

---

## DEBATES

---

### *Theory and Business History: New Approaches to Institutional Change*

Dianne Hutchinson

University of Melbourne

Stephen Nicholas

University of New South Wales

and Australian National University

In a typology of business history, Barry Supple identified the case study, entrepreneurship in the performance of the economy and comparative business structures as the three interrelated approaches employed by business historians.<sup>1</sup> Comparative business structures, pioneered by Alfred Chandler in a succession of books and articles on the rise of 'big business' in America,<sup>2</sup> is currently the dominant approach to business history. Chandler charted a pattern of large-firm growth in terms of entrepreneurial strategies of vertical integration and diversification which induced changes in the firm's internal

S. NICHOLAS gratefully acknowledges the Research Fellowship provided by The Research School of the Social Sciences, ANU during 1985 which allowed him to complete this paper.

<sup>1</sup> BARRY SUPPLE, "A Framework for British Business History" in B. SUPPLE, ed. *Essays in British Business History* (Oxford, 1977), p. 1-4.

<sup>2</sup> A. D. CHANDLER, "The Beginnings of Big Business in American Industry", *Business History Review*, XXXIII (1959); *Strategy and Structure. Chapters in the History of Industrial Enterprise* (Cambridge, Mass., 1962); "Management Decentralisation: An Historical Analysis" *Bus. His. Rev.* XXX (1956); "The Structure of American Industry in the Twentieth Century: A Historical Overview", *ibid.*, XLIII (1968); "The Development of Modern Management Structure in the U.S. and U.K." in L. HANNAH, ed., *Management Strategy and Business Development: An Historical and Comparative Study* (London, 1976); *The Visible Hand: The Managerial Revolution in American Business* (Cambridge, Mass., 1977). D. S. CHANNON *The Structure and Strategy of British Enterprise* (London, Macmillan, 1975); W. J. BURT, *The Australian Company* (London, Croom Helm, 1981).

structure. Within the comparative business structures approach, the first explicit theory of vertical integration was the transaction cost model utilised by Leslie Hannah in *The Rise of Corporate Economy*.<sup>3</sup> Transaction costs are the costs of gaining information, negotiating, monitoring and enforcing contracts within market-price institutions subject to risk and uncertainty. Hannah argued that 'the replacement of the invisible hand of the market by ...the more conscious integration and organisation of economic activities within the large firm' could be explained by '... costs of using the price mechanism — transactions costs — on which a firm can economise' and 'the distinguishing mark of the firm... is the supersession of the price mechanism'.<sup>4</sup> Subsequently, Chandler proposed in the *The Visible Hand* that '... modern multinunit business enterprise replaced small traditional enterprise where administrative co-ordination permitted greater productivity, lower costs and higher profits than co-ordination by the market mechanism'.<sup>5</sup> The advantage of administrative co-ordination in the firm, according to Chandler, was that by routinising transactions between units, the costs of these transactions were lowered.<sup>6</sup> Transaction costs, together with Chandlerian strategy-structure, form the 'internalisation' model which comprises the theory of comparative business structures. Given transactions costs, the hierarchical firm exists as an institution for economising on the market by internalising market functions.

The idea of internalising market functions in the hierarchical firm has become a major organising component in the theory of vertical integration since Ronald Coase's classic 1937 article on the nature of the firm.<sup>7</sup> Given externalities<sup>8</sup> in production or consumption, Coase argued that firms would exist and 'tend to expand until the costs of organising an extra transaction within the firm becomes equal to the costs of carrying out the same transaction by means of an exchange in the open market'.<sup>9</sup> The emphasis on transaction cost externalities as the reason for internalising market functions in firms has been strengthened by the concern that technological nonseparabilities<sup>10</sup>

<sup>3</sup> L. HANNAH, *The Rise of the Corporate Economy* (London, 1976), p. 3; "Introduction: Business Development and Economic Structure in Britain Since 1880" in Hannah, ed., *Management Strategy*, p. 3.

<sup>4</sup> HANNAH, *Rise*, p. 3.

<sup>5</sup> CHANDLER, *Visible Hand*, p. 6.

<sup>6</sup> *ibid.*, p. 7.

<sup>7</sup> R. COASE, "The Nature of the Firm", *Economica*, 4 (1937).

