business excellence. These texts extend the excellence project insofar as they reflect, variously, upon business practice in Europe and in the Far East; changing technology; product design; and, most recently, the role, status and position of women at work.

See Also

- ▶ Leadership
- ▶ Management Gurus
- ▶ Organizational Change

References

Carroll, D.T. 1983. A disappointing search for excellence. *Harvard Business Review* 61: 78–82.

Pagel, S., and R. Westerfelhaus. 2005. Charting managerial reading preferences in relation to popular management theory books. *Journal of Business Communication* 42: 420–448.

Watson, T. 2001. *In search of management: Culture, chaos and control in managerial work*. London: Thomson Learning.

Selected Works

Peters, T. 1987. *Thriving on chaos: Handbook for a management revolution*. London: Guild Publishing.

Peters, T. 1992. *Liberation management: Necessary disorganization for the nanosecond nineties*. London: Macmillan.

Peters, T. 1993. *The Tom Peters Seminar: Crazy Times Call for Crazy Organizations*. London: Macmillan.

Peters, T. 1994. *The pursuit of wow! Every person's guide to topsy turvy times*. London: Macmillan.

Peters, T. 1997. *The circle of innovation: You can't shrink your way to greatness*. London: Hodder & Stoughton.

Peters, T. 2003. *Re-imagine: Business excellence in a disruptive age*. London: Dorling Kindersley.

Peters, T. 2010. *The big little things: 163 ways to pursue excellence*. New York: HarperCollins.

Peters, T., and N. Austin. 1985. *A passion for excellence: The leadership difference*. Fontana: London.

Peters, T., and R. Waterman. 1982. *In search of excellence: Lessons from America's Best Run Companies*. New York: Harper & Row.

Pharmaceutical Industry

Luigi Orsenigo
IUSS University Institute for Advanced Studies, Pavia, Italy

Abstract

Pharmaceuticals have been a highly innovative, marketing-intensive, internationalized and heavily regulated industry, at least since the post-Second World War years. This industry has consistently shown a remarkable track record in innovativeness, economic and financial performance. Yet, throughout its whole history, pharmaceuticals have been characterized by harsh controversies: here classic market failures, economic trade-offs and social and ethical conflicts are brought to extremes. The history of pharmaceuticals has also been marked by major technological, organizational and institutional change. Indeed, some of the factors that have sustained the performance of the industry seem now to be weakening.

Definition The pharmaceutical industry develops, produces and markets drugs or pharmaceuticals licensed for use as medications. Pharmaceuticals are – and have always been – a large, high-growth, globalized, science-based and innovation-intensive industry. The patterns of competition in this industry are an extreme example of 'dynamic competition'. This industry has shown so far a remarkable track record in innovativeness, economic and financial performance and (although not completely uncontroversially) welfare.

Pharmaceuticals emerged in the late 19th century as part of the nascent chemical sector, especially in Germany and Switzerland, exploiting the competencies and knowledge in organic chemicals developed in the German institutions of advanced education (Murmann 2003). In the US, the industry developed somewhat later, largely relying on European technology. Until the Second War World, pharmaceuticals was not an R&D-intensive industry. For a basic bibliography on pharmaceuticals, see, among others, Pisano (1996), Henderson et al. (1999), Sutton (1998), Grabowski and Vernon (1994), Chandler (2005), Galambos and Sewell (1996), Galambos and Sturchio (1996), Gambardella (1995).

The Pillars of Industry Growth

The take-off of the pharmaceuticals after the Second War World was sustained first by the rapid growth of demand. Very few effective drugs were in existence and the space for new treatments and cures for most diseases was simply immense. Moreover, the growth of health care insurance in the US and the welfare state in most European countries provided a large, rich and organized market for drugs. The industry also began to invest heavily in sales efforts and marketing. Whilst until the 1930s drugs were sold and advertised mainly directly to patients, subsequent legislation introduced prescription drugs. Pharmaceutical companies started to contact prescribing physicians directly, building vast and sophisticated marketing forces. Marketing expenditures have continued to grow subsequently, and the introduction in the US of direct-to-consumer advertising for prescription drugs has further strengthened this trend. Marketing intensity was estimated to have reached between 20 % and 25 % in the US (Gagnon and Lexchin 2008).

The second pillar of industry growth was ▶ innovation. The war's crash programmes in penicillin and sulfa in the US and in the UK, and the subsequent expansion of public biomedical research, began to provide rich opportunities to discover new drugs. In the US in particular, mainly through the National Institute of Health (NIH), public funding of biomedical research soared.

Innovative activities in pharmaceuticals are extremely uncertain and costly. Despite tremendous scientific advances, knowledge about the causes of diseases and the mechanisms of action of potential drugs remains poor. Drug discovery has been long based on an approach known as 'random screening', whereby thousands of

compounds are screened in the search for potential therapeutic activity. Only a very small fraction of them showed promising potential and often discovery was the outcome of serendipity. In this respect, pharmaceutical R&D can be usefully represented as a lottery (Gambardella 1995; Sutton 1998).

Innovation has been further sustained by appropriability conditions. Pharmaceuticals is one of the few industries where ▶ patents are critical for ensuring the private appropriation of the benefits of research. Patent protection has been traditionally stronger in the US: until the 1960–1970s in most European countries (the UK excluded), only process patents were allowed on drugs. The patenting regime has become increasingly tight over recent decades, in particular since the 1980s, when legislation and court decisions drastically widened the domain of what is patentable and the protection granted to inventors. The TRIPS agreements have extended this 'propatent' attitude worldwide. Bitter controversies over the efficiency of such a strong regime have arisen.

Despite the 'blind' nature of the discovery process, R&D intensity and the rate of innovation started to soar and in the 1960 and 1970s hundreds of new chemical entities (NCEs) and several important classes of drugs were discovered. Rates of growth averaged well above 10 % from the 1950s until the 1980s. The profitability of the industry was so high that, in 1959, there was a call for the establishment of an investigation commission at the US Congress, that is, the Kefauver Committee (Comanor 1986). The economic and financial performance of the industry remained spectacular until the 1980s, and even since then it has remained remarkable.

Since the mid-1970s scientific advances in basic biomedical research have led to a deeper knowledge about the understanding of the diseases and of the mechanisms of action of drugs, thus opening the way for new techniques of research. The biotechnology revolution has been continuously adding new frontiers to innovation and it has led to 'industrialized R&D' (Pisano 2006), which offers the potential to understand and identify much more precisely the causes of diseases, to create new compounds, to screen them much more efficiently and to design rationally drugs with specific effects. (The encompassing term 'biotechnology' is used here to include also the advent in the 1990s of the so-called platform technologies (combinatorial chemistry, high-throughput screening and computational chemistry) as well as the dramatic progresses in the various '-omics', such as genomics and proteomics, bioinformatics, synthetic and structural biology, and so forth.)

Yet these enormous steps forward in scientific research do not yet seem to have produced the expected impact on innovation, which still remains a very uncertain and rather blind process. If anything, the productivity of R&D has been sharply falling. Since the mid-1990s, as the cost of bringing a drug to the market has been rising dramatically, the number of approved new chemical entities has been stagnating. In 2002, US R&D expenditures in pharmaceuticals were 30 times greater than in the early 1980s, while roughly the same number of drugs were approved annually.

Market Structure and the Firms' Organizational Forms

The high weight of sunk costs in R&D and marketing – coupled with strong patent protection – granted significant advantages to large, vertically integrated corporations, what is now termed 'Big Pharma'. The ability to run multiple parallel experiments on a huge scale (Nelson 1961), the gathering of libraries of potentially useful molecules and the development of organizational routines governing the R&D process became crucial capabilities conferring competitive advantages to large firms (Pisano 1996; Henderson et al. 1999). ▶ Vertical integration was also favoured by the need to minimize transaction costs and to integrate knowledge along the whole chain, from discovery to clinical trials, product approval, manufacturing and marketing. These factors also implied growth into foreign markets. The largest, highly R&D-intensive German, Swiss and American companies proceeded decisively

in international expansion, establishing networks of relations with local firms through licensing and commercialization agreements. A presence in foreign markets was also often necessary for complying with local regulation.

Thus a remarkably stable and relatively small group of German, Swiss, British and American firms has been consistently dominating the industry. Most were early entrants in the industry. However, the market share of these companies has been always lower than 10 % and only in very recent times has the current largest firm been able to pass this threshold, mainly through mergers and acquisitions.

Concentration has been limited by three main factors. The first is imitation. Despite strong 'isolating mechanisms' protecting the profits of innovative firms, 'inventing-around' existing molecules, or introducing new combinations among them or new ways of delivering them, constitutes a major component of firms' innovative activities, broadly defined. In France, Italy, Spain and Japan in particular, many firms have specialized in the production and marketing of products invented elsewhere. Second, the innovative process is characterized by extreme uncertainty and by the difficulty of leveraging the results of past innovative efforts into new products. Economies of scope and cumulativeness of technological advances are limited (Cockburn and Henderson 2001). Third, the pharmaceutical market results from the aggregation of many independent submarkets with little or no substitution between products. If the assets and knowledge that are necessary to gain market shares in one submarket cannot be easily used in different submarkets (as is the case not only for R&D but also for marketing), there is a high probability that different firms end up dominating different niches, without anyone of them being able to control the aggregate market.

Regulation

A crucial element shaping firms' strategies is regulation. Pharmaceuticals is heavily regulated. Indeed the market for drugs is plagued by all kinds of market failures (see Comanor 1986; Scherer 2000), ranging from monopolistic power to pervasive information asymmetries between producers, prescribing doctors and patients. Insurance aggravates these problems, since final users do not pay (or only partially) for the drug. A further set of reasons for regulation refers to cost containment. Even more important, a fundamental argument for regulation is based on equity and moral considerations and makes the analysis of the market to a large extent a social rather than a purely economic issue.

Regulation has taken a wide variety of forms and instruments. Most countries (with the notable exceptions of the US and Germany) have relied on price controls. Moreover, since the early 1960s and at least until the mid-1990s, the stringency of the drug approval processes has been drastically increasing, leading to substantial increases in R&D costs and to longer gestation times for new drugs. Yet it has also been argued that the creation of a stringent drug approval process in the US may have also helped create a strong competitive pressure favouring really innovative firm strategies. In fact, although the process of development and approval increased costs, it significantly increased barriers to imitation, even after patents expiry, thereby penalizing the less innovative firms (Thomas 1994).

Recent Developments

Since the early 2000s, the pharmaceutical sector has increasingly been showing symptoms of stress.

First, on the demand side, rising incomes, increasing prices of drugs and an aging population have caused soaring shares of pharmaceutical expenditure on total income and growing pressure on public outlays. In a period characterized by mounting concerns over budget deficits and – more generally – over the extension of public intervention in the economy, pharmaceutical expenditure has become a primary target for expense reduction. In many instances, the market for drugs has become a symbolic issue within the debate over the 'downsizing' of the welfare state.

Second, the 'biotechnological revolution' has implied a radical transformation in the required competencies for drug discovery and development which the existing firms painstakingly managed to accomplish. A significant process of de novo entry of new firms was observed for the first time after the Second World War. These companies were essentially university spin-offs backed by venture capital and sustained by an increasingly favourable patent regime. However, the new entrants did not displace the incumbents. They lacked the essential complementary assets (Teece 1986) which are necessary for extracting profits from innovation: since vertical integration in crucial downstream activities (like drug development, product approval by the regulatory authorities, manufacturing and marketing, etc.) was extremely expensive and required different specific competences and organizational structures, the new biotechnology firms were forced to sell their knowledge to larger, vertically integrated firms. The latter realized, in turn, that they could not rely solely on their internal knowledge to discover and develop new drugs. The prospect of the expiration of most key patents put pressure on attempts to discover and develop new blockbusters. Big companies reacted to this challenge through a wave of mergers and acquisitions and with processes of vertical disintegration, relying increasingly on small biotech companies and academia for new molecules, research techniques (but also clinical trials) through licences and collaboration agreements, a dense web of collaborative relationships and a vibrant market for technology developed among large incumbents specializing in downstream activities and new biotechnology companies. For detailed accounts of the emergence and development of the biotechnology industry, see Orsenigo (1989), Gambardella (1995), Pisano (1996), Henderson et al. (1999).

A further important development was the introduction of legislation favouring the diffusion of generics. In the US, the Waxman–Hatch Act in 1984 significantly reduced the safety control procedures for the generic drug bio-equivalent to branded products and allowed pharmacists to sell equivalent generics instead of branded products prescribed by doctors. Despite strong differences across countries, generics have been growing rapidly, accounting today for more than 50 % of the market in volume and around 10 % in terms of value.

The viability of the traditional business model (s) that have sustained the industry so far – that is, the large vertically integrated corporation and the specialized biotechnology companies – is increasingly being questioned. But no alternative model has yet appeared.

See Also

- ▶ Innovation
- ▶ Patents
- ▶ Research and Development (R&D) Investment
- ▶ Technological Change
- ▶ Vertical Integration

References

Balance, R. 1992. *The World's pharmaceutical industries: An international perspective on innovation, competition and policy, prepared for the United Nations Industrial Development Organization*. London: Edward Elgar.

Chandler, A.D. 2005. *Shaping the industrial century: The remarkable story of the modern chemical and pharmaceutical industries*. Cambridge, MA: Harvard Studies in Business History.

Cockburn, I., and R. Henderson. 2001. Scale and scope in drug development: Unpacking the advantages of size in pharmaceutical research. *Journal of Health Economics* 20: 1033–1057.

Comanor, W.S. 1986. The political economy of the pharmaceutical industry. *Journal of Economic Literature* 24: 1178–1217.

Gagnon, M., and J. Lexchin. 2008. The cost of pushing pills: A new estimate of pharmaceutical promotion expenditures in the United States. *PLoS Medicine* 5: 29–33.

Galambos, L., and J.E. Sewell. 1996. *Network of innovators: Vaccine development at Merck, Sharp & Dohme and Mulfor, 1895–1995*. Cambridge: Cambridge University Press.

Galambos, L., and J. Sturchio. 1996. The pharmaceutical industry in the twentieth century: A reappraisal of the sources of innovation. *History and Technology* 13: 83–100.

Gambardella, A. 1995. *Science and innovation in the US pharmaceutical industry*. Cambridge: Cambridge University Press.
Grabowski, H., and J. Vernon. 1994. Innovation and structural change in pharmaceuticals and biotechnology. *Industrial and Corporate Change* 3: 435–450.
Henderson, R., L. Orsenigo, and G.P. Pisano. 1999. The pharmaceutical industry and the revolution in molecular biology: Exploring the interactions between scientific, institutional and organizational change. In *The sources of industrial leadership*, ed. D.C. Mowery and R.R. Nelson. Cambridge: Cambridge University Press.
Murmann, J.P. 2003. *Knowledge and competitive advantage: The coevolution of firms, technology, and National Institution in the Synthetic Dye Industry, 1850–1914*. Cambridge: Cambridge University Press.
Nelson, R.R. 1961. Uncertainty, learning, and the economics of parallel research and development. *Review of Economics and Statistics* 43: 351–368.
Orsenigo, L. 1989. *The emergence of biotechnology*. London: Pinter Publishers.
Pisano, G.P. 1996. *The development factory: Unlocking the potential of process innovation*. Cambridge, MA: Harvard Business School Press.
Pisano, G.P. 2006. *Science business: The promise, the reality and the future of biotech*. Cambridge, MA: Harvard Business School University Press.
Scherer, F.M. 2000. The pharmaceutical industry. In *Handbook of health economics*, vol. 1, ed. A.J. Culyer and J.P. Newhouse. Amsterdam: Elsevier.
Sutton, J. 1998. *Technology and market structure: Theory and history*. Cambridge, MA: The MIT Press.
Teece, D.J. 1986. Profiting from technological innovation: Implication for integration, collaboration, licensing and public policy. *Research Policy* 15: 185–219.
Thomas, L.G. 1994. Implicit industrial policy: The triumph of Britain and the failure of France in global pharmaceuticals. *Industrial and Corporate Change* 3: 451–489.

Platform Innovation

Michael J. Leiblein
Ohio State University, Fisher College of Business, Columbus, OH, USA

Abstract

Platform innovation describes a form of extended competition that represents a large and growing share of global competition. In an increasing variety of settings, product and industry platforms are forcing managers to reconsider how they compete within an extended system of suppliers, complementors and rivals. This entry describes innovation in modular product and extended industry platforms, identifies distinct platform types in each of these domains and lists some recent research findings regarding actions that may help organizations to compete in these settings. Throughout, attention is devoted to describing some of the distinctive competitive challenges in platform markets.

Definition Platform innovation refers to changes in support structures that increase the effectiveness with which a group of activities may be performed. Product platform innovation entails changes to a common component or body of knowledge that may be redeployed across products. Industry platform innovation entails changes to infrastructure, standards and rules that enable transactions between multiple firms.

Introduction

In a business context, the term platform innovation refers to changes in the mechanisms or support structures that affect how a group or system of activities may be performed. The term is used in the product development and operations (e.g., Wheelwright and Clark 1992) and ▶ technology strategy (e.g., Cusumano and Gawar 2002; Eisenmann et al. 2006) fields to highlight how the organization of a set of activities affects innovation and competition. A platform thus entails the coordination or organization of a set of activities or components within a system through the use of common interfaces, tools or standards. A platform innovation changes the way these sets of activities or components interact.

References are often made to two broad families or forms of platform innovation. Product platform innovation entails changes to common components, subsystems or bodies of knowledge that may be redeployed across a series of products or services. Product platforms are often conceived of as architectures that define the arrangement of

functional elements, the mapping of these functional elements to physical components and the specification of interfaces among physical components (e.g., Ulrich 1995; Fisher et al. 1999). Classic examples include generations of automobiles, power tools and mobile phones (see Wheelwright and Clark 1992; Ulrich 1995). For instance, the Chrysler 'K-cars' of the 1980s and today's Toyota and Lexus lines share common designs and automotive chassis. Power tools often share common batteries and motors. Cameras and mobile phones often share common aesthetics, designs and lens components. Product platform innovation thus entails changes to elements of the product platform, typically across a family or series of products.

Industry platform innovation entails changes to hardware and software infrastructure, standards and rules that enable transactions between multiple firms or sides of a market (Gawar and Cusumano 2008). Industry platforms provide a mechanism that enables transactions between multiple sponsors and consumers through rules that govern the cost and access to information across component and technology interfaces. Notable industry platforms include PC operating systems such as Windows, Macintosh or Linux, which provide a protocol to link application developers with computer users; online data platforms such as Match.com and eHarmony, which provide a mechanism to link men and women; search advertising platforms such as Google, MSN and Yahoo, which provide a means to link searchers and advertisers; and stock exchange platforms such as NASDAQ or NYSE, which provide a method to link equity purchasers with listed companies (Cusumano and Gawar 2002; Eisenmann et al. 2006).

Although product and industry platforms consist of sets of subsystems and interfaces that coordinate activities and distribute value according to certain rules and protocols, they also entail substantive differences. Product platforms are generally developed and controlled within a single firm. They provide a means to alter the costs and benefits of a product family or series of generations of product families. Investing in a product platform innovation thus entails trading off the cost of an initial investment to design a shared architecture against the potential scope of economies associated with the reuse of components or design architectures. Ulrich (1995) and Baldwin and Clark (1999) explore these trade-offs and emphasize the benefits of reusing components and design frameworks in industries such as software and automobile manufacturing. The focus of product platform innovation is to improve the intrinsic value of a product family.

An industry platform often spans a group of firms within a business ecosystem. It not only generates value by coordinating complementary products, services and users but also by creating the potential for network effects (e.g., Katz and Shapiro 1994). Network effects refer to the benefits or costs conferred on others when an individual chooses to purchase a good. Positive network effects arise when one's utility is increased when others own or use the same product or complementary products or services. For instance, the value of a mobile phone service increases with the number of others who are members of that same network. Similarly, the value of a computer game or DVD player increases with the number of complementary games or movies that can be played on a device. Thus, the value of an industry platform is a function of both its stand-alone quality as well as the size of its actual or expected network of users. An implication is that competition occurs at two levels – over both the functional value of a standalone product and the ability to control the evolution of the industry platform.

Categorizing Product Platforms

The product platform innovation concept suggests an opportunity to identify and define particular forms of projects. Wheelwright and Clark (1992) identify and define three such projects – derivative projects, platform projects and breakthrough projects. In their framework, derivative projects refer to efficiency-enhancing improvements to existing products and minor enhancements to existing processes. These derivative extensions are often

low-risk projects that offer high margins and limited short-term growth. Platform projects are those involving 'mid-range' levels of product and process change. Often, innovations of this sort extend the core business and can be productively thought of as a vehicle for entry into new markets and continued growth. Finally, Wheelwright and Clark (1992) define breakthrough projects as those risky projects that promise significant change to existing products and processes. These innovations often provide radical new solutions to unmet customer needs and involve high competitive uncertainty. They may be 'game changers', however.

A product platform provides a number of potential advantages. Two frequently mentioned benefits are the ability to generate efficiency gains through the reuse of common components and to foster a degree of product variety by recombining modular components. There are other advantages, however. For instance, recognition that platform projects are one element of a portfolio of project types allows management to insure that lower- and higher-risk projects are judged on appropriate (and different) capital-budgeting thresholds, to assess risk and manage interdependencies across projects, and to achieve efficiencies by trimming or rebalancing an undisciplined portfolio of projects. Ultimately, as derivative projects tend to offer short-term cash flows and breakthrough projects tend to offer less certain but longer-term cash flows, it provides a means to manage the business over time.

One of the more interesting applications of the Wheelwright and Clark (1992) framework is the ability to state an intended resource allocation policy in terms of the percentage of resources devoted to derivative, platform and breakthrough projects, and evaluate their subsequent behaviour against that desired portfolio. For instance, firms that desire to implement a lower-risk competitive strategy would be expected to have portfolios that emphasized derivative projects, and firms attempting to implement a higher-risk strategy would have portfolios that placed greater emphasis on platform or breakthrough projects. Such an approach allows managers to compare the desired resource allocation policy based on a stated strategy against the actual resource allocation policy based on funded projects, to identify sources of misalignment and to re-evaluate the stated strategy or rebalance the actual investments.

Categorizing Industry Platforms

Industry platforms also offer opportunities to identify and define particular forms of innovation. In contrast to categorizing innovations in terms of the degree of product-market or process-technology change, industry platforms are often categorized in terms of access and control. Here, access refers to the degree to which information regarding the platform's rules, tools and standards are narrowly or broadly disseminated to the public, and control refers to the degree to which the property rights to these rules, tools and standards are narrowly or broadly held (Shapiro and Varian 1998; West 2003). For instance, *platform access* suggests distinctions between the Apple operating system, Skype voice-over internet protocol or New York Stock Exchange trading platforms where a single organization has access to the core information, and the WiFi (IEEE 802.11) standard for wireless connectivity, the Linux operating system or the Nintendo gaming system platforms where information is in the public domain. Similarly, *platform control* suggests distinctions between the Windows operating system, Adobe Acrobat portable document format (PDF) application software (not Adobe Reader freeware) and Monster.com platforms, where a single organization controls the property rights to the rules, tools and standards, and the DVD recording or CareerBuilder.com platforms, where multiple parties share these rights.

Researchers continue to identify and explicate the resource allocation and organizational policy implications associated with industry platform innovations. One stream of work has explored factors that affect the propensity for markets to 'tip' to a single platform (Katz and Shapiro 1994). For instance, markets are more likely to tip when there are supply-side scale economies and homogeneous demand-side preferences. Hossain

et al. (2011) show that platforms are more likely to tip when there are significant differences in vertical differentiation (or quality) but are much less likely to tip when they exhibit significant degrees of horizontal differentiation. The implication is that systematic differences in the structure of an ecosystem affect the likelihood that a single or multiple platforms will exist.

The potential for an industry to tip to a single platform has led to research that examines whether and how pricing, entry timing and access decisions influence the evolution of an industry platform. The primary thrust of this work is to determine whether and how active effort to increase the number of users in one's networks affects subsequent competition. For instance, Parker and van Alstyne (2005) explore how pricing models affect platform profits.

Gawar and Cusumano (2008) describe how technology design and information access, IP control and platform control, and firm scope and organizational policy may allow a platform leader to direct a platform's trajectory. Boudreau (2010) finds that granting greater levels of access to independent hardware developers in the handheld computing systems sector produces up to a fivefold acceleration in the rate of new product development. Zhu and Iansiti (2012) explore the relative importance of platform quality, indirect network effects and consumer expectations on the success of entrants in platform-based markets. Their work in the computer gaming industry suggests that whether an entrant can successfully enter a platform-based market depends on the extent to which consumers care about application variety and the extent to which consumers' value applications to be released in the future. This and related work is beginning to demonstrate the multiple levers managers may use to compete through the development of industry platforms.

See Also

- ▶ Closed vs Open Innovation
- ▶ Open Innovation
- ▶ Technology Strategy

References

Baldwin, C.Y., and K.B. Clark. 1999. *Design rules*, The power of modularity, vol. 1. Cambridge, MA: The MIT Press.

Boudreau, K. 2010. Open platform strategies and innovation: Granting access vs. devolving control. *Management Science* 56: 1849–1872.

Cusumano, M., and A. Gawar. 2002. The elements of platform leadership. *Sloan Management Review* 43: 51–58.

Eisenmann, T., G. Parker, and M. van Alstyne. 2006. Strategies for two-sided markets. *Harvard Business Review* 84: 92–101.

Fisher, M., K. Ramdas, and K. Ulrich. 1999. Component sharing in the management of product variety: A study of automotive braking systems. *Management Science* 45: 297–315.

Gawar, A., and M. Cusumano. 2008. How companies become platform leaders. *Sloan Management Review* 49: 28–35.

Hossain, T., D. Minor, and J. Morgan. 2011. Competing matchmakers: An experimental analysis. *Management Science* 53: 1913–1925.

Katz, M., and C. Shapiro. 1994. Systems competition and network effects. *Journal of Economic Perspectives* 8: 93–115.

Parker, G., and M. van Alstyne. 2005. Two-sided network effects: A theory of information product design. *Management Science* 51: 1494–1504.

Shapiro, C., and H.R. Varian. 1998. *Information rules: A strategic guide to the network economy*. Cambridge, MA: Harvard Business Review Press.

Ulrich, K. 1995. The role of product architecture in the manufacturing firm. *Research Policy* 24: 419–440.

West, J. 2003. How open is open enough? Melding proprietary and open source platform strategies. *Research Policy* 32: 1259–1285.

Wheelwright, S., and K. Clark. 1992. Creating project plans to focus product development. *Harvard Business Review* 70: 2–14.

Zhu, F., and M. Iansiti. 2012. Entry into platform-based markets. *Strategic Management Journal* 33: 88–106.

Platform Strategy

Geoffrey Parker[1] and Marshall Van Alstyne[2]
[1]Tulane University, New Orleans, LA, USA
[2]Boston University, Boston, MA, USA

Abstract

This article summarizes the platform strategy literature and is organized around launch strategies, governance and competition. A platform strategy is the mobilization of a networked

business platform to expand into and operate in a given market. A business platform, in turn, is a nexus of rules and infrastructure that facilitate interactions among network users. Platforms provide building blocks that serve as the foundation for complementary products and services. They also match buyers with suppliers, who transact directly with each other using system resources and are generally subject to network effects.

Definition A platform strategy is the mobilization of a networked business platform to expand into and operate in a given market.

A platform strategy is the mobilization of a networked business platform to expand into and operate in a given market (Cusumano and Gawer 2002). A business platform, in turn, is a nexus of rules and infrastructure that facilitate interactions among network users (Eisenmann et al. 2011). Stated differently, a platform is a published standard, together with a ▶ governance model, that facilitates third-party participation (Parker and Van Alstyne 2013). Platforms provide building blocks that serve as the foundation for complementary products and services (Cusumano and Gawer 2002; Gawer and Henderson 2007). They also match buyers with suppliers who transact directly with each other using system resources (Hagiu 2006; Hagiu and Wright 2011) and are generally subject to ▶ network effects (Eisenmann et al. 2009, 2011; Boudreau 2010). These definitions each have the property that reconfiguration of platform assets allows external parties to interact with each other and add value. Examples include operating systems, game consoles, payment systems, ride-sharing platforms, smart grids, healthcare networks and social networks.

Building on Thomson's (1967) typology of long-linked, mediating and intensive technologies, Stabell and Fjeldstad (1998) identified network platforms as one of three elemental configurations through which firms generate value. In traditional industries, bilateral exchanges follow a linear path as firms purchase inputs, transform them to add value, assemble components and subsystems into complete products and then sell the output. In platform industries, interaction follows a triangular relationship (Eisenmann et al. 2009) as parties first affiliate with the platform then connect or trade using platform resources. For example, on Airbnb, renters and hosts transact with one another but they use the platform to match with one another, to perform searches, enter into contractual agreements, transfer payment, acquire insurance and manage reputations to facilitate future transactions.

Platform firms share characteristics with platform products but operate in these triangular rather than linear markets. Both imply shared technology, reconfigurable elements and fixed costs that can be spread across multiple product types such as automobiles with common engines, transmissions and electronics (Cusumano and Nobeoka 1992). Firms, however, are also characterized by the network of value-adding relationships among users over and above the physical value of platform components. To manage and motivate these external relations, platforms must have rules that promote healthy participant interactions. Exchange platforms, in particular, require rules to address market failures, as noted below. Boundaries of this governance model distinguish the platform firm from the ecosystem in which it is embedded. Contributors to competing platforms, for example, are part of the ecosystem but need not abide by a focal platform's rules.

Network platforms further differ from product platforms because stronger network effects, switching costs and multi-homing costs create greater pressure for market concentration. Firms that compete in such markets need clearer guidelines by which to set strategy in order to harness these effects. Strategies to manage launch, openness and governance, as well as competition, follow below.

Launch Strategies

A central problem facing platforms subject to network effects is how to drive user adoption enough to reach critical mass (Evans and Schmalensee 2010). Network platforms often

have users of one type whose utilities depend on the presence of users of a different type (Parker and Van Alstyne 2000b, 2005; Rochet and Tirole 2003) as in the case of game developers and players or auction buyers and sellers. If there are not enough users of both types then the standalone value of the platform may not offer sufficient appeal to drive adoption. This leads to a 'chicken-and-egg' problem of launch and adoption (Caillaud and Jullien 2003; Evans and Schmalensee 2010). Firms have used a number of strategies to overcome this issue.

Subsidy

Platforms with substantial resources can entice users via subsidy to join the platform. Subsidies can be temporary penetration prices or permanent discounts and can take several forms. Direct cash transfers are possible, but this creates a moral hazard problem where users might accept the subsidy without using the platform. One solution is to offer subsidies, such as technical support, that only have value when consumed with the platform (Parker and Van Alstyne 2005). Subsidies may take the form of free information, which has zero marginal cost (Parker and Van Alstyne 2000a, 2005), but can also work for certain performance characteristics (Anderson et al. 2014) or even physical goods. Larger launch subsidies generally decline after platform usage reaches critical mass. For example, when Sony launched the PlayStation 3, it distributed the console well below marginal cost. Liu (2010) notes that this strategy was common in the video game console market over multiple generations. The subsidy is recovered by later taxing the sale of game complements sold by game developers. Even in equilibrium, however, prices below marginal cost are common among search and matching platforms in order to keep one side of the market on board the platform.

Seeding and Marquee Users

Platforms typically launch with complements that give their interactions value. On two-sided platforms, a 'seeding' strategy solves participation on one side of the network by offering users of that type enough value that they adopt (Gawer and Henderson 2007; Boudreau 2012). The platform sponsor can either develop complements on its own or it can work with partners and offer them incentives to produce seed interactions for the new platform. Seed interactions must be provided until both sides of the market reach critical mass, at which point transactions volume becomes self-sustaining. Financial service providers have used this approach, offering their own products, before opening their platform to third-party financial instruments (Hagiu and Eisenmann 2007). The lead firm must also decide whether seed content will substitute for or complement subsequent content provided by partners (Hagiu and Spulber 2013).

A launch strategy closely related to seeding is to coax marquee users onto the platform. For example, when Microsoft launched the Xbox video game platform, it brought Electronic Arts (EA) to its platform by offering incentives that included category exclusivity (Eisenmann et al. 2009). Microsoft courted EA because its existing strength with users ensured that it could bring a large number of users to the platform. EA fans might have been unwilling to join a platform that did not include it. SAP offered similar category exclusivity to ADP for payroll processing when it launched its cloud services platform (Parker and Van Alstyne 2013).

Adobe managed a highly successful marquee seeding strategy by convincing the US Federal government to issue tax documents in its proprietary portable document format (PDF). Putting documents online in an unalterable format dramatically cut printing and postage costs while creating a prospective user base equal in size to the US tax base (Tripsas 2001).

Micro-Market Launch

One effective strategy restricts launch to a small community in order to generate strong, albeit bounded, network effects. eBay started as a market for Pez sweet dispensers (Evans 2003). Diner's Club targeted restaurants and patrons on Manhattan Island (Evans and Schmalensee 2010). Facebook launched exclusively among Harvard undergraduates before expanding to all '.edu' and then '.com' domains (Ellison et al. 2007).

The idea is that a more sharply defined community will experience stronger network effects if a substantial fraction of the community adopts a particular platform product or service. Once adoption takes place within that community, the platform can be opened to adjacent groups, pulling new users onto the platform. In addition, this strategy has the advantage of allowing the platform to build capacity in stages. Launching into an adjacent market from an established platform is the process of 'platform envelopment' described under competition below.

Piggybacking

Small companies that lack a user base of their own may seek to borrow users from another network. This is one of the strategies used to launch PayPal, which piggybacked on eBay before it was acquired (Penenberg 2009). PayPal set up software bots on eBay to buy and sell merchandise, insisting that the other side of the transaction use its payment system. Buying and selling at market prices also made this a low capital cost strategy. Similarly, the room reservation service Airbnb launched by integrating into Craigslist without securing permission to do so. The payment service square launched on top of the iPhone and Android platforms, connecting financially through existing credit card networks. Both firms piggybacked on existing networks without having to create new demand.

Governance

A governance model includes rules for participation, interaction and resolution of conflict. Participation rules govern openness, defining who can affiliate with the platform. Interaction rules govern behaviour on the platform, division of surplus, privileges and responsibilities.

Open Platform Business Models

Many scholars have focused on the firm's decisions with respect to sharing intellectual property and opening its systems to external firms and individuals (Edwards 2001; Chesbrough 2003; Eisenmann et al. 2009; Boudreau 2010). Shapiro and Varian (1999) and West (2003) describe the tension between adoption, which calls for more openness, and appropriation, which calls for more control. The levels of openness and threat of subsequent appropriation can significantly affect participation and investment incentives of platform partners (Parker and Van Alstyne 2009).

Rules allowing unrestricted openness allow partners easy access to, and usually exit from, the platform. Rules restricting access limit participation partners by number or type and often lock partners into longer-term relationships (Kauffman and Mohtadi 2004). Types of partners can include (i) users or consumers, (ii) developers or suppliers, (iii) platform providers and (iv) platform sponsors (Eisenmann et al. 2009). Mobile phone platforms provide a representative illustration with callers as users, application developers as suppliers, an app store as the provider (who serves as point of contact) and the intellectual property rights holder as the sponsor (who decides the rules of the governance model). A two-sided model would consider how the platform manages buyers and suppliers. A multi-sided model might also consider hardware manufacturers and telecommunications firms (West and Mace 2010).

Rules for openness often include the issue of compatibility and multi-homing (Rysman 2009). As a condition of participation, an ecosystem partner may be required to affiliate exclusively with one platform – that is, a single-home. More open rules allow partners to affiliate with competing platforms – that is, multi-home. Evidence suggests that multi-homing of applications hurts sales of a given platform (Landsman and Stremersch 2011). If resource heterogeneity works to the advantage of one platform, then multi-homing of platform partners can reduce that advantage as capabilities spread to competing platforms (Sun and Tse 2009). Platforms thus prefer that ecosystem partners single-home and offer novel content on their own platform exclusively. Platform entrants, however, hope to attract popular content that is already resident on an incumbent platform. Entrants thus seek to persuade partners of incumbent platforms to multi-home. A platform sponsor may also choose to

centralize key decisions to limit negotiation costs and optimize the health of the ecosystem. Apple's decision to standardize and cap prices in its iTunes store provides an example (Rochet and Tirole 2003; Ye et al. 2012).

Open strategies can achieve 'permissionless innovation' (Cerf 2012; Parker and Van Alstyne 2013) whereby third parties add new value to the platform without having to negotiate with the platform owner. Recognizing the risk of appropriation by the lead firm, complementors can be reluctant to share their technology or invest in the platform. In response, the lead firm opens the platform and offers a default contract whereby the complementor receives longer lead time before facing competition. Cisco and SAP have both used this strategy to induce third-party investment. The best innovations might then be absorbed into the platform in the future as the lead firm builds, buys or partners to secure new features for its platform. Empirically, complementors that have stronger intellectual property (IP) rights have more successful initial public offerings (IPOs) and more easily resist having their innovations bundled by the platform owner (Huang et al. 2013).

A platform sponsor may share control when it lacks sufficient resources to act alone in pursuit of its objectives (Gawer and Henderson 2007). Sponsors also share control to increase adoption, improve innovation and reallocate resources (Jacobides and Billinger 2006). Sharing control, however, increases risk that the sponsor fails to capture platform value. Risk of platform fragmentation also rises (Yoo et al. 2012). Rules, both technical and social, must then be used to maintain coherence and curb rent seeking by platform partners (Garud et al. 2008). Platform maturity can also drive regulation. At launch when market share is small, governance tends to be less permissive in participation but more permissive in behaviour in order to encourage key partners to innovate and explore. At maturity when market share is large, governance tends toward tighter control over behaviour but looser control over participation as revenue comes less from third-party innovation and more from rent extraction (O'Reilly 2010).