<sup>8</sup> Externalities are side-effects of production or consumption which confer a benefit or a cost on other producers or consumers for which no payment or compensation is made.

<sup>9</sup> COASE, "Firm", p. 395.

<sup>10</sup> Technological nonseparabilities are technological interdependencies such as successive processes which follow immediately in time and place.

by themselves do not account for vertical integration in production.<sup>11</sup> Econometric tests of technologically optimal firm size by John James and Jeremy Attack suggest that production economies by themselves do not explain the size and concentration in American manufacturing at the end of the nineteenth century.<sup>12</sup> Moreover, economies of scale embedded in the firm's production function could not account for the firm's ability to internalise additional functions such as marketing. Not surprisingly, business historians found the internalisation approach a powerful tool for explaining large firm growth.

However, the internalisation approach has not been without its critics. There has been a tendency to justify the hierarchical firm on efficiency grounds by the simple expedient of hypothesising high enough transaction costs in the marketplace and low enough costs of internal firm administration.<sup>13</sup> Making a similar criticism, Oliver Williamson admitted that the Coasean model tended to be tautological and that the transaction cost approach had not been operationalised, nor was it obvious how to do so.<sup>14</sup> Further, the market-hierarchy internalisation paradigm is static. It fails to explain the transition from one mode (external or internal) to the other. A related problem is the existence of intermediate institutions including agents, long-term supply contracts, licences and franchises which are neither arm's length markets nor hierarchical firms. The market-hierarchy internalisation approach fails to integrate these institutional arrangements into the growth of the large multiunit/multiproduct firm. While economists have recently modified the theory of the firm to include these institutional arrangements,<sup>15</sup> business historians have been slow to utilise this new institutional microeconomic work. Employing these new theories of the firm would allow business

<sup>11</sup> O. WILLIAMSON, "The Vertical Integration of Production: Market Failure Considerations", *American Economic Review*, LXI (1971), pp. 112-3.

<sup>12</sup> JOHN JAMES, "Structural Change in American Manufacturing", *Journal of Economic History*, XLIII (1983), pp. 450-1; J. ATTACK, "Industrial Structure and the Emergence of the Modern Business Industrial Corporation", *Explorations in Economic History* 22 (1985) 29-52.

<sup>13</sup> S. CHEUNG, "The Fable of the Bees: An Economic Investigation", *Journal of Law and Economics*, XVIII (1975), p. 33; A. ALCHIAN and H. DEMSEY, "Production Information and Economic Organisation" *A.E.R.*, 62 (1972), p. 783.

<sup>14</sup> O. E. WILLIAMSON, "The Modern Corporations: Origins, Evolution and Attributes", *Journal of Economic Literature*, XIX (1981), p. 1546.

<sup>15</sup> The new microtheory literature on contracting and transacting is referenced in O. WILLIAMSON, "Credible Commitments: Using Hostage to Support Exchange", *A.E.R.*, 73 (1983), pp. 520-1; WILLIAMSON, "Modern Corporation"; R. CAVES, "Corporate Strategy and Structure", *Journal of Economic Literature*, XVIII (1980).

historians to fulfil their time honoured role as testers of economic theory.<sup>16</sup> More importantly, the theory would allow historians to provide a better understanding of the growth and development of large firms and their effect on the economy.

This paper develops the recent theoretical work on alternative institutional modes as structures for mediating transactions. Specifically, intermediate modes are integrated into the analysis of the rise of multiunit, multiproduct firms, extending and modifying the market-hierarchy dichotomy of the internalisation approach. In analysing vertical integration as a process of large-firm growth, Section I develops the theory for the reasons firms transact in the market while Section II analyses the choice of the mode of transacting. The model is shown to be dynamic in Section III. Section IV generalises the model to explain product and geographical diversification, while the concluding section suggests a research agenda.