Regulation and Private Ordering

The need for regulation arises from the fact that platforms facilitate exchange. Analogous to standard exchange markets, platforms are subject to market failures stemming from factors such as network effects, information asymmetry, uninsured risks and congestion. Regulation combines contractual, technical, informational and economic instruments to minimize these market failures (West 2003; Boudreau and Hagiu 2009; Evans 2012). Strategic importance follows from the ability of the platform to create greater wealth and thereby win markets via users' voluntary participation in the open platform. In order to capture the benefits of ecosystem growth, the platform must impose certain regulations on the user participants. This need was anticipated by Teece (1986) in his seminal work on the conditions necessary for firms to profit from technological innovation.

Positive externalities, as in the case of software developers attracting users to an operating system, have been addressed by offering price subsidies to the group generating beneficial spillovers (Parker and Van Alstyne 2000b; Rochet and Tirole 2003). Negative externalities, as in the case of on-platform traffic congestion, has been addressed with congestion pricing (Evans 2012). Information asymmetry, as in the case of counterfeit goods on auction platforms or insider trading on stock platforms, has been addressed using penalties, arbitration and exclusion. Platforms address missing transactions born of information asymmetry via improved search and matching (Evans 2012). Game platforms and social networks have addressed a 'lemons problem' of low quality driving out high quality by using technological lock-out mechanisms, quality review, reputation systems and 'bouncer's rights' to exclude based on low quality goods or bad behaviour (Strahilevitz 2006; Boudreau and Hagiu 2009). Financial platforms have absorbed users' risk of fraud by offering insurance (Evans et al. 2006).

Externalities, especially two-sided and multi-sided, are endemic in platforms due to participation by multiple types of users who attract or repel one another. Platforms internalize these externalities to create social value, which substantially

drives pricing and monetization decisions (Parker and Van Alstyne 2000a, b, 2005; Rochet and Tirole 2003; Nocke et al. 2007). Rules that divide the pie differently change the size of the pie (Evans 2012).

Regulation by the platform can potentially be superior to regulation by state or federal governments. The mechanism for such contracts is articulated in the law and economics literature on 'private ordering', which is governance via private contract that seeks to achieve welfare gains higher than that provided by a system of public laws (Eisenberg 1976). Due to information asymmetry and one-size-fits-all regulation, private ordering by firms can yield better results than uniform laws (Williamson 2002). Platforms often have considerably greater visibility into user behaviour than public regulators, providing an opportunity to sanction behaviour earlier and with greater accuracy (Evans 2012).

A platform firm may choose to absorb developers' ideas into the core system, while at the same time making these new technologies available for partners to build upon. Absorbing new features and publishing them has the goal of fostering higher rates of user adoption and developer innovation by promoting R&D spill-overs (Parker and Van Alstyne 2013). Sharing ideas across a developer pool fosters a knowledge externality analogous to that which increases the productive capacity of a region (Audretsch and Feldman 1996; Edwards 2001). In 3D printing, for example, designers can build on concepts of other designers, a process made feasible by requiring innovations to be reusable. Knowledge spillovers recursively increase the output of ecosystem partners through an iterative cycle of recombination, reabsorption and republication.

Competition

Competition occurs at three levels of a platform ecosystem. It exists from one platform to another, as in the video game console battles of Sony, Microsoft and Nintendo (Evans et al. 2006). Competition can also exist between a platform and its partners as in the case of Microsoft appropriating such partner innovations as browsers, multi-threading, streaming media and instant messaging into its operating system (Jackson 1999; Nalebuff 2004). Finally, competition can exist among partners each vying for position within a focal platform, as in the case of two games reaching for the same consumers on the same console (Boudreau and Hagiu 2009; Markovich and Moenius 2009). Strategy becomes vastly more complex as firms consider dynamic interactions of a multi-layered ▶ business ecosystem (Teece 2012).

Competition between platforms tends towards winner-take-all concentration in the context of large demand or supply economies of scale, high multi-homing costs and the absence of niche specialization (Eisenmann et al. 2006). Demand economies of scale, equivalently network effects, favour the firm with stronger feedback as users attract users. Higher multi-homing costs reduce a user's ability to straddle two platforms, forcing a choice of one platform. Niche specialization can facilitate survival despite network effects and multi-homing costs as in the case of Apple's retreat into graphic design for desktop operating systems during the 1990s.

Platform-to-platform competition is one of the principal drivers of openness as each platform seeks to recruit more allied developers (Chesbrough 2003; West 2003). Competition among platforms in turn interacts with governance as this motivates the division of rents and the level of R&D spillovers that drive innovation (Parker and Van Alstyne 2013). Strategy also suggests courting large marquee partners who can bring technology or large blocks of users (Cusumano and Gawer 2002; Eisenmann et al. 2006) and, at the same time, deny these resources to competing platforms.

Large platforms can enter smaller markets protected by network effects and switching costs via the process of 'platform envelopment' (Eisenmann et al. 2006, 2011). An attacking platform bundles product features that exist on a target platform, that are new to the attacking platform and that exhibit a high degree of user overlap. Because the attacking platform already has a larger user base, the problem of mobilizing one side of the network is substantially solved.

Stronger network effects of the larger network typically win market share. For example, when Microsoft enveloped Real's audio streaming technology into Windows, Microsoft enjoyed more than 90 % market share in desktop operating systems. Each new release of Windows caused Real Audio to lose streaming market share among content consumers and content creators because neither would pay the incremental cost of Real's now duplicate functionality (Eisenmann et al. 2011). Platform entrants can also overcome an incumbent's advantage in network effects if quality is sufficiently great or users' willingness to wait for a stream of new quality is sufficiently high (Zhu and Iansiti 2012). Entrants can also seek to avoid direct platform competition by identifying distinctive and underserved user segments (Suarez and Kirtley 2012).

Managing the partner-to-platform competition and partner-to-partner competition is an even more dynamic problem owing to the need to cooperate as well as compete. Eisenmann (2008) draws an important distinction between the platform 'sponsor' and the platform 'provider'. A sponsor controls rights to the technology and designs the platform rules. The provider has the direct customer relationship. For example, Microsoft and Google sponsor their respective Windows and Android platforms while HP and Samsung provide the hardware that customers use to experience the platform. Platform sponsors often engender competition among providers who deal with customers to increase the affiliated user base. Greater openness at the sponsor layer does, in fact, accelerate the rate at which platform providers ship new products (Boudreau 2010). Cusumano and Gawer (2002) provide a number of strategies that platforms use to motivate and cope with external complementors: (i) platform standards should remain open in order that complementors continue to invest, (ii) the platform owner should not play favourites with news or surprise complementors with changes in strategy, (iii) the interests of partners should be treated fairly relative to interests of the leading platform firm. Intel, for example, created a separate internal division to represent goals of partners at a level equal to that of other internal divisions, (iv) platforms can share risk by investing along with partners in uncertain innovations, (v) the platform should promote the long-term financial health of partners, especially smaller ones. Boudreau and Hagiu (2009) further advise platform firms to, (vi) devise rules that balance partner inducements to participate against self-enrichment and (vii) limit the financial competition among complementors to promote investment but encourage prestige competition among individuals to promote effort. A platform may allow increased competition among supply-side partners if it can separately charge for access on its demand side. But if free entry occurs on the demand side, platforms prefer to limit supply-side competition (Armstrong 2006). If users on one side of the market single-home but suppliers on the other side multi-home, then a 'competitive bottle-neck' exists. Suppliers compete away their profits such that platform decisions typically favour the single-homing users (Armstrong 2006). Individual suppliers can lose buyers vis-à-vis one another but still gain overall if their collective investments cause their host platform to win market share (Markovich and Moenius 2009).

Platforms prefer to limit competition from transactions that partners take off-platform. For example, the 'no surcharge' rule forbids merchants from charging buyers higher prices for use of the platform's credit card relative to cash or other cards (Rochet and Tirole 2002; Wright 2003). Similarly, app stores impose a 'most favoured nation' rule that forbids developers from charging less when selling directly or through competing platforms. Platform strategy resembles traditional strategy in the manner that three-dimensional chess resembles the standard game (Eisenmann et al. 2011). Firms negotiate dynamic multi-layered trade-offs from platform to platform, from platform to partner, and from partner to partner.

See Also

- ▶ Business Ecosystem
- ▶ Governance
- ▶ Innovation Networks

- ▶ Network Effects
- ▶ N-Sided Markets
- ▶ Open Innovation
- ▶ Platform Innovation
- ▶ Winner-Take-All Markets

References

Anderson, E.G., G.G. Parker, and B. Tan. 2014. Platform performance investment in the presence of network externalities. *Information Systems Research* 25: 152–172.

Armstrong, M. 2006. Competition in two-sided markets. *RAND Journal of Economics* 37: 668–691.

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

Boudreau, K. 2010. Open platform strategies and innovation: Granting access versus devolving control. *Management Science* 56: 1849–1872.

Boudreau, K.J. 2012. Let a thousand flowers bloom? An early look at large numbers of software app developers and patterns of innovation. *Organization Science* 23: 1409–1427.

Boudreau, K., and A. Hagiu. 2009. Platform rules: Multisided platforms as regulators. In *Platforms, markets, and innovations*, ed. A. Gawer. Cheltenham: Edward Elgar.

Caillaud, B., and B. Jullien. 2003. Chicken and egg: Competition among intermediation service providers. *RAND Journal of Economics* 34: 309–328.

Cerf, V. 2012. Remarks at the digital broadband migration: The dynamics of disruptive innovation: Internet speculations. *Journal on Telecommunications and High Technology Law* 10: 21.

Chesbrough, H.W. 2003. *Open innovation: The new imperative for creating and profiting from technology.* Boston: Harvard Business Press.

Cusumano, M.A., and A. Gawer. 2002. The elements of platform leadership. *MIT Sloan Management Review* 43: 51–58.

Cusumano, M.A., and K. Nobeoka. 1992. Strategy, structure and performance in product development: Observations from the auto industry. *Research Policy* 21: 265–293.

Edwards, S. 2001. Openness, productivity and growth: What do we really know? *The Economic Journal* 108: 383–398.

Eisenberg, M.A. 1976. Private ordering through negotiation: Dispute-settlement and rulemaking. *Harvard Law Review* 89: 637–681.

Eisenmann, T. 2008. Managing proprietary and shared platforms. *California Management Review* 50: 31–53.

Eisenmann, T., G. Parker, and M. Van Alstyne. 2006. Strategies for two-sided markets. *Harvard Business Review* 84: 92–101.

Eisenmann, T., G. Parker, and M. Van Alstyne. 2009. Opening platforms: How, when and why? In *Platforms, markets, and innovations*, ed. A. Gawer. Cheltenham: Edward Elgar.

Eisenmann, T., G. Parker, and M. Van Alstyne. 2011. Platform envelopment. *Strategic Management Journal* 32: 1270–1285.

Ellison, N.B., C. Steinfield, and C. Lampe. 2007. The benefits of Facebook 'friends': Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication* 12: 1143–1168.

Evans, D.S. 2003. Antitrust economics of multi-sided platform markets. *Yale Journal on Regulation* 20: 325–381.

Evans, D.S. 2012. Governing bad behavior by users of multisided platforms. *Berkeley Technology Law Journal* 27: 1201–1250.

Evans, D.S., and R. Schmalensee. 2010. Failure to launch: Critical mass in platform businesses. *Review of Network Economics* 9: 1–26.

Evans, D.S., A. Hagiu, and R. Schmalensee. 2006. *Invisible engines: How software platforms drive innovation and transform industries*. Cambridge, MA: The MIT Press.

Garud, R., S. Jain, and P. Tuertscher. 2008. Incomplete by design and designing for incompleteness. *Organization Studies* 29: 351–371.

Gawer, A., and R. Henderson. 2007. Platform owner entry and innovation in complementary markets: Evidence from Intel. *Journal of Economics & Management Strategy* 16: 1–34.

Hagiu, A. 2006. Pricing and commitment by two-sided platforms. *RAND Journal of Economics* 37: 720–737.

Hagiu, A., and T. Eisenmann. 2007. A staged solution to the catch-22. *Harvard Business Review* 85: 25–26.

Hagiu, A., and D. Spulber. 2013. First-party content and coordination in two-sided markets. *Management Science* 59: 933–949.

Hagiu, A., and J. Wright. 2011. *Multi-sided platforms*. Harvard Business School Working paper No. 12-024.

Huang, P., M. Ceccagnoli, C. Forman, and D.J. Wu. 2013. Appropriability mechanisms and the platform partnership decision: Evidence from enterprise software. *Management Science* 59: 102–121.

Jackson, T. P. 1999. *Us v. microsoft: Findings of fact*. Technical Report, Civil Action No. 98-1232.

Jacobides, M.G., and S. Billinger. 2006. Designing the boundaries of the firm: From make, buy, or ally to the dynamic benefits of vertical architecture. *Organization Science* 17: 249–261.

Kauffman, R.J., and H. Mohtadi. 2004. Proprietary and open systems adoption in e-procurement: A risk-augmented transaction cost perspective. *Journal of Management Information Systems* 21: 137–166.

Landsman, V., and S. Stremersch. 2011. Multihoming in two-sided markets: An empirical inquiry in the video game console industry. *Journal of Marketing* 75: 39–54.

Liu, H. 2010. Dynamics of pricing in the video game console market: Skimming or penetration? *Journal of Marketing Research* 47: 428–443.
Markovich, S., and J. Moenius. 2009. Winning while losing: Competition dynamics in the presence of indirect network effects. *International Journal of Industrial Organization* 27: 346–357.
Nalebuff, B. 2004. Bundling as an entry barrier. *Quarterly Journal of Economics* 119: 159–187.
Nocke, V., M. Peitz, and K. Stahl. 2007. Platform ownership. *Journal of the European Economic Association* 5: 1130–1160.
O'Reilly, T. 2010. Government as a platform. In *Open government*, ed. Daniel Lathrop and Laurel Ruma. Sebastopol: O'Reilly Media.
Parker, G., and M. Van Alstyne. 2000a. Information complements, substitutes, and strategic product design. *Proceedings of the Twenty-First International Conference on Information Systems, Association for Information Systems*, 13–15.
Parker, G., and M. Van Alstyne. 2000b. Internetwork externalities and free information goods. *Proceedings of the Second ACM conference on Electronic Commerce, Association for Computing Machinery*, 107–116.
Parker, G., and M. Van Alstyne. 2013. Innovation, openness, and platform control. Working paper. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1079712. Accessed 5 May 2014.
Parker, G., and M. Van Alstyne. 2005. Two-sided network effects: A theory of information product design. *Management Science* 51: 1494–1504.
Parker, G., and M. Van Alstyne. 2009. Six challenges in platform licensing and open innovation. *Communications & Strategies* Q2: 17–36.
Penenberg, Adam L. 2009. *Viral loop: From Facebook to Twitter, how today's smartest businesses grow themselves*. New York: Hyperion.
Rochet, J.-C., and J. Tirole. 2002. Cooperation among competitors: Some economics of payment card associations. *RAND Journal of Economics* 33: 549–570.
Rochet, J.-C., and J. Tirole. 2003. Platform competition in two-sided markets. *Journal of the European Economic Association* 1: 990–1029.
Rysman, M. 2009. The economics of two-sided markets. *Journal of Economic Perspectives* 23: 125–143.
Shapiro, C., and H.R. Varian. 1999. *Information rules: A strategic guide to the network economy.* Boston: Harvard Business Press.
Stabell, C.B., and Ø.D. Fjeldstad. 1998. Configuring value for competitive advantage: On chains, shops, and networks. *Strategic Management Journal* 19: 413–437.
Strahilevitz, L.J. 2006. Information asymmetries and the rights to exclude. *Michigan Law Review* 104: 1835–1898.
Suarez, F., and J. Kirtley. 2012. Dethroning an established platform. *Sloan Management Review* 53: 35–41.
Sun, M., and E. Tse. 2009. The resource-based view of competitive advantage in two-sided markets. *Journal of Management Studies* 46: 45–64.
Teece, D.J. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* 15: 285–305.
Teece, D.J. 2012. Next generation competition: New concepts for understanding how innovation shapes competition and policy in the digital economy. *Journal of Law Economics and Policy* 9: 97–118.
Thomson, J. 1967. *Organizations in action*. New York: McGraw-Hill.
Tripsas, M. 2001. Adobe systems, inc. *Harvard Business School Publishing*. Case number 9-801-199, 1–29.
West, J. 2003. How open is open enough? Melding proprietary and open source platform strategies. *Research Policy* 32: 1259–1285.
West, J., and M. Mace. 2010. Browsing as the killer app: Explaining the rapid success of Apple's iPhone. *Telecommunications Policy* 34: 270–286.
Williamson, O.E. 2002. The lens of contract: Private ordering. *American Economic Review* 92: 438–443.
Wright, J. 2003. Optimal card payment systems. *European Economic Review* 47: 587–612.
Ye, G., R.L. Priem, and A.A. Alshwer. 2012. Achieving demand-side synergy from strategic diversification: How combining mundane assets can leverage consumer utilities. *Organization Science* 23: 207–224.
Yoo, Y., R.J. Boland, K. Lyytinen, and A. Majchrzak. 2012. Organizing for innovation in the digitized world. *Organization Science* 23: 1398–1408.
Zhu, F., and M. Iansiti. 2012. Entry into platform-based markets. *Strategic Management Journal* 33: 88–106.

Polycentric Staffing

Dawn E. Eash
Berkeley Research Group, LLC, Emeryville, CA, USA

Definition 'Polycentric staffing' is an organizational structure wherein foreign subsidiaries are locally managed by host-country nationals while corporate positions are likewise filled with home-country nationals.

'Polycentric staffing' is an organizational structure wherein foreign subsidiaries are managed by host-country nationals and corporate positions are filled with home-country nationals. The term 'polycentric' was originally coined by Perlmutter (1969) to describe companies with a host-country orientation as opposed to ethnocentric (home-country orientation) or geocentric (world orientation) policies.

Heenan and Perlmutter (1979) later identified a fourth 'regiocentric' orientation. Polycentric staffing policies have historically been prevalent within multinational organizations although more recent research suggests a decreasing downward trend (Cooper et al. 2007). Polycentric staffing can be the optimal staffing approach when there is significant national asymmetry (e.g., cultural differences, unique consumer bases, or specialized host-country regulations).

Advantages of Polycentric Staffing

Host-country nationals who are familiar with the language, culture, local business practices and political situation of their countries arguably ease interaction with employees, clients, and local officials. Second, employing host-country nationals could provide cost savings relative to paying for expatriate training, moving expenses, and higher salaries depending upon the host country of interest. Moreover, hiring local managers limits expatriate failure (i.e., the event where the expatriate does not successfully complete her assignment and thus returns to the home country) and the associated costly expenses (Cooper et al. 2007). In addition, host countries may have hiring quotas for ▶ multinational corporations (MNCs) that a polycentric staffing approach would help to satisfy.

Disadvantages of Polycentric Staffing

On the other hand, hiring local managers instead of transferring parent country nationals can lead to the isolation of foreign operations without adequate interaction and oversight from corporate headquarters. This gap can be exacerbated by cultural and political differences, language barriers and occasional cross-cultural interaction with other foreign subsidiaries. Over time this could create difficulty in aligning consistent competencies, goals and activities within the company. While this approach may align with a company's overall strategy, it will inevitably lead to a less unified business overall.

Any gap between headquarters and the foreign subsidiary can also cause undiagnosed duplication of efforts which can waste both time and money. Similarly, isolating home-country nationals from parent-country nationals reduces the opportunity for internal synergy and knowledge transfers. Thus the MNC's effort to adapt to a host country's culture and language may cause needless duplication and complexity while sacrificing the MNC's global progress and/or policies (Perlmutter 1969).

Alternatives to Polycentric Staffing

There are three other principal international staffing approaches: ethnocentric, geocentric and regiocentric staffing. Ethnocentric staffing is an international staffing approach in which foreign subsidiaries are managed by parent-country nationals. ▶ geocentric staffing is an international staffing policy in which the best person is appointed to manage foreign subsidiaries regardless of nationality. Regiocentric staffing policy is defined as the practice of transferring managers on a regional basis as a compromise between pure polycentric, ethnocentric or geocentric approaches.

Conclusions

Polycentric staffing is useful when the advantages of having a host-country national's cultural expertise outweigh the disadvantages of the foreign subsidiary functioning as an autonomous business entity. To the extent that an MNC is pursuing a multinational strategy in foreign countries with strong barriers to entry, polycentric staffing in its pure form may be appropriate. However, as the global marketplace and communication technology expand, approaches to staffing policy have also expanded to include flexible and virtual staffing approaches (Farndale et al. 2010). Accordingly, MNCs should select and combine the staffing policies that best achieve their overall globalization approach.

See Also

- ▶ Geocentric Staffing
- ▶ Home–Host Country
- ▶ Multinational Corporations

References

Cooper, D., L. Doucet, and M. Pratt. 2007. Understanding 'appropriateness' in multinational organizations. *Journal of Organizational Behavior* 28: 303–325.

Farndale, E., H. Scullion, and P. Sparrow. 2010. The role of the corporate HR function in global talent management. *Journal of World Business* 45: 161–168.

Heenan, D., and H. Perlmutter. 1979. *Multinational organization development*. New York: Addison-Wesley.

Perlmutter, H. 1969. The tortuous evolution of the multinational corporation. *Columbia Journal of World Business* 4: 9–18.

Porter, Michael E. (Born 1947)

Kathryn Rudie Harrigan
Columbia University, Business School,
New York, NY, USA

Abstract

Michael Porter's early interpretations of research findings from industrial organization economics transformed how the world thinks about strategy. Subsequent publications offered original frameworks and prescriptions for improving corporate as well as national economic performance.

Biographical Notes

Michael E. Porter, Bishop William Lawrence University Professor at Harvard University, heads the Institute for Strategy and Competitiveness at the Harvard Business School. In addition to his MBA, degrees in aeronautical engineering (Princeton) and economics (Harvard) equipped Porter to publish on strategy topics in *American Economic Review*, *Quarterly Journal of Economics*, *Review of Economics and Statistics* and other journals that increased the perceived gravity of the field of strategic management within academic institutions. His friendship with Richard E. Caves and A. Michael Spence motivated Porter to create a comparative advantage framework that resulted in national competitiveness assignments for him throughout the world. In 1983 Porter and five other entrepreneurs founded the Monitor Group consulting firm to tackle these and other strategy assignments.

How Porter Changed Strategic Management

Porter's major contribution to strategic management was his syntheses of theories from industrial organization economics, evolutionary economics and international economics in forms that were useful to managers for thinking about strategy formulation. Strategy research in the 1970s embraced the Learned–Christianson–Andrews–Guth framework which treated the exogenous, competitive environment as a black box; meanwhile industrial organization economists were using case studies to unravel the laws of the marketplace and the normative structure-conduct-performance paradigm to evaluate industry outcomes. Briefly, their frameworks focused on industries instead of individual firms.

Porter's *Competitive Strategy* (1980) integrated his doctoral-level research on bilateral bargaining power and customer switching-cost barriers (Porter 1976) with prevailing economic theories concerning entry barriers, non-price competition and cross-elasticity of substitutes to create the Five Forces model of industry profitability potential. Like the industrial organization economics studies before Porter, the Five Forces model predicted the average profitability of an industry. But with its emphasis on considering the uniqueness of a particular firm's strategic posture (or that of the ▶ strategic groups with which it competed), the possibility was created of devising a strategy that outperformed the industry's average profitability. Such a firm was one that achieved ▶ competitive advantage.

While dissecting what types of strategic postures might enable a firm to outperform an industry's average profitability, Porter (1980) introduced three generic strategies that might be appropriate to use in certain types of industry structures. For example, the 'differentiation' strategy could be used when products and services were not commodity-like. The ▶ focus strategy could be used where particular customer demands were underserved. The '▶ cost leadership' strategy was the most durable of Porter's generic approaches to competition and was expanded upon by numerous researchers that explored the cost-reduction possibilities of a set of activities that were necessary to serve customers well (Porter 1996).

Porter (1980) recognized the dynamic nature of competition in his framework of competitive evolution that characterized (1) embryonic environments where multiple product configurations (and technical standards) coexisted and competition focused on developing product demand, (2) emerging environments where customer's expectations favoured product configurations that were becoming the dominant approaches to competition and firms focused on satisfying customer demand that often grew faster than vendors' abilities to supply products and services, (3) established environments where demand was mature and competitors focused on taking market share from each other, and (4) end-game environments where competitors pondered how long to continue serving remaining customers. Porter's interpretation of evolutionary economics suggested how an industry's profitability potential changed as the elements of his five forces model evolved, as well as how exogenous forces changed the attractiveness of investment opportunities within them.

Where *Competitive Strategy* (Porter 1980) examined relationships between vertically related business units (aka vertical integration), Porter's *Competitive Advantage* (1985) emphasized the internal value-chain decisions undertaken by a single business unit (or undiversified firm) to improve a customer's willingness to pay for their products and services. Paramount among the decisions covered in developing the value-chain framework are choosing which actions can be taken to reduce costs in anticipation of customers' inevitable expectations of receiving increasing value over time from their vendors. Such analyses pre-dated most other studies of how firms might best use outsourcing and strategic alliances in their mix of chosen activities.

In *Competitive Advantage* (Porter 1985) the decomposition of a business unit's value chain facilitates Porter's expanded enquiry into the activity set that firms engage in to attain competitive advantage and suggests a range of organizational forms by which the firm's value chain might be completed. Porter's discussion of the drivers by which firms gain cost advantages illustrates the critical need to make trade-offs in strategy – for example, how differentiation undercuts efforts to reduce costs – in order to attain their goal of creating value for their customers.

In *Competitive Advantage* (1985) Porter provides a basis for determining how interrelationships between business units can best exploit the shared services that a diversified firm might provide from a central facility in their quest to create value for customers. His application of the value chain to realizing synergies in diversification strategies was not fully developed until Porter added consideration of geographic location to his overarching framework with *The Competitive Advantage of Nations* (1990). Although many of the concepts introduced in *Competitive Strategy* (Porter 1980) are taught in business school classes, the concepts introduced in *Competitive Advantage* (Porter 1985) have influenced corporate business activities and opportunities for consulting engagements more significantly.

Porter's enquiry into *The Competitive Advantage of Nations* (1990) introduced the Strategic Diamond Model and his prescription for creating clusters of competence for regional economic development. Porter's Diamond Model expands upon the Ricardian theory of a nation's comparative advantage in trade and suggests how nurturing domestic, industrial clusters of competence through government assistance may improve a country's comparative advantage. The Strategic Diamond Model significantly influenced the economic development policies of several nations, as well as thinking about international strategy for multinational corporations. Briefly, the economic

cluster strategy combined Porter's ideas about transactions between firms within vertically related industries with his ideas about coordination among a firm's geographically diverse business units. These ideas were embraced by policymaker because the theory identified those traits which made geographic venues relatively more or less attractive sites for locating companies' facilities.

Although Porter has influenced thinking about urban communities, environmental policy and the role of corporations in society, his framework with Teisberg in *Redefining Health Care* (2006) has achieved the greatest public policy consideration. Porter's proposal for health care delivery has dramatic implications for the structure by which health care providers, employers, governments, health care plan administrators and other actors ensure the well-being and productivity of a firm's employees.

See Also

- ▶ Competitive Advantage
- ▶ Competitive Strategy
- ▶ Cost Leadership
- ▶ Exit Barriers
- ▶ Five Forces Framework
- ▶ Focus Strategy
- ▶ Generic Strategy
- ▶ Global Strategy
- ▶ Management Gurus
- ▶ Mobility Barrier Permeability
- ▶ Strategic Groups

Selected Works

Porter, M.E. 1976. *Interbrand choice, strategy and bilateral market power*, Harvard economic studies, vol. 146. Cambridge, MA: Harvard University Press.

Porter, M.E. 1980. *Competitive strategy: Techniques for analyzing industries and competitors*. New York: Free Press.

Porter, M.E. 1985. *The competitive advantage: Creating and sustaining superior performance*. New York: Free Press.

Porter, M.E. 1990. *The competitive advantage of nations*. New York: Free Press.

Porter, M.E. 1996. What is strategy? *Harvard Business Review* 74: 61–78.

Porter, M.E., and E.O. Teisberg. 2006. *Redefining health care: Creating value-based competition on results*. Boston, MA: Harvard Business School Press.

Porter, M.E., R.E. Caves, A.M. Spence, and J.T. Scott. 1977. *Studies in Canadian industrial organization*. Toronto: Canadian Commission on Corporate Concentration.

Porter, M.E., R.E. Caves, and A.M. Spence. 1980. *Competition in an open economy: A model applied to Canada*, Harvard economic studies, vol. 150. Cambridge, MA: Harvard University Press.

Portfolio Planning: A Valuable Strategic Tool

Lance R. Newey[1] and Shaker A. Zahra[2]
[1]University of Queensland, Business School, St Lucia, QLD, Australia
[2]University of Minnesota, Minneapolis, MN, USA

Definition Portfolio planning is the decision process that translates strategies into investment decisions in the form of projects that will deliver the short-, medium- and long-term performance of a company's strategic goals.

What Is It?

The successful execution of a company's strategy requires transforming plans into concrete action by selecting different activities and allocating resources for them (Bower and Gilbert 2005). Portfolio planning is the decision process that translates strategies into investment decisions in the form of projects that will deliver the short-, medium- and long-term performance of a company's strategic goals (Cooper et al. 2001; Patterson 2008). Typically, there are two levels of portfolio planning: strategic and tactical. The strategic level involves deciding the relative emphasis between investments across short-, medium- and long-term time horizons. The tactical level then determines the specific new business/product projects in which the company should invest in each of these strategic time horizons. As strategies are implemented in an uncertain world, risk is diversified in portfolio

planning by investing in projects with different time horizons (some short term, some long), degree of product change (e.g., incremental product modifications versus more radical technology forecasting), degree of risk (some low, some high) and type of markets (e.g., existing customers, new customers, new geographic areas). Portfolio planning thus helps build sustainable value by making use of multiple skills that include forecasting and options management, as well as business and product development.

Why Do Companies Need Portfolio Planning?

Companies that do not practise portfolio planning often show inadequate attention when extending and defending core businesses of today or building entirely new emerging businesses and longer-term options. These companies often take on too many new product projects, and many of these projects lack clear connection to strategy (Christensen 1997; Cooper 2011). As a result, their staffs are overstretched, projects fail to finish and reach the market on time, and the firm's costs begin to outpace revenue growth. Portfolio planning helps to avoid these problems by instilling discipline and focus. It also ensures the effective alignment between strategy formulation, resource allocation and project selection. Reaching this alignment requires communication and collaboration across organizational units, thereby improving the odds of the successful execution of the strategy. Portfolio planning enables key decision-makers to articulate and share their views of the firm's environment, internal capabilities and resources, and organizational bottlenecks that might stifle strategy implementation. Portfolio planning fosters coherence in a company's strategic decisions and the actions they follow in their respective market arenas.

How Do Companies Undertake Portfolio Planning?

Frequently, companies start with a growth strategy, which must work out how much of the planned growth will come from the existing business and how much needs to come from new businesses and products (Baghai et al. 2000). Staff conducting portfolio planning then select the key new business and product development projects in which their company will invest in pursuit of the growth strategy. As these new projects are surrounded by uncertainty, each individual project proceeds through a series of stages punctuated by gates. As better evidence comes in, business and product development projects that appear promising remain in the portfolio while others may be abandoned or put on hold. Portfolio planning is thus a dynamic decision process that is often performed several times a year, as senior management continually update and evaluate how well their overall new business and/or product portfolio is working (Cooper 2011). The need for a portfolio approach to strategy execution occurs at business and ▶ corporate strategy levels.

Business Strategy Level

Each of a company's individual business units typically plans its portfolio. Business units start by deciding how much of their budget will be spent on defending/growing the current business. A portion is then allocated to new product projects. To plot all these investments, business units often develop a pie chart with three segments, either short term, medium term, long term, or product change, geographic markets and product categories (Cooper et al. 2000). Each segment of the pie chart is known as a strategic bucket – a bucket of money which is allocated to that segment (Cooper et al. 2001). In each segment, by using both qualitative and quantitative criteria, the company scores the new product opportunities that it perceives are available to it. The projects are then rank-ordered from best to worst and the available bucket of money is allocated to projects until the bucket is totally allocated. These allocations reflect the company's chosen corporate strategy.

Corporate Strategy Level

At the corporate level, companies need to determine the right mix of whole business units that should make up the company. Looking to the future, companies need to make investments in

longer-term options, such as in multiple competing technologies (Baghai et al. 2000). As these technologies mature and develop the right option for further investment is clarified. If it is a totally new technology, it has the potential to replace/complement the existing businesses in the corporate portfolio. They move out of being a technology option and are now invested in as an emerging business. When fully developed in terms of scale and return, the emerging business may become a core business, and so the firm learns how to dynamically evolve its businesses and products.

Making Portfolio Planning a Success

Clearly, portfolio planning can be a valuable tool in mapping a company's strategic choices and corresponding investments. To be successful, however, companies need to ensure the collaboration and input of different units and ensure mutual understanding while safeguarding against excessive bureaucracy (Newey and Zahra 2009). Attention should focus on creating and maintaining synergies among the various strategies to be followed, thereby generating value. Top management should also make sure that portfolio planning is based on accurate and realistic predictions about the firm's opportunities, capabilities and resources.

See Also

- ▶ Business Development
- ▶ Business Strategy
- ▶ Corporate Strategy
- ▶ Resource Allocation Theory
- ▶ Strategic Risk Management

References

Baghai, M., S. Coley, and D. White. 2000. *The alchemy of growth: Practical insights for building the enduring enterprise*. New York: Basic Books.

Bower, J.L., and C.G. Gilbert. 2005. *From resource allocation to strategy*. New York: Oxford University Press.

Christensen, C.M. 1997. *Improving the product development process at Kirkham Instruments Corporation. HBS Note 9-697-058*. Boston: Harvard Business School Press.

Cooper, R.G. 2011. *Winning at new products: Creating value through innovation*. New York: Basic Books.

Cooper, R.G., S.J. Edgett, and E.J. Kleinschmidt. 2000. New problems, new solutions: Making portfolio management more effective. *Research Technology Management* 43: 18–33.

Cooper, R.G., S.J. Edgett, and E.J. Kleinschmidt. 2001. *Portfolio management for new products*, 2nd ed. New York: Basic Books.

Newey, L.R., and S.A. Zahra. 2009. The evolving firm: How dynamic and operating capabilities interact to enable entrepreneurship. *British Journal of Management* 20: S81–S100.

Patterson, M.L. 2008. New product portfolio planning and management. In *The PDMA handbook of new product development*, ed. K. Kahn. Hoboken: Wiley.

Positioning

Sijun Wang
Loyola Marymount University, Los Angeles, CA, USA

Abstract

Positioning concept has evolved in two distinctive yet related research streams – brand positioning and strategic positioning. Strategic positioning sets the basic direction for the development of the brand positioning. Brand positioning can be based on various brand associations. Research to date has provided various space graphics and mathematical modelling techniques to locate the optimal brand position in a target market. A firm's strategic positioning was found to be positively associated with its long-term performance. Firms are advised to pursue a match between market demands and capability profiles to build positional advantages in the market.

Definition Positioning can be directed either at a firm's offerings (brand, product and service) or at the firm's overall strategic standing. The former is referred to as brand positioning, defined as the act of designing the company's offering and image to

occupy a distinctive place in the mind of the target market, while the latter as strategic (market) positioning, defined as the competitive market standing of a firm against its competitors.

Historical Developments

Positioning concept has evolved from ▶ market segmentation, targeting and ▶ market structure changes during the 1960s and the early 1970s. Ries and Trout (1981) popularized the term through their work and it was later adopted by marketing strategists (e.g., Kotler 2000). Technically speaking, positioning can be directed towards product, a service, a ▶ brand, a company, an institute, an idea or even a person. Two distinctive yet related streams of literature have centred on brand (product/service) positioning and strategic (market) positioning (DiMingo 1988). Strategic positioning sets the basic direction for the development of the brand positioning.

Brand (Product/Service) Positioning

Kotler (2003: 308) defines brand positioning as 'the act of designing the company's offering and image to occupy a distinctive place in the mind of the target market. The end result of positioning is the successful creation of a customer-focused value proposition, a cogent reason why the target should buy the product.'

In principle, companies can position their brands on an infinite number of brand associations. Fuchs and Diamantopoulos (2010) classified the positioning bases into five types of brand-positioning strategies, namely concrete attributes (e.g., horsepower and air conditioning for automobiles), abstract attributes (e.g., style, sporty and fast acceleration), direct (functional) benefits (e.g., cost reduction, ease of use), indirect (experiential/symbolic) benefits (e.g., making the driver feel younger) and surrogate positioning (designed to create consumer associations about external aspects of the brand (e.g., user image)).

Essentially, the aim of brand (product/service) positioning is to locate a brand in a target customer's 'perceptual space', in relation to its competitors. Various perceptual mapping techniques have been utilized to operationalize the perceptual distance between two brands. According to Keon (1983), four analytic methods are adopted to evaluate or identify a brand's positioning in the market: multidimensional scaling (DeSarbo et al. 1997), factor analysis, discriminant analysis (Huber and Holbrook 1979) and multi-attribute composition models (conjoint analysis is commonly used in this category) (Green and Srinivasan 1990). Gwin and Gwin (2003) suggest that each method has strengths and limitations and that no single method outperforms the others in all positioning situations.

Strategic Positioning

Strategic (market) positioning refers to the competitive market standing of a firm against its competitors (Porter 1979). The relationship between positioning and the long-term growth of the firm is asserted by Porter (1996) and empirically supported by several studies (e.g., Brooksbank 1994).

A debate on the relevance of the two 'opposing' perspectives on positioning strategy has carried on throughout the 1990s; the market-orientation perspective suggests that superior market performance is realized through maintaining diligence in relation to market opportunities, industry structure, market intelligence and the delivery of unique offers (Grant 1991), while the resource-based view suggests that strong market performance is primarily realized through a focus on the utilization of historically developed resources and assets (Collis and Montgomery 1995). The debate has led to a view that goes some way to reconciling the two approaches. This is reflected in the contribution of Hooley et al. (1998), who state that by giving equal weight to market demands and capability profiles when selecting targets and implementing positioning strategies, firms can ensure an enduring match between their offerings and their markets. Extending the concept of strategic market positioning, the positioning concept has also been

extended to various domains, including tourist destinations (Crompton et al. 1992) and nations (e.g., Kotler et al. 1997).

See Also

- ▶ Brand
- ▶ Market Segmentation
- ▶ Market Structure

References

Brooksbank, R. 1994. The anatomy of marketing positioning strategy. *Marketing Intelligence & Planning* 12: 10–14.

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

Crompton, J., P. Fakeye, and C. Lue. 1992. Positioning: The example of the low Rio Grande Valley in the winter long stay destination market. *Journal of Travel Research* 31: 20–26.

DeSarbo, W.S., M.R. Young, and A. Rangaswamy. 1997. A parametric unfolding procedure for incomplete, non-metric preference/choice set data in marketing research. *Journal of Marketing Research* 34: 499–516.

DiMingo, E. 1988. The fine art of positioning. *Journal of Business Strategy* 9: 34–38.

Fuchs, C., and A. Diamantopoulos. 2010. Evaluating the effectiveness of brand-positioning strategies from a consumer perspective. *European Journal of Marketing* 44: 1763–1786.

Grant, R.M. 1991. The resource-based theory of competitive advantage: Implications for strategy formulation. *California Management Review* 33: 114–135.

Green, P., and V. Srinivasan. 1990. Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing* 54: 3–20.

Gwin, C.F., and C.R. Gwin. 2003. Product attributes model: A tool for evaluating brand positioning. *Journal of Marketing Theory and Practice* 11: 30–42.

Hooley, G., A. Broderick, and K. Moller. 1998. Competitive positioning and the resource-based view of the firm. *Journal of Strategic Marketing* 6: 97–115.

Huber, J., and M.B. Holbrook. 1979. Using attribute rating for product positioning: Some distinctions among compositional approaches. *Journal of Marketing Research* 16: 507–515.

Keon, J.W. 1983. Product positioning: Trinodal mapping of brand images, ad images, and consumer preference. *Journal of Marketing Research* 20: 380–392.

Kotler, P. 2000. *Market management*. Upper Saddle River: Prentice Hall.

Kotler, P. 2003. *Marketing management*, 11th ed. Englewood Cliffs: Prentice Hall.

Kotler, P., S. Jatusripitak, and S. Maesincee. 1997. *The marketing of nations: A strategic approach to building national wealth*. New York: Free Press.

Porter, M.E. 1979. How competitive forces shape strategy. *Harvard Business Review* 57: 137–145.

Porter, M.E. 1996. What is strategy? *Harvard Business Review* 74: 61–78.

Ries, A., and J. Trout. 1981. *Positioning: The battle for your mind*. New York: McGraw-Hill.

Post-Acquisition Management

Harbir Singh and Patia Mcgrath
University of Pennsylvania, Wharton School of Business, Philadelphia, PA, USA

Abstract

Although post-acquisition management marks the last stage of the ▶ acquisition process, it represents the first substantive opportunity to create value from the transaction. An effective post-acquisition process requires preparation in the very early stages of transaction, when the strategies to identify and complete the transaction are under development. It is thus imperative to have a clear understanding of the types of decisions that must be made during the post-acquisition management period, along with their associated ramifications for performance. The extent of integration between the acquirer and acquired firms represents a crucial choice and consequent focus of executives responsible for post-acquisition management. Theoretically, the ▶ resource-based view of the firm helps to surface key aspects of the integration dilemma. The knowledge-based view of the firm reveals that a firm's development and deployment of post-acquisition management acquisition capability can play a prominent role in achieving acquisition success.

Definition The post-acquisition management stage of the acquisition process follows the completion of the transaction. In this phase, executives of the newly combined firm make decisions in the areas of strategy, structure, systems and people

with the objective of achieving the transaction's goals and realizing its value.

In spite of their popularity and prevalence, ▶ acquisitions frequently fall short of their initial promise. Their disappointing results are reflected across a variety of performance measures, such as shareholder value, profitability and patent generation. Some estimates place acquisition failure rates as high as (if not higher) than 50 % (cf. Datta 1991; Marks and Mirvis 2001; McKinsey & Co. 2010). Numerous causes of these strikingly poor outcomes have been identified, with the likes of ill-fated strategic intentions, executive distraction, acquirer overpayment, cultural incompatibility and talent turnover commonly heading the list. Notably, many of these sources of variation in acquisition performance manifest during the final phase of the acquisition process, which is known as post-acquisition management. Accordingly, scholarly interest in postacquisition management has been (and continues to be) intense. Although much work remains, meaningful progress has been made towards decoding post-acquisition management's inherent complexities, and, consequently, furthering understanding of acquisition value generation.

Critical Choices in Post-Acquisition Management

Executives face a litany of consequential choices in the post-acquisition management stage, which may be broadly categorized into strategy, structure, systems and people. Strategy encompasses those decisions that will help actuate the promised value of the acquisition, be it through efficiency improvements, market power gains, or new competitive dynamics. Pragmatically, loss of strategic continuity between the pre- and post-acquisition stages is as damaging to firms as it is unfortunately repeated by them. Strategic initiatives and projects, which would ideally span multiple functions and be highly visible to the organization, can be used to not only bridge and align strategy across the pre- and post-acquisition periods, but also build momentum and firm-wide engagement with the new entity.

Next, choices concerning structure and systems must be made. Decisions around structure include those factors that involve the formal organization, especially its design and reporting relationships. Structure typically follows strategy, and thus serves as an enabler and enforcer of the strategy underpinning the acquisition and resultant organization. Structural decisions are notorious for their unintended repercussions; if proactive preparation is neglected, a seemingly simple choice to merge the acquired firm's R&D group into that of the acquirer could, for example, detrimentally affect the fulfillment of an innovation-based motivation for the acquisition. Systems decisions pertain to information, control and incentive processes. Across these systems, identification and tracking of strategically relevant metrics is an important priority. Leading indicators of performance (such as employee satisfaction and customer retention) are equally important to lagging indicators (such as return on investment and profitability), since leading indicators provide early signals of the acquisition's performance and could reveal the need for course-corrections.

Decisions about people comprise the fourth major category of post-acquisition management choices. The importance of this category should not be underestimated; many accounts of post-acquisition management place people at the heart of success or failure in acquisitions (Marks and Mirvis 2001). Colloquially stated, people-oriented decisions address 'who stays and who goes' in the combined entity and, for those who stay, whether and how acculturation should ensue. These choices must be navigated with great thoughtfulness and caution. Although realization of acquisition 'cost synergies' often involves employee downsizing, losing managerial talent from the acquired firm could cause the new entity to not only fail to achieve the expected value of the acquisition, but also not capitalize upon unanticipated benefits that might have arisen through the exchange of new ideas, surprisingly valuable technologies, and other unforeseen advantages (Graebner 2004). Further, realistic and frequent communications with employees about the acquisition and its implementation can help to build trust and collaboration between employees in the combined firm, while

mitigating the uncertainty and confusion that acquisitions invariably cause (Schweiger and DeNisi 1991; Ranft and Lord 2002).

Challenges and choices pertaining to people in post-acquisition management also include issues pertaining to culture. An acquisition of one organization by another presents the opportunity for 'culture clash' – or, more severely, debilitating conflict – between the two entities from the perspective of organizational (and, in the case of cross-border acquisitions, national) culture. The more incongruence that exists between the acquirer and acquired firm in terms of their acculturation preferences, the greater the resultant acculturative stress will result (Nahavandi and Malekzadeh 1988). These issues directly impact the success of the acquisition's implementation, and post-acquisition management represents a crucial time for proactively recognizing and attenuating the effects of such dangers. Furthermore, when addressed correctly, cultural differences can represent a valuable opportunity. Cross-border acquisitions, for instance, can provide the acquirer with access to a diversity of valuable routines and repertoires not previously available to the acquiring firm, with their novelty increasing with greater cultural distance (Morosini et al. 1998).

The Integration Decision

Perhaps the most fundamental and powerful issue in post-acquisition management is integration. Decisions concerning the degree and scope of integration between the acquirer and the acquired organization form a common underlying thread that is woven through the strategy, structure, systems and people categories. Indeed, integration is such a central issue on the post-acquisition management agenda that this stage is oftentimes dubbed 'post-acquisition integration'.

Viewing the integration decision through the lens of the resource-based view of the firm brings the dilemma facing managers into sharp relief. The resource-based view begins with the premise that there is heterogeneity among firms in terms of the resources – both tangible and intangible – that they control. Sustained competitive advantage is rooted in these resources and the strategies that firms can develop to leverage them (Wernerfelt 1984; Barney 1991; Peteraf 1993). When a firm acquires an organization, it is essentially acquiring a bundle of resources that may then be combined with its own (Wernerfelt 1984; Capron et al. 1998). Synergies are created when the resources of the acquirer and the acquired organization reinforce or complement each other to create greater value than would have been possible separately.

This therefore surfaces the characteristic tension of the integration decision. In order to build and fully capitalize on these potential synergies, an acquirer may believe that it must integrate the acquired firm. Integration would facilitate coordination, especially in cases where tacit, complex knowledge needs to be transferred from the acquired organization. Yet integration could stymie the hoped-for advantages from the acquired firm. Loss of the acquired organization's autonomy could be harmful on several fronts, such as through disruption of its routines, instigation of employee confusion and turnover, and impairment of its innovation capabilities. Tradeoffs posed by integration are well illustrated in technology acquisitions, such as the archetypal case of a large firm acquiring a smaller entrepreneurial organization for its technical prowess. While integration assists with the transfer of complex technical knowledge, integration can, at the same time, damage the innovative spirit and performance of the acquired organization that initially motivated the acquisition (Graebner 2004; Puranam et al. 2006, 2009).

Notably, the integration decision is not a binary one. Executives may opt for approaches anywhere along the spectrum from full integration to complete autonomy. A typology developed by Haspeslagh and Jemison (1991) weighs the need for strategic interdependence against the need for organizational autonomy to formulate integration recommendations. Integration approaches range from absorption (high interdependence, low autonomy) to symbiotic (high interdependence, high autonomy) to preservation (low interdependence, high autonomy); a 'holding' approach is recommended for cases where the need for interdependence and autonomy are both low. Furthermore, the integration decision may need to be

dynamic, so that the combined firm has the flexibility to respond to the ongoing challenges posed by post-acquisition management. For instance, there are benefits to transitioning from an initial autonomous approach to an integrative one as the technological trajectory of the acquired firm's product unfolds (Haspeslagh and Jemison 1991; Puranam et al. 2006). Additionally, the speed of integration is an important aspect of the decision, and the nature of its impact hinges upon the relatedness between the acquirer and the acquired organization. Speed of integration is beneficial under the combined condition of high product market and geographic relatedness and low internal relatedness (i.e., strategic orientation, management style and performance); speed of integration is detrimental when these conditions are reversed (Homburg and Bucerius 2006).

Lastly, assessment of the integration decision should include non-structural considerations. While purely formal, structurally-based design mechanisms are habitually used to drive integration, these blunt instruments are not the only option. One alternative is for the combined firm to leverage any 'common ground' that may exist or be fostered between the two entities (Kale et al. 2009; Puranam et al. 2009). Common ground encompasses the likes of shared knowledge, practices, values, and language. Should sufficient common ground be present between the acquirer and the acquired organization, this shared platform promotes informal coordination. In this way, common ground could complement or even substitute for structurally based integration procedures, thereby allowing the combined firm to reap the benefits of integration while avoiding some of its drawbacks.

Achieving Acquisition Success: The Role of Post-Acquisition Management Capability

Based on prior research on post-acquisition management, it is apparent that firms vary in their capability to effectively achieve the projected outcomes of their acquisition transactions. Recent work has focused on acquisition management capability as a means to unpack the drivers of interfirm variation in achieving post-acquisition outcomes. Much of this work on post-acquisition management capability is rooted in a knowledge-based perspective of the firm. The knowledge-based view of the firm, which is typically considered as related to the domain of ▶ dynamic capabilities, advocates knowledge as being the fundamental source of firm competitive advantage (Kogut and Zander 1992; Grant 1996). This knowledge, which may reside at the individual or organizational levels, may be recombined and integrated to give rise to organizational capabilities (Nelson and Winter 1982; Zollo and Winter 2002). In addition to post-acquisition management (and, even more specifically, integration) capability (Zollo and Singh 2004), firms may also develop capabilities directed towards other stages of the acquisition process and additional modes of corporate development (Kale et al. 2002; Sapienza et al. 2006; Capron and Mitchell 2009). More specifically, capabilities would be driven both by tacit, experience-based knowledge and by codified knowledge. Acquisition management capability has been shown to include a combination of tacit knowledge, embodied in the skills of key individuals involved in the process, and codified knowledge, embodied in routines that are invoked in the strategy phase, the due diligence phase, and the post-transactional phase of the process. These routines would include a cross-functional orientation in the due diligence phase, as well as specific routines in the post-transactional phase to coordinate various functions (such as human resources, operations and customer relations).

Some organizational capabilities may be tacitly developed through the firm's repetition and accumulation of its relevant experiences (Levitt and March 1988). However, this 'learning-by-doing' approach is insufficient for post-acquisition management capability development. Instead, due to the complexity of post-acquisition management, explicit codification of the firm's associated collected knowledge is a precondition for improving acquisition performance (Zollo and Singh 2004). Higher levels of codification – which may be instituted through such mechanisms as training manuals, information systems, or even

checklists – lead to higher levels of acquisition performance, especially in the face of more extensive integration between the acquirer and acquired firm. Additionally, routinization of these codified practices will further advance the development of post-acquisition management capability and performance improvement, as routinization helps the organization to anticipate the implementation challenges of post-acquisition management. Procedures relating to any of the categories of strategy, structure, systems and people are candidates for codification and routinization.

Conclusion

The resource-based, evolutionary and knowledge-based views of the firm provide useful perspectives by which to address the challenges and opportunities presented by post-acquisition management. While the post-acquisition management stage of the acquisition process poses many complex choices and tradeoffs, especially concerning the degree and scope of integration, it also offers palpable possibilities for value creation from the acquisition. The elements of acquisition management capability are becoming clearer, as is its potentially positive impact on both strategy and in its role in driving outcomes after completion of the transaction. By proactively building post-acquisition management capability, the combined firm can better position itself to reap the rewards of the acquisition transaction.

See Also

- ▶ Acquisition Strategy
- ▶ Dynamic Capabilities
- ▶ Organizational Learning
- ▶ Resource-Based View

References

Barney, J. 1991. Firm resources and sustained competitive advantage. *Journal of Management* 171: 99–120.

Capron, L., and W. Mitchell. 2009. Selection capability: How capability gaps and internal social frictions affect internal and external strategic renewal. *Organization Science* 20: 294–312.

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.

Datta, D.K. 1991. Organizational fit and acquisition performance: Effects of post-acquisition integration. *Strategic Management Journal* 12: 281–297.

Graebner, M.E. 2004. Momentum and serendipity: How acquired leaders create value in the integration of technology firms. *Strategic Management Journal* 25: 751–777.

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

Haspeslagh, P.C., and D.B. Jemison. 1991. *Managing acquisitions: Creating value through corporate renewal*, vol. 416. New York: Free Press.

Homburg, C., and M. Bucerius. 2006. Is speed of integration really a success factor of mergers and acquisitions? An analysis of the role of internal and external relatedness. *Strategic Management Journal* 27: 347–367.

Kale, P., J.H. Dyer, and H. Singh. 2002. Alliance capability, stock market response, and long-term alliance success: The role of the alliance function. *Strategic Management Journal* 23: 747–767.

Kale, P., H. Singh, and A.P. Raman. 2009. Don't integrate your acquisitions, partner with them. *Harvard Business Review* 87: 109–115.

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

Levitt, B., and J.G. March. 1988. Organizational learning. *Annual Review of Sociology* 14: 319–340.

Marks, M.L., and P.H. Mirvis. 2001. Making mergers and acquisitions work: Strategic and psychological preparation. *Academy of Management Executive* 15: 80–92.

McKinsey & Co. 2010. Perspectives on merger integration, *June* 2010, 1–52.

Morosini, P., S. Shane, and H. Singh. 1998. National cultural distance and cross-border acquisition performance. *Journal of International Business Studies* 29: 137–158.

Nahavandi, A., and A.R. Malekzadeh. 1988. Acculturation in mergers and acquisitions. *Academy of Management Review* 13: 79–90.

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

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

Puranam, P., H. Singh, and M. Zollo. 2006. Organizing for innovation: Managing the coordination-autonomy dilemma in technology acquisitions. *Academy of Management Journal* 49: 263–280.

Puranam, P., H. Singh, and S. Chaudhuri. 2009. Integrating acquired capabilities: When structural integration is (un)necessary. *Organization Science* 20: 313–328.

Ranft, A.L., and M.D. Lord. 2002. Acquiring new technologies and capabilities: A grounded model of acquisition implementation. *Organization Science* 13: 420–441.

Sapienza, H.J., E. Autio, G. George, and S.A. Zahra. 2006. A capabilities perspective on the effects of early internationalization on firm survival and growth. *Academy of Management Review* 31: 914–933.

Schweiger, D.M., and A.S. DeNisi. 1991. Communication with employees following a merger: A longitudinal field experiment. *Academy of Management Journal* 34: 110–135.

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

Zollo, M., and H. Singh. 2004. Deliberate learning in corporate acquisitions: Post-acquisition strategies and integration capability in US bank mergers. *Strategic Management Journal* 25: 1233–1256.

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

Potential Competition

L. G. Thomas III
Emory University, Goizueta Business School, Atlanta, GA, USA

Abstract

This entry places the concept of potential competition in industrial organization economics. Late twentieth-century economic analyses have brought theoretical rigour and empirical evidence for this concept. These years have also seen the emergence of alternative concepts in strategic management, and changes in the nature of competition among firms that relatively favours these alternate concepts. The importance and range of business activity that the concept of potential competition is expected to explain has thus narrowed, even as this concept has become more formal, more proven and more readily applicable.

Definition Potential competition for incumbent firms in an industry alters market outcomes before or without actual entry. The economic concept of potential competition is narrower than the mere possibility of entry by new firms. If there is no pre-entry effect, there is no potential competition, and only post-entry or actual competition occurs, if any at all.

Competition is consequential interaction among firms. Possible new firms in an industry may provide competition for incumbent firms before or without their actual entry. Alternately, possible new firms may have no pre-entry effect. Only the former case is potential competition.

Potential competition has been an analytical tool for industrial organization economics. This approach offers an explanation for the differential performance of firms and industries through analysis of market structure. Formally, incumbent firms exploit structural features of industries to earn above-normal profits. In particular, incumbents strategically deter entry of new firms, and entry deterrence has been the focal preentry effect studied with analysis of potential competition. Note that the immediate details of potential competition are in service of the larger agenda of understanding patterns of firm performance.

Incumbent Commitment and Potential Competition

The modern theory of industrial organization economics identifies credible commitment by incumbent firms as the mechanism linking possible new firms and pre-entry effect (Tirole 1988: ch. 8; Gilbert 1989). If an incumbent firm can undertake pre-entry action with irreversible post-entry impact, then potential competitors may reduce their scale of entry or be deterred from entry altogether.

Spence (1977) examines capacity expansion as a commitment mechanism. By increasing durable preentry capacity, incumbent firms lower future cost and price. If the future price is sufficiently low, entry is discouraged. The irreversible nature of incumbent action is central to any effect of potential competition. If the incumbent firm can resell the excess capacity and fully recover that investment after entry, then it is no longer credible

that the incumbent will maintain this higher level of capacity and lower prices should entry occur.

Fudenberg and Tirole (2000) examine the customer base of a product with network externalities (as occurs with computer operating systems). The incumbent firm sets a low current price, expanding its current and future customer base.

Empirical Evidence for Potential Competition

The most prominent evidence for the existence and nature of potential competition is provided by the US airline, pharmaceutical and computer software industries. Peteraf (1995) examines domestic US city-pairs, with airlines serving one of these cities but not the city-pair itself. Potential entry by a traditional airline into a city-pair unexpectedly increases prices, while that by a newer low-cost firm lowers prices. Goolsbee and Syverson (2008) confirm the latter finding, examining potential entry by low-cost Southwest Airlines into city-pair markets. They find large impacts of both potential and actual competition from Southwest, with over half the impact of actual competition realized before entry. They identify building customer loyalty through lower prices as the method of incumbent commitment.

Grabowski and Vernon (1992) examine the impending entry of US generic drugs before patents on incumbent drugs expire. They find, unexpectedly, that potential competition increases incumbent prices, as the incumbent segments the market into price-sensitive and price-insensitive components. The former will be conceded to entrants with patent expiry. Ellison and Ellison (2011) also find that prices rise with potential competition. They further find that incumbent prices continue rising in markets where actual entry is limited, though they fall in markets with significant actual entry. Again, demand-side mechanisms drive pre-entry actions by market incumbents.

Hall et al. (2003) estimate that potential entry forces Microsoft to price operating system software (Windows) and productivity software (Office) at 40 % of the pure monopoly price with no entry. The commitment mechanism is again on the demand side, with consumer switching costs and network externalities. The lower price limits the customer base for potential competitors, successfully deterring their entry.

This evidence validates the basic concept of potential competition and the underlying mechanism of commitment. However, these empirical effects are often in a different direction from those predicted by canonical models, in particular because consumer demand is more complicated in practice than in the original theory.

Practical Limits for the Concept of Potential Competition

Analytical models highlight one mechanism at the expense of others. Models of potential competition focus on rivalrous interactions among homogenous firms/customers with static technology/market definition. The concept of potential competition sits uneasily in modern strategic management, where analyses focus instead on heterogeneous firms and dynamic competition.

The decades that have seen formalization of the theory and demonstration of empirical effect for potential competition have also seen increases in differences among firms and increases in the pace of innovation. These immediate gains for the concept are thus offset in part by a changed environment that reduces the range of economic activity explained by it.

For example, the US airline industry shows steady decline in inflation-adjusted average fares, emergence and success of two distinct business models (hub/spoke network and point-to-point), large financial losses for most firms, and significant process technology innovations based on outsourcing, labour relations and information technology (Borenstein and Rose 2013). Even though this industry provides perhaps the strongest evidence for potential competition, strategic entry deterrence of potential competition explains an arguably very small part of the actual evolution of this industry and the performance of firms therein.

See Also

- ▶ Dynamic Capabilities
- ▶ Industrial Organization

References

Borenstein, S., and N.L. Rose. 2013. How airline markets work . . . or do they? Regulatory reform in the airline industry. In *Economic regulation and its reform: What have we learned?* ed. N.L. Rose. Chicago: University of Chicago Press.

Ellison, G., and S.F. Ellison. 2011. Strategic entry deterrence and the behavior of pharmaceutical incumbents prior to patent expiration. *American Economic Journal: Microeconomics* 3: 1–36.

Fudenberg, D., and J. Tirole. 2000. Pricing a network good to deter entry. *Journal of Industrial Economics* 48: 373–390.

Gilbert, R.J. 1989. The role of potential competition in industrial organization. *Journal of Economic Perspectives* 3: 107–127.

Goolsbee, A., and C. Syverson. 2008. How do incumbents respond to the threat of entry? Evidence from the major airlines. *Quarterly Journal of Economics* 123: 1611–1633.

Grabowski, H.G., and J.M. Vernon. 1992. Brand loyalty, entry, and price competition in pharmaceuticals after the 1984 Drug Act. *Journal of Law and Economics* 35: 331–350.

Hall, R.E., J. Royer, and M. Van Audenrode. 2003. Potential competition and the prices of network goods: Desktop software. Working paper, Stanford University.

Peteraf, M.A. 1995. Sunk costs, contestability, and airline monopoly power. *Review of Industrial Organization* 10: 289–306.

Spence, A.M. 1977. Entry, capacity, investment, and oligopolistic pricing. *Bell Journal of Economics* 8: 534–544.

Tirole, J. 1988. *The theory of industrial organization*. Cambridge, MA: The MIT Press.

Power

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

Abstract

Power is the ability to prevail in contested situations, be the situations competitions for promotions or conflicts over decisions. Sources of power include personal attributes such as energy and empathy, which can be developed, as well as structural position and the actions people choose, such as asking for things and highlighting one's own accomplishments. Power has consequences for the power holder, including being subject to more scrutiny and also the ability to monetize and otherwise gain advantage from power.

Definition Power is typically defined as the capability of producing an intended effect – on others' attitudes or behaviours, on the environment, on groups or organizations. The greater the effect that can be produced, the greater the power.

Power is the ability to get one's way in contested situations. In organizations, there is often disagreement about what objectives to pursue or what weight to give conflicting goals. There is even more likely to be disagreements about how to achieve objectives – what actions should be taken and which have the best chance of success. Organizations are also arenas for competition for promotion because, as hierarchical structures, there are inevitably fewer high-status than lower-status or entry-level positions. Consequently, there are career tournaments in which losing at one stage of the competition severely adversely affects the person's ability to compete, let alone prevail, in subsequent promotion competitions. Power thus reflects a person's ability to prevail in the conflicts about what objectives to pursue and the best strategies to achieve those objectives, and also the competition for higher-status jobs. In fact, one way to assess power is to see which individuals or departments are best able to obtain higher-level jobs, prevail in competition for scarce budgetary and other resources, and are able to get their way in choices about strategy and structural arrangements. A department's level of reporting – whether it reports to the CEO or, if not, how many levels down for instance – would be one indicator of power, as would the desirability of the department's office locations (Pfeffer 1992).

Conditions Under Which Power Is Used

There are a set of conditions that must be present for power to be an important construct for

understanding decisions. *All* these conditions must occur for power processes to be important for understanding behaviour.

The first condition is disagreement – about what to do or how to do it. In some sense promotion competitions represent one form of such disagreement. Most people believe that they deserve promotions more than their competitors – in part because of the above average effect, which finds that people see themselves as above average on positive dimensions such as merit and ability (e.g., Chambers and Windschitl 2004). Without disagreement, there is no contested decision for power to resolve.

The second condition is interdependence – people's need for the cooperation, or at least the assent, of others. Without interdependence, disagreement doesn't matter because there is no need to reach a joint decision or take joint action. Interdependence is partly a function of resource scarcity. When resources are plentiful, there is less interdependence because every department or individual can obtain pretty much what it wants and needs. When resources are scarce, allocations become much more zero sum, thereby increasing interdependence among social actors. Some studies show that the effect of power on resource allocations is higher when resources are tighter (Salancik and Pfeffer 1974). Interdependence is also a consequence of decisions concerning organizational design. Self-contained units or tasks obviously confront less interdependence than jobs and departments that are more team-oriented, in matrix-like structures, with shared resources and decision-making. Therefore, many of the current team-oriented and more participative organizational structures that involve more people in decisions make power more important and more likely to affect decision-making, not less.

The third condition for the use of power is the importance of the decision. People are unlikely to use their power, in part because they won't care very much, with regard to unimportant decisions. Robert Caro's (2002) biography of Lyndon Johnson's rise in the Senate related how Johnson built a power base by doing the small, mundane things such as scheduling uncontroversial bills and providing information. Because these were seemingly small tasks, no one was likely to compete with Johnson to do them and, in fact, prior to his arrival in the role of senate minority leader, the activities mostly had been left undone.

Sources of Power

One important source of individual-level power is the personal qualities individuals possess. Studies of genius, which is often defined as exceptionally high levels of performance, consistently demonstrate that, in domains ranging from athletics to art to medicine and science, outstanding performance is related more to practice, experience and receiving expert guidance or coaching than to innate individual differences. There is no reason to believe this situation is any different for building the qualities that help people acquire power.

One such quality is energy. The ability and willingness to work long hours provides time that others may not have. And energy can inspire others to become allies, in part because they assess those with energy as more likely to win and people like to be associated with winners – to bask in their reflected glory.

A second quality is the ability to take others' perspectives, to see the world as others might see it. This permits a person to understand where others are coming from, anticipate their objections and even their countermoves, and negotiate over common interests, which at a minimum requires understanding of what those common interests are.

A third quality is persistence and resilience. Although not typically emphasized in the heroic stories we like to tell ourselves about major figures, almost everyone suffers opposition and setbacks. Even Apple's iconic CEO, Steve Jobs, was forced out of the company for a time in the 1980s, and many highly successful senior executives have been forced from a job at some point. Whether people have the personal fortitude to re-enter difficult arenas and persist even in the face of opposition frequently determines their ultimate success.

A fourth quality is the ability and willingness to tolerate conflict. Most individuals are conflict

averse, which gives those that are not an advantage. Being perceived as difficult can work to one's advantage. That is because warmth and competence, as dimensions of interpersonal perception, are often perceived to be negatively correlated. Consequently, warm people are sometimes perceived as lacking in strength or competence.

There is a long list of personal attributes that provide power. Some of these characteristics emerge from the literature on personality, and some from other sources (e.g., Judge et al. 2002). Recent research suggests that being nice is negatively related to achieving positions of power, particularly for men, while conscientiousness is positively related to career success.

Another source of power is an individual's structural position. There are two dimensions to a person's position that are consequential for power. The first is the individual's department or subunit. Some subunits have more power, because they control more resources than other departments do, deal with more critical contingencies, or cope with more important uncertainties in the company's environment compared to other units. If an individual is located in such a subunit the person is favoured because that unit is better able to get its way – including in contested decisions such as who will be promoted and who will get the largest salaries or pay increases (Sheridan et al. 1990). There is downside to being located in the most powerful unit, as well – because such units are seen as the route to the top, and often offer higher wages for people with equal skills, they attract more talent and, as a consequence, more competition. Therefore, an even better structural position is to be located in a unit that is not yet but is on its way to becoming powerful. That permits someone to gain the advantages of being located in the 'right' place without confronting as much competition from talented others.

The second dimension of structural position concerns a person's location in the network of communication. Sociologist Ronald Burt (2005) has consistently shown that people in a position to bridge structural holes – to provide a link between two or more units that would benefit from transacting but which are not currently doing so – can profit in their careers. Another network-related construct is centrality – people at the centre of communication networks are, by definition, central in the information flow and therefore will have knowledge that others lack. Because of this central position, they will also be able to influence what information is disseminated and therefore, what decisions are made. Studies of industrial purchasing decisions, including decisions to purchase certain computers, show the power of being in the centre of, and therefore to some extent in control of, the flow of information.

Strategies for Building and Using Power

There are numerous strategies for building and using power (e.g., Pfeffer 2010). Some of the strategies rely on the insight that power is conferred or granted by others, so their impressions of the target profoundly affect whether or not that individual is perceived as powerful. Impressions depend much less on the content of what people say and much more on how they say it and how they appear as they say it.

Because of the heuristic association between power and confidence – powerful people are presumed to be confident and self-assured – acting with confidence and assurance can confer power to the individual. Adopting a powerful pose – such as being expansive and taking up space – not only causes individuals to feel more confident and assured, it actually affects their blood chemistry, decreasing the level of cortisol, a stress hormone, and increasing the level of testosterone.

Speaking with power can entail speaking loudly. Powerful people interrupt; the less powerful are interrupted. Speaking without notes conveys a mastery of the subject matter that using notes does not.

Power provides those that have it with the right to be angry and even rude. Therefore, people who display anger rather than sadness or remorse (e.g., Tiedens 2001) and who even behave in slightly rude or inappropriate ways are seen to have power.

If you can get someone to do you a favour or provide you with help, that individual will be

committed to you from having expended the effort to be helpful and because, as a consequence of the interaction providing the assistance, the person will be more closely associated with you. Therefore, one strategy for building power and support is to make requests. Research shows that people are often reluctant to ask for what they want and need. Part of this reluctance may derive from cultural expectations of self-reliance, some may come from a fear of looking weak, and some of the hesitation may derive from the fear of rejection. Therefore, people overestimate the difficulty of obtaining help from others and fail to ask for things as much as they might (Flynn and Lake 2008).

Another strategy for acquiring power is to activate the norm of reciprocity and do things that are helpful for others. Those things can be small tasks that make their lives easier and help them with seemingly minor aspects of their job. They do not need to be substantial or dramatic actions.

Because we like those who are similar to ourselves, and are much more likely to comply with requests from similar others (Burger et al. 2004), building a perception of similarity with others helps provide power over them. This requires finding out what one has in common or even creating commonalities through, for instance, shared experiences.

The Consequences of Power

There are numerous consequences of power, some positive, others less so. One consequence of power is increased scrutiny and visibility. And it is not just job-relevant personal traits that garner attention – people feel perfectly free to watch and comment on the cars, the houses, the holidays, the money-spending habits and the personal associates and spouses of those in power. Research shows that doing one's job in the limelight, particularly if that job requires learning and adaptation, is taxing and adds an extra burden to those in power.

Another consequence of power is its effect on the psychology of the power holder (Keltner et al. 2003). Many studies show that those in power become self-focused and directed towards fulfilling their own needs and goals. Consequently, the powerful are more likely to behave with fewer inhibitions, believing that the rules don't really apply to them. They are less sensitive to the needs and preferences of others, particularly those with less power. Those in power see outcomes as being more under their control and tend to take more credit for good outcomes. Simply put, power tends to set in motion behaviours of inattentiveness and disinhibition that often bring about the eventual downfall of, or at least problems for, those with power.

A third important, albeit seldom discussed, effect of power is its addictive quality. Those in power are the centre of attention and the focus of much activity. Once in a less powerful position, the attention wanes and the frenetic pace of life does too. This change requires acclimation. One of the reasons that people resist stepping down from powerful positions, even when they would appear to have more than enough financial resources to do so, is this fear of losing the limelight and the attention of others.

Power also has positive consequences for those who have it. Power and political skill produce career success (e.g., Ng et al. 2005). Power and status can be monetized. Because people like to be associated with power and the powerful, those in power can sell their personal appearances and books as well as exclusive access to themselves and their network. Many politicians and generals have used this principle to make substantial sums of money after leaving the positions that brought them power and status and the visibility that often accompanies such positions.

Power permits people to accomplish organizational and social change. The literature is filled with references to inertia and resistance to change. That suggests that in order to get things moving in a different direction, some force is needed. Power is such a force.

And power, to the extent that it has implications for one's control over the conditions of a person's work, is predictive of longevity. There is a fairly large epidemiological literature on the

effects of job control – the ability to control the pace, timing and content of work, which is different from whether or not one works hard – on the incidence of cardiovascular disease and mortality from same. The underlying logic is clear – an inability to control the conditions of one's work life is customarily experienced as being stressful, and the harmful effects of stress for both physical and mental health are well known. Consequently, the evidence suggests that mortality follows a status-graded hierarchy – those in higher-level positions are healthier and live longer.

In spite of power's importance in organizations and social life, power remains the organization's last dirty secret. Relatively few companies train their emerging leaders in power, even though career derailments are common and the inability to master power dynamics is a prominent cause of such derailments. And the ▶ leadership literature is largely silent on the topic of power. That is because power implies individuals acting on the basis of their own self-interests. Of course, self-interest seeking is a fundamental assumption of economic theory and it is generally recognized that individuals evaluate or filter decisions in terms of their effects on themselves.

See Also

- ▶ Leadership
- ▶ Machiavellianism
- ▶ Resource Dependence

References

Burger, J.M., N. Messian, S. Patel, A. del Prado, and C. Anderson. 2004. What a coincidence! The effects of incidental similarity on compliance. *Personality and Social Psychology Bulletin* 30: 35–43.

Burt, R.S. 2005. *Brokerage and closure: An introduction to social capital*. New York: Oxford University Press.

Caro, R.S. 2002. *Master of the senate*. New York: Knopf.

Chambers, J.R., and P.D. Windschitl. 2004. Biases in social comparative judgments: The role of nonmotivated factors in above-average and comparative optimism effects. *Psychological Bulletin* 130: 813–838.

Flynn, F.J., and V.K.B. Lake. 2008. If you need help, just ask: Underestimating compliance with direct requests for help. *Journal of Personality and Social Psychology* 95: 128–143.

Judge, T.A., J.E. Bono, R. Ilies, and M.W. Gerhardt. 2002. Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology* 87: 765–780.

Keltner, D., D.H. Gruenfeld, and C. Anderson. 2003. Power, approach, and inhibition. *Psychological Review* 110: 265–284.

Ng, T.W.H., L.T. Eby, K.L. Sorensen, and D.C. Feldman. 2005. Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology* 58: 367–408.

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

Pfeffer, J. 2010. *Power: Why some people have it – And others don't*. New York: Harper Business.

Salancik, G.R., and J. Pfeffer. 1974. The bases and use of power in organizational decision making: The case of a university. *Administrative Science Quarterly* 19: 453–473.

Sheridan, J.E., J.W. Slocum Jr., R. Buda, and R.C. Thompson. 1990. Effects of corporate sponsorship and department power on career tournaments. *Academy of Management Journal* 33: 578–602.

Tiedens, L.Z. 2001. Anger and advancement versus sadness and subjugation: The effect of negative emotion expressions on social status conferral. *Journal of Personal and Social Psychology* 80: 86–94.

Prahalad, C. K. (1941–2010)

Vijay Govindarajan
Dartmouth College, Tuck School of Business, Hanover, NH, USA

Abstract

The article considers the career of the business management theorist C. K. Prahalad through a discussion of three of his main works. Following discussion of the future-oriented nature of his thinking, the article considers the first of Prahalad's books, *The Multinational Mission*, in which he argued that multinational companies should aim to achieve a balance between global scale and local demands. In what is perhaps the best known of his titles, *Competing for the Future*, he challenged much of the

conventional wisdom about competitive strategy, highlighting concepts such as the resource-based view of the firm and strategic intent. Finally, in *Fortune at the Bottom of the Pyramid*, Prahalad turns his attention to how entrepreneurs can address the needs of the poorest in society, who may – contrary to expectations – prove to be the fastest-growing consumer segment. A short conclusion highlights that Prahalad has practised what he has preached, putting into effect some of his key business messages.

Three beliefs shaped the thinking of C. K. Prahalad (hereafter 'CK'): business school research must have managerial relevance; business is a powerful positive force for the greater good; and much of extant management theory is not capable of explaining business realities. These beliefs led to three distinguishing characteristics in CK's scholarly contributions. First, CK opened up new fields instead of exploiting the same concept throughout his career. This is illustrated in some of his groundbreaking ideas: ▶ strategic intent, core competencies, bottom of the pyramid, customer co-creation. None of his books were written with the same co-author. Second, he was a contrarian thinker. His work is full of counter-intuitive insights and fresh thinking. Third, CK had a strong bias towards managerial action. Most researchers confirm hypotheses using a wide range of samples. By definition, such research implies that the majority of companies have embraced the idea. CK was interested in next practices, not current best practices. He observed a few companies who are leaders and captured their practices.