## I

Goods require several successive stages of production before they are placed in the hands of the final consumer. At the simplest level, these stages include the manufacturing, transportation and wholesaling of raw materials, the manufacture of the final product, and the distribution, marketing and retailing of the product. While vertically integrated firms combine two or more stages of production and distribution, separate stages can also be carried out by independent firms. As George Stigler pointed out as early as 1951, the cost functions of any two stages may be complementary, rival or independent.<sup>17</sup> Complementarity exists where the sum of the two cost functions of independently performed stages of production is greater than the costs of performing both stages jointly. The standard example of technological complementarities between successive stages is heat savings captured by vertically integrating iron and steel-making. In 1975 Kenneth Arrow identified informational complementarities between successive stages as a further important factor in vertical integration.<sup>18</sup> For example, in distributing goods firms develop knowledge of market demand allowing the more efficient scheduling of production. Alternatively, in producing a good the firm acquires knowledge about input requirements leading to more efficient purchasing of

<sup>16</sup> For a similar point see H. WILLIAMSON, "Business History and Economic History", *Journal of Economic History*, XXVI (1966), p. 415.

<sup>17</sup> G. STIGLER, "The Division of Labour is Limited by the Extent of the Market", *Journal of Political Economy*, 59 (1951), pp. 185-193.

<sup>18</sup> K. ARROW, "Vertical Integration and Communication", *Bell Journal of Economics*, 6 (1975), pp. 173-83.

raw materials. These informational complementarities between successive stages of production have not been fully recognised by historians as factors rivalling technological complementarities in explaining vertical integration. Below, we spell out how these informational complementarities force vertical integration.

All knowledge-related complementarities are firm-specific assets. Knowledge which yields R&D advantages, while the best-known firm-specific asset, is unlikely to be the only important type of knowledge owned by the firm. Knowledge related to distribution, marketing, product quality and differentiation, input requirements and scheduling are important but largely neglected forms of firm-specific assets. Firm-specific assets yield a return when the firm either sells the asset in return outright to another party or uses the asset itself to make and sell goods and services. Outright sale denies the firm further use of the asset for a lump-sum payment. When the firm uses the asset itself, for example knowledge related to pre-sale assembly, credit financing or after-sale service of durable goods, the firm earns a quasi-rent in the form of superior profits. It is this case with which we are concerned.

Knowledge advantages display public-good characteristics. They can be utilised by other firms, in alternative non-competing applications, without substantially reducing the value of their original application. This characteristic gives rise to the appropriability problem.<sup>19</sup> By appropriability we mean the ability of a private firm which is the originator of firm-created knowledge to obtain for itself the pecuniary value (rent) of the asset to society. In short, appropriability depends on the ability of firms to exercise property rights over knowledge. The ability to earn quasi-rents to knowledge is affected by the transactions costs related to the recognition, disclosure and embodied nature of knowledge. Recognition costs include costs of searching for second parties with whom to trade knowledge, negotiating the terms and specifying the contract covering knowledge use and finally monitoring the agreement, to ensure the terms are observed. The disclosure problem is related to second-party use. Before knowledge can be sold, it must be disclosed for the other party to ascertain its value. Once disclosed the other party has acquired the knowledge without cost and has no incentive to compensate the seller of the knowledge. Finally, the knowledge may be embedded in human capital requiring the transfer of individuals to the second party before the knowledge

<sup>19</sup> This section relies on STEPHEN MAGEE, "Information and the Multinational Corporation: An Appropriability Theory of Direct Foreign Investment" in J. N. BHAGWATI, ed., *The New International Economic Order* (Cambridge, Mass., 1977), pp. 318-21; H. G. JOHNSON, "The Efficiency and Welfare Implications of the International Corporation" in C. P. KINDLEBERGER, ed., *The International Corporation* (Cambridge, Mass., 1970); D. TEECE, "Economics of Scope and the Scope of the Enterprise", *Journal of Economic Behaviour and Organization*, 1 (1980).

can be fully utilised. Thus transacting in knowledge is both costly and risky for the firm.