This article highlights the three distinguishing characteristics of CK's work through three of his books.

The Multinational Mission

Until the early 1980s, global strategy was viewed as a choice between global scale and local responsiveness. Global scale implies standardization (Prahalad and Doz 1987). Local responsiveness implies differentiation. Michael Porter's pioneering work (1980) posited that competitive strategy has to make a choice between differentiation and cost leadership. Multinational strategy, too, was viewed as making a similar choice.

CK argued that multinationals do not have to choose between scale and localization. In fact, an effective strategy must optimize both. CK introduced the now famous 'integration-responsiveness' (I-R) grid and showed how multinationals can 'have their cake and eat it too'. First, he argued that multinationals should develop strategies at the business level, not at the corporate level. He showed that three of the Corning businesses – electronics, TV products and cookware – were at different places in the I-R grid. Second, even within the business, it is best to break down the overall value chain into its components and ask where one can derive scale benefits and where it is good to localize. This fine-grained approach assures the business the benefits of differentiation and cost leadership. Third, CK argued that the organizational architecture has to be re-crafted to support the multinational mission to balance local demands and global scale.

Global strategy researchers in the past 25 years have been strongly influenced by CK's thinking (e.g., Bartlett and Ghoshal 1988; Ghemawat 2007).

Competing for the Future

Perhaps CK is best known for his book *Competing for the Future*(Hamel and Prahalad 1994**)**. This is a groundbreaking work which challenged much of conventional wisdom. Let me give three examples:

- Mike Porter's work (1980) focused on product/market imperfections and how a firm can create competitive advantage through erecting entry barriers to sustain such imperfections. CK's position was that there are also factor market imperfections and a firm can create lasting advantage by building core competencies that distinguish it from others. CK's work gave

momentum to the ▶ resource-based view of the firm and the dynamic capabilities literature (Helfat et al. 2007).

- Behavioural theories have posited that goals that are achievable are the best motivators. Unrealistic goals are argued to be a demotivator (Maciariello and Kirby 1994). CK introduced the concept of strategic intent – goals that are bold and unrealistic. John F. Kennedy's 'man on the moon' is an example of an unrealistic goal. A bold intent has the potential to produce breakthrough innovations. Why? Because people are drawn to a bold and challenging goal. Deep inside, we feel uplifted by the thought of climbing a mountain in a way we are not by the idea of scaling a molehill. Performance is a function of expectations, since we rarely exceed our expectations or outperform our ambition.
- Marketing scholars have argued that high market share leads to high profitability due to experience curve effects (Kerin et al. 1990). CK provided many examples where small, resource-challenged firms were unseating incumbents: Honda versus General Motors; CNN versus CBS; Canon versus Xerox; Wal-Mart versus Sears. CK challenged the prevailing theories about the market power of incumbents and the advantages of market share.

Fortune at the Bottom of the Pyramid

Development economists have long focused on poverty, and NGOs and governments have devoted considerable efforts and resources to the problem, yet we have not redressed it (Prahalad 2005). CK offered a fresh and unique approach. He argued that poverty alleviation is not a problem for charity, it is a problem for commerce. He presented a framework showing how the private sector and entrepreneurship can serve the bottom of the pyramid (BOP) and make profits in the process. Not only can the world's poor be relevant customers, but they represent the faster-growing customer segment. Meeting their needs requires breakthrough innovations that have to scale – something that corporations know how to do. There are precedents: the Unilever subsidiary Hindustan Lever has transformed the distribution model for rural India, thereby bringing the enormous rural population of that country into the consumer base. According to CK, we will have succeeded when business views BOP as a mega opportunity, with billions in profits at stake.

Concluding Thoughts

CK practised what he preached. There are two views of strategy. One argues that the firm should match its resources with external opportunities. Another, the one CK espoused, is that the firm should expand its resource base to meet its ambition. Most doctoral students' lifestyles are constrained by meagre scholarships. As a doctoral student, CK took on a job and consulted with corporations so that he could expand his resource base to live more like an executive. His life is his message.

See Also

- ▶ Innovation
- ▶ Multinational Corporations
- ▶ Resource-Based View
- ▶ Strategic Intent

References

Bartlett, C.A., and S. Ghoshal. 1988. *Managing across borders: The transnational solution*. Boston: Harvard Business School Press.

Ghemawat, P. 2007. *Redefining global strategy crossing borders in a world where differences still matter*. Boston: Harvard Business School Press.

Hamel, G., and C.K. Prahalad. 1994. *Competing for the future*. Boston: Harvard Business School Press.

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 Publishing.

Kerin, R.A., V. Mahajan, and P.R. Varadarajan. 1990. *Contemporary perspectives on strategic market planning*. Boston: Allyn & Bacon.

Maciariello, J.A., and C.J. Kirby. 1994. *Management control systems: Using adaptive systems to attain control*. Englewood Cliffs: Prentice Hall.
Porter, M.E. 1980. *Competitive strategy: Techniques for analyzing industries and competitors*. New York: Free Press.

Selected Works

Prahalad, C.K., and Y.L. Doz. 1987. *The multinational mission: Balancing local demands and global vision*. New York: Free Press.
Prahalad, C.K. 2005. *The fortune at the bottom of the pyramid*. Upper Saddle River: Wharton School of Publishing.

Predatory Pricing

Justin Tumlinson
Institute for Economic Research, Munich, Germany

Definition Predatory pricing is a two-stage strategy of (1) temporarily setting prices low enough to force rivals to exit, deter their entry or discipline them into accepting lower market share, in expectation of (2) subsequently using gained market power to raise prices long enough to recoup first-stage losses.

Predatory pricing is distinguished from *competitive pricing* by the second stage – predatory pricing is not identified by the intended harm to *less efficient* rivals (the natural consequence of competition) but, rather, the long-run damage done to consumers through the ultimate raising of prices *above competitive levels* and the displacement of *more efficient* rivals.

Legalities

Predatory pricing is actionable under many countries' antitrust laws. The courts' primary challenge is to determine whether or not price cuts are predatory or simply competitive. Since economic theory suggests that perfect competition drives prices to ▶ marginal cost but not lower, Areeda and Turner argue that prices below *marginal cost* are predatory (Areeda and Turner 1975). Since one cannot generally observe marginal costs (i.e., the incremental cost of the last unit produced), they advance the following test, historically applied in US antitrust cases: firms (with significant market power) setting prices below *average variable costs* indicate predatory intent.

Cases date from at least 1911, when the US Supreme Court dissolved the Standard Oil Company under the Sherman Antitrust Act of 1890, in part for predatory price discrimination. At least 123 cases were tried in federal court over the following 60 years, resulting in 95 convictions (Koller 1971). However, in 1993 the US Supreme Court decided, A plaintiff must prove (1) that the prices complained of are below an appropriate measure of its rival's costs and (2) that the competitor had a reasonable prospect of recouping its investment in below cost prices' (Brooke Group Ltd. v. Brown and Williamson Tobacco Corp. 1993). Owing to the difficulty of satisfying these conditions, the Federal Trade Commission has not successfully prosecuted a case since.

Economic Theory

While laymen find predatory pricing intuitive, many economists dispute the rationality of the practice, except under extraordinary circumstances. John McGee was among the first to express doubt in a 1958 analysis of the Standard Oil case. First, he argued that when prices are below marginal costs, large firms incur losses faster than smaller rivals. Thus, firms with significant market power rarely initiate price wars. Second, although predators *may* (coincidentally) have deep pockets, they must outlast both consumers and the tolerance of financial markets – consumers must not be able to stockpile goods during the price reduction, and financiers must be unwilling to support rivals through the price war. This latter condition seems improbable, as the premise is that the victor will enjoy a

long-term, profitable increase in market power. Hence, a price war could be protracted and risky. Third, rivals could scale down production until prices return to profitable levels. Moreover, any gains from acquired market power are likely to be short-lived as new rivals enter, perhaps even acquiring the assets of previously vanquished rivals at fire-sale prices. Finally, mergers offer a more attractive mechanism to gain market power, as there is no associated period of losses. This *price theoretic* reasoning has influenced current judicial scepticism of predatory pricing.

However, more recently, *game theoretic* models have identified conditions under which predatory pricing may be rational (Milgrom and Roberts 1990). These models hinge on asymmetric information – predators must be better informed, say about costs or market demand, than prey. In one variation, the predator drops prices to signal low costs. Although the prey realizes that the predator could be 'lying' about its cost structure, it reasons that the probability that the predator's costs are genuinely low is sufficient to induce exit or prevent entry. However, predators must not be able to credibly reveal true costs; otherwise, true low-cost producers would always disclose their costs, preventing bluffing by less efficient predators. In other words, neither the market nor prey can distinguish between predatory and competitive pricing strategies. Thus, occurrences should be rare, and, when they occur, devising a test to identify them is non-trivial. Indeed, economists cannot agree on any clear-cut, demonstrable cases of predatory pricing. Not only are cost-based tests inadequate, but a poorly designed test risks deterring the very competition that punitive sanctions would seek to preserve. Finally, rational predation does not even necessarily imply consumers are worse off. Given the strong conditions required to sustain rational predation, the difficulty of detecting it, the likelihood of inadvertently inhibiting competition and the ambiguity of its welfare implications, few economists advocate statutes to prevent it. Likewise, few business strategists recommend attempting predatory pricing.

See Also

- ▶ Game Theory
- ▶ Marginal Cost
- ▶ Marginal-Cost Pricing

References

Areeda, P., and D.F. Turner. 1975. Predatory pricing and related practices under section 2 of the Sherman Act. *Harvard Law Review* 88: 697–733.

Brooke Group Ltd. v. Brown & Williamson Tobacco Corp. 1993. (92-466), 509 U.S. 209.

Koller, R.H. 1971. The myth of predatory pricing: An empirical study. *Antitrust Law and Economics Review* 4: 105–123.

McGee, J.S. 1958. Predatory price cutting: The standard oil (N.J.) case. *Journal of Law and Economics* 1: 137–169.

Milgrom, P., and J. Roberts. 1990. New theories of predatory pricing. In *Industrial structure in the new industrial economics*, ed. G. Bonanno and D. Brandolini. Oxford: Clarendon Press.

Pre-emption and Entry Timing

Xavier Martin
Tilburg University, CentER, Tilburg, The Netherlands

Abstract

Entry timing research examines how firm performance varies, possibly non-monotonically, with the order (also known as order of entry) or elapsed time since first entry into a new market. While the pre-emption literature in economics focuses on assumptions for a first entrant to monopolize a market, contemporary strategy scholarship emphasizes more subtle and contingent entry timing effects – with recognition of the importance of endogenous firm characteristics (e.g., complementary assets) and choices (e.g., modes of expansion or entry), and industry contingencies (e.g., knowledge tacitness). Groundbreaking applications of entry timing concepts pertain to international management as well as innovation contexts.

Definition Although pure pre-emption – the idea that the first entrant can monopolize a market – once received much theoretical interest, contemporary strategy scholarship emphasizes more subtle and contingent entry timing effects instead – that is, conditions under which a firm may benefit from being an early, as opposed to intermediate (or late), entrant, with increasing recognition of the importance of endogenous firm characteristics and choices.

Contemporary strategic management research addresses the time-bound, evolutionary processes whereby firms gain, maintain and lose distinctive competitive positions. In this context, an intriguing extreme solution could be found in early game-theoretic work about the possibility of pre-empting a market – that is, obtaining a durable monopoly by the mere fact of being the first entrant. Gilbert and Newbery (1982) derived an influential model whereby an incumbent can secure a future monopoly via 'sleeper' patenting, and concluded that monopoly may hence arise without market failure. However, they and subsequent researchers described multiple caveats: foremost, uncertainty may hinder both the assessment of rival threat and the conduct of preemptive R&D; and heterogeneity of the market space or bundling of complementary products may prevent hermetic pre-emption, especially with low exit costs for potential entrants (for a review see Reinganum 1989). Besides lacking empirical contents, this literature has largely faded because of two other limitations. It hinges on the artefact of perfect patents, and thus only has a bearing on (some) industries dominated by technological competition – and even then, early pre-emption scholars noted that patents may not actually secure monopoly (Katz and Shapiro 1987); and, while many pre-emption models hinged on incumbents having superior incentives, mechanically they revolved around which firm had lower costs or superior ability to accelerate innovation – effectively, factors better explained in resource-based and dynamic capabilities theorizing (Choi 1996; Lieberman and Montgomery 1998; Ofek and Sarvary 2003). Nevertheless, the pre-emption literature has served to inform subsequent strategy research on the pros (and cons) of early market entry.

The strategy literature on entry timing has grown to address a broader question: Under what conditions does a firm stand to benefit by entering a market early, rather than at an intermediate or late stage? This question is typically addressed by examination of the ranking of cumulative entries, that is, order of entry (OoE) effects. Lieberman and Montgomery's (1988) seminal paper set the ground for such research by identifying not just potential sources of first-mover advantage (FMA) (technological leadership, pre-emption, buyer switching costs), but also of first mover disadvantages (free-riding by followers, resolution of technological or market uncertainty, technological discontinuities, incumbent inertia); and by pointing to the importance of studying mechanisms that enhance (moderate) FMA, and to the endogeneity of first-mover opportunity given differential proficiency (and luck).

Empirically, strategy research has shown that any FMA when moving into a new product market is contingent: Mascarenhas (1992) found a first-mover effect, but subject to erosion over time (i.e., weak pre-emption); while Makadok (1998) found early-mover effects, but without consistent difference between first and second to fifth movers, and subject to erosion with cumulative rival entries. Paralleling the pre-emption literature, much FMA research dealt with the effect of technological innovation. The context of technology remains fruitful for FMA research, but another prime context for entry timing research is international expansion, as that offers the opportunity to model entry decisions and outcomes in multiple countries over time and thus control for both observed and unobserved firm characteristics (Martin et al. 2007). International business research has investigated OoE effects across the range of entry ranks, unlike the technology literature, and has shown that the propensity for foreign entry (Martin et al. 1998) and the resulting performance (subsidiary survival: Mitchell et al. 1994) are positively associated with an intermediate OoE (neither leader nor laggard) relative to other foreign entrants. Strategy scholars typically

attribute these effects to information spillovers, negated beyond some point by competitive pressure.

Extant research further highlights contingencies on OoE effects. The instances below, at the firm, industry and technology (knowledge resource) level respectively, illustrate the importance of this approach. Most prominently, firm-level complementary assets (Teece 1986; Mitchell 1989) can compensate for lateness in entry, so OoE strategy should be a function of complementary assets too. At the industry level, OoE effects vary with the pace of market evolution (Agarwal et al. 2002; Suarez and Lanzolla 2007). Finally, features of knowledge (especially its tacitness) determine the time to imitation (Zander and Kogut 1995), and thus the durability of an early-entry effect. OoE research offers plentiful research opportunities yet. Among them are risk-return profiles of OoE strategies (Robinson et al. 1994); quality trade-offs in speeding entry and lower OoE (Salomon and Martin 2008); and, critically given the contingencies above, the distinction between OoE and calendar timing (Martin et al. 1998) and their potentially interacting effects, especially on corporate performance. Further research on entry timing conditions is thus well warranted.

See Also

- ▶ Competitive Advantage
- ▶ Industrial Organization
- ▶ Innovation Strategies
- ▶ Intellectual Capital

References

Agarwal, R., M.B. Sarkar, and R. Echambadi. 2002. The conditioning effect of time on firm survival: An industry life cycle approach. *Academy of Management Journal* 45: 971–994.

Choi, J. 1996. Preemptive R&D, rent dissipation, and the leverage theory. *Quarterly Journal of Economics* 111: 1153–1181.

Gilbert, R.J., and D.M.G. Newbery. 1982. Preemptive patenting and the persistence of monopoly. *American Economic Review* 72: 514–526.

Katz, M.L., and C. Shapiro. 1987. R&D rivalry with licensing or imitation. *American Economic Review* 77: 402–420.

Lieberman, M.B., and D.B. Montgomery. 1988. First-mover advantages. *Strategic Management Journal* 9: 41–58.

Lieberman, M.B., and D.B. Montgomery. 1998. First-mover (dis)advantages: Retrospective and link with the resource-based view. *Strategic Management Journal* 19: 1111–1125.

Makadok, R. 1998. Can first-mover and early-mover advantages be sustained in an industry with low barriers to entry/imitation? *Strategic Management Journal* 19: 683–696.

Martin, X., A. Swaminathan, and W. Mitchell. 1998. Organizational evolution in an interorganizational environment: Incentives and constraints on international expansion strategy. *Administrative Science Quarterly* 43: 566–601.

Martin, X., A. Swaminathan, and L. Tihanyi. 2007. Modeling international expansion. In *Research methodology in strategy and management*, vol. 4, ed. D.J. Ketchen and D.D. Bergh. Amsterdam: Elsevier.

Mascarenhas, B. 1992. Order of entry and performance in international markets. *Strategic Management Journal* 13: 499–510.

Mitchell, W. 1989. Whether and when? Probability and timing of incumbents' entry into emerging medical subfields. *Administrative Science Quarterly* 34: 208–230.

Mitchell, W., J.M. Shaver, and B. Yeung. 1994. Foreign entrant survival and foreign market share: Canadian companies' experience in United States medical sector markets. *Strategic Management Journal* 15: 555–567.

Ofek, E., and M. Sarvary. 2003. R&D, marketing, and the success of next generation products. *Marketing Science* 22: 355–370.

Reinganum, J.F. 1989. The timing of innovation: Research, development, and diffusion. In *Handbook of industrial organization*, vol. 1, ed. R. Schmalensee and R.D. Willig. New York: North-Holland.

Robinson, W.T., G. Kalyanaram, and G.L. Urban. 1994. First mover advantages from pioneering new markets: A survey of empirical evidence. *Review of Industrial Organization* 9: 1–23.

Salomon, R., and X. Martin. 2008. Learning, knowledge transfer, and technology implementation performance: A study of time-to-build in the global semiconductor industry. *Management Science* 54: 1266–1280.

Suarez, F., and G. Lanzolla. 2007. The role of environmental dynamics in building a first mover advantage theory. *Academy of Management Review* 32: 377–392.

Teece, D.J. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing, and public policy. *Research Policy* 15: 285–305.

Zander, U., and B. Kogut. 1995. Knowledge and the speed of transfer and imitation of organizational capabilities: An empirical test. *Organization Science* 6: 76–92.

Price Control

Carl Danner[1], Edward F. Sherry[2] and Elizabeth Olson[2]
[1]Berkeley Research Group, Chicago, IL, USA
[2]Berkeley Research Group, Emeryville, CA, USA

Abstract

Price controls involve government setting prices for products, goods or services at a non-market-determined level. In competitive industries, price controls typically have economically harmful effects but may be imposed for political reasons, especially during economic disruptions. The longer that controls on competitive prices are maintained, the more consequential and complicated the resulting economic problems are likely to become. Price controls for utilities and like services are a common and far less problematic governmental response to the market power that exclusive service franchises may create.

Definition Price controls are one form of governmental regulation of the economy, under which laws or rules set maximum (or, in some cases, minimum) prices, whether for particular products or for products more generally, especially in times of high inflation or significant economic disruptions.

Price controls are one form of governmental regulation of the economy, under which laws or rules set maximum (or, in some cases, minimum) prices, whether for particular products (such as crude oil or petrol, residential rents, pharmaceuticals and agricultural products) or for products more generally, especially in times of high inflation or significant economic disruptions (such as wartime price controls). As such, price controls can often either substitute for, or complement, other forms of regulation. This article focuses on price controls in otherwise predominantly free-market economies; the issue of price setting in planned economies (such as the former Soviet Bloc), although significant, is not treated here.

Price controls in competitive contexts come in two major forms: restrictions on the *minimum* prices that can be charged for particular goods or services, and restrictions on the *maximum* prices that can be charged. The latter are much more common. As a matter of economics, price controls are generally understood to stimulate surpluses or shortages if they (respectively) raise or reduce prices from market-clearing levels (Mankiw 1998: 112–121). Another context involves oversight of utilities or other businesses given exclusive franchises by government, where authorities may determine authorized service prices to try to limit the exercise of monopoly power stemming from the grant of exclusivity or an underlying 'natural monopoly' technology.

As a general matter, economists find little justification for controlling prices of competitively provided goods and services. Politically, such controls are often justified by arguments about possible defects in market processes. In the US, minimum price restrictions were adopted for many agricultural commodities during the Great Depression to protect small farmers against the claimed price-depressing effects of uncoordinated excess supply, and such restrictions persist today in the form of various agricultural price supports and import restrictions. Anti-dumping legislation can similarly be thought of as a form of minimum price regulation, at least with respect to imports from overseas alleged to be subsidized by producing countries. Limits on maximum prices can be motivated by concerns about shortages of important goods and services. In some instances, governments use their buying power to effectively impose price controls. One well-known example involves pharmaceuticals, since in many European countries and Canada the government health service is the primary purchaser of prescription drugs. Another example from the US involves Medicare and Medicaid limits on reimbursement for various forms of medical care.

Price controls are frequently adopted in response to a perceived short-term need. However, such controls sometimes are extended for lengthy periods, effectively becoming permanent, because their beneficiaries may have political influence that outweighs the influence of the

parties who bear the resulting costs. One well-known example in the US involves rent control in New York City.

Controls on competitive prices have sometimes been held out as illustrating 'the law of unintended consequences'. For example, sustained residential rent control tends to discourage landlords from investing in maintenance or making improvements, thereby reducing the available stock of rental housing over time and exacerbating the limitations of housing supply that led to the imposition of rent control in the first place. Some tenants may stay in apartments oversized for their needs if their rent is kept low, decreasing the effective capacity of the housing stock. Landlords dissatisfied with returns from rent-controlled housing may withdraw their rental units from the market (for example by converting them into condominiums), which may reduce the supply of housing and perhaps lead to a further expansion of regulation to limit such conversions.

A governmental bureaucracy typically administers a price control programme. In the US, state public utility commissions do so for franchise monopoly industries such as electrical and gas utilities, usually through a cost-of-service or rate-of-return regulatory model (Kahn 1988: 1–62). During the late 1970s and early 1980s, many US federal agencies formerly charged with transport price regulation, such as the Interstate Commerce Commission (which formerly regulated trucking and railroad freight rates) and the Civil Aeronautics Board ('CAB') (which formerly regulated airfares), were largely stripped of their powers over industry pricing – in part due to the analysis of economists that price-setting was inefficient and harmful to consumers in structurally competitive industries (Breyer 1982: 197–239). Economic performance subsequently improved markedly in those industries, including sharp reductions in service prices (Morrison and Winston 1999).

Price-setting regulatory agencies often have a mandate to act in the public interest or the interest of consumers, although it is widely recognized that the agencies may end up being significantly influenced by, and in some instances even 'captured' by, the entities that they are charged with regulating (Stigler 1971). Such agencies need to acquire data about factors (such as costs and profit margins) to consider when setting the level of controlled prices, and so typically can compel regulated entities to submit information and records.

For some vertically integrated firms that hold monopoly franchises, economists and others have argued for separating out those aspects of the industry for which competition is feasible. One example involves electricity, where local distribution of electricity to the end-user (often seen as involving natural monopoly characteristics) has in recent years largely been separated from less-regulated markets for electricity generation and long-distance energy transmission in a number of countries (Joskow 2008). In such instances, price regulation may persist at the interfaces between the regulated and deregulated sectors of the industry, especially in connection with 'interconnect charges', allowing entrants access to incumbent's equipment needed for the final connection to the end-user.

Price controls can also lead to what economists call rent-seeking behaviour, under which firms or individuals that benefit have an economic incentive to seek to impose or maintain price controls, bar competitive entry or seek other restrictive rules. For example, regulated incumbent firms, especially those subject to a 'universal service' requirement, may object if a new entrant seeks to supply only the most profitable segments of the market, leaving the highcost, low-profit segments to be served by the incumbents. However, entrants using new technology also target attractive customers first, and regulators can have difficulty protecting price-regulated incumbents without simultaneously limiting innovation. Similarly, firms owning long-lived plant and equipment will object if changes to the price control mechanism leave them with investments on which they fear they cannot earn an adequate return.

Price controls can be especially problematic where some but not all prices are controlled, as such a situation can distort the *relative* prices of different goods and services, and thus distort both demandside purchase decisions and supply-side investment and production decisions, all of which

can affect resource-allocation decisions more generally. This can occur because many prices are effectively uncontrollable for many reasons (e.g., as determined in world markets beyond the regulator's authority, or as affected by exogenous factors such as weather conditions and agricultural prices, etc.), or because they are either effectively unregulated or are overseen by a different entity with a different agenda (such as wage controls).

Price controls also can lead to efforts to control product characteristics since producers may have opportunities to modify them to maintain profits (or avoid losses) in ways that undermine the intended objectives of the controls. This can make price controls problematic in the context of technological change, as new products with new features become available, as older products are withdrawn from the market and as new production technologies become available. Because in many industries (especially high-tech industries) the locus of competition largely focuses on *feature*-based competition along *non*-price dimensions rather than on *price* competition with a *fixed* set of product characteristics, impacts on product characteristics can be a significant source of economic distortions in some contexts.

For example, the US Civil Aeronautics Board was criticized for setting many airfares above efficient market-clearing levels, which led airlines to compete for passengers on a non-price basis with features such as the quality of airline food and beverages, and excess capacity in the form of point-to-point service and large wide-body aeroplanes (thereby increasing the number of empty seats, and thus space available to each passenger). Airfares fell substantially after deregulation as capacity utilization rose and air travel (especially leisure travel) increased dramatically, albeit with complaints that the quality of service had declined relative to the old price-regulated regime. In other instances, price controls can induce firms to degrade quality, or to eliminate or reduce the supply of lower-priced, lower-margin product lines in favour of higher-margin products.

Because price controls reduce or eliminate the ability of firms to equate supply and demand via price adjustments, some alternative mechanism (such as rationing or queuing) is often adopted to do so. For example, during the Second World War in the US, limits on maximum prices, administered by the Office of Price Administration, were supplemented by rationing of various consumer products, including food, clothing, petrol, tyres and many other products. Economists usually find these alternative mechanisms to be wasteful; examples include the time spent in queuing, or an inability for administrative measures to match scarce goods to their highest-value uses.

Price controls are often justified in public debates as a mechanism for dealing with what are seen as 'excess profits' and/or 'price gouging'. The argument is frequently made that, in the absence of price controls, consumers will be treated unfairly and suppliers will obtain an unfair 'windfall'. Such arguments often arise in two contexts: monopoly franchises (discussed above), and what are seen as temporary disruptions to normal market circumstances. As an example of the latter, in response to the Arab oil embargo in 1973, President Nixon imposed price controls on petrol in the US, causing motorists to have to queue up for petrol. Some states adopted a rationing mechanism whereby drivers could only buy petrol on an every-other-day basis, depending upon their licence plate number. Supply shocks or shortages can pit economic and political logic directly against one another, as the temporarily elevated prices needed to direct a scarce good to those who most need it also create the elevated seller profits viewed as illegitimate.

See Also

- ▶ Logic of Consequences and Logic of Appropriateness
- ▶ Marginal Product
- ▶ Marginal-Cost Pricing

References

Breyer, S. 1982. *Regulation and its reform*. Cambridge, MA: Harvard University Press.

Joskow, P.L. 2008. Lessons learned from electricity market liberalization. *The Energy Journal* 29: 9–42. *The future*

of electricity: Papers in honor of David Newbery, Special issue.
Kahn, A. 1988. *The economics of regulation: Principles and institutions*. Cambridge, MA: The MIT Press.
Mankiw, N.G. 1998. *Principles of economics*. Fort Worth: The Dryden Press.
Morrison, S., and C. Winston. 1999. Regulatory reform of US intercity transportation. In *Essays in transportation economics and policy: A handbook in honor of John R. Meyer*, ed. J. Gomez-Ibanez, W.B. Tye, and C. Winston. Washington, DC: Brookings Institution Press.
Stigler, G. 1971. The theory of economic regulation. *Bell Journal of Economic Management Science* 2: 3–21.

Price Discrimination

Rajeev Bhattacharya
Berkeley Research Group, Boston, MA, USA

Abstract

Price discrimination requires sufficient separability of customers, sufficiently high costs of arbitrage and sufficient ▶ market power. It involves transferability of the good and/or transferability of demand. It can be categorized as first degree (or perfect), second degree (or self-selection), or third degree (multimarket). Its impact on consumer surplus is ambiguous.

Definition A supplier practises price discrimination if it charges different prices for different units of essentially the same good, with essentially the same marginal cost of supply, to either the same customer or to different customers.

Successful price discrimination requires the following:

- The seller must be able to sufficiently separate customers (e.g., discounts for students or seniors).
- The seller must be able to prevent resale, or at least make resale very costly, across segments (e.g., textbooks or pharmaceutical drugs are sold at very different prices in different countries, even when the suppliers have no redistribution intentions).
- The seller must have some amount of ▶ market power.

Essentially, two types of arbitrage can defeat price discrimination. One type involves *transferability of the good* from the low-paying to the high-paying consumer – in such a case, only one kind of customer pays the fixed component of a possible two-part tariff. Typically, transaction costs provide limits on the level of such arbitrage that is possible – medical treatment, travel and utilities provide examples where such transaction costs are quite high.

The second such type of arbitrage involves *transferability of demand*, where the producer uses self-selection devices to match a type of customer to the price.

To prevent arbitrage involving transferability of the good, the supplier tries to *reduce* the spectrum; to prevent arbitrage involving transferability of demand, the supplier tries to *enhance* the spectrum. As an example of the latter, consider the oft-quoted example from Dupuit (1849), quoted in Tirole (2000: 150):

> It is not because of the few thousand francs which would have to be spent to put a roof over the third-class carriages or to upholster the third-class seats that some company or other has open carriages with wooden benches . . . What the company is trying to do is prevent the passengers who can pay the second-class fare from traveling third-class; it hits the poor, not because it wants to hurt them, but to frighten the rich . . . And it is again for the same reason that the companies, having proved almost cruel to third-class passengers and mean to second-class ones, become lavish in dealing with first-class passengers. Having refused the poor what is necessary, they give the rich what is superfluous.

First-Degree (or Perfect) Price Discrimination

Each customer is charged the maximum willingness and ability to pay for every unit of the good; as a result, the entire consumer surplus is appropriated by the supplier. However, it is possible

that some segments that were not served under uniform pricing can be served under price discrimination.

Second-Degree (or Self-Selection) Price Discrimination

The supplier uses consumer behaviour to 'self-select' consumers into appropriate market segments. Examples include volume discounts that self-select consumers into less elastic (e.g., single individual) and more elastic (e.g., family) segments.

Firms also offer different peak and off-peak prices on mobile telephony, for instance, with the intention of self-selecting calls into business (less elastic) and pleasure (more elastic) categories – the critical point here is that, for most instances, the marginal cost of a peak call is about the same as that of an off-peak call.

Airlines (and some train companies) offer different prices for business and economy classes. Within these classes, discounts are offered for non-refundable tickets, Saturday-night stays or advanced purchase. The objective of this differential pricing is to self-select travel into less-elastic (e.g., business) and more-elastic (e.g., vacation) categories – for example, a business traveller will be reluctant to stay a Saturday night at the destination and will want flexibility. Despite the extra frills of business class, the marginal cost of an additional business class passenger is not very different from that of an additional economy class passenger, especially as a proportion of the fixed costs involved. And the marginal cost of an additional passenger is nearly identical for the different categories within each class.

Firms often offer different prices to current customers versus switching customers for essentially the same reasons.

Third-Degree (or Multimarket) Price Discrimination

The supplier uses observable signals related to a consumer's demand and charges prices based on these signals. Examples include student/senior discounts at cinemas. Students and seniors typically have a more elastic demand, and the status is verifiable.

It is a perception that women have a less elastic demand for dry-cleaning than men; as a result, women pay more for essentially the same dry-cleaning services. Women's clothes are usually distinguishable from men's clothes, so preventing arbitrage between the customer segments is not difficult.

It is well documented that candidates who visit a college campus prior to admissions decisions have a less elastic demand for that particular institution. Such visits are typically coordinated through the college, so the college knows which candidates have visited. Similarly, a student from a poorer background is likely to have a more elastic demand for education at a college. It is more difficult to separate out the economic categories, given the incentives to under-report a family's economic circumstances. However, it is not the case that a candidate can arbitrage a college aid package with another candidate – transferability of the good is not a factor here. These reasons explain, to an extent, the differentials in financial aid packages offered to candidates, over and above the merit-based differentials.

Academic journals typically have a sliding scale for subscriptions (e.g., a low rate for students, a higher rate for academics and a still higher rate for libraries). Some journals also have different rates depending on the country of the subscriber, even though there might not be redistribution issues. These differentials are because these segments have different elasticities of demand for journal subscriptions. Coca-Cola's 'smart' vending machines, which charge different prices depending on outside temperatures, fall into this category as well.

Marginal revenue equals marginal cost holds for each market segment, and the inverse elasticity rule holds for each market segment; that is, for each segment i (price in segment i − marginal cost)/(price in segment i) $= -1/$(own-price elasticity of demand in segment i). In other words, the supplier should charge more in market segments with less elastic demand.

Price discrimination reduces welfare if it does not increase total output. If the total output was to remain the same or decrease under price discrimination, the marginal rate of substitution would differ across customers and, therefore, there would be lower welfare under price discrimination than under uniform monopoly pricing. In other words, for price discrimination to be welfare-increasing, it is a necessary condition that total output be higher under price discrimination.

For the special case of linear demand functions, if we were to impose the additional condition that all markets would be served under price discrimination, then welfare would be lower under price discrimination. In the absence of the additional condition of all markets being served under price discrimination, it is easy to visualize scenarios where price discrimination would lead to a Pareto improvement. The welfare effects of price discrimination are, therefore, ambiguous.

The Robinson–Patman Act in the United States, though rarely used currently against price discrimination, applies to price discrimination and injury to ▶ competition in sales of commodities of like grade and quality in commerce. Such price discrimination can be legally justified through cost differentials or through meeting a competitor's price.

See Also

- ▶ Arbitrage and Its Limits
- ▶ Competition
- ▶ Market Power
- ▶ Market Segmentation

References

Tirole, J. 2000. *The theory of industrial organization*. Cambridge, MA: The MIT Press.

Further Reading

Besanko, D., D. Dranove, M. Shanley, and S. Schaefer. 2010. *Economics of strategy*. Hoboken: Wiley.
Devinney, T. (ed.). 1988. *Issues in pricing*. Lexington: Lexington Books.
Dolan, R., and H. Simon. 1996. *Power pricing*. New York: Free Press.
Gellhorn, E. 1986. *Antitrust law and economics*. St Paul: West Publishing Company.
Nagle, T., and R. Holden. 2002. *The strategy and tactics of pricing*. Upper Saddle River: Prentice Hall.

Price Leadership

Edward F. Sherry
Berkeley Research Group, Emeryville, CA, USA

Definition A firm is a price leader if other firms in the industry tend to follow the leader's pricing behaviour, whether price increases or price decreases.

A firm is a price leader if other firms in the industry tend to follow the leader's pricing behaviour, whether price increases or price decreases.

Formal agreements between leaders and followers are typically prohibited by the relevant antitrust or competition laws. However, in determining whether an empirical pattern of price announcements by one firm, followed by parallel price movements by the firm's competitors, does or does not violate the antitrust law can be a complex matter. Such behaviour is often seen as a form of tacit collusion, although it can also be argued that such patterns may reflect nothing more than industry-wide recognition of changes in competitive circumstances that would result in parallel price movements even in the absence of any collusive agreement.

In some industries, notably the airline industry, some firms may announce future price changes in advance, and wait to see if other firms in the industry follow suit. If they do, the price change may be implemented, but if other firms do not follow, the initial price move may be rescinded. Determining whether or not such behaviour amounts to tacit collusion can be a difficult task.

Game theorists sometimes study price leadership under the heading 'Stackelberg competition' (see Fudenberg and Tirole 1991). In such

models, the standard assumption is that one firm 'moves first', and other firms respond in a privately optimal fashion to the first mover's behaviour.

One interesting question in price leadership models is how a firm becomes the price leader. It is often assumed that the price leader is one of the largest, if not the largest, firm in the industry. However, there have been instances in which a smaller firm is seen as the price leader, with larger firms following its lead. To some extent this may be explained as a particular variant of a 'coordination game' (see Schelling 1960), although how firms arrange to coordinate their conduct behind that of a particular, smaller leader remains a difficult theoretical question.

See Also

► Futures Markets

References

Fudenberg, D., and J. Tirole. 1991. *Game theory*. Cambridge, MA: The MIT Press.
Schelling, T. 1960. *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Price Taking

Edward F. Sherry
Berkeley Research Group, Emeryville, CA, USA

Definition A buyer is a price taker when he believes he is unable to affect the price he pays for goods or services. Alternatives to price taking, which are more popular in some cultures than others, include bargaining and haggling over price.

A buyer acts as a price taker in situations where he believes that he is unable to affect the price that he must pay for the goods or services that he purchases. In perfectly competitive markets, the usual explanation is that buyers have no *incentive* to buy at a price above the market-clearing price, as they can buy all that they want at that price; and that sellers have no *ability* to buy at a price below the market-clearing price, because sellers can sell all their output to other buyers who are willing to buy at that price.

In mainstream textbook-level economics, it is usually assumed that buyers act as price takers except in oligopsonistic markets, in which large buyers have some degree of market power.

That said, in many cultures there is a significant degree of bargaining or haggling over price, even when the buyer is quite small relative to the overall size of the market as a whole. This can sometimes lead to a clash of cultures, in which individuals accustomed to such bargaining find themselves in situations in which such bargaining is uncommon and may be seen as unusual or even rude (or vice versa).

In many if not most transactions in urbanized Western societies, the seller (e.g., a retail store or wholesaler) sets a 'posted' price and does not haggle over price (with the possible exception of standardized discounts, such as formulaic quantity discounts or prompt-payment discounts). But even in such cultures, in some industries – notably the markets for new and used cars, markets for non-fungible goods such as antiques, markets for customized goods and services, and markets for surplus or obsolescent goods – one often observes a significant degree of negotiation over price. In such industries, buyers who come from a culture in which haggling is not common often find themselves uncomfortable dealing with sellers for whom such bargaining is an everyday occurrence.