While knowledge may be generated as a by-product of the firm's operations, firms only allocate scarce resources to create knowledge when they can solve the appropriability problem and capture the quasi-rents to the knowledge through exclusive use. The State recognises the appropriability problem by administering a patent system which gives legal guarantees to exclusive property rights in firm-created knowledge. However, the patent system reduces, but does not eliminate, the costs of policing exclusive property rights to knowledge. Many types of knowledge, including some types of technical knowledge, product differentiation, input requirement information, marketing know-how or specialised management systems are not patentable. In these cases, the costs of enforcing property rights to firm-created knowledge is borne exclusively by the firm. For instance, when a firm generates knowledge on the most effective method of demonstrating to consumers a durable product, it is difficult to sell this knowledge to independent distributors. Vertical integration into distribution would then be the least costly way of earning quasi-rents from the knowledge. Conversely, independent distributors may generate knowledge related to the design, production and sale of the firm's products. Transaction costs may mean that the firm has virtually to integrate forward into retailing to gain the knowledge derived by distributing.

## II

Business historians employing the internalisation model have depicted the firm and the price-market system as the two alternative institutions for economising on transaction costs. However, over a decade ago G.R. Richardson broke the market-hierarchy dichotomy by rejecting the idea that firms were islands of planned co-ordination in a sea of market relations.<sup>20</sup> Richardson argued that the firm and the market are at opposite poles of a continuum of alternative institutional arrangements. The middle ground is occupied by range of intermediate modes, including agency agreements, long-term supply contracts, licences and franchises, whose distinguishing feature is a contract to govern behaviour of at least one party. These intermediate modes are clearly differentiated both from allocation by the market-price system and allocation by fiat in the hierarchical firm. Market contracting through the price system does not involve the promise of future conduct and is subject to potential risks of price changes, supply availability, demand shortfall and depreciated quality of service. The firm is a mode of transacting where em-

<sup>20</sup> G. R. RICHARDSON, "The Organisation of Industry", *Economic Journal*, 82 (1972).

employees contract to follow the directions of management within limits. Intermediate modes are contracts between independent firms involving assurances of future conduct relating to price and quality and quantity of intermediate or final goods and services.

The alternative institutional approach which recognises a range of transactional modes is not simply of theoretical interest. The growth of large vertically integrated firms was not accomplished by a sudden shift from the market to the firm. Intermediate modes, such as long-term supply contracts, agencies or licensing agreements, formed important phases in the growth process of integrated firms. For example, Alfred Chandler's detailed case studies of vertical integration in American industry recognise intermediate arrangements such as agents, licensing and franchising. However, constrained by his reliance on the internalisation approach, Chandler provides no detailed explanation of the shift from these intermediate modes to vertical integration. More serious is the neglect by British business historians of the existence of intermediate arrangements, and the subsequent failure to discuss the replacement of agents and franchisees by depots and branch offices. Chandler expressed the received view that '[British] family firms, however, rarely set up a network of branch offices as did their American counterparts... Only in the 1920's did Cadbury, Crosse and Blackwell, and others begin to build a network of warehouses and depots'.<sup>21</sup> The available historical evidence does not support the received view.

British firms utilised a range of intermediate modes for distribution and retailing. Archival evidence points to networks of agents, warehouses, depots and branch offices as a precursor to forward integration, in much the same pattern as occurred in the American economy. The agency system was widespread in the engineering, food, textile, metal product and chemical industries before 1900. Like their American counterparts, British agricultural engineers, including Bentall, Marshall, Ransomes and Fowler employed agents on commission to assemble, stock, display and repair farm machinery.<sup>22</sup> In the interwar period Gourcock Rope Company had a network of 319 agents including farmers, machinery owners, agricultural engineers, corn factors and saddlers, to sell twine.<sup>23</sup> The Linen Thread Company, Peak Frean, Pilkingtons, Huntley and Palmer, Courtaulds and Reckitts<sup>24</sup> controlled an agency distribution

<sup>21</sup> A. CHANDLER, 'The Growth of the Transnational Industrial Firm in the United States and the United Kingdom: A Comparative Analysis', *Economic History Review*, XXXIII (1980), p. 401.

<sup>22</sup> Agency contracts deposited at Institute of Agricultural History, University of Reading, Ransomes, *Agency Books* AD7/48-49; MARSHALL, *Agency Term Books*, No. 103; Fowler Agencies C61/9 and Essex Record Office, Bentall Agreements, DF1/16.

<sup>23</sup> University of Glasgow Archives, Gourcock Rope Company, UGD42/34/3 *Home Organisation*.