This raises the interesting sociological/economic question: why is bargaining or negotiation common in some cultures and/or in some industries, but not in others? To a significant extent, the former appears to be largely a cultural issue. But from an economic perspective one would predict that the willingness to haggle would depend on the transaction costs associated with engaging in such negotiations, especially as a measure relative to the amount of money, potential profit, and potential consumer surplus involved in the potential transaction. Negotiations take time and effort

for both the buyer and the seller, which is effectively wasted if the parties are unable to come to an agreement. Haggling appears to be more common in societies in which time is cheap and money is scarce.

Haggling is also sometimes (though incorrectly) seen as a 'zero-sum' game, in which the seller's desire for a higher price is thought to be diametrically opposed to the buyer's desire for a low price. The fallacy here is that there may be mutually beneficial gains from trade, which may only be achieved if the parties are able to reach agreement, and the negotiation process may be necessary in order to determine *whether* such a mutually beneficial deal exists. Against this must be offset the time and effort associated with haggling.

A somewhat different explanation involves ▶ Principal-Agent issues, in which selling firms are reluctant to delegate negotiating authority to low-level, front-line employees, because of concerns that the firm's interests and the employee's interests are not fully aligned, especially when employees are compensated on the basis of sales generated rather than profits. This helps explain the common practice in which hierarchically organized sellers often delegate some limited authority to make price concessions to lower-level employees, reserving the authority to make larger price concessions to individuals higher up in the hierarchy, and the situation in which principals are more willing to haggle than to authorize their agents/subordinates to do so.

Another explanation looks to micromarket factors and strategic considerations. Rather than there being 'the' single market-clearing price traditionally assumed in textbook discussions to be known to all market participants, in practice there may be a range of prices charged by different sellers for what is effectively the same commodity, and buyers may either engage in (costly) search for the best price or incur the transaction costs of bargaining with a seller hoping to get a better price.

Moreover, while any given buyer may be small relative to the overall market, that buyer's business may be important to, or account for a significant fraction of, any given seller's business, especially in the short term. When the seller's 'posted' price is above the seller's incremental cost of supplying the good in question to a particular buyer, the seller risks foregoing the profit margin on such sales if he insists on not discounting and thereby loses the sale, unless he can be assured that another buyer is 'waiting in the wings' willing to purchase at the posted price. This in turn depends on the uniqueness and/or fungibility of the product in question, on the 'thickness' of the markets at issue, and on the transparency of prices.

That said, sellers are also reluctant to grant discounts too readily, for fear that buyers will come to expect such discounts. Rational sellers may wish to develop a reputation for being unwilling to negotiate if doing so results in higher overall profits, especially once the transaction costs of haggling are taken into account.

See Also

- ▶ Principal Agent
- ▶ Transaction Cost Economics

Principal Agent

David Bardolet
Bocconi University, Milan, Italy

Definition Principal-agent conflicts arise when a party (the agent) is compensated for performing a task that is useful to another party (the principal) but costly to the agent, and that contains elements that are difficult to observe due to asymmetric information, uncertainty or risk.

Principal-agent (agency) theory is a staple in economics (Stiglitz 2008), organization theory and political science. It originally described the conflicts between owners and managers of large organizations as well as the mechanisms that might curb the managers' opportunistic behaviour, such as equity ownership (Jensen and

Meckling 1976), efficient capital and labour markets (Fama 1980) and effective boards of directors (Fama and Jensen 1983). Later, this view expanded to include other relationships such as employer–employee or buyer–supplier, and focused on formalizing the most efficient contract alternatives under different scenarios of uncertainty, information or risk attitudes.

In standard economic theory, a party in need of certain goods or services could hire and pay another party to deliver them. ▶ Agency theory identifies and tries to solve two types of problem that arise from that relationship. First, the goals of the agent and the principal might differ, leading the agent to pursue her own self-interest in performing the contracted task. Crucially, it is difficult for the principal to know what the agent is actually doing (or capable of doing), which increases the incentives for the agent to behave opportunistically. For instance, a divisional manager in a multi-business firm might misrepresent the true value of her unit's investment opportunities in order to secure unmerited resources at the risk of lowering overall firm performance. Second, the principal and the agent might have different attitudes towards risk, which might lead them to prefer different actions in light of the task. A risk-averse manager, for example, might reject business opportunities that are otherwise good for the firm's owners, which generally have a more risk-neutral profile. Thus, principal-agent models assume self-interest and bounded rationality among actors as well as the existence of information asymmetry and goal conflicts.

The principal-agent approach is contained to different degrees in several mainstream organization views (Eisenhardt 1989), which facilitates the integration of agency theory into those views. For instance, transaction cost theory assumes self-interested and boundedly rational individuals having conflicting goals in a contractual relationship. The difference is that transaction cost theory is concerned about choosing an efficient governance form to manage those relationships while principal-agent theory focuses on finding the incentives that would align the agent's behaviour with the principal's interest. Similarly, behavioural theories of the firm view agency conflicts as part of the organizational environment that managers face, although their main focus is on the necessary managerial capabilities for coordination and adaptation (rather than formal contracts) and satisficing behaviour (rather than utility maximization). Political models of the firm also assume self-interest and goal conflict but they differ in that they resolve those conflicts through power mechanisms such as negotiation and not through individual incentives.

Principal-agent theory has made two key contributions to management theory. One is its recognition of information as a valuable organizational variable, which has highlighted the role of organizational processes (such as budgeting, financial reporting and corporate board monitoring) that provide that information. Firms with better monitoring/reporting systems can control managerial self-serving behaviour and achieve better outcomes in their allocation of resources and employee compensation. An alternative to reducing information asymmetry through (costly) organizational monitoring systems is to use contract incentives. One of the most popular applications of the principal-agent model focuses on employee compensation and its ability to align the diverging interests of principals (owners, employers) and agents (managers, employees). Agency theory conceives two basic alternatives for employee compensation: paying a fixed salary or paying on the basis of some observed output of the worker or company. Principal-agent theory identifies a fundamental trade-off between financial incentives and organizational information systems, and shows that performance-based pay will be more attractive to firms when they have less information about the employee's effort or skill level. The more accurately the firm monitors its employees, the less it will rely on performance-based compensation and vice versa. Other factors that scholars have identified as influencing compensation contracts include outcome uncertainty and measurability, time and task programmability.

The second contribution is the consideration of risk attitude differences in the contracts between principal and agent. In firms, agents often take too little or too much risk with respect to the

preferences of a generally risk-neutral principal, a situation that might require the principal to pay a premium to correct the distortion. Agency theory postulates that such distortion will increase in situations of high outcome uncertainty (innovation processes, technology-intensive industries, etc.).

See Also

- ▶ Agency Problems
- ▶ Agency Theory
- ▶ Asymmetric Information
- ▶ Incentive Design
- ▶ Opportunism

References

Eisenhardt, K.M. 1989. Agency theory: An assessment and review. *Academy of Management Review* 14: 57–74.

Fama, E. 1980. Agency problems and the theory of the firm. *Journal of Political Economy* 88: 288–307.

Fama, E., and M. Jensen. 1983. Separation of ownership and control. *Journal of Law and Economics* 26: 301–325.

Jensen, M., and W. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3: 305–360.

Stiglitz, J. 2008. Principal agent. In *The new Palgrave dictionary of economics*, ed. S.N. Durlauf and L.E. Blume. New York: Macmillan.

Prisoner's Dilemma

Elif Ketencioglu and Dan Lovallo
The University of Sydney, Faculty of Economics and Business, Sydney, NSW, Australia

Definition A fundamental problem in game theory, the Prisoner's Dilemma shows that players may not choose cooperative behaviour even when it is in their best interest to do so.

The Prisoner's Dilemma problem was inspired by laboratory experiments at the ▶ rand corporation in 1950s. Some of the analysts argued that human behaviour may not always yield in Nash equilibrium (Nash 1950), where players will not benefit from changing their strategies while the others keep theirs constant. Instead, in experiments in which two players repeated a game 100 times, Flood (1952) and Dresher (1961) observed cooperation. Later on, Albert Tucker, a former professor of John Nash, invented an anectode in which two people were arrested for a crime; he used this story in his lectures at the Psychology Department at Stanford University to present the cooperative behaviour equilibrium.

In the primary version of Prisoner's Dilemma anecdote, two people have been arrested for a crime and put into isolated cells. Since the police lack insufficient information to convict either suspect, the suspects are required to give consequent testimonies against each other. The suspects are given the option to confess (defect) or remain silent (cooperate), and they are told that if their confession incriminates the other suspect they will be released with a reward, provided that the other suspect's confession does not incriminate them. If neither confesses, both will be released due to insufficient evidence, and no rewards will be given. The dilemma arises from the fact that whatever the other player does, each one is better off when they confess as opposed to remaining silent. This particular problem of ▶ game theory has gained considerable interest since it addresses the fundamental conflict between what is a rational choice for an individual in a group versus the group as a whole.

In its simplest form, the payoff matrix of a two-player Prisoner's Dilemma game with cardinal payoffs is described as below:

		Player 1	
		C	D
Player 2	C	R_r, R_C	S_r, T_C
	D	T_r, S_C	P_r, P_C

As can be seen, the two players have two possible moves: to cooperate (C) or defect (D), corresponding with remaining silent or confessing

in the mentioned anecdote above. The order of payoffs is $T>R>P>S$, where R is the reward payoff received by each player when they cooperate and P is the punishment received by each player when they defect. T is the temptation that each player receives if he defects and the other cooperates, and S is the sucker payoff received by the player who cooperates when the other player defects. In this simple form, it is assumed that the payoff values are the same for each player (symmetric game) and that the payoffs have an ordinal structure. Further, standard game theory assumes that each player is rational and that there is no private information, that is, each player knows only his or her own payoffs.

In order to see the dilemma, assume that the column player cooperates; then the row player will choose to defect, since $T_r > R_r$. The same applies when the column player defects; then the row player will again choose to defect, since $P_r > S_r$. It is also worth noting that the outcome (D, D) of both players defecting is the game's only strong Nash equilibrium; that is, it is the only outcome from which each player could only do worse by unilaterally changing their move. This equilibrium is preserved regardless of whether the game is played sequentially or simultaneously. Flood and Dresher's interest in their dilemma seems to have stemmed from their view that it provided a counterexample to the claim that the Nash equilibria of a game constitute its natural 'solutions'.

In an iterated game, the Prisoner's Dilemma scheme is played repeatedly where each player can change strategies according to their experiences from previous rounds. Standard economic theory states that despite the incentive to punish the opponent, players will continue to perform non-cooperative behaviour. However, results from Aumann (1959) show that in an iterated game with indefinite rounds, players will choose to cooperate.

See Also

- ▶ Game Theory
- ▶ RAND Corporation

References

Aumann, R. 1959. Acceptable points in general cooperative n-person games. In *Contributions to the theory 23 of games IV*, Annals of Mathematics Study 40, ed. R.D. Luce and A.W. Tucker. Princeton: Princeton University Press.

Dresher, M. 1961. *The mathematics of games of strategy: Theory and applications*. Upper Saddle River: Prentice Hall.

Flood, M. 1952. *Some experimental games: Research memorandum*. Santa Monica: RAND Corporation.

Nash, J. 1950. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences* 36: 48–49.

Process-Oriented Strategic Theory

Catherine A. Maritan
Syracuse University, Whitman School of Management, Syracuse, NY, USA

Abstract

Research in strategic management is often categorized as concerning either strategy content or strategy process. Strategy process deals with one or more aspects of how strategies are developed, implemented or changed. Strategy process theories and models outline patterns in decisions or actions over time, and address mechanisms and paths that shape and govern strategies. There are both theories of strategy-making and theories that address linkages between the development or implementation of strategies and the creation of competitive advantage.

Definition Process-oriented strategic theory refers to models describing or explaining how strategies are developed, implemented and changed. Strategy process models outline patterns in decisions or actions over time, and address mechanisms and paths that shape and govern strategies.

Research in strategic management is often categorized as concerning either strategy content or strategy process. Although this dichotomy

between what is done (content) and how it is done (process) has been acknowledged as being a false one (Schendel 1992), and there have long been calls for integrating process and content research (e.g., Jemison 1981; Zajac 1992), separate bodies of literature have developed. The domain of strategy process research is broad, and there exist a variety of process models that draw on different underlying theoretical perspectives and logics. There are even different notions of what the term 'process' means in strategy research (Van de Ven 1992). However, what all this work has in common is that it addresses one or more aspects of how strategies are developed, implemented or changed (Chakravarthy and White 2002). Strategy process models outline patterns in decisions or actions over time, and address mechanisms and paths that shape and govern strategies.

While more general theories of decision-making and organizational change have informed our understanding of strategy process, those are not theories of strategy-making per se. However, there are theories and models that address strategy-making processes explicitly. Many, though not all, were developed based on field observations in organizations, and capture complexities of interrelationships among organizational structures, systems, actors and contexts. There are also process theories that link strategy processes with performance outcomes and competitive advantage. Some of the significant process theories and models of strategy are described briefly below.

Process Theories of Strategy-Making

One of the most influential process theories is Henry Mintzberg's (1978) theory of ▶ emergent strategy. Reacting to the then dominant view in the literature, Mintzberg argued that strategies are not necessarily first formulated through a rational decision process and then implemented as planned; some strategies can emerge. He defined strategy as 'a pattern in a stream of decisions' (Mintzberg 1978: 935), and distinguished intended strategy, which can be identified a priori and may or may not come to fruition, from realized strategy, which can be recognized only after patterns of behaviour and activity can be observed. Realized strategies are a combination of deliberate strategies intended by top management and subsequently carried out, and emergent strategies, which were carried out despite not having been intended. Development of the emergent strategy concept was significant because it led to a broader perspective of strategy-making that takes into account aspects of an organization's structure and context as potential drivers of emergence. While deliberate and emergent strategies are theoretically distinct, in practice strategies are likely to combine elements of both (Mintzberg and Waters 1985), and we see this combination in many of the major process theories of strategy-making.

Another theory of how strategies are developed is ▶ James Brian Quinn's (1980) process of ▶ logical incrementalism. In this view, strategies emerge in an incremental, iterative manner as executives make series of smaller decisions involving partial commitments rather than large, long-term commitments to a fully developed strategy for the entire organization. The underlying premise is that it is not possible to predict all the major events and forces that will shape the future of an organization, so it is preferable to deal with individual issues and problems relating to different aspects of the organization using the best information available at the time. This approach accommodates changing circumstances and assumptions, and permits experimentation, evaluation and learning during the process. In contrast to other incremental approaches to management such as muddling through (Lindblom 1959), this type of incrementalism is described as being logical because 'it is a conscious, purposeful, and proactive . . . practice' (Quinn 1980: 58).

Although many strategy process models, such as logical incrementalism, can be described as evolutionary in the dictionary definition sense of the term, there are strategy theories that employ the logic of variation, selection and retention from evolutionary biology. A prominent example is Burgelman's intra-organizational ecological model of strategy-making that views an organization as 'an ecology of strategic initiatives that

emerge in patterned ways' (Burgelman 1983a, 1991: 240). According to this model, there are two types of initiatives: induced initiatives that come out of the existing strategy and autonomous initiatives that fall outside the scope of current strategy. Autonomous initiatives are emergent and are often the result of actions and decisions by managers at different levels and in different parts of the organization; they introduce variation into the set of initiatives. There is a selection process through which resources are allocated to some initiatives but not others. Retention, in this model, refers to incorporating the results of pursuing the selected initiatives into the organization's strategy. If autonomous initiatives are selected and retained, they can introduce significant changes to the strategy. Thus, Burgelman provides a theoretical explanation of a mechanism by which emergent strategies resulting from autonomous initiatives can contribute to an organization's realized strategy.

While resource allocation is an important element of Burgelman's ecological model, Noda and Bower (1996) present a strategy-making model with resource allocation at its centre. They developed a theory of strategy-making based on the Bower-Burgelman (Bower 1970; Burgelman 1983b) model of the resource allocation process. That model describes a complex multi-stage process in which managers at different hierarchical levels of a firm play distinct roles in resource allocation, and incorporates cognitive processes of individuals, social processes involving relationships among individuals and groups, and political processes by which individuals or groups exert power or influence. Noda and Bower conceptualize strategy-making as iterated processes of resource allocation, and propose that understanding how series of resource allocation decisions are made over time provides insights into how strategies are formed. Comparing the resource commitments made by firms competing in the same industry, beginning with similar endowments and facing similar opportunities, they found that differences in the firms' resource allocation processes resulted in divergent business development experiences. Noda and Bower's model speaks to both Mintzberg's notion of emergent strategy in terms of the capturing multi-level and interrelated managerial activities leading to outcomes that were not necessarily planned, and to Quinn's incrementalism in terms of the iterative nature of series of resource allocation decisions that incorporate experimentation and learning.

Most process theories of strategy-making tend to incorporate roles of multiple actors in different parts of an organization and integrate cognitive, political and social processes, either implicitly or explicitly; however, there are some models that emphasize particular roles or particular processes more than others. For example, Floyd and Wooldridge (2000) focus on middle-level managers, and develop a process model of their role in strategy-making. Pettigrew (1973, 1977) and Narayanan and Fahey (1982) define organizations as political entities, and offer a view of strategy-making as a political process in which individuals and groups make demands and mobilize power around their demands. Strategies emerge from these internal dynamics. Kaplan (2008) emphasizes cognitive processes, and presents a model of strategy-making under uncertainty as framing contests.

Strategy Processes and Competitive Advantage

Strategy process research has sometimes been criticized for traditionally focusing more on process outcomes than on strategy outcomes (Chakravarthy and White 2002). More recently, scholars have proposed process theories that address linkages between the development or implementation of strategies and the creation of competitive advantage.

Hart presents and tests (Hart 1992; Hart and Banbury 1994) a framework of strategy-making processes that integrates multiple models on the basis of the roles played by participants to derive a set of strategy-making modes. Taking a contingency perspective, he argues that particular modes will lead to superior performance outcomes in particular environmental contexts, and that higher performing firms are able to combine modes.

Winter and Szulanski (2001) offer a theory of replication as strategy. Unlike the models discussed thus far, this is not a strategy-making process model; it is a process theory that addresses a type of expansion and growth. Replication as strategy is based on Nelson and Winter's (1982) theory of evolutionary economics and grounded in the economics of information. A firm replicates itself, meaning it creates a large number of similar operations to deliver a product or service. This is accomplished through a repeated knowledge transfer process, which involves determining what aspects of the business model are replicable and worth replicating, and developing a template for transferring the necessary core knowledge to new operations. Winter and Szulanski link replication not only to process outcomes but to the creation of sustainable competitive advantage. A replication strategy can be difficult or costly for a competitor to imitate, because the firm has superior access to a working model of the replication template and has learned from experience.

The capability to replicate a business model is referred to as a dynamic capability in Winter and Szulanski's process theory. Dynamic capability theory (Teece et al. 1997; Eisenhardt and Martin 2000) more broadly can also be characterized as a process theory of strategy (Helfat et al. 2007; Teece 2007). ▶ Dynamic capabilities provide an organization with the capacity to create, extend or modify the resource base it uses to compete, and managerial and organizational processes provide their underpinnings (Teece 2007). Under certain circumstances dynamic capabilities may lead to superior performance and competitive advantage, and the processes associated with their development and deployment contribute to those performance outcomes. In contrast to many strategy process theories and models, dynamic capability theory bridges strategy content and process by addressing both the 'what' and the 'how' of resource and capability change (Maritan and Peteraf 2007).

See Also

- ▶ Dynamic Capabilities
- ▶ Emergent Strategy
- ▶ Logical Incrementalism
- ▶ Organizational Ecology
- ▶ Resource Allocation Theory
- ▶ Strategic Decision-Making
- ▶ Strategic Implementation

References

Bower, J.L. 1970. *Managing the resource allocation process*. Boston: Harvard University Graduate School of Business Administration.

Burgelman, R.A. 1983a. A model of the interaction of strategic behavior, corporate context, and the concept of strategy. *Academy of Management Review* 8: 61–70.

Burgelman, R.A. 1983b. A process model of internal corporate venturing in the diversified major firm. *Administrative Science Quarterly* 28: 223–244.

Burgelman, R.A. 1991. Intraorganizational ecology of strategy making and organizational adaptation: Theory and field research. *Organization Science* 2: 239–262.

Chakravarthy, B.S., and R.E. White. 2002. Strategy process: Forming, implementing and changing strategies. In *Handbook of strategy and management*, ed. A. Pettigrew, H. Thomas, and R. Whittington. London: Sage.

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

Floyd, S.W., and B. Wooldridge. 2000. *Building strategy from the middle*. Thousand Oaks: Sage.

Hart, S.L. 1992. An integrative framework for strategy-making processes. *Academy of Management Review* 17: 327–351.

Hart, S.L., and C. Banbury. 1994. How strategy-making processes can make a difference. *Strategic Management Journal* 15: 251–269.

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.

Jemison, D.B. 1981. The importance of an integrative approach to strategic management research. *Academy of Management Review* 6: 601–608.

Kaplan, S. 2008. Framing contests: Strategy making under uncertainty. *Organization Science* 19: 729–752.

Lindblom, C.E. 1959. The science of muddling through. *Public Administration Review* 19: 79–88.

Maritan, C.A., and M.A. Peteraf. 2007. Dynamic capabilities and organizational processes. In *Dynamic capabilities: Understanding strategic change in organizations*, ed. C.E. Helfat, S. Finkelstein, W. Mitchell, M.A. Peteraf, H. Singh, D.J. Teece, and S.G. Winter. Malden: Blackwell.

Mintzberg, H. 1978. Patterns in strategy formation. *Management Science* 24: 934–948.

Mintzberg, H., and J.A. Waters. 1985. Of strategies, deliberate and emergent. *Strategic Management Journal* 6: 257–272.
Narayanan, V.K., and L. Fahey. 1982. The micro-politics of strategy formulation. *Academy of Management Review* 7: 25–34.
Nelson, R.R., and S.G. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.
Noda, T., and J.L. Bower. 1996. Strategy making as iterated processes of resource allocation. *Strategic Management Journal* 17: 159–192.
Pettigrew, A.M. 1973. *The politics of organizational decision-making*. London: Tavistock.
Pettigrew, A.M. 1977. Strategy formulation as a political process. *International Studies of Management & Organization* 2: 78–87.
Quinn, J.B. 1980. *Strategies for change*. Homewood: Richard D. Irwin.
Schendel, D.E. 1992. Introduction to the winter 1992 special issue: Fundamental themes in strategy process research. *Strategic Management Journal* 13(winter special issue): 1–3.
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.A. Shuen. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18: 504–534.
Van de Ven, A.H. 1992. Suggestions for studying strategy process: A research note. *Strategic Management Journal* 13(summer special issue): 169–191.
Winter, S.G., and G. Szulanski. 2001. Replication as strategy. *Organization Science* 12: 730–743.
Zajac, E.J. 1992. Relating economic and behavioral perspectives in strategy research. *Advances in Strategic Management* 8: 69–96.

Product Champion

Ram Rao and Yu Wang
University of Texas at Dallas, School of Management, Richardson, TX, USA

Abstract

Product champions play an important role in the process of ▶ innovation and ▶ new product development. We define product champions and describe how they contribute to innovation by driving projects through the 'valley of death'. Due to the informal nature of the product champion's role, strategically managing their behaviour poses a potential challenge. We explore some strategies firms can adopt to effectively foster and manage championing behaviour.

Definition Product champions emerge informally within firms and organizations during the process of innovation and new product development. They passionately identify with new ideas and express enthusiasm and confidence about the success of the innovation. They also persist under adversity, and get the right people involved to move the process of innovation and commercialization through critical stages.

Product Champion: Who They Are and What They Do

Product champions emerge informally within firms and organizations during the process of ▶ innovation and ▶ new product development. They passionately identify with new ideas, express enthusiasm and confidence about the success of the innovation, persist under adversity, and get the right people involved to move the process of innovation and commercialization through critical stages.

Schon (1963) was the first to introduce the concept of product champion. He proposed that champions are needed to break the resistance to change in large organizations. The idea that increasing organizational complexity calls for champions to sponsor innovations is well accepted (Maidique 1980; Frey 1991; Shane 1994). What is interesting is that champions are not uniquely identified with functional areas in a firm or with a certain place in the firm's hierarchy. They can arise from different functional areas such as marketing, R&D or production (Markham et al. 1991; Markham and Aiman-Smith 2001), and also from higher or lower levels in an organization (Maidique 1980; Pinchot 1985; Morita et al. 1986; Day 1994). They can be found across a wide range of firms, irrespective of the firms' level of technology, the markets the firms operate in (B2B or B2C), and whether or not they have in place formal new product development

processes (Markham and Griffin 1998; Markham and Aiman-Smith 2001).

Empirical research consistently supports the notion that champions play a key role in moving technologies from the laboratory to the market (e.g., Markham et al. 1991; Markham and Griffin 1998; Howell and Shea 2001, 2006; Lichtenthaler and Ernst 2009). In particular, it has been argued that product champions are especially important in keeping projects alive when neither the 'technology-push' from the lab nor the 'need-pull' from the market is strong enough to activate an innovation lifecycle (Beath and Ives 1988). Markham (2002) refers to the gap between discovery and commercialization as a 'valley of death'. To drive a project across this valley of death, the champion must accomplish eight steps (excerpted and modified from Markham 2002: 33):

1. Discover that the research has commercial value;
2. Manifest the discovery as a product;
3. Communicate the potential through a compelling business case;
4. Acquire resources needed to establish potential;
5. Use resources to reduce risk;
6. Seek approval of the project for formal development;
7. Translate the project into the criteria used for approval;
8. (After approval) develop and launch the product (not done exclusively by the champion).

What Makes Product Champions

Past research has identified characteristics of product champions: their personalities, leadership behaviours and influence tactics (Schon 1963; Chakrabarti 1974; Howell and Higgins 1990; Shane 1994; Howell and Shea 2001; Roure 2001; Howell 2005; Howell et al. 2005). Successful champions are often found to be inspiring, charismatic leaders who are risk-taking, enthusiastic about innovation, driven by their vision, knowledgeable about the company, technology and market, persistent under adversity, well connected internally and externally, and politically astute. Although champions are not assigned to their roles but rather emerge in an informal way, they identify with an innovation idea and promote it as a cause to a degree that goes far beyond the requirement of their job. Champions typically use informal as well as formal channels within their firm. They are proficient in a wide variety of influence tactics that they rely on to get the right people on board for ▶ product innovation

Can Product Champions be Fostered?

There is little consensus on whether potential champions can be trained to become more effective or if they can be identified. Schon (1963: 85) claims that 'It is extremely difficult in practice for top management to admit the need for such a man, since the implication in doing so is that something is wrong with … the organization's "climate for creativity" … In fact, there is evidence that these men cannot be hired and "developed" the way some others can.' Similarly, Chakrabarti (1974) argues that the role of product champion is an informal and non-routine one that primarily depends on the individual's choice and initiative. Champions can also be hard to manage, because they seek autonomy from organizational norms and rules and frequently circumvent organizational hierarchy (Shane 1994). Hence, top management may not necessarily have either the incentive or the means to identify and coach champions.

On the other hand, some researchers have developed measures of effective championing behaviour (e.g., Howell et al. 2005), based on which, assessing candidates' 'champion potential' may be possible. Howell (2005) offers a few ideas about how to identify and coach potential champions, although how to implement the ideas in practice is unclear. In fact, Markham and Aiman-Smith (2001) list the following topics as unexamined issues regarding championing:

- Training people to be champions
- Managerial support of champions
- Getting champions to do what you want them to do

– Rewards for championing behaviours
– Career implications for champions.

Regardless of whether product champion positions can be assigned or not, there is no doubt that the emergence and effectiveness of champions hinges upon the organizational environment created by management philosophy and ▶ innovation strategies. Initial empirical evidence (Markham and Griffin 1998; Lichtenthaler and Ernst 2009) suggests that product champions are more prevalent in supportive environments offered by firms with strategies that emphasize innovativeness, and that demonstrate their support for new product development. Having formal new product development processes in place is not necessarily a substitute for champions. Instead, champions and formal new product development processes seem to go together in supporting innovation (Markham and Aiman-Smith 2001). Hence, to foster championing behaviour, it is critical for firms to establish a supportive environment while openly recognizing the contributions of champions within the organization.

Potential Caveats

Although champions improve the performance of projects by obtaining resources to keep projects alive, reducing the cycle time and boosting a team's shared belief in innovation success, extant empirical evidence is not entirely consistent on how champions influence products' market success and the performance of the firm (e.g., Markham 1998; Markham and Griffin 1998; Howell and Shea 2001, 2006). A few case studies suggest that although championing behaviour has primarily positive effects, it may come with caveats (e.g., Royer 2003; Mirza et al. 2008).

In particular, some researchers underscore the need for firms to prevent escalation of commitment to failing projects. Royer (2003: 53) states that 'When it reinforces others' perceptions and desires, collective belief is often contagious and can easily spread among the various decision makers who control a project's fate', and often the original true believer is the project champion. Therefore, misplaced faith in the projects can result in 'blindness to signs of failure', which causes firms to keep allocating valuable time and recourses to failing projects. Royer concludes that the value of both product champion and 'exit champion' (i.e., 'someone who is able to pull the plug on a project before it becomes a money sink') should be appreciated. Practical advice from Royer includes:

– Put in place a well-defined review process (p. 50)
– Assemble teams not entirely composed of people who are enthusiastic about the project – include sceptics as well (p. 55)
– Directly involve exit champions in the project. (p. 55).

Along the same lines, Boulding et al. (1997) suggest that to effectively reduce commitment to a losing course of action, firms need to have predetermined decision rules in place or introduce a new decision-maker at the time of the stop/no stop decision.

See Also

▶ Innovation
▶ Innovation Strategies
▶ New Product Development
▶ Product Innovation

References

Beath, C. M., and B. Ives. 1988. The information technology champion: Aiding and abetting, care and feeding. In *Proceedings of the Twenty-First Annual Hawaii International Conference on System Sciences* (vol. 4), 115–123.

Boulding, W., R. Morgan, and R. Staelin. 1997. Pulling the plug to stop the new product drain. *Journal of Marketing Research* 34: 164–176.

Chakrabarti, A.K. 1974. The role of champion in product innovation. *California Management Review* 17: 58–62.

Day, D.L. 1994. Raising radicals: Different processes for championing innovative corporate ventures. *Organization Science* 5: 148–172.

Frey, D. 1991. Learning the ropes: My life as a product champion. *Harvard Business Review* 69: 46–52.

Howell, J.M. 2005. The right stuff: Identifying and developing effective champions of innovation. *Academy of Management Executive* 19: 108–119.
Howell, J.M., and C.A. Higgins. 1990. Champions of technological innovation. *Administrative Science Quarterly* 35: 317–341.
Howell, J.M., and C.M. Shea. 2001. Individual differences, environmental scanning, innovation framing, and champion behavior: Key predictors of project performance. *Journal of Product Innovation Management* 18: 15–27.
Howell, J.M., and C.M. Shea. 2006. Effects of champion behavior, team potency, and external communication activities on predicting team performance. *Group & Organization Management* 31: 180–211.
Howell, J.M., C.M. Shea, and C.A. Higgins. 2005. Champions of product innovations: Defining, developing, and validating a measure of champion behavior. *Journal of Business Venturing* 20: 641–661.
Lichtenthaler, U., and H. Ernst. 2009. The role of champions in the external commercialization of knowledge. *Journal of Product Innovation Management* 26: 371–387.
Maidique, M.A. 1980. Entrepreneurs, champions, and technological innovation. *Sloan Management Review* 21: 59–76.
Markham, S.K. 1998. A longitudinal examination of how champions influence others to support their projects. *Journal of Product Innovation Management* 15: 490–504.
Markham, S.K. 2002. Moving technologies from lab to market. *Research-Technology Management* 45: 31–42.
Markham, S.K., and L. Aiman-Smith. 2001. Product champions: Truths, myths and management. *Research-Technology Management* 44: 44–50.
Markham, S.K., and A. Griffin. 1998. The breakfast of champions: Association between champion and product development environment, practices and performance. *Journal of Product Innovation Management* 15: 436–454.
Markham, S.K., S.G. Green, and R. Basu. 1991. Champions and antagonists: Relationships with R&D project characteristics and management. *Journal of Engineering and Technology Management* 8: 217–242.
Mirza, M., A.Pinsonneault, and R. Tamblyn. 2008. Progress and peril in the championing process. Paper presented at Proceedings of the Fourteenth Americas Conference on Information Systems in Toronto, 14–17 August.
Morita, A., E. Reingold, and M. Shimomura. 1986. *Made in Japan: Akio Morita and Sony.* New York: E. P. Dutton Publishers.
Pinchot III, G. 1985. *Intrapreneuring*. New York: Harper & Row.
Roure, L. 2001. Product champion characteristics in France and Germany. *Human Relations* 54: 663–682.
Royer, I. 2003. Why bad projects are so hard to kill. *Harvard Business Review* 81: 49–56.
Schon, D.A. 1963. Champions for radical new inventions. *Harvard Business Review* 41: 77–86.
Shane, S.A. 1994. Are champions different from non-champions? *Journal of Business Venturing* 9: 397–421.

Product Innovation

Sara Beckman
University of California Berkeley, Haas School of Business, Berkeley, CA, USA

Abstract

Product innovation is critical to firm performance as the primary source of organic growth, and a primary means of satisfying customers and users, particularly in an 'experience-based economy' where creating and delivering outstanding customer experiences is required. Product, or experience, ► innovation engages the entire firm from the senior managers to the front-line employees who deliver the experience, and requires a great deal of discipline to execute the many elements of the product innovation process. Generally, it is best accomplished by a small, empowered team of individuals who can act nimbly, while still representing the broader interests of the organization.

Definition Product innovation allows organizations to develop and deliver new or enhanced solutions for their customers, which in turn fuels organic growth. It includes basic capabilities in understanding customer and user needs, technology and solution development, and execution of new ideas throughout the organization.

Before a discussion of how companies approach product innovation, it is important to provide context for the use of the word 'product'. Companies no longer focus their product innovation efforts solely on physical solutions (such as the iPhone), but on a broader range of elements of the customer experience (such as the Apple store, the iTunes ecosystem). While some organizations still focus on extracting relatively undifferentiated commodities for which they charge market-based prices, many organizations have moved through to making goods, then delivering services and finally to staging customer experiences which are highly differentiated and command premium

pricing (Pine II and Gilmore 1998). In this definition of 'product innovation' we take the word 'product' to refer to the full range of commodity, goods, services and experience-related ▶ innovation activities.

Businesses derive a significant percentage of their sales and profits from the introduction of new products or solutions. Estimates range from a third (Griffin 1997) to a half (Cooper 2001) of current sales derived from the introduction of new products. Successful products are shown to drive a significant portion of profits as well.

There is, however, a wide range of types of new products that a company might deliver (Olson 1995):

- New-to-the-world products: products that are both new to the company developing them and to the marketplace using them
- Line extensions: products that are new to the marketplace, but not to the company
- Me-too products: products that are new to the company but not to the marketplace
- Product modifications: existing products that have been modified or enhanced, and thus are neither new to the company or to the marketplace.

Product innovation is a complex activity that engages all the functions in a firm. Senior management starts the process by setting strategy for the organization, which drives choices as to which products will be developed and when, and may also include upfront technology development or sourcing decisions. At the core of product innovation work is a ▶ new product development process executed by a new product development team. Finally, there is increasing investment in information technology to support product innovation. The key constructs that drive successful product innovation are: a high-quality new product process; a clear, well-communicated new product strategy for the company; adequate resources for new products; senior management commitment to new products; an entrepreneurial climate for product innovation; senior management accountability; strategic focus and synergy; high-quality development teams; and cross-functional teams (Cooper and Klenschmidt 1995).

Product Innovation Planning

A firm's ▶ business strategy drives its choices as to what types of innovation projects it will take on. The amount to be invested in product innovation is often assessed as a percentage of revenue, and is set on the basis of desired competitive positioning and stage of the industry lifecycle. That investment is then allocated among the variety of possible product innovation projects that might be undertaken by the organization in a portfolio planning process that often results in a product roadmap detailing which products will be brought to market when and what the interdependencies among them are. Senior management plays a critical role in guiding product innovation by providing resources, showing commitment and establishing an innovation culture.

New Product Development Process

The new product development process guides the work of the company in creating and commercializing solutions. Although processes vary widely in the details of their implementation, there are standard views of the process (Roschuni 2013). At a high level of abstraction, the product innovation process is seen as one of analysis, synthesis and evaluation (Asimow 1962) that closely parallels the general problem-solving process (Simon 1969). This view, integrated with experiential learning theory (Kolb 1984), yields a design-based view of the process with four stages: observation (to better understand the context for which the innovation is being created), framing (to abstract from the observation work a different perspective of the problem to be solved), imperatives (which translate the frames into the required outputs of the innovation effort) and solutions (the artefacts that embody the imperatives) (Barry and Beckman 2007). More concrete descriptions of the process detail stages such as (Ulrich and Eppinger 2008):

- Planning
- Concept development
- System-level design
- Detail design
- Testing and refinement
- Production ramp-up.

These steps are often captured in a 'stage-gate' model that designates both the activities (in the stages) that are to be accomplished as well as the milestones that must be met (at each gate) for the project to move forward (Cooper 2001). More fluid approaches to the process were originally described in a spiral model of software development (Boehm 1988) and have evolved to become known as agile development methodologies. In line with increased emphasis on customer experience design, product innovation methods are becoming more soundly based in deep empathy for customers and on embedding customer understanding throughout the innovation process (Brown and Katz 2009).