<sup>24</sup> University of Glasgow Archives, Linen Thread Company. Executive Commit-

system in the XIXth century. These agency networks were supplemented and then replaced by warehouses, depots and branch offices. By 1895 Peak Frean, the London biscuit maker, had depots in Bristol, Leeds, Cardiff, Liverpool, Newcastle and Edinburgh - each controlled by managers with two clerks and supporting teams of home travellers and agents.<sup>25</sup> Rather than only beginning to develop its own distribution network after World War 1, Peak Frean expanded its depots to 30 by 1921.<sup>26</sup> In textiles, Alex Morton Sons and Co. had established showrooms at Carlisle, London, Manchester, Birmingham, Newcastle and Glasgow and offices in Manchester, London and Bradford before 1900 a practice followed by other textile firms including Lawthian Drake, William Jones, Coates and English Sewing Cotton.<sup>27</sup> By 1880 the large carpet makers had warehouses in the main provincial centres as a base for their commercial travellers and agents.<sup>28</sup> Gourock's 319 agents were serviced by warehouses in Liverpool and London by 1880 and in Lowestoft, Yarmouth, Plymouth and Hull by 1919.<sup>29</sup> Lever in soap, Wall Paper Manufactures, Borthwick in meat, Schweppes in mineral water and fruit drinks, W. & T. Avery in precision weighing machines and Lister in farm appliances blended depots, warehouses, branch offices with an agency network.<sup>30</sup>

The neglect of these intermediate arrangements is not unique to work on

tee Minutes, UGD143.1/5; University of Reading Archives, PEAK FREAN, *Depots*, PF26/1; D. C. COLEMAN, *Courtaulds: An Economic and Social History*, Vol. 1, (Oxford, 1969), p. 193; T. BARKER, "A Family Firm Becomes a Public Company: Changes at Pilkington Brothers Limited in the Inter-War Years" in L. HANNAH, ed., *Management Strategy*, p. 85; B. RECKITT, *The History of Reckitt & Sons Ltd.* (London, 1958); COLEMAN, *Courtaulds*, Vol. 1, p. 184; J. SALMON, "The Development and Organisation of J. LYONS & Co. Ltd.", LSE Seminar on Problems in Industrial Administration, no. 314, (1962/3), p. 6.

<sup>25</sup> PEAK FREAN, *Depots*, PF26/1.

<sup>26</sup> *ibid.*

<sup>27</sup> MORTON SUNDOUR Ltd., National Register of Archives for Scotland, GD326/409/2 *Report by Urawick, Orr and Partners; Across the Globe 1848-1848* (Manchester, 1948), no page numbers; *Cotton Waste* (Private c. 1922), no page numbers.

<sup>28</sup> J. NEVILLE BARTLETT, *Carpeting the Millions: The Growth of Britain's Carpet Industry*, (Edinburgh, 1978), p. 90.

<sup>29</sup> GOUROCK, *Home Organisation*, UGD42/34/3. Also G. BLAKE, *The Gourock* (private, 1963), p. 51.

<sup>30</sup> C. WILSON, *The History of Unilever*, Vol. 1 (London, 1954), p. 52; *WPM: The Pattern of a Great Organization*, (Manchester, 1949); G. HARRISON, *Borthwicks: A Century in the Meat Trade, 1863-1863*, (London, 1913), p. 97; F. C. HOOPER, "Development and Organisation of Schweppes Ltd", LSE Seminar on Problems in Industrial Administration, no. 120, (1952/3), p. 1; P. H. MILLS, "Development and Organisation of W. & T. Avery Ltd", LSE Seminar on Industrial Administration, No. 144, (1953/4), p. 3; D. E. EVANS, *Lister's: The First Hundred Years*. (Gloucester, 1979), p. 31

Britain. They are little discussed in Australian and European business history. Franchises, licenses, agents and other intermediate arrangements employed by firms to attenuate transaction costs are pre-vertical integration stages in the growth of the firm. By focusing on the costs of each arrangement, including markets and hierarchies, the alternative institutional analysis predicts when and where a particular mode of transacting is the most efficient way of organising economic activity.

All modes involve costs. Oliver Williamson has emphasised the costs of internal management between functional activities within the firm.<sup>31</sup> Such costs are attenuated by specific internal structures which establish information flows necessary to monitor and assess the performance of each of the stages carried out within the firm.<sup>32</sup> When choosing the market there are the costs of search and negotiations and, most important, costs due to the loss of any potential quasi-rent which cannot be appropriated in a pure price-market exchange. When an intermediate arrangement is chosen, specification of future behaviour allows quasi-rents to be appropriated. However, intermediate modes involve their own costs, including the costs of monitoring and enforcing the arrangement and opportunism.<sup>33</sup>

The costs of intermediate modes depend on the frequency of transacting. Discrete transactions, for example the sale of a machine to a firm with which one never again expects to transact, will not generate an intermediate mode. However, for occasional and recurrent transactions, firms frequently rely on a range of intermediate arrangements, including merchants, agents, travellers, franchises and licensing. Secondly, the nature of the transaction determines whether transaction-specific investments are required. These are investments in specialised physical and human capital, goodwill and brandname, which are unique to transacting in a particular firm's product and have low or zero salvage value. The return on transaction-specific investments depends, in

<sup>31</sup> O. WILLIAMSON, *Markets and Hierarchies: Analyses and Antitrust Implications*, (N. Y., 1975), p. 132-8.

<sup>32</sup> *ibid.*, pp. 140-54.

<sup>33</sup> The general theoretical literature includes O. WILLIAMSON, "Transaction-Cost Economics: The Governance of Contractual Relations", *J. Law & Econ.*, XXII (1979); B. KLEIN, "Transaction Cost Determinants of 'Unfair' Contractual Arrangements", *A.E.R.* 70 (1980); M. JENSEN and W. MECKLING, "Theory of the Firm: Managerial Behaviour, Agency Costs and Ownership Structure", *Journal of Financial Economics* 3 (1980); B. KLEIN and K. B. LEFFLER, "The Role of Market Forces in Assuring Contractual Performance", *Journal of Political Economy* 89 (1981); P. RUBIN, "The Theory of the Firm and the Structure of the Franchise Contract", *J. Law & Econ.*, XXI (1978); B. KLEIN, R. CRAWFORD and A. ALCHIAN, "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process", *ibid.*; O. WILLIAMSON, "Credible Commitments", pp. 519-40.

part, on the performance of the licensee, agent, merchant, or franchisee.<sup>34</sup> For example, the return to MacDonald's brandname depends on each franchisee supplying a standardised quality product. Any franchisee who fails to maintain product quality depreciates the returns to MacDonald's investment in brandname advertising, 'blowing' MacDonald's reputation. Generally, the more frequent the transactions between the partners and the more the nature of the transaction requires transaction-specific investments, the greater the transaction costs and the more highly specialised the mode of transacting.

Finally, intermediate modes are subject to opportunism.<sup>35</sup> In contractual relationships, opportunism is defined as the maximisation of one party's wealth at the expense of the other party. Opportunism arises in contractual arrangements entered into under uncertainty because knowledge of the other party is incomplete and it is costly to precisely specify behaviour in advance for all possible contingencies.<sup>36</sup> In addition, contractual performances are costly to meter, especially when it involves the input of human energy into task competition, such as working a sales area.<sup>37</sup> Both the costs of knowing and specifying agent, licensee or franchisee behaviour, and the costs of monitoring performance are borne by the firm. The greater the risk of opportunistic behaviour, which would reduce the wealth of the firm and the higher the costs of policing opportunism, the more costly the mode.

There are various methods for attenuating opportunism. Spot checks by inspectors on agents and franchisees and selling by commission are common monitoring and incentive practices. Contracts may require the agent to invest in transaction-specific capital, such as buildings or machine repair facilities, which 'lock' the agent into the product of a particular manufacturer. In the case of building and machine repair, the agent's transaction-specific investments are necessary for product marketing. However, transaction-specific investment such as signs and standardised buildings, for example, MacDonald's 'golden arches', mainly serve to 'bond' the franchisee to the principal. The withdrawal of business when opportunism is detected results in pecuniary loss not only from the withdrawal of future business but from the capital loss from transaction-specific investments tailored to the unique products of the other contractual party. Although firms attempt to reduce opportunism, no contractual arrangement totally eliminates the risk of such behaviour. The greater the firm's own investment in transaction-specific assets, the more vulnerable the firm is to opportunism and the more likely the firm is to re-

<sup>34</sup> KLEIN, "Vertical Integration", pp. 298-302; WILLIAMSON, "Transaction-Cost", p. 234.