Product Innovation Team

The product innovation team is at the heart of the execution of product innovation. Companies construct product innovation teams in a variety of ways including functionally structured teams, project organizations, lightweight project matrix organizations and heavyweight project matrix organizations (Hayes et al. 1988). Best practice, however, employs a 'core team' of five to seven individuals, each of whom represents one of the functions that is critical to the innovation effort (e.g., product marketing, operations, ▶ research and development). This core team is an accountable, cross-functional set of individuals that is specifically selected to make decisions on behalf of the product innovation effort for their particular functional area. In doing so, they act as extensions of the executives in those functions. They work together to ensure a 'general management' outcome for the customer and the company, making informed, cross-functional trade-offs and managing programme risk effectively throughout (Creech 2013).

Conclusion

The design of superior customer experiences, and ultimately of customer transformations, will be the focus of innovation efforts by companies and other organizations for some time to come. While some of the elements of the 'product' innovation process will certainly evolve to better conduct experience or transformation design, critical activities will still have to be performed: product innovation will always require senior management support, providing both direction and resources; it will require increasingly customer-focused processes, grounded in deep empathy for the customer's situation and generatively creating alternative solutions; and it will depend on cross-disciplinary teams of people working together virtually as well as physically. Rapid evolution of technology will provide an increasingly wide range of ways to innovate, as well as ways to create ever-changing customer experiences.

See Also

- ▶ Business Strategy
- ▶ Business-to-Consumer (B2C) Marketing
- ▶ Corporate Strategy
- ▶ Innovation
- ▶ New Product Development
- ▶ Organizational Culture
- ▶ Portfolio Planning: A Valuable Strategic Tool
- ▶ Research and Development (R&D) Organization

References

Asimow, M. 1962. *Introduction to design*. Englewood Cliffs: Prentice Hall.

Barry, M., and S.L. Beckman. 2007. Innovation as a learning process: Embedding design thinking. *California Management Review* 50: 25–56.

Boehm, B.W. 1988. A spiral model of software development and enhancement. *Computer* 21: 61–72.

Brown, T., and B. Katz. 2009. *Change by design: How design thinking transforms organizations and inspires innovation*. New York: Harper Business.

Cooper, R.G. 2001. *Winning at new products: Accelerating the process from idea to launch*, 3rd ed. New York: Perseus Publishing.
Cooper, R.G., and E.J. Klenschmidt. 1995. Benchmarking the firm's critical success factors in new product development. *Journal of Product Innovation Management* 12: 374–391.
Creech, J. C. 2013. Interview in person in Berkeley, CA, conducted by S. L. Beckman, on leading teams and high performance organizations, 31 January.
Griffin, A. 1997. *Drivers of NPD success: The 1997 PDMA report*. Chicago: Product Development & Management Association.
Hayes, R.H., S.C. Wheelwright, and K.B. Clark. 1988. *Dynamic manufacturing: Creating the learning organization*. New York: Free Press.
Kolb, D.A. 1984. *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs: Prentice Hall.
Olson, E.W. 1995. Organizing for effective new product development: The moderating role of product innovativeness. *Journal of Marketing* 59: 48–62.
Pine II, B.J., and J.H. Gilmore. 1998. Welcome to the experience economy. *Harvard Business Review* 74: 97–105.
Roschuni, C. N. 2013. Communicating design research effectively. Ph.D. thesis, University of California, Berkeley.
Simon, H.A. 1969. *The sciences of the artificial*. Cambridge, MA: MIT Press.
Ulrich, K.T., and S.D. Eppinger. 2008. *Product design and development*. New York: McGraw-Hill.

Product Market Strategy

Anoop R. Menon and Dennis A. Yao
Harvard University, Business School, Boston, MA, USA

Abstract

Product market strategy is the collection of choices, actions and activities of a firm that determines how it positions itself in its product markets, and allows it to achieve and maintain a competitive advantage. This entry examines product market strategy from the perspective of positioning, using the value-based strategy framework.

Definition Product market strategy is the collection of choices, actions and activities of a firm that determines how it positions itself in its product markets, and allows it to achieve and maintain a competitive advantage.

A product market strategy addresses the following questions, among others: What product(s) do we produce and sell? What customers and segments do we aim for? Which geographies do we aim for? When do we enter the market? How do we produce and sell the product? What activities associated with the product do we undertake? (Zott and Amit 2008). These questions have to be answered while taking into account the product market strategies of the other players such as rivals, complementors and regulators.

Framework of Willingness to Pay and Cost

A useful representation for thinking about product market strategies is the value-based ▶ business strategy framework (Brandenburger and Stuart 1996). This representation (shown in Fig. 1) succinctly links the value chain to value creation and value capture. As an example of how this representation of value works, consider a consumer product. The activities of the participants in the value chain combine to create value, which is represented by the amount the end-consumer is willing to pay for this product ('willingness to pay' or WTP) minus the cost of creating this value. This overall created value is captured in part (or in whole) by the various participants in the value chain. Which participant captures what share of the overall value depends on the level of competition at each level of the value chain, and sometimes by the outcome of bargaining. For example, if the entire supply chain is extremely competitive the consumer will receive overall value while the firms in the value chain will receive only normal profits. One of the most interesting features of the Brandenburger and Stuart added-value representation is that both value creation and value capture can be represented by it.

From this added-value perspective, a strategy for a particular market can be seen as a set of choices which determine how a firm creates and

Product Market Strategy, Fig. 1 The added-value framework (Brandenburger and Stuart 1996)

captures value. At a micro level, each activity that a firm engages in can be analysed in terms of its effect on willingness to pay and cost. An increase in willingness to pay usually entails some increase in costs. Similarly, decreases in cost will sometimes result in a decrease in willingness to pay. At a macro level, a firm's product market strategy can, for example, be described in terms of its emphasis on increasing a consumer's willingness to pay for a product or on reducing the cost of producing this product (or service).

Positioning

The choice of how to create and capture value can be thought of as a positioning choice. Along these lines Porter (1980) identifies three generic strategies – differentiation, cost leadership and focus – for competing in an industry. Porter's approach is useful for thinking about product market strategies. A firm that chooses a differentiation strategy orients its organizational activities towards increasing its customers' willingness to pay. Such strategies would emphasize, for example, increasing the function of a product or service, branding, post-sales service or ease of ordering. In contrast, a firm that chooses a cost leadership strategy will focus on reducing the costs of producing a given product or service. Components of a cost leadership strategy would include reducing production costs by taking advantage of economies of scale or learning. Both differentiation and cost leadership are strategies applied across the entire industry. Focus strategies, on the other hand, target a particular product niche to tailor the product more precisely to the customer preferences, thereby increasing the willingness to pay and/or reducing costs.

The business success of such strategies depends, of course, not only on value creation but on value capture. Under a differentiation strategy, value capture is enhanced because the products produced by differentiating firms are not perfect substitutes, but, from the firm's perspective, are ideally seen by some customers as offering a somewhat unique value. The customer value created by a cost leadership strategy is typically embodied in a relatively lower price for a given willingness to pay. Firms can capture value by using such a strategy partly because some customers will prefer the net value offered through a lower price over the net value offered by other firms in the market. Value capture through a cost leadership strategy is possible particularly when there are production economies that effectively limit the number of competing low-cost suppliers (i.e., the possible economies exist in quantities that are large proportions of the total market) or where a firm has dynamic capabilities that allow it to continually outperform its competitors on cost reductions.

If alignment of the activities of a firm increase the relative efficacy of a strategy, then, as Porter (1996) argues, a pure play position would be superior to a mixture of a WTP and cost leadership strategy. For example, it may be difficult to generate a culture of cost reduction in one part of the organization while maintaining a culture of customer service in another part. Typically, the activity systems associated with increasing willingness to pay are different from the activity systems useful for reducing costs. Hence, Porter argues that firms whose strategies are 'stuck in the middle' will not compete well with those following one of these two generic strategies.

In arguably unusual circumstances there may be a link in the activity systems between quality

and cost that allow a firm to achieve a dual advantage even while serving the entire industry. Here, for example, a high level of production might allow a firm to lower costs through economies of scale, while also giving the firm some advantage regarding quality or product innovation that raises a consumer's willingness to pay. A large number of sales might, for example, allow the manufacturer to gain more field experience, which translates into increased reliability or quality as well as greater knowledge of characteristics desired by consumers. Unfortunately, because some links between quality and cost reductions can always be found even when the links are not first order, dual advantage will sometimes be used to justify strategies that in actuality lack strategic focus.

Activity systems that can simultaneously increase willingness to pay and decrease cost, while rare for firms serving the entire market, are more common among firms pursuing focused strategies. The level of 'focus' of a strategy becomes another dimension of positioning but one which is not captured in the one-dimensional added-value representation which is oriented around a particular customer segment.

Another dimension to be considered in the positioning decision is the question of timing – when to enter the market (Lieberman and Montgomery 1988). Being first to the market with a novel product confers many competitive benefits, such as the ability to build a reputation or lock-in key resources, but at same time exposes the firm to significant risks, potentially allowing a 'fast-follower' to come in with a much more fine-tuned product that captures the market. The typology of generic strategic positions (prospectors, defenders, analysers, reactors) proposed by Miles and Snow (1978) incorporates this dimension as well as the willingness and ability of firms to explore and develop new markets. Positioning can also include choices about which portfolio of markets to be in.

Rivals and Complementors

Porter's generic strategies can be seen as the two primary paths to achieve ► competitive advantage. Because competitive advantage depends on a firm's relative performance versus its close competitors, the success of a product market strategy depends not only on a firm's own choices but also on the choices of firms outside the firm's direct value chain. The most obvious category of such firms is rivals. Conventionally, one thinks of rivals as affecting value capture through competition using their own product market strategy. But the added-value representation makes it clear that value capture need not be limited to direct market competition, but also extends to activities through which firms attack the competitive advantage of their rivals (Brandenburger and Stuart 1996). A cost leadership strategy aims to increase the cost advantage of the firm over its rivals. Conventionally, such an advantage is pursued through internal cost reductions. But cost advantage can also be pursued through activities designed to increase the costs of rivals (Salop and Scheffman 1987). Similarly, a firm can attack the consumers' willingness to pay for the rival's product. Examples of such actions include negative advertising and the creation of switching costs (Brandenburger and Stuart 1996).

Another category of interrelated strategies involves complementors, which are firms whose actions increase the value of the focal firm's product or service (see, e.g., Brandenburger and Nalebuff 1996). The actions of complementor firms may have a significant effect on the overall willingness to pay of a consumer who receives value when two or more individual products are used together. A good example is how software applications increase the value of computer hardware and vice versa. In this extreme case, one product is useless without the other. Thus, differentiation strategies will sometimes depend on the strategies of other firms. In some industries an important part of product market strategy will be how one's activities are designed to affect the strategies of others.

Non-market Considerations

Finally, an alternative strategy used by many firms, especially in developing markets, is a

relationship and influence strategy with respect to government and non-governmental organizations. Here, government relationships are briefly addressed.

Markets exist under government-mandated and enforced rules. These rules can be neutral across firms in a market, but may also favour firms with particular characteristics (e.g., domestic or foreign), or even particular firms within a group of firms with similar characteristics. Particularly in the latter circumstance, the success of a firm's product market strategy may depend as much on its relationship with key governmental decision makers as on its basic product market strategy. That is, a key competitive advantage of some firms will be their capability to navigate government processes, or perhaps their relationship with key government actors who can use the political process to give that firm a unique advantage in the relevant market. This advantage may originate outside the focal market, but government can directly alter the relative willingness to pay or cost among firms within the product market; hence, thinking about product market position without taking into account these non-market strategies would cause firms to miss an important 'positioning' that allows them to capture value in a product market.

Dynamics

Our discussion of product market strategies focuses on ideal positions. Other interesting questions include whether a firm should change its position and, if it chooses to do so, the path along which change should take place. A firm might have, for example, a collection of resources or capabilities that supports both a differentiation and a cost leadership position and which may partially explain why the firm has survived in the market thus far. Does the firm have a better chance of survival by quickly moving to a purer play strategy? Or perhaps the firm might be able to identify and then follow a ▶ focus strategy in which its resources are better aligned. These important questions are beyond the scope of this short article.

A Caveat

The discussion so far has assumed that firms under consideration are profit maximizers. This assumption allowed us to boil down the complex interactions between the various players and their implications for the focal firm onto a single dimension – the added-value representation. But we acknowledge that profit maximization need not be the only objective of firms. In such cases, a broader definition of value should be used, but similar considerations of the interaction between the value creation and capture by the different actors can still be applied profitably. Miles and Snow (1978) argued for four types of firms: prospectors, who proactively identify and develop novel markets and products, focusing on multiple flexible technologies, product-based management and decentralized control; defenders, who aggressively protect their current product market, investing in technological efficiency, functional management structure and centralized control; analysers, who lie between the two prior types, carefully exploring new options while maintaining its core skills, products and customers; and the reactors, which are firms that are the residual, who 'missed the bus'.

See Also

- ▶ Business Strategy
- ▶ Competitive Advantage
- ▶ Competitive Strategy
- ▶ Focus Strategy
- ▶ Generic Strategy

References

Brandenburger, A.M., and B.J. Nalebuff. 1996. *Co-opetition*. New York: Doubleday.

Brandenburger, A.M., and H. Stuart. 1996. Value-based business strategy. *Journal of Economics & Management Strategy* 5: 5–25.

Lieberman, M.B., and D.B. Montgomery. 1988. First-mover advantages. *Strategic Management Journal* 9: 41–58 (summer special issue).

Miles, R.E., and C.C. Snow. 1978. *Organizational strategy, structure and process*. New York: McGraw-Hill.

Porter, M.E. 1980. *Competitive strategy: Techniques for analyzing industries and competitors*. New York: Free Press.
Porter, M.E. 1996. What is strategy? *Harvard Business Review* 74: 61–78.
Salop, S.C., and D.T. Scheffman. 1987. Cost-raising strategies. *Journal of Industrial Economics* 36: 19–34.
Zott, C., and R. Amit. 2008. The fit between product market strategy and business model: Implications for firm performance. *Strategic Management Journal* 29: 1–26.

Profit

Giovanni Valentini
Bocconi University, Milan, Italy

Abstract

The field of strategic management deals with fundamental questions about how company profits emerge and persist. After a short definition of the term, we will therefore discuss how to empirically measure firms' profit, discuss the key determinants of the observed variance in firms' profit, and analyse whether these differences in profits are persistent.

Definition Profit is defined as the difference between the revenues a firm receives and (all) the cost it incurs.

The field of strategic management deals with the fundamental question of how company profits emerge and persist. Profit is defined as the difference between the revenues a firm receives and the cost it incurs.

In the scientific literature, a distinction is often made between two types of profit: accounting profit and economic profit. Accounting profit is obtained by subtracting accounting costs from revenues, whereas economic profit considers economic costs. Accounting costs appear in accounting statements (i.e., income statements, balance sheet) and are based on historical costs, whereas economic costs also reflect the opportunity costs of any activity or resource (including capital); that is, the value of the best alternative foregone when carrying out a given activity or using a resource.

In formal economic theories of firm behaviour, aimed at analysing why firms take their decisions and what distinguishes good decisions from bad ones when choosing among competing alternatives, economic profits are usually emphasized. Conversely, when assessing firms' past performance, comparing performance across firms of different industries, or to evaluate the financial viability of a firm, the informed used of accounting profit is actually helpful (Besanko et al. 2007).

More generally, Grant (2010) suggests that economic profit might have two main advantages over accounting profit as a performance measure. First, it sets a more challenging performance discipline for managers. At many capital-intensive companies seemingly healthy profits disappear once the cost of the capital is taken into account. Second, using economic profit improves the allocation of capital between the different businesses of the firm by taking into account the real costs of more capital-intensive businesses.

In both cases, a central assumption of traditional economic models is that firms (are expected to) maximize expected profits. Although part of the literature has relaxed the assumption of *maximization* (see, e.g., Cyert and March (1992)), another question is whether profit is or should be the only objective of organizations and if companies are expected to have responsibilities other than the immediate interests of their owners. Here we abstract from these themes and consider that if research in strategic management is aimed at understanding and explaining how company profits may emerge and persist over time, three questions become relevant:

- How can firms' ability to generate profit be measured?
- What determines the observed variance in firms' profits?
- Are differences in company profits persistent?

In strategy literature, firms' (ability to generate) profit has been measured in a number of ways. Following Schmalensee's (1989) classification, the many measures that have been used fall into four main classes.

A first stream suggested that firms' ability to sustain profits depends on their ability to hold price above long-run average costs, where costs should include also the competitive return on the capital employed. A possible measure along these lines is the ratio of *excess profit* to sales revenue. Second, several studies have employed accounting rates of return on assets or equity. Increases in leverage make the residual return to equity more variable, and in competitive capital markets investors should generally be paid higher average returns to compensate for the higher risk. Rate of returns on assets, on the other hand, mainly reflect operating results, and not capital structure decisions. Third, the so-called price–cost margin (to simplify: revenue-variable cost/revenue) has been used; this measure can generally be computed for more narrowly defined industries than accounting rate of returns. Under competitive conditions, the price–cost margin should equal the required rental on assets employed per dollar of sales. Finally, measures that employ the market value of a firm's securities are often attractive, because, under the assumption of capital market efficiency, they should reflect all available information about the firm's future profitability. Within this class, Tobin's q is a frequently used measure, and it is defined as the market value of a firm to the replacement cost of its tangible assets.

Since specific cases and data availability vary, and since researchers cannot calculate all these measures, the key issue becomes to understand whether these families are equally valuable to measure firms' profitability, which in turn highlights the need for understanding the correlation between these variables. Schmalensee (1989) contends that correlations among accounting rates of return are high, and that studies investigating industry structure on firms' performance present results that are usually not sensitive to the specific measures employed. Correlations of accounting rates of return with the price–cost margin and with measures based on market values are lower, and regression results may actually depend on which type of measure is used.

The second issue is what determines the documented variance in firm profits. To answer this question, after the initial studies based on the ▶ structure–conduct–performance paradigm, researchers have attempted to break down the variance in firms' profitability into components associated with year, industry, corporate-parent and business-specific effects (e.g., Rumelt 1991; McGahan and Porter 1997).

McGahan and Porter (2002) summarize the main findings of this stream of research and provide additional evidence. In particular, they describe how business-specific effects are more important than other effects. Yet the relative importance of year, industry, corporate-parent and business-specific effects differ across different sectors of the economy. Moreover, these individual effects are not necessarily independent. For example, several studies show that industry and corporate-parent effects are simultaneously determined, in that the choice of industry by diversifying corporate parents is related to industry performance. In broad terms, this literature has confirmed the limitations of the structure–conduct–performance models: there is evidence of feedback and co-evolution between the industry, corporate-parent and business-specific effects.

But are the observed differences in profits persistent? In his seminal study, Mueller (1986) shows that firms with abnormally high levels of profitability tend to decrease in profitability over time. By contrast, firms with abnormally low levels of profitability tend to experience an increase in profitability over time. However, the profit rates of the abnormally profitable firms and abnormally unprofitable firms do not seem to converge to a common mean: firms that start out with high profits converge, in the long run, to rates of profitability that are higher than the rates of profitability of firms that start with low profits (Mueller 1986). These results imply that market forces are a threat to superior profits, as suggested by standard economic theory, but only up to a point. Other forces appear to protect profitable firms and allow them to sustain their ▶ competitive advantage.

McGahan and Porter (2003) examine the emergence and sustainability of abnormal profits among business that were part of US public corporations between 1981 and 1994, and that reported financial results for at least 6 years.

Their analysis reveals some additional broad regularities (McGahan and Porter 2003: 101):

- Industry effects are more important than business-specific and corporate-parent effects in the sustainability of high performance. Business-specific effects are more important than industry and corporate-parent effects in the emergence and sustainability of low performance, as well as in the emergence of high performance.
- Industry and corporate-parent effects are more important on average to high performance than to low performance. Business-specific effects are more important on average to low performance than to high performance.
- On average, high performance is preceded by high performance, whereas low performance is preceded by average performance.
- High and low performance erode at about the same rate.

However, the results of Geroski and Jacquemin (1988) suggest the results of these analyses might be contingent on some institutional factors. More specifically, they study the evolution of profits of 134 firms from three different European countries. They show that, in contrast to the results of France and West Germany, the UK stands out as a country in which profits above and below the norm persist enduringly in a relatively large number of cases. What is more, although the authors find there are associations between various structural traits of firms, industry characteristics and the persistence of success, it remains difficult to find factors which are systematically associated with either the persistence or the predictability of profits. In their study, countrywide factors have turned out to be more discriminating than firm- or industry-specific ones.

See Also

- ▶ Agency Problems
- ▶ Bounded Rationality
- ▶ Competitive Advantage
- ▶ Performance Measures
- ▶ Structure–Conduct–Performance
- ▶ Variance Decomposition

References

Besanko, D., D. Dranove, M. Shanley, and S. Schaefer. 2007. *Economics of strategy*, 4th ed. New York: Wiley.

Cyert, R.M., and J.G. March. 1992. *A behavioral theory of the firm*, 2nd ed. Cambridge, MA: Blackwell.

Geroski, P.A., and A. Jacquemin. 1988. The persistence of profits: A European comparison. *The Economic Journal* 98: 375–389.

Grant, R.M. 2010. *Contemporary strategy analysis*, 7th ed. Chichester: Wiley.

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

McGahan, A.M., and M.E. Porter. 2002. What do we know about variance in accounting profitability? *Management Science* 48: 834–851.

McGahan, A.M., and M.E. Porter. 2003. The emergence and sustainability of abnormal profits. *Strategic Organization* 1: 79–108.

Mueller, D.C. 1986. *Profits in the long run*. Cambridge: Cambridge University Press.

Rumelt, R. 1991. How much does industry matter? *Strategic Management Journal* 12: 167–185.

Schmalensee, R. 1989. Inter-industry studies of structure and performance. In *Handbook of industrial organization*, ed. R. Schmalensee and R.D. Willig. Amsterdam: Elsevier.

Profit Centres

Gordon Walker
Southern Methodist University, The Edwin L Cox School of Business, Dallas, TX, USA

Abstract

A profit centre is an operating unit in a diversified corporation and is so named because it makes a profit or loss on sales in an external product market (Anthony and Govindarajan 2006). The apposite comparison is with a cost centre, which incurs expenses but no revenues to cover them. A profit centre acts and is treated like a stand-alone business, and its management behaves and is evaluated in this context. Profit centres are considered 'decentralized' within the firm, whereas, since their managers have less discretion, cost centres are thought of as 'centralized'.

Definition A profit centre is an operating unit in a diversified corporation and is so named because it makes a profit or loss on sales in an external product market.

Profit centres are almost the same as two related types of organizational unit: product divisions and strategic business units (SBUs). A product division is formed to manage a single product or product line and usually contains the functions required, such as product engineering, manufacturing, marketing and sales. It would be very unusual if a product division were not also a profit centre, but profit centres need not be product divisions. An SBU is a special form of profit centre developed by General Electric (GE) in the 1960s to improve business planning and accountability. GE overlaid SBUs on to an existing set of product divisions, both dividing them up and aggregating them, to improve line management's control over the resources it needed to compete effectively. Better control allowed SBU managers to produce more realistic plans, but this realism reduced their wiggle room in performance reviews, a result corporate management wanted and had foreseen. Many firms currently use the term SBU without reference to GE's defining criteria, in which case the SBU is just another term for a profit centre.

Thus there are two major operating reasons for creating profit centres: better decision-making by business managers and improved management accountability. But there is also a third reason: top management can analyse its portfolio of businesses more effectively when they are profit rather than cost centres. Since the company is an aggregate of the businesses, an analysis of business unit contributions to the corporation's financial return is simplified when each unit reports profits (or losses), not budget variances.

Profit Centres and Business Diversification

Profit centres are found in corporations with more than one line of business. Chandler's (1962) history of the rise of the multi-divisional firm details the shift from functional to product division structures in US corporations. His account is basically a description of how profit centres emerged as a solution to the challenges of managing multiple businesses. Other authors have described the global diffusion of this form of corporate structure (Stopford and Wells 1972; Egelhoff 1988).

In Chandler's account, decentralized product divisions are more efficient because they allow coordinated decision-making by grouping together the activities relevant for each business (e.g., engineering, manufacturing, marketing). Also, the corporate office can hold the managers of the product divisions responsible for the performance of the business since they control the resources needed and report a profit and loss statement, as in a free-standing company. In this way, diversified firms perform better if their businesses are organized as profit centres (product divisions) rather than cost centres (functional divisions).

Some firms turn cost centres into profit centres to create de facto diversification. Common examples are service operations, such as human resources, IT and logistics. In these cases, the unit's level of expertise should be high enough to compete in external product markets, even if it only charges a fee for its services to internal customers, who are free to buy from outside vendors.

Profit Centres and Resource Allocation: The Internal Capital Market

In a diversified organization, profit centres receive funds from the firm through an internal capital market (Gertner et al. 1994). Internal capital markets may be superior to external sources of capital, but only under specific conditions (Liebeskind 2000). First, the profit centre should be 'capital constrained', in the sense that some promising projects would not be funded externally. This is a necessary condition for the internal capital market to substitute for external capital. Second, the parent may have trade secrets or proprietary information in the unit that it wants to protect from external investors.

Many observers believe, however, that internal politics constrain corporations from making

effective investment decisions for their profit centres (Milgrom and Roberts 1988; Wulf 2002). Moreover, even if politics were absent, cognitive biases may distort the firm's resource allocation decisions. For example, one frequently observed tendency is to provide capital to profit centres whether their cash flows are growing or not. 'Corporate socialism' of this kind may be a result of politics or decision rules favouring equity among the profit centres or both (Billet and Mauer 2003; Vieregger 2013).

Profit Centres and Inter-Unit Transfers

Profit centres frequently transfer goods and services to each other through an internal *product* market. These transfers are like inputs from external suppliers and are commonly benchmarked against them. Since profit centres are evaluated in terms of profitability, there may be conflicts between serving inside and outside customers and between buying inside or outside the firm. From the buyer's viewpoint, sourcing from an inside division is justified if it increases the buyer's performance more than an outside vendor would. But an internal supplier may not be cooperative. Rather, it may subsidize outside sales by shaving its investment levels for sales internally. In this case, the parent firm must resolve the conflict between the competing interests of the internal buyer and supplier.

The complexity of this trade-off is illustrated by the problem of pricing inter-profit centre transfers. Prices between units are logically called transfer prices and may be mandated or negotiated. Robert Eccles (1985) has identified four types: mandated market price, mandated full cost, exchange autonomy and dual pricing.

Mandated Market Price

Here corporate policy ties the in-house buyer to the in-house supplier, and the supplier, not the buyer, is clearly dominant in the relationship. The supplier has the same price for both internal and external sales. The buyer cannot buy the good or service outside the parent corporation (it is mandated) although the supplier may sell to external customers. Since there is no relative cost advantage for the buyer from input prices, it is hard to justify this type of transfer price if the buyer's market position is based on lower costs.

Mandated Full Cost

The buyer now dominates. The supplier sells to the buyer at full cost and thus has to balance internal sales at cost and external sales at market price. This tension may well degrade the degree of cooperation the supplier offers the buyer over time. Unlike market-based transfer prices, full-cost prices can be used to support a buyer strategy of cost leadership.

Dual Pricing

Dual pricing involves two transfer prices: full cost to the buyer and market price to the supplier. Although this policy seems beneficial to both sides, it is inherently unstable because it undermines the integrity of the corporation's management control systems. In some situations the buyer and supplier can both report profits while the corporation is losing money. The method is therefore useful as a temporary fix when a few critical transactions require special attention.

Exchange Autonomy

This scheme applies when transactions are not mandated between the buyer and supplier. In this case, transactions are infrequent, and the price could be based on cost or mark to the market depending on the circumstances. This type of internal supply relationship is therefore used to handle ad hoc transactions.

These four types of transfer price show that there is no overarching solution to the internal pricing problem. The strategies of both the buying and supplying profit centres must be considered. Also, it is important to consider external market prices as benchmarks for transactions within the firm.

Profit Centres and Performance Evaluation

There are three major objectives in choosing a performance metric for profit-centre reporting:

(1) to enable effective decision-making by the unit manager; (2) to allow effective management performance appraisal; (3) to lead to an improvement in *corporate* economic performance. The two best-known and used performance metrics are return on investment (ROI) and residual income. ROI is measured as the division's net income over total capital invested in operations. Residual income is computed as net income generated by a project minus a capital charge, typically the weighted average cost of capital (as a percentage) multiplied by the capital invested (Solomons 1985).

On which ▶ performance measures should the manager base his decision? Since, in some cases, using ROI alone to measure division performance can reduce shareholder value, residual income is the preferred metric (see Balachandran 2006). Making decisions that contribute to shareholder value is necessary since shareholder returns are, in theory, the best long-term predictors of the market value of the firm.

See Also

- ▶ Market Price
- ▶ Multinational Corporations
- ▶ Performance Measures

References

Anthony, R.N., and V. Govindarajan. 2006. *Management control systems*, 12th ed. Burr Ridge: Irwin/McGraw-Hill.

Balachandran, S. 2006. How does residual income affect performance? The role of prior performance measures. *Management Science* 52: 383–394.

Billet, M., and D. Mauer. 2003. Cross-subsidies, external financing constraints, and the contribution of the internal capital market to firm value. *Review of Financial Studies* 16: 1167–1170.

Chandler, A.D. 1962. *Strategy and structure*. Cambridge, MA: The MIT Press.

Eccles, R. 1985. *The transfer pricing problem*. Lexington: Lexington Books.

Egelhoff, W.G. 1988. Strategy and structure in multinational corporations: A revision of the Stopford and Wells model. *Strategic Management Journal* 9: 1–14.

Gertner, R., D.S. Scharfstein, and J.C. Stein. 1994. Internal versus external capital markets. *Quarterly Journal of Economics* 109: 1211–1130.

Liebeskind, J. 2000. Internal capital markets: Benefits, costs and organizational arrangements. *Organization Science* 11: 58–76.

Milgrom, P., and J. Roberts. 1988. An economic approach to influence activities in organizations. *American Journal of Sociology* 94: 154–179.

Solomons, D. 1985. *Divisional performance: Measurement and control*, 2nd ed. New York: Markus Weiner.

Stopford, J.M., and L.T. Wells Jr. 1972. *Managing the multinational enterprise*. New York: Basic Books.

Vieregger, C. 2013. Do firms really allocate capital so inefficiently? Working paper, Washington University.

Wulf, J. 2002. Influence and inefficiency in the internal capital market: theory and evidence. Working paper, Department of Management, Wharton School, University of Pennsylvania.

Profiting from Innovation

David J. Teece
Berkeley Research Group, LLC, Emeryville, CA, USA
Haas School of Business, University of California, Berkeley, Berkeley, CA, USA

Abstract

Profiting from ▶ innovation is a theory that accounts for marketplace outcomes between innovators and follow-on rivals. Almost all innovations require complementary investments. The weaker the ▶ appropriability regime applicable to an innovation, and the weaker the market position of the innovator with respect to providers of complements, the harder it will be for the innovator to build a long-term advantage without pursuing corrective measures such as vertical integration. The timing of commercialization is also important because the earlier in the lifecycle of the industry an entry is made, the more financial resources will be required to survive.

Definition Profiting from innovation is a theory that addresses the issue of why pioneers in markets for innovative goods are often overtaken by later arrivals. A key element is the recognition that innovations cannot stand on their own, and that the position of the innovator with respect to

intellectual property, the issue of imitation and the ownership of certain complements can limit the innovator's profit potential, even in the presence of a good strategy.

Profiting from ▶ innovation is a theory that addresses the issue of why pioneers in markets for innovative goods are often overtaken by later arrivals. It is both a normative theory of strategy and a predictive theory of how the benefits from a focal innovation are likely to be distributed between the innovator, customers, imitators, suppliers and the owners of ▶ complementary asset (Teece 1986, 2006).

Many Innovators Do Not Derive Sustainable Advantage

Innovators and the firms that are the first to commercialize a new product or process in the market do not always profit the most from their innovation. Sometimes a fast second entrant or even a slow third will capture most of the market.

For example, the technology behind the computerized axial tomography (CAT) scanner, now a standard medical diagnostic tool, was developed in the late 1960s by a senior scientist/engineer at EMI Ltd, a diversified UK-headquartered entertainment and electronics conglomerate. Although EMI brought the technology to market relatively quickly, introducing a commercial model in the United States in 1973, 8 years later it had dropped out of the scanner business, leaving the market to later entrants (Teece 1986: 286).

The EMI story is far from unique. The earliest vendors of microcomputers for home use (R2E, CTC, MITS, Commodore) are all but forgotten today. Xerox (in its PARC laboratory) and Apple invented the graphical user interface, but Microsoft Windows dominates the PC market with its follow-on version. Apple's iPod was not the first portable digital music player, but it has a commanding position in the category today. Merck was a pioneer in cholesterol-lowering drugs (Zocor), but Pfizer, a late entrant, secured a superior market position with Lipitor.

Yet, in other cases, a first-mover advantage seems to apply. Genentech was a pioneer in using biotechnology to discover and develop drugs, and 30 years later was the second largest biotechnology firm, right up to its acquisition by Hoffmann-La Roche in 2009. Intel co-invented the microprocessor and still has a leading market position 40 years later. Dell pioneered a new distribution system for personal computers and, despite recent challenges and many would-be imitators, remains one of the world's leading PC vendors. Toyota's much studied 'Toyota Production System' has provided the company with a source of competitive advantage for decades, contributing to its becoming the world's biggest car manufacturer in 2008.

The Profiting from Innovation Framework

A framework that endeavours to account for why some pioneers thrive and others vanish was first introduced in a 1986 *Research Policy* entry by ▶ Teece, David J. (born 1948). This spawned a body of work that has come to be known as the theory of profiting from innovation (PFI).

The essence of the argument is that almost all innovations require complementary investments. Relative features of the innovation (and the ease of imitation) and complements can account for how the profits flowing from the innovation are distributed amongst the innovators, rivals, complementors, suppliers and consumers. The theory posits that profits in a business ecosystem tend to migrate to the 'bottleneck' asset – that is, the asset that is hardest to replicate. This could be the innovator's intellectual property; more often than not, though, it is a complementary asset. It may or may not be owned or controlled by the innovator. The theory represents the current culmination of theorizing about innovation that began with Joseph Schumpeter and transitioned through work by Kenneth Arrow to the current formulation (Winter 2006).

In its original formulation, the framework integrated three concepts: ▶ appropriability, industry evolution and complementarity. Additional

concepts, such as system integration and industry structure, have subsequently been introduced to increase the framework's explanatory power.

Appropriability

Appropriability is the extent to which the innovator can capture the profits generated by the innovation. The degree of capture is influenced by characteristics of the technology and the legal environment, and by the ownership of complementary assets that are needed to bring the innovation to market. These determine the strength of the innovation's ► appropriability regime.

An appropriability regime is 'weak' when innovations are difficult to protect, as when they can be easily imitated and/or legal protection of intellectual property is ineffective. Appropriability can be 'strong' when innovations are easy to protect because knowledge about them is tacit and/or they are well protected legally. Regimes differ across fields of endeavour, not just across industries and countries.

Appropriability regimes change over time, and the regime applicable to a given innovation can be influenced by firms (Pisano and Teece 2007). For example, a firm with a strong position in downstream complementary assets might decide it is in its interest to weaken the upstream appropriability regime, as in the case of Google making its Android operating system available at no cost to gain advantage in the sale of mobile search advertising.

It is vital for firms to recognize that ► patents, which may have strategic value beyond the direct profit goals discussed here, rarely confer strong appropriability, beyond special cases such as new drugs, chemical products and rather simple mechanical inventions (Levin et al. 1987). Many patents can be 'invented around' at modest cost (Mansfield et al. 1981; Mansfield 1985). Nevertheless, a small subset of patents is often very valuable, particularly if they are pioneering patents in a commercially significant area.

However, the legal and financial requirements for upholding a patent's validity, or for proving its infringement, are high. Validity is never firmly established until a patent has been upheld in court.

In some industries, particularly where the innovation is embedded in processes, trade secrets are a viable alternative to patents, which are especially ineffective at protecting process innovation. Trade secret protection is possible in cases where a firm can put its product before consumers and the public and still keep the underlying technology secret. Many industrial processes, including semi-conductor fabrication, are of this kind.

Industry Evolution

In the early stages of an industry's development, product design is often the basis for competition (Abernathy and Utterback 1978). After considerable trial and error by rival companies, one design, or narrow class of designs, begins to dominate the market. Pioneering innovators must be prepared with considerable financial resources while the market uncertainty is being resolved.

The establishment of standards is a critical stage in the evolution of an industry. When standard-setting is a formal process, an innovating firm can solidify the demand for its technology by offering its technology as part of the official standard. When standard-setting is left to the market, each new entrant will have the opportunity to modify or imitate the pioneering innovator's product (or process) while trying to make its own design the de facto industry standard and leave the pioneer at a disadvantage.

Many of the newer growth industries that rely on the Internet or on telecommunications networks bring an important caveat to the 'latecomer advantage' view of industry evolution. Most network-based industries are characterized by mechanisms of positive feedback – including positive adoption externalities, increasing returns to scale and switching costs – that provide a built-in advantage for early entrants. Nevertheless, later entrants, such as Google in the case of search engines, can still become the category leader by offering a better technology or user experience.

Complementary Assets

Successful commercialization of an innovation almost always requires that technical knowledge be used in conjunction with other assets or capabilities such as marketing, manufacturing, after-sale service, distribution and software. Other necessary complements may include a host of ▶ intangible assets, such as a viable business model, customer relationships, reputations and organizational culture. If an innovator is slow to realize the importance of these assets/capabilities, does not have them or cannot easily contract to access them, it is likely to lose out to an imitator that is strong in these areas.

EMI's CAT scanner, for example, was a sophisticated machine that required a high level of customer training, support and servicing. EMI had none of these capabilities, could not easily contract for them and was slow to realize its strategic vulnerability (Teece 1986: 298). Competitors like GE with more experience selling complex healthcare equipment (along with the important complements of an experienced sales and marketing organization and a good reputation) were able to work around EMI's intellectual property and get into the market quickly with improved versions.

EMI's situation, in which the appropriability regime for its innovation had weaknesses and the absence of specialized assets left it compromised, is a common one. In these circumstances, the innovator must decide whether to contract for the supply of a critical capability (potentially creating a rival), build the capability internally (thus sacrificing flexibility), or find a joint venture partner to share the risk and rewards.