<sup>35</sup> JENSEN and MECKLING, "Theory of the Firm", p. 308; KLEIN, "Vertical Integration", pp. 303-07; O. WILLIAMSON, *Markets and Hierarchies* Chapters 1-2.

<sup>36</sup> JENSEN and MECKLING, "Theory of the Firm", pp. 308-9.

<sup>37</sup> KLEIN, "Transaction Cost", p. 356.

place intermediate arrangements by vertical integration. For example, a firm would not build a retail outlet on land subject to a short-term rental contract. Oil refiners, particularly vulnerable to opportunism due to large transaction-specific investments, are more likely to vertically integrate into oil wells and pipelines than enter into contracts for oil and pipeline services.

### III

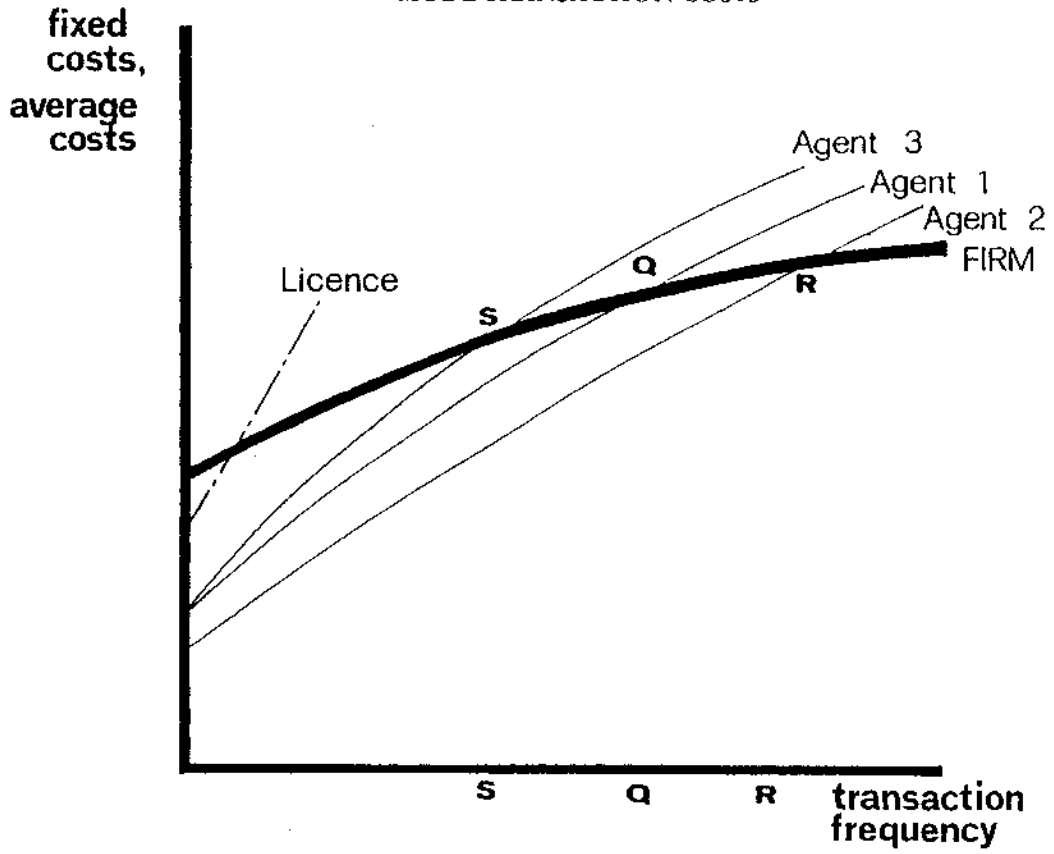
The alternative institutional approach is dynamic. The transition between modes depends on the transaction frequency, nature of the product, opportunism and mode learning.<sup>38</sup> While recognising that specific historical factors impinge on the process of vertical integration, the basic elements in mode transition are sketched using the example of a switch from intermediate modes to vertical integration within the hierarchical firm. In Figure 1 the transaction cost function for each mode is drawn so that the fixed costs of agents is lower than licences, while those of the firm are the greatest. The figure also assumes that total costs increase with the frequency of transacting. In Figure 1, agents are preferred to licensing at all frequencies of transacting, since the agency cost function lies below that for licensing. Further, the more frequent the transacting, *ceteris paribus*, the more likely a mode with a high fixed but low variable cost will be selected. For example, the firm is more efficient than agent 2 when the frequency of transacting is greater than R. In Figure 1, R marks the switch point between modes, where the transaction cost function for agent 2 and the firm intersect.