System Integration

Since the profiting from innovation framework was introduced, purchasing and partnering arrangements with domestic or offshore enterprises have become everyday occurrences. Many intermediate goods and services that were once hard to access in numerous industries are now available 'off the shelf'. The global transfer of technological know-how and capabilities through the investment and trading activities of multinational firms has helped to further spread know-how and capabilities across the globe.

In this altered landscape, the 'system integration' function – those capabilities required for business enterprises to orchestrate global resources – remains in scarce supply. With innovation occurring in different parts of the supply chain, the system-level innovator must decide which technologies/features to incorporate into its products, and then make those elements work together in a product that is useful and attractive to customers. This is especially important when the innovation is systemic in nature, meaning that a change in any component will require changes elsewhere in the system (Teece 1984).

The danger of failing to understand the need of system integration capabilities was evident in Boeing's experience with its 787 Dreamliner, which was over 3 years behind schedule when it began production in September 2011. Boeing, against the advice of some of its engineers, decided to rely far more than ever before on a global array of suppliers to develop parts for its new plane. This was seen as a cost- and risk-sharing measure; but Boeing reportedly failed to build sufficient internal monitoring capacity. Because some suppliers lacked the capabilities to develop parts of the necessary quality, the entire project experienced years of (very costly) delay. In the end, Boeing had to step in and help its suppliers reach the required level of competence (Kesmodel 2011).

Implications for Strategy

The theory of profiting from innovation provides a valuable template for guiding strategy formation and business model selection by innovators. Each element of the framework –the stage of industry evolution, the appropriability regime, the necessary complementary assets – requires careful analysis and reflection by itself. The framework, as elaborated in the initial 1986 entry and elsewhere, also provides guidance for understanding the interactions of these elements. For example,

complementary assets (and hence the firm's internal investments and external contracting relationships) play a more important role in industries where a ▶ dominant design has already emerged.

The theory is particularly relevant to the design of the innovator's business model with respect to whether to integrate a complementary component or activity or to contract for it (Jacobides et al. 2006). Making this decision correctly is one of the most critical steps for securing the innovator's profitability (Teece 2010). It requires that the innovator correctly assesses the firm's existing capabilities and/or its ability to develop new ones in a timely, cost-effective manner.

In certain cases, internal supply (i.e., [vertical] integration) may be worth pursuing even if it looks unattractive from a cost or time-to-market perspective. One such strategic reason is that the complement is co-specialized with the innovation (or, worse, the innovation is specialized to the complement but not the reverse). The dependence creates a potential hold-up problem that could allow an external supplier to extract a large share of profits (Williamson 1985).

An example of this hold-up problem is Intel's ability to sustain high prices (and profits) for its microprocessors vis-à-vis the computer companies that depend on it. If, during the initial development of its PC, IBM had asked its internal chip division to develop a microprocessor, then it would have entered the market later, but would probably still have dominated thanks to its reputation with business customers and its marketing muscle, while being able to deny its imitators access to a key input. More importantly, it would have captured much of the profits that it unwittingly delivered to Intel.

Another situation in which building internal supply capabilities makes sense is when the focal innovation creates a new industry and no existing suppliers have the required capabilities in place to provide the complement in sufficient quality or quantity. In such cases, strategic or other considerations could make it counterproductive to spend time convincing a potential supplier of the value of making the necessary investments. This was, for example, the logic behind vertical investments by the major industrial firms that emerged in the late nineteenth century (Chandler 1990). Companies exploiting new products (like sewing machines) or processes (like meat packing) often chose to integrate upstream into materials or other inputs, and downstream into marketing and distribution to attain the desired level of throughput.

Contracting for components or complements can reduce operating costs and risks, but it also entails strategic hazards. One of these is the risk of technology leakage (unintentional or otherwise) to competitors who are not part of the contract. A subtler hazard in such a relationship is the inability to pace or direct the evolution of a supplier's proprietary technology (De Figueiredo and Teece 1996). Microsoft, for example, develops certain applications that run on its Windows operating system, competing in some cases with independent software vendors who must rely on Windows for their development environment. Microsoft's ability to pace its upstream operating system technology, and its ability to use its intimate knowledge of that technology in its applications software, helped it to become one of the dominant players in applications.

In the presence of such hazards, controlling the path of learning and innovation sometimes requires vertical integration. But integration of a complementary product or function is generally a last resort that is most likely to be necessary when the innovator is disadvantageously positioned with respect to the complement. Recognizing the presence of such bottlenecks and developing an appropriate repositioning strategy is the key to profiting from innovation.

See Also

- ▶ Appropriability
- ▶ Capturing Value from Advantages
- ▶ Complementary Asset
- ▶ Dominant Design
- ▶ Imitability
- ▶ Innovation
- ▶ Intangible Assets
- ▶ Make-or-Buy Decisions: Applications to Strategy Research

- ▶ Network Effects
- ▶ Organizational Design
- ▶ Outsourcing
- ▶ Patents
- ▶ Systemic Innovation
- ▶ System Integrators
- ▶ Teece, David J. (Born 1948)

References

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

Chandler Jr., A. 1990. *Scale and scope: The dynamics of industrial capitalism*. Cambridge, MA: Belknap/Harvard University Press.

De Figueiredo, J.M., and D.J. Teece. 1996. Mitigating procurement hazards in the context of innovation. *Industrial and Corporate Change* 5: 537–559.

Jacobides, M.G., T. Knudsen, and M. Augier. 2006. Benefiting from innovation: Value creation, value appropriation and the role of industry architectures. *Research Policy* 35: 1200–1221.

Kesmodel, D. 2011. Boeing examines supply chain for weak links. *Wall Street Journal*, 30 December. Available at http://professional.wsj.com/article/SB10001424052970204058404577111091095438300.html. Accessed 12 Feb 2013.

Levin, R.C., A.K. Klevorick, R.R. Nelson, S.G. Winter, R. Gilbert, and Z. Griliches. 1987. Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity* 3: 783–831.

Mansfield, E. 1985. How rapidly does new industrial technology leak out? *Journal of Industrial Economics* 34: 217–223.

Mansfield, E., M. Schwartz, and S. Wagner. 1981. Imitation costs and patents: An empirical study. *Economic Journal* 91: 907–918.

Pisano, G.P., and D.J. Teece. 2007. How to capture value from innovation: Shaping intellectual property and industry architecture. *California Management Review* 50: 278–296.

Teece, D.J. 1984. Economic analysis and strategic management. *California Management Review* 26: 87–110.

Teece, D.J. 1986. Profiting from technological innovation. *Research Policy* 15: 285–305.

Teece, D.J. 2006. Reflections on profiting from innovation. *Research Policy* 35: 1131–1146.

Teece, D.J. 2010. Business models, business strategy and innovation. *Long Range Planning* 43: 172–194.

Williamson, O.E. 1985. *The economic institutions of capitalism*. New York: Free Press.

Winter, S. 2006. The logic of appropriability: From Schumpeter to Arrow to Teece. *Research Policy* 358: 1100–1106.

Promotions

Chakravarthi Narasimhan
University of Washington in St Louis, Olin Business School, St Louis, MO, USA

Abstract

Traditionally, promotional strategies include various forms of communications with the target consumers such as mass media advertising, targeted communications (*advertising*), monetary and non-monetary incentives provided to the channel intermediaries (*trade promotion*) and monetary and non-monetary incentives provided directly to consumers (*consumer promotion*) of the product. In most consumer product and service markets the promotional expenditure of brand is a significant budget item and, consequently, managing this is of critical importance for manufacturers of products and services. Since advertising is dealt with elsewhere in this encyclopedia, this entry is confined to research on trade and consumer promotions, and, in particular, to monetary incentives.

Definition The term 'promotions' as used in marketing describes the set of strategies used to stimulate awareness, interest and, ultimately, the sales of a brand or a product.

Traditionally, promotional strategies include various forms of communications with the target consumers such as mass media advertising, targeted communications (*advertising*), monetary and non-monetary incentives provided to the channel intermediaries (*trade promotion*) and monetary and non-monetary incentives provided directly to consumers (*consumer promotion*) of the product. In most consumer product and service markets the promotional expenditure of ▶ brand is a significant budget item and, consequently, managing this is of critical importance for manufacturers of products and services. Since advertising is dealt elsewhere in this encyclopedia, this entry is confined to research on trade and

consumer promotions, and, in particular, to monetary incentives.

In marketing, the terms trade and consumer promotions are generally used to refer to monetary incentives provided to channel members and consumers respectively. Generally, these are temporary incentives aimed at stimulating consumer demand. It is important to note that retailers may also promote a specific brand or product to its consumers that could be the result of a trade promotion received or without the receipt of such incentive. Academic research has focused on exploring the strategic rationale of manufacturers to offer such temporary promotions and the retailers' response to such incentives, and, using secondary data, quantifying the magnitude of the impact of different types of promotions.

Impact of Promotion on Sales

When a manufacturer lowers the retail price by offering a monetary incentive either directly (through a coupon or a rebate), or through a retailer, the impact on product demand is easy to predict. When retail price is lowered sales or volume goes up. However, it is important to understand exactly what drives this increase in volume. The fact that sales go up when a deal or a promotion is offered could be due to several factors. Consumers could switch from another brand (*brand switching*), consumers who would have normally bought the product at some time in future decide to buy now (*purchase acceleration*), consumers buy more than the usual volume either to consume more today (increase in *consumption*) or to inventory and consume it sometime in future (*stockpiling*) or both. There have been number of studies that have documented these basic effects (for a review see Neslin 2002; Van Heerde and Neslin 2008). Gupta (1988) found that in the coffee category brand switching accounted for 84 % of sales increase, acceleration 14 % and stockpiling only 2 %. Bell et al. (1999) find the average across 13 categories of brand switching to be 75 %. Van Heerde et al. (2003) make a case for decomposing promotion sales not through elasticities, as in the above studies, but through raw sales. Doing this, they find that the brand-switching effect is only 33 % and the rest is due to quantity effect. Using a structural model of optimizing consumers, Chan et al. (2008) find that stockpiling accounts for 44 % of the promotion bump, and consumption and brand switching equally accounts for the remainder. Apart from these effects, promotion in any given week could affect sales in future periods through its impact on consumer behaviour. Early research (Dodson et al. 1978) provided some evidence, in a laboratory setting, that buying on promotion could lead to a lower likelihood of buying in the next period. But subsequent research, using field experimental data, has disputed this claim (Davis et al. 1992), and others (see e.g., Neslin and Shoemaker 1989) have argued that even if a lower rate were to be observed other factors such as segmentation may be at work. Similarly, expected future prices or promotions (Krishna 1994; Gonul and Srinivasan 1996) could have an impact on current purchases. For example, expecting a lower price tomorrow consumers may postpone purchasing today; this has been termed purchase deceleration (for an empirical demonstration see Chan et al. (2008)).

Why Temporary Promotions?

There are various reasons why a manufacturer or retailer may offer incentives to consumers either directly or through the channel intermediaries. Tables 1 and 2 illustrates some possible reasons.

Regardless of the type of promotion there are several intriguing questions. Why are price incentives provided at periodic or aperiodic intervals? Second, if promotions induce ex post segmentation among users, why is this optimal for the firm ex ante to induce such segmentation ex post? A large number of analytical models have examined the rationale of firms offering some form of promotion, in monopoly and competitive contexts, and in selling directly to the consumers or selling through an intermediary such as a retailer. Here is a brief overview of some of these studies.

Promotions, Table 1 Examples of promotional incentives

Type	To trade (T) or consumer (C)	Comment
Coupons/ rebates	C	Must have a coupon or rebate to get a discount, both have a limited life, stricter compliance for rebates; may involve purchase of multiple units
Off invoice	T	% off every case bought by a retailer over a limited time
Free case	T	1 free case with the purchase of X cases, like a quantity discount
Bill back	T	% off every case sold by the retailer over a limited time
Weekly sales	C	Administered by the retailer but often supported by trade promotions. These vary from simple discount on a per unit to buy one get one (BOGO) free, etc.
Free good, bonus packs	C	Generated by manufacturers and sold through retailers' bonus packs, a small amount of complementary goods, etc.

Promotions, Table 2

Rationale	Examples	Comment
Liquidate excess inventory	End of the season sale, newer models/products	Timing is somewhat predictable
Market research	Any promotion, coupon in an advertisement	To learn about price sensitivity, to fine-tune advertising appeals
Introducing new product	Coupons, price discounts	To obtain shelf space, reduce cost of trial
Pass along cost reductions	Gas prices	If volume expands enough due to a lower price
Price discrimination	Coupons, some temporary promotions	Charge different prices to different individuals based on differences in valuation, information set, hassle cost, inventory cost etc.

Blattberg et al. (1981) consider a retailer selling to two segments of consumers, one with low inventory cost (h_L) and another with high inventory cost (h_H). The retailer's holding cost is $h_R > h_L$. All consumers have same constant consumption rate c and buy $Q \geq c$, trading off inventory cost and the current price. The retailer, having bought a fixed quantity, finds it optimal to shift his inventory cost to the low inventory cost consumers by periodically reducing the retail price, hence a promotion every T periods. Low inventory cost consumers buy from deal to deal and high inventory cost consumers buy every period. Jeuland and Narasimhan (1985) take a slightly different track and assume that the retailer is an efficient economic unit with a zero holding cost but faces two types of consumers (i =1,2) with a consumption rate that varies by price as $X_i = \alpha_i - \beta * P$, where P is the retail price. The crucial assumption they make and justify with examples is that α_i and consumers' holding costs are correlated. That is, the segment with a higher consumption rate has a higher holding cost. Given this set-up they show that retailers would offer a great deal to price discriminate between the two segments. The lower inventory cost segment not only buys more but expands its consumption because it buys on a deal yielding a different set of implications from Blattberg et al. (1981). Narasimhan (1988a) considers a market where a monopolist faces a loyal segment and a segment of consumers who haven't yet tried the product. Loyal consumers buy the product every period as long as it is less than a reservation price r. The non-purchasers will try the product only if the price $P \leq r$, and a fraction of them will repeat purchase at the reservation price r for an additional period. He shows that the monopolist will discount the product every T period to attract a chunk of the potential market. In Narasimhan (1984) the role of cents of coupons is examined. Noting that (1) the redemption rate of coupons is

often in the low single digit, (2) not all consumers use coupons, and therefore coupons induce ex post segmentation, (3) using coupons seems to involve hassle and time costs, he constructs a model of utility-maximizing consumers who face both the traditional budget constraint and constraint on the total hours available to them. He shows that the amount of coupon usage varies across households with identical preferences but varying opportunity cost of time. He also shows that the consumers who self-select to use coupons are consumers with more elastic demand. This explains why it is profit-maximizing for a firm to induce ex post heterogeneity in coupon usage. Using panel data across several grocery categories he finds strong support on the relationship between coupon usage and demand elasticity as well as across products. He obtains limited support for socio-economic-demographic variables to proxy for hassle and time costs.

All the models described above rely on either ex ante heterogeneity (segmentation) among consumers or self-selection by rational consumers that lead to ex post segmentation in explaining the existence of promotions. The above articles are representative of promotion models in a single-firm context. We next turn to models that explore firms' incentives to offer temporary discounts in a competitive context.

Narasimhan (1988b) offers a model of promotion in a competitive context that relies on heterogeneity among consumers in their preferences for brands in a competitive marketplace. He considers a ▶ duopoly market, with each firm offering one product. A fraction of consumers is loyal to one (α_i) and the remaining fraction $\beta = 1 - \alpha_1 - \alpha_2$ consists of switchers. The switchers prefer one brand over the other at equal prices but will be willing to switch to the less preferred brand if it were to be cheaper than the preferred brand by δ. The two firms set prices simultaneously. Narasimhan shows that, depending on the size of the loyal segments (α_i) and the switching premium, δ, we obtain interesting pricing strategies, including a constant price with no promotion, and periodic promotion. In the general case, there is no equilibrium in pure strategies. He interprets mixed strategies as a realization of promotions and shows that the size of the loyal segments (α_i) and the switching difference drive the comparative statics on the price promotions such as which firm is likely to promote more, and who, on average, is likely to offer a deeper discount. Raju et al. (1990) adopt a similar framework to Narasimhan except that they do not assume extreme loyalty and model a switching difference among brands. Like Narasimhan, they also obtain a mixed-strategy equilibrium that is interpreted as price promotions. Lal (1990) considers an infinitely repeated game in which two national brands marketing through a retailer compete for switchers who are willing to buy either one of the national brand or a store brand marketed by the retailer. He shows that a collusive equilibrium could arise where each national brand takes a turn in promoting its product in order to entice the switchers away from the store brand. Rao (1991) considers competition between a national brand and a private label where firms choose a price first and then a promotion strategy, that is, promotion frequency and depth. He shows that there is an equilibrium where a national brand promotes and a store brand does not. For an example of empirical testing of such strategies, see Villas-Boas (1995). More recently, Freimer and Horsky (2008) model consumer purchases as following a first order Markov and show that, in a competitive context, firms would alternate in promotion. For an empirical application of these concepts see Villas-Boas and Villas-Boas (2008). Lal and colleagues (1996) show that, in the presence of the retailer, a manufacturer would offer temporary promotions even when he knows that the retailer could forward buy, draining some of the profits. Lal and Villas-Boas (1998) model trade and consumer promotion in a competitive setting. For a recent review of trade promotions, see Narasimhan (2009).

See Also

- ▶ Brand
- ▶ Duopoly
- ▶ Price Discrimination

References

Bell, D.R., J. Chiang, and V. Padmanabhan. 1999. The decomposition of promotional response: An empirical generalization. *Marketing Science* 18: 504–526.

Blattberg, R.C., G.D. Eppen, and J. Liebeman. 1981. A theoretical and empirical evaluation of price deals for consumer nondurables. *Journal of Marketing* 45: 116–129.

Chan, T., C. Narasimhan, and Q. Zhang. 2008. Decomposing promotional effects with a dynamic structural model of flexible consumption. *Journal of Marketing Research* 45: 487–498.

Davis, S.J., J. Inman, and L. McAslister. 1992. Promotion has a negative effect on brand evaluations: Or does it? Additional disconfirming evidence. *Journal of Marketing Research* 29: 143–148.

Dodson, J.A., A.M. Tybout, and B. Sternthal. 1978. Impact of deals and deal retraction on brand switching. *Journal of Marketing Research* 15: 72–81.

Freimer, M., and D. Horsky. 2008. Try it, you will like it: Does consumer learning lead to competitive price promotions? *Marketing Science* 27: 796–810.

Gonul, F., and K. Srinivasan. 1996. Estimating the impact of consumer expectations of coupons on purchase behavior: A dynamic structural model. *Marketing Science* 15: 262–279.

Gupta, S. 1988. Impact of sales promotions on when, what, and how much to buy. *Journal of Marketing Research* 25: 342–355.

Jeuland, A.P., and C. Narasimhan. 1985. Dealing temporary price cuts by seller as a buyer discrimination mechanism. *Journal of Business* 58: 295–308.

Krishna, A. 1994. The impact of dealing patterns on purchase behavior. *Marketing Science* 13: 351–373.

Lal, R. 1990. Manufacturer trade deals and retail price promotions. *Journal of Marketing Research* 27: 428–444.

Lal, R., and J.M. Villas-Boas. 1998. Price promotions and trade deals with multiproduct retailers. *Management Science* 44: 935–949.

Lal, R., D. John, C. Little, and J.M. Villas-Boas. 1996. A theory of forward buying, merchandising, and trade deals. *Marketing Science* 15: 21–37.

Narasimhan, C. 1984. A price discrimination theory of coupons. *Marketing Science* 3: 128–147.

Narasimhan, C. 1988a. A model of discounting for repeat sales. In *Issues in pricing: Theory and research*, ed. T. Devinney. Lexington: Lexington Books.

Narasimhan, C. 1988b. Competitive promotional strategies. *Journal of Business* 61: 427–449.

Narasimhan, C. 2009. In *Trade promotions, handbook of pricing research in marketing*, ed. V. Rao. Northampton: Edgar Elgar.

Neslin, S.A. 2002. Sales promotion. In *Handbook of marketing*, ed. B. Weitz and R. Wensley. London: Sage.

Neslin, S.A., and R.W. Shoemaker. 1989. An alternative explanation for lower repeat rates after promotion purchases. *Journal of Marketing Research* 26: 205–213.

Raju, J.S., V. Srinivasan, and R. Lal. 1990. The effects of brand loyalty on competitive price. *Promotional Strategies* 36: 276–304.

Rao, R.C. 1991. Pricing and promotions in asymmetric duopolies. *Marketing Science* 10: 131–144.

Van Heerde, H.J., and S.A. Neslin. 2008. Sales promotions models. In *Handbook of marketing decision models*, ed. B. Wierenga. New York: Springer.

Van Heerde, H.J., S. Gupta, and D.R. Wittink. 2003. Is 75% of the sales promotion bump due to brand switching? No, only 33% is. *Journal of Marketing Research* 40: 481–491.

Villas-Boas, J.M. 1995. Models of competitive price promotions: Some empirical evidence from the coffee and saltine crackers markets. *Journal of Economics and Management Strategy* 4: 85–107.

Villas-Boas, J.M., and S.B. Villas-Boas. 2008. Learning, forgetting and sales. *Management Science*, 1951–1960.

Property Rights and Strategic Management

Nicolai J. Foss
Copenhagen Business School, Center for Strategic Management and Globalization, Frederiksberg, Denmark

Abstract

The economic analysis of property rights was pioneered in the 1960s, and has had some impact on strategic management theory over the last decade. These ideas redefine resources as endogenous outcomes of transaction cost economizing, link transaction costs and value creation and capture, and highlight the role of contracting in competitive strategy.

Definition Property rights have had a considerable influence on strategic management. These rights are being seen increasingly as another factor of production, and the ability to place a value on them has important implications for the strategic management of an organization.

Introduction

Applied micro and industrial organization economics have for a long time been the dominant

foundations of strategic management theory. Even though psychological and sociological perspectives have become increasingly important in theory development and empirical work in the field, the understanding of the central phenomena – ▶ value creation and appropriation, sustained competitive advantage – is fundamentally based on economics. Many of the foundational theories (e.g., the economics of uncertainty and information and human capital theory, as well as the first stabs at contract theory) emerged in the 1960s. One of the important breakthrough theories of the 1960s was property rights theory (PRT) (aka the economics of property rights), developed, first, by economists such as Armen Alchian, Ronald Coase and Harold Demsetz, and, subsequently, by, among others, Yoram Barzel, Eirik Furubotn and John Umbeck. PRT has had considerable influence on the development of a number of fields in economics, notably law and economics, economic history, the theory of the firm, contract economics, and resource and agricultural economics as well as, more indirectly, corporate governance and industrial organization theory (Foss 2010). Its impact on strategic management is smaller and more recent. Indeed, only around a dozen papers in strategic management are currently explicitly based on PRT (including Kim and Mahoney 2002, 2005; Foss 2003; Foss and Foss 2005, 2008; De Avila Monteiro and Zylbersztajn 2012; Foss et al. 2013), although many more papers make use of property rights ideas (along with other theories).

Property Rights Economics

Property Rights

PRT is basically an extension of neoclassical economics, in the sense that: (1) the utility-maximization hypothesis is applied to all choice; (2) all the constraints implied by the prevailing structure of property rights and transaction costs are considered; and (3) the organizational and institutional implications of (1) and (2) are highlighted.

The unit of analysis in PRT is the property right. As part of his critique of the Pigouvian tradition in welfare economics, Coase (1988: 155) notes that one reason for its failure to come fully to grips with ▶ externality issues is its 'faulty concept of a factor of production', which, according to Coase, should be seen not as a physical entity but as a right to perform certain actions (note how this can be extended to the notion of a resource; cf. Foss and Foss 2005). These rights are property rights. Demsetz (1964), Alchian (1977) further defined property rights as individuals' rights to the use, income and transferability of assets, and discussed the rights' relation to property law. An outcome of this discussion was that property rights can be analysed conceptually in isolation from legal considerations (as a result, some scholars now talk in terms of 'economic rights', e.g., Barzel 1997).

A fundamental idea in Coase (1988) is that transactions involve the exchange of property rights (rather than goods and services per se). As Coase explained, property rights to a resource can be partitioned in various ways. This led to a highly sophisticated analysis of how the property rights associated with an asset impact individual incentives, because property rights are fundamentally about who should bear the consequences of choices involving the relevant resource. In line with such ideas, Alchian and Allen (1969: 158) offered a highly compact definition of property rights as the 'expectations a person has that his decision about the uses of certain resources will be effective'. In other words, property rights are about the expected value of control over resources.

Transaction costs can be defined in terms of property rights, namely as the costs of delineating, protecting and capturing control over resources in use and in exchange. A famous benchmark case obtains when transaction costs are zero: Coase (1988) shows that if transaction costs are zero – so that any property right can be costlessly delineated and protected – any allocation of property rights results in the same pattern of economic activities under which maximum value is created from the use of resources. The presupposes that: (1) property rights to all possible uses of resources

are delineated; (2) all property rights are priced; and (3) all property rights can be traded – all at zero cost. Maximizing agents will have incentives to trade property rights so that resources end up being allocated to those uses where they make the maximum contribution to value creation.

Applications to Strategy

The most direct applications of the property rights model to strategy research are those by Kim and Mahoney (2002, 2005), and Foss and Foss (2005). Overall, these authors argue that PRT can further the ▶ resource-based view by providing insight into the link between transaction costs, and the value that can be created and appropriated by resource owners. The starting point lies in the Coase theorem: without transaction costs, maximum value will be created.

Foss and Foss (2005) argue that resources are not exogenously given (as in Barney 1991), but are endogenous outcomes of transaction cost minimization. Their starting point is that resources have multiple attributes (uses, services, functionalities). Some of these are worth defining, protecting and exchanging; others are not. Since agents confront different transaction costs of, for example, defining resource attributes, resources end up being economically heterogeneous. Foss and Foss (2005) also show that reducing transaction costs can be a source of value creation in a resource-based context. For example, superior contracts, sorting systems, reward systems and so on create value not just because they provide superior incentives, but also because they reduce value dissipation in the form of transaction costs. In a related manner, Kim and Mahoney (2002) explain how a resource-based analysis of value creation must take into account how property rights can internalize externalities (which increases value creation), and they exemplify this view using a case of oilfield unitization. A later paper by the same authors explains the contribution of PRT along other economics approaches to the understanding of strategic management (Kim and Mahoney 2005). De Avila Monteiro and Zylbersztjan (2012) provide a summary of these ideas, and apply them to a study of royalty collection. Further explorations of PRT in the context of strategic management include entrepreneurship and competitive strategy.

Foss and Foss (2008) apply PRT ideas to the study of strategic, firm-level entrepreneurship, showing how transaction costs and property rights interact in creating path-dependent entrepreneurial opportunities. The extent to which property rights can be enforced and the matrix of transaction costs determine the search directions and intensity of entrepreneurs.

Foss et al. (2013) argue that the PRT can inform the foundations of competitive rivalry research. Most of this research is based on a market failure paradigm where superior profitability is caused by monopolistic distortions in product markets. PRT suggests that such inefficiencies can only persist in the presence of transaction costs. If transaction costs are sufficiently low, buyers and other victims of monopolization can bribe the monopolist to supply the competitive output. In other words, the presence of transaction costs that block welfare-enhancing trades is necessary for competitive strategies based on market power to succeed.

Conclusion

While so far there have been only a handful of articles that are explicitly based on PRT foundations, there is reason to believe that PRT will gain in influence in strategic management theory, as it addresses core strategy issues – what are the sources of value creation? How are value creation and appropriation related? What is the economic nature of resource protection? What is the cost of strategizing? and so on – and does so in a novel and insightful manner. PRT yields new insight in the phenomena of strategic management because of its highly 'micro' unit of analysis, namely the individual property right; its emphasis on resource as having multiple attributes to which property rights can, depending

on transaction costs, be defined; and its consistent use of transaction costs everywhere in the analysis.

See Also

- ▶ Externality
- ▶ Resource-Based View
- ▶ Transaction Cost Economics
- ▶ Value

References

Alchian, A.A. [1965] 1977. Some economics of property rights. In his *Economics Analysis of Property Rights*. Indianapolis: Liberty Press.

Alchian, A.A., and W.R. Allen. 1969. *Exchange and production: Theory in use*. Belmont: Wadsworth.

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

Barzel, Y. 1997. *Economic analysis of property rights*, 2nd ed. Cambridge: Cambridge University Press.

Coase, R.H. [1960] 1988. The problem of social cost. In *The firm, the market and the law*. Chicago: University of Chicago Press.

De Avila Monteiro, G.G., and D. Zylbersztajn. 2012. A property rights approach to strategy. *Strategic Organization* 10: 366–383.

Demsetz, H. 1964. The exchange and enforcement of property rights. *Journal of Law and Economics* 7: 11–26.

Foss, N.J. 2003. The strategic management and transaction cost nexus: Past debates, central questions, and future research possibilities. *Strategic Organization* 1: 139–169.

Foss, N.J. 2010. Transaction costs and property rights. In *Handbook of transaction cost economics*, ed. P.G. Klein and M. Sykuta. Aldershot: Edward Elgar.

Foss, K., and N.J. Foss. 2005. Value and transaction costs: How the economics of property rights furthers the RBV. *Strategic Management Journal* 26: 541–553.

Foss, K., and N.J. Foss. 2008. Understanding opportunity discovery and sustainable advantage: The role of transaction costs and property rights. *Strategic Entrepreneurship Journal* 2: 191–207.

Foss, K., N.J. Foss, and P.G. Klein. 2013. A transaction cost approach to strategizing. Working paper, Department of Strategic Management and Globalization, Copenhagen Business School.

Kim, J., and J.T. Mahoney. 2002. Resource-based and property rights perspectives on value creation: The case of oil field unitization. *Managerial and Decision Economics* 23: 225–245.

Kim, J., and J.T. Mahoney. 2005. Property rights theory, transaction costs theory, and agency theory: An organizational economics approach to strategic management. *Managerial and Decision Economics* 26: 223–242.

Prospect Theory and Strategic Decision-Making

Galia Rosen-Schwarz and Zur Shapira
New York University, Stern School of Business, New York, NY, USA

Abstract

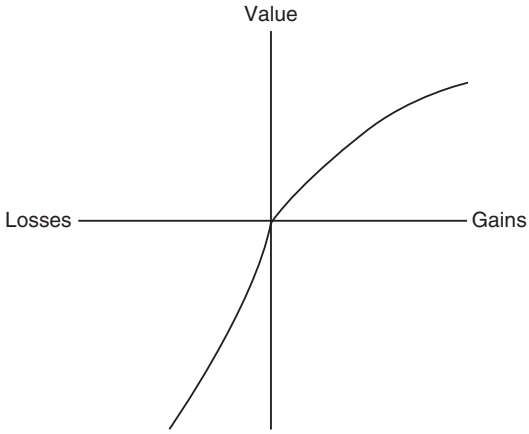
Prospect theory is one of the most influential theories of decision-making under risk. The theory draws on psychology and has been proposed as an alternative to expected utility theory. It substitutes a value function for the utility function, and a decision weights function for probabilities. Its main ideas are that changes in wealth are more important than absolute levels of wealth for describing risk-taking, and that decision weights is a non-linear function of probability. We describe the theory's main assumptions and its effects on strategy research and suggest a possible direction for its use in future research.

Definition Prospect theory is a descriptive theory of choice under risk that has been proposed as an alternative to expected utility theory. Like expected utility, it is framed mathematically as a bilinear pattern, but it substitutes an S-shaped value function for the utility function and decision weights for the probability measure.

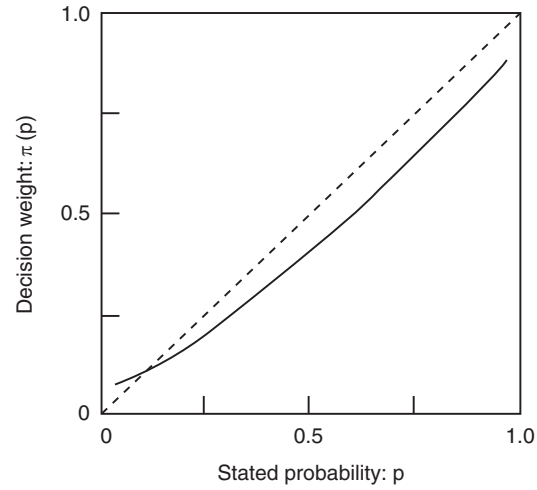
Prospect theory has been one of the most influential theories of decision under risk since it was first proposed by Kahneman and Tversky (1979). We describe its main components and review strategy research that has used it to analyse managerial and organizational decisions under risk.

The Value Function

The immediate association a researcher has with prospect theory is its S-shaped value function. This function has three properties: (1) it describes value as changes from a certain reference point; (2) it is marked by diminishing sensitivity – that



Prospect Theory and Strategic Decision-Making, Fig. 1 A hypothetical value function



Prospect Theory and Strategic Decision-Making, Fig. 2 A hypothetical weighting function

is, the difference in value between the same outcomes is smaller the further away they are from the reference point; and (3) it is concave for gains and convex for losses, and steeper for losses than for gains (see Fig. 1).

The Weighing Function

The second function replaces the probability element in expected utility with decision weights. According to Kahneman and Tversky (1979: 280), 'Decision weights are not probabilities, they do not obey the probability axioms and they should not be interpreted as measures of degree or belief.' Instead, 'Decision weights measure the impact of events on the desirability of prospects, and not merely the perceived likelihood of these events.' The function is non-linear: it overweights small probabilities and underweights moderate and high probabilities (see Fig. 2).

Editing and Evaluation

A major element of prospect theory is the two-phase choice process. In the first phase, prospects are edited so as to simplify the subsequent evaluation. Editing consists of several operations such as coding, combination, segregation and cancelation. Kahneman and Tversky (1979) pointed out that people usually perceive outcomes as either gains or losses and not as final wealth states, as is assumed under expected utility theory. They also noted that while gains and losses are perceived relative to a neutral reference point, 'the location of the reference point, and the consequent coding of outcomes as gains or losses, can be affected by the formulation of the offered prospects' (p. 274). The idea of formulation was later developed into the notion of framing (Tversky and Kahneman 1981), which became a cornerstone of their choice paradigm.

More Recent Developments in Prospect Theory

Tversky and Kahneman (1992) provided a new version of prospect theory to account for the idea of rank dependency, that is, the idea that outcome values and decision weights are not perfectly independent of each other. For example, Rottenstreich and Hsee (2001) demonstrate how very unlikely events that are affect-rich (such as a kiss from a film star, arguably a very valuable outcome) lead to feelings of hope, while almost-certain chances of loss lead to fear of loss. Together, these two effects lead to an added curvature of the decision-weighting function.

Prospect Theory in the Strategy Literature

Theories of risk-taking in strategy have frequently interpreted Kahneman and Tversky's prospect theory (1979) to explain (i) choices of decision makers in organizations and (ii) firm-level differences in risk behaviour. Scholars have used prospect theory's value function's asymmetry for gains and losses to predict risk seeking by organizations with low performance relative to a reference point, and risk avoidance by organizations with high performance relative to a reference point (e.g., Bowman 1980, 1982; Fiegenbaum and Thomas 1986, 1988; Fiegenbaum 1990; Jegers 1991).

Interpretations of prospect theory for the study of risk-taking by decision makers and groups within organizations have appeared in a variety of contexts that can generally be referred to as the behavioural perspective (cf., Audia and Greve 2006). The behavioural perspective has been applied widely in the last decades by researchers of strategic change and managerial ▶ decision-making. Its central argument is that decision makers use an aspiration level to evaluate performance and that the performance relative to the aspiration level influences their inclination to take further risks and make changes. The perspective is based on psychological processes of risk perception and preference suggested by prospect theory (Kahneman and Tversky 1979) and organizational search processes (Cyert and March 1963). Examples include the following.

Bowman's Paradox

Bowman (1980, 1982) found that the risk-return relationship in firms' performance is negative rather than positive, as economics and finance theories suggest. He suggested that, consistent with prospect theory, the risk-seeking tendency of poorly performing firms may explain this paradox. This paradox led to the development of a research stream known as the 'Bowman's paradox'. A comprehensive review of the Bowman's paradox literature is found in work by Nickel and Rodriguez (2002). Fiegenbaum and Thomas (1988) attempted to explain Bowman's risk-return paradox by complementing it with research from the behavioural theory of the firm. The research emphasized the role of targets in analysing risky choices. Observing the risk-return relationship across a sample of US public firms, the authors found a negative risk-return association for firms with returns below target levels and a positive association for firms with returns above their targets. Return and risk were measured by return on equity (ROE) and variance in ROE, respectively. Measures of the risk-return association were captured by Spearman rank-order correlation coefficients.

Strategic Reference Point Theory

Based on prospect theory, Fiegenbaum et al. (1996) suggest a mechanism of organizational decision-making targeted at obtaining strategic alignment between organizational competencies and the external environment. Organizations behave in a risk-seeking mode when their performance is below their target and in a risk-averse mode when their performance is above their target (Fiegenbaum and Thomas 1988). By signaling organizational priorities, decision makers focus the attention of organizational members on particular goals and objectives, thus defining the strategic reference point for the firm (Fiegenbaum et al. 1996).

Behavioural Agency Theory

Building on prospect theory and agency theory views, Wiseman and Gomez-Mejia (1998) constructed a behavioural agency model of managerial risk-taking. The authors argue that prospect and agency theories are complementary and that combining them may improve the power of agency-based models to predict and explain executive risk-taking behaviour. The model suggests that executive risk-taking varies across and within different forms of monitoring and that agents may exhibit risk-seeking as well as risk-averse behaviours. The authors examine key aspects of incentive alignment and monitoring control, and develop propositions on how the decision and

risk-bearing attributes associated with these aspects influence risky choices.

Alternatives Behavioural Views of Organizational Risk-Taking

Alternative views to prospect theory have been proposed to explain organizational risk-taking. We briefly discuss two examples, the first related to prospect theory, the second not based on formal models of risk analysis.

The Variable Risk Preferences Model

March and Shapira's (1992) variable risk preferences model proposes that rather than focusing their attention on a single reference point, as prospect theory suggests, decision makers shift their focus of attention between two alternative reference points: a dynamic aspiration level and fixed survival point. Both reference points have different implications for risk-taking. In addition, the resources at the disposal of managers (or firms) make a difference as well. When focused on aspirations, decision makers are expected to behave in a similar way to the prediction of the value function suggested by prospect theory; they are expected to be risk seeking when accumulated resources are below target and risk averse when their resources are above target. However, when focused on survival, decision makers are hypothesized to be risk averse.

The model proposes a more complex pattern than the one implied by prospect theory, due to the combined effects of resources and focus of attention. These two variables are potentially independent; that is, decision makers can focus on either reference point regardless if they are below or above their target. Therefore, when decision makers are below their target, they are likely to behave in a risk-seeking manner if they are focusing on their aspiration level and in a risk-averse manner if they focus on the survival point.

Escalation of Commitment

Escalating commitment to a previously chosen but losing course of action is usually attributed to the decision-makers' need to maintain the illusion that they have not made an erroneous decision. The individual, when motivated by a need for justification, seeks to appear competent in previous as opposed to future actions (Staw 1980). Such behaviour runs against the rational principle of sunk cost. Whyte (1986: 311) proposed that 'Prospect theory suggests a different explanation for this commonly observed tendency-escalating commitment is seen as an artifact of the framing of decisions. As a result, escalating commitment may occur in a much wider variety of circumstances than is suggested by the view that it is a product of self-justification motives.' This view has guided much of the later work on escalation of commitment (Sharp and Salter 1997).

Yet the two theories focus on different phenomena. Prospect theory centres on single choices among gambles and proposes a generic model to describe such behaviour. In contrast, the escalation of commitment focuses on a series of decisions where each decision depends on the previous decision (and its outcome). Proponents of this model see their bases in social psychology rather than in behavioural decision-making.

On the Application of Prospect Theory to Strategy Research

Bromiley (2010) argues that prospect theory has not been followed properly in strategic management research. For instance, the analysis of choices between gambles in isolation, while ignoring current wealth, is different from the situation where managers take risks. He points out also that in the S-shaped value function of prospect theory, risk tendencies decline with distance from the reference point, whereas some strategy researchers found the opposite in some cases. He also notes that strategy researchers ignored the role of the decision weighting function in affecting risk-taking.

Many aspects of Bromiley's critique are correct, yet in evaluating prospect theory, one has to praise its amazing insights regarding risk-taking. The theory attempted to provide an alternative to utility theory given evidence suggesting that it was not accurately describing choice behaviour. Prospect theory is based on the psychology of

perception and, as such, it naturally focuses on the individual decision maker. As one of its creators told the second author, they were interested in describing pure attitudes towards money, hence their focus on either only positive or only negative gambles. In addition, they tried to come up with an account of risk-taking that was general; therefore, they did not incorporate context into their model. It is clear that strategic risk-taking is embedded in context and that such decisions are mostly of the mixed-gamble type. Prospect theory contributed tremendous insight to strategy by pointing to the effects of reference points on risk attitudes. In this respect it complemented the behavioural theory of the firm's emphasis on the role of aspiration levels. Future research in ▶ strategic decision-making should continue to integrate the insights from these two great theories for a better understanding of organizational risk-taking.

See Also

- ▶ Behavioural Strategy
- ▶ Choice Modelling
- ▶ Decision-Making
- ▶ Strategic Decision-Making

References

Audia, P.G., and H.R. Greve. 2006. Less likely to fail: low performance, firm size, and factory expansion in the shipbuilding industry. *Management Science* 52: 83–94.

Bowman, E.H. 1980. A risk/return paradox for strategic management. *Sloan Management Review* 21: 17–31.

Bowman, E.H. 1982. Risk seeking by troubled firms. *Sloan Management Review* 24: 33–42.

Bromiley, P. 2010. Looking at prospect theory. *Strategic Management Journal* 31: 1357–1370.

Cyert, R., and J. March. 1963. *A behavioral theory of the firm*. Englewood Cliffs: Prentice Hall.

Fiegenbaum, A. 1990. Prospect theory and the risk return association: An empirical examination of 85 industries. *Journal of Economic Behavior & Organization* 14: 187–203.

Fiegenbaum, A., and H. Thomas. 1986. Dynamic and risk measurement perspectives on Bowman's risk-return paradox for strategic management: An empirical study. *Strategic Management Journal* 7: 395–407.

Fiegenbaum, A., and H. Thomas. 1988. Attitudes toward risk and the risk-return paradox: Prospect theory explanations. *Academy of Management Journal* 31: 85–106.

Fiegenbaum, A., S. Hart, and D. Schendel. 1996. Strategic reference point theory. *Strategic Management Journal* 17: 219–235.

Jegers, M. 1991. Prospect theory and the risk-return relation: Some Belgian evidence. *Academy of Management Journal* 34: 215–225.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47: 263–291.

March, J.G., and Z. Shapira. 1992. Variable risk preferences and the focus of attention. *Psychological Review* 99: 172–183.

Nickel, M.N., and M.C. Rodriguez. 2002. A review of research on the negative accounting relationship between risk and return: Bowman's paradox. *Omega* 30: 1–18.

Rottenstreich, Y., and C. Hsee. 2001. Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science* 12: 185–190.

Sharp, D., and J.S. Salter. 1997. Project escalation and sunk costs: A test of the international generalizability of agency and prospect theories. *Journal of International Business Studies* 28: 101–121.

Staw, B.M. 1980. Rationality and justification in organizational life. In *Research in organizational behavior*, vol. 2, ed. B. Staw and L. Cummings. Greenwich: JAI Press.

Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211: 453–458.

Tversky, A., and D. Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5: 297–323.

Whyte, G. 1986. Escalating commitment to a course of action: A reinterpretation. *Academy of Management Review* 11: 311–321.

Wiseman, R.M., and L.R. Gomez-Mejia. 1998. A behavioral agency model of managerial risk-taking. *Academy of Management Review* 23: 133–153.

Prospectors, Defenders, Analysers

Raymond E. Miles[1] and Charles C. Snow[2]
[1]University of California Berkeley, Berkeley, CA, USA
[2]Pennsylvania State University, Smeal College of Business, University Park, PA, USA

Definition The Miles–Snow typology (prospectors, defenders and analysers) has been widely used in the field of strategic management as well as in marketing. It applies to the business-level strategies of a firm, and it includes the capabilities,

structures and processes necessary for the implementation of each strategy type.

At the time of our original research in the 1960s and 1970s, the field of strategic management did not yet exist as a formal academic discipline. The Miles and Snow (1978) typology helped to establish the field by integrating the literature on managerial strategizing and planning with the literature on organization theory. Our original empirical studies included firms in an industry undergoing transformation (college textbook publishing), two industries adapting to new technologies (microelectronics and hospitals), and a stable industry (food processing). In each industry, we identified patterns in firms' strategic choices. Prospectors were first to the market with new technologies and products; analysers were adept at product enhancement and commercialization; and defenders were efficient producers of products in market segments that are predictable and expandable. Further, our studies examined the relationship between a firm's strategy and its management philosophy (Miles 1975) and capabilities (Snow and Hrebiniak 1980). Last, we investigated how the strategy types become part of multi-firm networks and collaborative communities of firms (Snow et al. 2011).

Since the 1980s, scholars have examined the Miles–Snow typology's validity and reliability (Shortell and Zajac 1990), the effectiveness of the strategy typology compared with other prominent typologies (Doty et al. 1993), the functional attributes and performance of the strategy types in different industries and countries (Hambrick 1983; DeSarbo et al. 2005), the relationship of each strategy type to the firm's marketing orientation (Slater et al. 2006) and the extent of the typology's use (Zahra and Pearce 1990). New academic disciplines benefit from early typological development (Tiryakian 1968), and the Miles–Snow typology contributed in this way to the field of strategic management.

See Also

► Innovation Strategies

References

DeSarbo, W.S., C.A. Di Benedetto, M. Song, and I. Sinha. 2005. Revisiting the Miles and Snow strategic framework: Uncovering interrelationships between strategic types, capabilities, environmental uncertainty, and firm performance. *Strategic Management Journal* 26: 47–74.

Doty, D.H., W.H. Glick, and G.P. Huber. 1993. Fit, equifinality, and organizational effectiveness: A test of two configurational theories. *Academy of Management Journal* 36: 1196–1250.

Hambrick, D.C. 1983. Some tests of the effectiveness and functional attributes of Miles and Snow's strategic types. *Academy of Management Journal* 26: 5–26.

Miles, R.E. 1975. *Theories of management: Implications for organizational behavior and development*. New York: McGraw-Hill.

Miles, R.E., and C.C. Snow. 1978. *Organizational strategy, structure, and process*. New York: McGraw-Hill.

Shortell, S.M., and E.J. Zajac. 1990. Perceptual and archival measures of Miles and Snow's strategic types: A comprehensive assessment of reliability and validity. *Academy of Management Journal* 33: 817–832.

Slater, S.F., E.M. Olson, and G.T.M. Hult. 2006. The moderating influence of strategic orientation on the strategy formation capability–performance relationship. *Strategic Management Journal* 27: 1221–1231.

Snow, C.C., and L.G. Hrebiniak. 1980. Strategy, distinctive competence, and organizational performance. *Administrative Science Quarterly* 25: 317–336.

Snow, C.C., Ø.D. Fjeldstad, C. Lettl, and R.E. Miles. 2011. Organizing continuous product development and commercialization: The collaborative community of firms model. *Journal of Product Innovation Management* 28: 3–16.

Tiryakian, E.A. 1968. Typologies. In *International encyclopedia of the social sciences*, vol. 16, ed. D.L. Sills. New York: Macmillan.

Zahra, S.A., and J.A. Pearce. 1990. Research evidence on the Miles–Snow typology. *Journal of Management* 16: 751–768.

Psychic Distance

Jan Johanson[1] and Jan-Erik Vahlne[2]
[1]Uppsala University, Uppsala, Sweden
[2]University of Gothenburg, Gothenburg, Sweden

Abstract

Psychic distance is a behavioural concept capturing the uncertainty of decision makers due to lack of knowledge about foreign markets. This uncertainty is a consequence of overall social

and economic factors in the foreign markets and of the experience of the decision makers. It has been used with success to explain firms' market selection and entry, but is less useful for explaining management in foreign markets.

Definition Psychic distance is defined as the subjectively perceived distance to a given country.

The Early Background to Psychic Distance

That distance impacts on economic exchange was clear to Adam Smith. That the 'psychic distance' as a specific dimension of distance could explain the geographical distribution of trade was pioneered by Beckermann (1956), who found that countries trade mainly with their neighbours. Beckermann did not give a precise definition of the concept but the term psychic distance indicates that it is a matter of subjective perception. Vahlne and Wiedersheim-Paul (1973), using the concept to explain the spatial distribution of Swedish exports, defined psychic distance as the sum of factors preventing flows of information between markets. In operationalizing the theoretical concept differences between countries in factors such as level of economic development, level of education, culture, language and cultural differences were measured. Psychic distance was shown to have a stronger explanatory value than physical distance measured as tariffs and transport costs. A subsequent study showed that psychic distance had a strong impact also on the pattern of location of sales subsidiaries by Swedish multinational firms, starting close to the home market and then moving gradually further away from home in psychic distance terms (Hörnell et al. 1973). Together with market size, psychic distance contributed to understanding of the pattern of internationalization of firms (Johanson and Wiedersheim-Paul 1975).

The reason behind the interest in psychic distance was that this was believed to explain why decision makers would be far from well informed about a foreign market when taking decisions to enter that market. Following Penrose (1959) it was assumed that markets are heterogeneous and cannot be understood without direct market experience that cannot be easily transmitted between markets. Thus, decisions were taken under uncertainty. International business transactions were influenced not only by objective economic realities but by bounded rationality – behavioural factors.

An early consequence of the psychic distance concept is that it placed attention on the role of knowledge in ▶ international business and, in particular, the problems associated with lack of knowledge about foreign markets. The empirical findings of the influence of psychic distance, mentioned above, were important input in the conceptualization of the internationalization process of the firm as a gradual process of experiential knowledge development (Johanson and Vahlne 1977).

Consequences of Psychic Distance

Over the years, the concept has been used with some success to explain firms' foreign market selection and entry (Dow 2000; Child et al. 2002; Ellis 2008). But it has also been used without success in a number of other aspects of international business. One of the reasons can be traced to its use in the early Uppsala studies. As the name of the concept indicates it is meant to capture a psychological dimension of distance, it stresses a subjective element in international ▶ decision-making. In contrast, the Uppsala group used a set of objective indicators to measure psychic distance. Thus, they introduced an ambiguity in the concept that has influenced much of later research and, consequently, its results. In the following we present some studies that have used psychic distance with different results.

A first important study was conducted by Kogut and Singh (1988), who demonstrated that entry mode choice was influenced by cultural distance between home and target countries. Their measure of cultural distance was based on the cultural dimensions formed by Hofstede (1980) – individualism, uncertainty avoidance, power distance and masculinity/femininity. They argued that the cultural distance was almost the same as psychic distance but had the advantage that it made it possible

to measure distances between all pairs of countries in the same way, thus allowing them to conduct large-scale studies of the effects of psychic distance. Following Kogut and Singh, a number of researchers studied the effects of distance using the cultural distance scales as a proxy for psychic distance. In a study of foreign direct investments Benito and Gripsrud (1992) found that the decisions were based on rational location choices rather than cultural learning processes.

In a study of Canadian retail firms' internationalization O'Grady and Lane (1996) made some interesting observations. As expected, they found that the retailer firms began in the United States, which is the closest and, in many ways, the most similar country to Canada. Unexpectedly, however, they found that almost 80 % of the firms failed. Thus, while psychic distance influenced firms' market entry it did not result in successful performance in the market. In particular, the research indicated that starting internationalization in a country close to the home market may result in poor performance. They called this the psychic distance paradox. Their explanation for the paradox was that the perceived similarity of the foreign country may make the actors unprepared for the differences and unable to learn about it. Although all the managers acted as if there was almost no psychic distance to the American market, there were individual differences in performance. Managers with experience of the American market were more successful, and American managers were still more successful, in handling the American subsidiary.

Distance factors have been attractive explanations of international business activities. But, as indicated above, the results of research have been confusing. An important step was taken by Nordström (1990), who developed a direct measure of psychic distance based on the original view that psychic distance is the subjectively perceived distance to a given country. After a presentation for managers of the original conception of psychic distance, he let respondent managers set the psychic distance from their home country to different foreign countries on a scale from 1 to 100. With some minor variations, this definition and the corresponding measure has been used by several researchers (Håkansson and Ambos 2010).

This step was important because it demonstrated that there is a fundamental difference between psychic distance and cultural distance. Cultural distance as defined by Hofstede is a difference between two countries. It is objective and consequently independent of the specific actor. Psychic distance is a subjective perception of an actor. It captures the uncertainty due to lack of information about a foreign country.

In order to reduce the confusion surrounding the concept, Håkansson and Ambos (2010) conducted a large-scale study of the psychic distance between a number of countries and factors that influence it. They used the method employed by Nordström (1990). Managers in the 25 largest countries (measured by their absolute GDP in 2001) set the psychic distance to the other countries in the group. In this way the researchers could specify the average psychic distance on a scale from 0 to 100 between all pairs of countries in the group according to the managers.

Antecedents of Psychic Distance

The analysis of the data shows that the following all have some impact on the psychic distance to the target country: (1) cultural distance, (2) geographical distance, (3) common language, (4) political rivalry, (5) differences in economic development, (6) economic development of target country, (7) relative governance quality of target country and (8) gross domestic product of target country.

The central finding of the study is that psychic distance and cultural distance are related but different. This means that they should not be used as substitutes for each other. An interesting finding is that geographical distance has a strong influence on psychic distance. This implies that psychic distance is stable over time. As expected, common language is important. The economic development of the target country has a strong influence on the psychic distance, which implies that the distance is not symmetric.

As a summary, psychic distance is a behavioural concept capturing the uncertainty of decision makers due to lack of knowledge about foreign markets. This uncertainty is a

consequence of overall social and economic factors in the foreign markets and of the experience of the decision makers. It has been used with success to explain firms' market selection and entry, but is less useful for explaining management in foreign markets.

See Also

- ▶ Decision-Making
- ▶ International Business

References

Beckermann, W. 1956. Distance and the pattern of intra-European trade. *Review of Economics and Statistics* 28: 31–40.

Benito, G., and G. Gripsrud. 1992. The expansion of foreign direct investments: Discrete rational location choices or a cultural learning process? *Journal of International Business Studies* 23: 461–476.

Child, J., S.H. Ng, and C. Wong. 2002. Psychic distance and internationalization. *International Studies of Management and Organization* 32: 36–56.

Dow, D. 2000. A note on psychological distance and export market selection. *Journal of International Marketing* 8: 51–64.

Ellis, P.D. 2008. Does psychic distance moderate the market size-entry relationship? *Journal of International Business Studies* 39: 351–369.

Håkansson, L., and B. Ambos. 2010. The antecedents of psychic distance. *Journal of International Management* 16: 195–210.

Hofstede, G. 1980. *Cultures consequences: International differences in work-related values*. Beverly Hills: Sage.

Hörnell, E., J.-E. Vahlne, and F. Wiedersheim-Paul. 1973. Ekonomiskt avstånd och etablering. In *Export och utlandsetableringar*, ed. E. Hörnell, J.-E. Vahlne, and F. Wiedersheim-Paul. Uppsala: Almqvist & Wiksell.

Johanson, J., and J.-E. Vahlne. 1977. The internationalization process of the firm: A model of knowledge development and increasing foreign market commitments. *Journal of International Business Studies* 8: 25–34.

Johanson, J., and F. Wiedersheim-Paul. 1975. The internationalization of the firm: Four Swedish cases. *Journal of Management Studies* 12: 305–322.

Kogut, B., and H. Singh. 1988. The effect of national culture on the choice of entry mode. *Journal of International Business Studies* 19: 411–432.

Nordström, K.A. 1990. *The internationalization process of the firm*. Stockholm: Stockholm School of Economics.

O'Grady, S., and H.W. Lane. 1996. The psychic distance paradox. *Journal of International Business Studies* 27: 309–333.

Penrose, E. 1959. *The theory of the growth of the firm*. Oxford: Basil Blackwell.

Vahlne, J.-E., and F. Wiedersheim-Paul. 1973. Ekonomiskt avstånd: model och empirisk prövning [Economic distance: model and empirical investigation]. In *Export och utlandsetableringar [Exports and foreign establishment]*, ed. E. Hörnell, J.-E. Vahlne, and F. Wiedersheim-Paul. Uppsala: Almqvist & Wiksell.

Psychological Basis of Quality Decision-Making

Charlan Jeanne Nemeth
University of California Berkeley, Department of Psychology, Berkeley, CA, USA

Abstract

Effective strategic management requires analysis, decisions and actions by an organization to create and sustain competitive advantage. Good decisions are clearly desirable, but whether a decision is good or not is a judgement call, often after the fact, and is itself subject to bias. What is less open to debate is the process that leads to accuracy or quality decision-making. This requires not just access to available information but proper processing, interpretation and integration of that data. Consideration of multiple options and perspectives is critical at all stages and there are many reasons why people do not do that. Poor decisions come from ineffective information search, selective bias in processing the information, a lack of alternatives being considered, a failure to examine the risks of the preferred choice and a rush to judgement (Janis and Mann 1979). In short, the selection, interpretation and integration of information are biased.

Definition Decision-making involves choosing among alternatives based on the goals and values of the person or group making the decision. Good (or quality) decision-making involves a thorough analysis of available information and a consideration of alternatives in an unbiased manner.

In this brief article we consider bias both at the individual and the group level. The overarching perspective is that there are psychological reasons that constrain and bias thought and psychological antidotes that can improve it. The former are more numerous and well documented than the latter, but in both individual and group decision-making, the influences are predictable, pervasive and profound, resulting in defective decision-making sometimes and to better and even creative decision-making at other times.

Bias in Individual Judgement and Decision-Making

The discipline of psychology has long recognized that the 'rational man' model of economics has its limits. People do not attend to, nor do they interpret, information in a rational manner, at least not in the logical or statistical sense. Rather than being a matter of intelligence or sophistication, people show bias and use shortcuts in their reasoning. Some of the biases are motivational. Some are more cognitive, a way of thinking in the face of the enormity and complexity of the information.

The Motivational Biases

The classic work on dissonance (Festinger 1957) pointed to important motivational reasons for bias in the selection and interpretation of information, even (and perhaps especially) after decisions are made. The simple idea is that people want their cognitions to be consistent and, if they are not, the mind works to make them so. Thus, the fact that I lied is inconsistent with my being an honest person and if I lie for a small reward or reason, the cognitions are even more inconsistent. A way of making them more consistent is to believe the lie – it becomes a truth. Thus, you get the counter-intuitive finding that lying leads to a change of attitude in the direction that the lie was true and the smaller the reward the more the attitude change. Going through 'hell week' (or a version of it in the laboratory) doesn't make you hate the organization; rather it leads to more empathy and bonding with the group.

Pertinent to decision-making, substantial research shows that we search selectively for information that confirms our initial belief. If the information is ambiguous, we interpret it in a way that supports that belief. Consider a 'foul' in a game or the question of 'who won' a political debate. Once a decision is made things get worse. Having taken a decision, people are even more convinced it was the right one and derogate the path not chosen (Cooper 2007).

The Cognitive Biases

In recent years, emphasis has been placed on 'bias' that is cognitive in nature. The assumption is that the bias has to do with the ways in which we think and process complex information.

Some 'biases' are due to the fact that most people are not good statisticians. One example is the *gambler's fallacy*. If you throw heads six times in a row, most people assume that tails is more likely on the next throw without realizing that the die has no memory so it remains a 50/50 chance. Another is the '*representativeness heuristic*' (Tversky and Kahnemann 1974). People judge the likelihood of something happening by matching it to a prototype. Such snap judgements often ignore other pertinent information such as base rates.

This tendency to ask 'what goes with what' results in confidence when the information matches expectations. While confidence is generally considered an asset, it can be problematic. Termed the 'overconfidence effect', people's subjective confidence in their judgements is often greater than the objective accuracy. In some quizzes, for example, people rate their answers as 99 % certain but are wrong 40 % of the time. Investors are often overconfident of their strategies, ignoring evidence to the contrary. While such confidence can lead to action, it often impedes the consideration of alternatives (Nemeth 1997; Nisbett and Ross 1980).

The 'availability heuristic' refers to the fact that the ease with which information comes to mind makes it seem more likely or more frequent. When asked 'Which is more common – words that start with the letter k or words that have k as

the third letter', most assume the former when in fact, there are more than twice as many words where k is the third letter. This is related to the 'vividness effect'. Just one vivid case study trumps extensive information on the topic, Think of the appeals for contributions to alleviate hunger. The story of one charming little girl who lives off the sale of salvage is far more effective than extensive statistics on food supply and child mortality. A single colourful testimonial is more effective than any survey (see, for a general view, Kahneman et al. 1982).

The confirmation bias exacerbates selective bias because people test their hypotheses by confirmation rather than disconfirmation (Wason 1960), Given a sequence of 2, 4, 6, most assume the rule is ascending even numbers and then 'test' that by giving sequences such as 10, 12, 14 which confirm the assumption, rather than a sequence such as 1, 2, 3 which could disconfirm it. Thus, they tend not to find the correct rule, which is any sequence in ascending order.

Framing of alternatives (Tversky and Kahnemann 1974) is particularly revealing. When alternatives are framed positively, we are risk averse. We prefer saving 500 lives to a 1 in 3 probability that 1,500 will be saved. When framed negatively, we take the risk. We prefer a 1 in 3 probability that 1,500 will die to a certainty that 500 will die. Similar findings are found for investment strategies.

Attempts to reduce such biases in the service of decision-making have included mechanisms of getting people to reassess the 'shortcut' or to consider alternatives. This includes education about how biases operate, training, the technique of 'considering the alternative' and inviting dissenting viewpoints (Bazerman and Moore 2008; Nemeth 2012).

Bias in Group Decision-Making

Most research on group decision-making has documented incidents of failure (Janis 1972). The Challenger accident is one dramatic example and has become a case study of what *not* to do. On 28 January 1986, this shuttle exploded shortly after liftoff and all seven crew members were killed, including Christa McAuliffe, a civilian teacher-in-space. NASA's scientific research and operations experienced years of setback. What makes this of particular concern is that NASA was forewarned of a problem.

Thiokol, the subcontractor responsible for the shuttle's 'O' rings, expressed concerns that the rings might not seal properly in cold temperatures. They recommended that the shuttle not be launched until the outside temperature reached 53 °F, a rate not forecast for several days. This recommendation was not followed and is illustrative of the symptoms of bad decision-making, including pressure on the dissenter. Famously, Thiokol's chief engineer was asked to 'take off his engineering hat and put on his management cap', which essentially argued for organizational goals to take precedence over safety considerations. People assumed the majority were in agreement and when asked if there were any objections, there was silence.

Illustrative of what has become known as groupthink, this is characterized by a highly cohesive and insulated group whose leader has a distinct preference and where there is pressure such as time or external mandates. Under such circumstances, groups often express overconfidence, an illusion of invulnerability and conformity. There is a lack of consideration of alternatives and a resistance to reconsidering the initial preference. Repeatedly, and over many cultures, research has confirmed the power of the leader and of majority opinion to gain compliance. However, this is often at the cost of accuracy and good decision-making. Attempts to counter this have included breaking into subgroups, using a role play such as devil's advocate, calling on the opinion of outsiders and having the leader refrain from stating his preferences. Subsequent empirical work suggests that the directed leadership is more of a problem than cohesiveness, while some techniques such as devil's advocate do not work (Nemeth and Goncalo 2005).

Work over the past two decades has demonstrated perhaps a more powerful phenomenon of majority viewpoints, namely that people *think* from the perspective of the majority to the exclusion of other considerations. People choose to

take into account information that supports the majority view and avoid the full array of available information; they adopt strategies utilized by the majority to the exclusion of other useful strategies they would normally use; and they show less originality of thought (Nemeth 2012).

By contrast, there is evidence that dissent liberates people; they are more independent and resistant to conformity. Perhaps more importantly, dissent, even when wrong, stimulates thought that is divergent in form. People attend to information on all sides of the issue, consider multiple alternatives and show creativity. As a result, decisions are better and more creative (Nemeth 2012; Van Dyne and Saavera 1996).

Reflections

The research literature in both individual and group decision-making points to the biases in information processing and the interpretation of that information. While some form of 'shortcut' may be needed given the complexity of the issues and time pressures, there is evidence that thought is constrained by either individual 'biases' or by a majority view in groups. People restrict the information they read and the alternatives they consider and often end up making inaccurate or poor decisions. In the group setting, these biases are exacerbated given the tendency to assume that the majority is correct and the fear of disapproval should one persist in a minority viewpoint. Yet it is dissent that appears to be the antidote. At an individual level, we can be our own dissenter. At the group level, it means inviting dissent. This does not mean an easy role-playing technique such as devil's advocate but rather inviting contrasting, authentic viewpoints.

See Also

- ▶ Bounded Rationality
- ▶ Heuristics and Biases and Strategic Decision-Making
- ▶ Organizational Ambidexterity
- ▶ Strategic Decision-Making

References

Bazerman, M., and D. Moore. 2008. *Judgment in managerial decision making*, 7th ed. New York: Wiley.

Cooper, J. 2007. *Cognitive dissonance: 50 years of a classic theory*. London: Sage.

Festinger, L. 1957. *A theory of cognitive dissonance*. Stanford: Stanford University Press.

Janis, I.L. 1972. *Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes*. Oxford: Houghton Mifflin.

Janis, I.L., and L. Mann. 1979. *Decision making: A psychological analysis of conflict, choice and commitment*. New York: Free Press.

Kahneman, D., P. Slovic, and A. Tversky. 1982. *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Nemeth, C.J. 1997. Managing innovation: When less is more. *California Management Review* 40: 59–74.

Nemeth, C.J. 2012. Minority influence theory. In *Handbook of theories in social psychology*, ed. P. Van Lange, A. Kruglanski, and T. Higgins. New York: Sage.

Nemeth, C., and J.A. Goncalo. 2005. Influence and persuasion in small groups. In *Persuasion: Psychological insights and perspectives*, ed. T.C. Brock and M.C. Green. London: Sage.

Nisbett, R.E., and L. Ross. 1980. *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs: Prentice Hall.

Tversky, A., and D. Kahnemann. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185: 1124–1131.

Van Dyne, L., and R. Saavera. 1996. A naturalistic minority influence experiment: Effects on divergent thinking, conflict and originality in work groups. *British Journal of Social Psychology* 35: 151–168.

Wason, P.C. 1960. On the failure of to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology* 12: 129–140.

Public Policy: Strategy in the Public Interest

Anita M. McGahan[1] and Peter G. Klein[2]
[1]University of Toronto, Rotman School of Management, Toronto, ON, Canada
[2]University of Missouri, Columbia, MO, USA

Abstract

The entry discusses the strategy of designing and implementing public policies intended to promote public interests. It begins by recognizing that public interests are complex,

multifaceted, ambiguous and implemented in weak selection environments. However, theories from strategic management, originally formulated to deal with private-sector issues, can illuminate the trade-offs associated with different policy options, the relationships between organization and behaviour within public agencies, and the potential gains from policies that align the interests of diverse individuals and groups. It reflects on historical lessons from different disciplines – in business, military and politics – and suggests areas for future research.

Definition The public interest is frequently described in terms such as 'common well-being', 'general welfare', 'the common good' and 'sustainable shared values'. We focus in this entry on public interests that are widely acknowledged as privately worthwhile but not fully aligned with the individual private interests of each member of the relevant community. Strategy, in this context, is the management of organizational trade-offs in pursuit of the public interest.

Scholarly research on public policy and public administration is being revitalized by applying strategic management theories, initially developed to study profit-maximizing firms, to organizations pursuing public interests (Mahoney and McGahan 2007; Kivleniece and Quelin 2012). In this entry, we briefly review several areas of inquiry in this domain and suggest avenues for further research.

Antecedents

The field of strategic management has roots in several fields related to public interests, such as the operation of the military. Sun Tzu's *Art of War*, Machiavelli's *The Prince*, Clausewitz's *On War* and Schelling's *Strategy of Conflict* are classic texts offering insights for private managers from political and military situations – for ill or good. The related management fields of logistics, procurement and operations research have military origins as well (Fortun and Schweber 1993).

Yet despite these roots, research in strategic management within business schools has focused primarily on private companies, mostly large corporations, defining its domain as the determinants of superior company performance. Until recently, the predominant approaches in the field of strategic management model strategic decision makers as private actors pursuing private goals. Agency theory, transaction cost economics, property rights approaches and capabilities theories consider alternative units of analysis and objectives, but each generally assumes that organizational goals are private, straightforward and measurable.

The idea that pursuing private interests also generates public benefits goes back at least to Bernard Mandeville (1714) and is famously associated with classical liberal thinkers such as Adam Smith (1776), Hayek (1948), and Friedman (1972). Arrow's (1950) 'impossibility theorem' suggests that a singular 'public interest' is impossible to identify except under specific conditions. Without a clear objective, the strategic pursuit of public goals appears to be the province of politics rather than strategic management or ▶ organization theory.

However, as the strategic management has increasingly focused on challenges of resource trade-offs, innovation dynamics and the management of knowledge, the field turned towards intra-organizational measures of performance not uniquely tied to financial profit (Perry 1988; Stewart 2004; Rynes and Shapiro 2005). Similarly, extensions of sociological and macro-organizational approaches to the internal dynamics of organizations do not depend specifically on the goals of organizations per se. Thus, the insights regarding organizational dynamics developed on corporations were readily applicable to organizations such as non-profit firms, government agencies, philanthropies and non-governmental organizations (NGOs) that explicitly pursue public, social or non-market objectives. At the same time, the demand for insight on the effective stewardship of public resources was compounded: global challenges such as climate change and persistent, desperate poverty combined with fiscal challenges in health care, education and public safety generated

unprecedented need for insights on the effective deployment of scarce public resources. As these two forces develop together, the momentum for research in this area is significantly increased.

Concepts

Defining public interests is a central topic in the field of political science. Strategic management research recognizes that public interests are complex, multifaceted, ambiguous and implemented in weak selection environments. Two major approaches are prevalent. The first is to take a particular definition of public interest as given and analyse how strategic relationships develop in pursuit of the interest (Ferlie 1992; Andrews et al. 2011). The second is to discuss the impact of various public policies on strategy without any reference to a particular public interest (Mair and Marti 2006; Klein et al. 2010).

Strategic management theory illuminates both organizational dynamics and the impact of policy. Agency problems, property rights and transaction costs are as relevant to public organizations as to those pursuing private interests. Agency issues, a core issue in political science, are equally central to strategic management. Transaction cost economics can illuminate the efficient boundaries of public organizations and how these boundaries evolve in response to shifts in environmental conditions, including the costs of fundamental resources. The resource-based view, capabilities theories and evolutionary approaches provide insight into the existence, sustainability and appropriability of public organizations in the deployment of co-specialized assets. Extensions from behavioural and Penrosean analyses generate important insights on tacit capabilities, organizational routines, absorptive capacity, diversification and organizational growth. We have gone as far as to argue that the application of theories from strategic management to public sector problems will elevate scholarship at the core of the field of management.

Besides applying mature theories, constructs and approaches from strategic management to the public sector and the public interest, strategy scholars are also devising new ways to understand the public–private interface, helping to advance strategic management theory itself. Relationships between private and public actors and organizations can be characterized as a governance design problem (Baum and McGahan 2012; Kivleniece and Quelin 2012). The co-evolution of public and private capabilities is a primary area of research, and promises to elucidate and expose important ideas such as the benefits and costs of subcontracting (Cabral et al. 2010, 2013).

Frontier

Frontier issues related to this topic include those already mentioned: theoretical insights on organizational dynamics, governance and boundaries; empirical questions regarding how public and private organizations co-evolve; and studies of the interaction between public and private organizations.

Another important area of study, now drawing attention from strategic management scholars, deals with innovative organizational mechanisms for reconciling private and public interests, not only for 'mundane' issues such as financial market innovation (Faerman et al. 2001), but also for more vexing social problems such as climate change, epidemic infectious disease and persistent destitute poverty. Measuring and managing spillover benefits and costs, aggregating heterogeneous interests and conceptualizing investment opportunities across the commons is emerging as an important area of both theoretical and empirical enquiry (Kramer and Porter 2011; Klein et al. 2012).

Finally, the field of strategic management is beginning to conceptualize normative recommendations for public policy. One early area of research in this domain deals with the implications of public–private interaction for regulation, regulatory capture and innovation (Klein et al. 2012). How narrowly should public agencies' mandates be defined? What types of interventions are necessary to resolve the economic crisis in the European Union? Which specific types of investments should occur, and in what

order? How can multilateral agencies, governments and corporations develop capabilities in anticipation of future public and private needs, and then organize optimally to fulfill them? These questions lie at the heart of this burgeoning area in the field of strategic management.

See Also

- ▶ Organization Theory
- ▶ Penrose, Edith T. (1914–1996)
- ▶ Strategic Decision-Making
- ▶ Strategic Implementation
- ▶ Strategic Objectives

References

Andrews, R., G.A. Boyne, J. Law, and R.M. Walker. 2011. Strategy implementation and public service performance. *Administration & Society* 43: 643–671.

Arrow, K.J. 1950. A difficulty in the concept of social welfare. *Journal of Political Economy* 58: 328–346.

Baum, J.A.C., and A.M. McGahan. 2012. *Outsourcing war: Private military companies and command-and-control capabilities after the Cold War*. Working paper, University of Toronto.

Cabral, S., P. Furquim, and S. Lazzarini. 2010. Private operation with public supervision: Evidence from hybrid modes of governance in prisons. *Public Choice* 145: 281–293.

Cabral, S., P. Furquim, and S. Lazzarini. 2013. Private entrepreneurs in public services: A longitudinal examination in outsourcing and statization of prisons. *Strategic Entepreneurship Journal* 7: 6–25.

Faerman, S.R., D.P. McCaffrey, and D.M. Van Slyke. 2001. Understanding inter-organizational cooperation: Public–private collaboration in regulating financial market innovation. *Organization Science* 12: 372–388.

Ferlie, E. 1992. The creation and evolution of quasi markets in the public sector: A problem for strategic management. *Strategic Management Journal* 13: 79–97.

Fortun, M., and S. Schweber. 1993. Scientists and the legacy of World War II: The case of operations research (OR). *Social Studies of Science* 23: 595–642.

Friedman, M. 1972. *Capitalism and freedom*. Chicago: University of Chicago Press.

Hayek, F.A. 1948. *Individualism and economic order*. Chicago: University of Chicago Press.

Kivleniece, I., and B.V. Quelin. 2012. Creating and capturing value in public–private ties: A private actor's perspective. *Academy of Management Review* 37: 272–299.

Klein, P.G., J.T. Mahoney, A.M. McGahan, and C. Pitelis. 2010. Toward a theory of public entrepreneurship. *European Management Review* 7: 1–15.

Klein, P.G., J.T. Mahoney, A.M. McGahan, and C. Pitelis. 2012. A property rights approach for a stakeholder theory of the firm. *Strategic Organization* 10: 304–315.

Kramer, M., and M.E. Porter. 2011. Creating shared value. *Harvard Business Review*, 2–17.

Mahoney, J.T., and A.M. McGahan. 2007. The field of strategic management within the evolving science of strategic organization. *Strategic Organization* 5: 79–99.

Mair, J., and I. Marti. 2006. Social entrepreneurship research. *Journal of World Business* 41: 36–44.

Mandeville, B. 1714. *The fable of the bees: Or, private vices, public benefits*. London: J. Roberts.

Perry, J.L. 1988. The public–private distinction in organization theory. *Academy of Management Review* 13: 182–201.

Rynes, S.L., and D.L. Shapiro. 2005. Public policy and the public interest: What if we mattered more? *Academy of Management Journal* 48: 925–927.

Smith, Adam. 1776. *An inquiry into the nature and causes of the wealth of nations*. London: W. Strahan and T. Cadell.

Stewart, J. 2004. The meaning of strategy in the public sector. *Australian Journal of Public Administration* 63: 16–21.