As the nature of the product becomes more complex, requiring assembly, demonstration and after-sale service, the fixed costs and the need for transaction-specific investments increase for all modes. Depending on how these costs affect each mode's transaction cost function, the change can advance (as shown by the shift from R to Q in Figure 1) or postpone the switch point between institutional arrangements. Third, the increase in specific investments related to the firm's product increases the potential for opportunistic behaviour. The increased costs of monitoring the agent and the increased cost of agent opportunism is shown in Figure 1 by the pivoting of the cost function for agent 1 to agent 3, which means a switch occurs at S, a lower transaction frequency.

Finally, through contracting with an intermediate mode the firm accumu-

<sup>38</sup> This section relies on P. BUCKLEY and M. CASSON, "The Optimal Timing of a Foreign Investment", *Economic Journal* 91 (1981), pp. 75-87 and S. NICHOLAS "Agency Contracts, Institutional Modes and the Transition to Foreign Direct Investment by British Manufacturing Multinationals Before 1835", *J. Econ. Hist.*, XLIII (1983), pp. 675-86.

Figure 1  
MODE TRANSACTION COSTS



lates a stock of knowledge on servicing a particular market. Market learning is mode specific, depending on the type of mode and the length of time the mode is utilised. For example, market familiarisation is more complete when using an intermediate arrangement where sales reports, inventory stock information and visits by the firm's employees form a feature of the mode. The effect of market learning pivots the total cost function of the firm downwards, replacing the agent by the firm at a lower transaction frequency. Clearly, the interaction of the four variables yield complex scenarios for the transition between alternative arrangements for transacting.

#### IV

These results for vertical integration can easily be extended to explain geographic and product diversification. Knowledge advantages have been at the centre of the explanation both of cross-border expansion by multinational firms and product diversification by domestic corporations. The theory of foreign direct investment posits that firm-specific advantages in technology, know-how, marketing, management skill, product differentiation and brand-name, yield quasi-rents when exploited through production in overseas markets. Frequently, these know-how advantages, labelled scope economies, can be employed in two or more geographic locations without reducing the pecuniary value of the advantage in its original application. For example, marketing knowledge related to one market can often be applied to other, different markets. In a world of risk and uncertainty, there are costs of gaining information, bargaining and negotiating and monitoring and enforcing agreements to utilise this know-how across markets, therefore, multinational firms choose contractual arrangements which economise on these transactions costs. This hypothesis is supported by the evidence that international firms pass through successive stages of increasingly sophisticated organisational form. For example, American and British multinationals first exported through merchant houses, then employed agents, next they established their own fully-owned sales branch before finally making an overseas plant investment.<sup>39</sup> These stages formed the typical path of overseas expansion historically, and today's multinationals follow much the same pattern.

Product diversification based on transaction costs and scope economies is congruent with the analysis of geographic diversification. For example, British Tobacco diversified into a range of snack food and drink industries in Australia from 1964 on the basis of scope economies related to its market seg-

<sup>39</sup> S. NICHOLAS, "British Multinational Investment before 1939", *Journal of European Economic History*, XI (1982), pp. 605-30; M. WILKINS, *The Maturing of Multinational Enterprise* (Cambridge, Mass., 1979), pp. 417-22; 432-7.

mentation expertise developed in cigarettes.<sup>40</sup> Firms with product scope economies select intermediate modes, including franchises and licensing which are alternatives to diversification, to economise on transaction costs.

Knowledge complementarities provide a basis for understanding vertical integration and diversification. But the new theory of the firm is in fact more general than the cases focused on here. It includes the related explanation of backward integration based on specialised assets, such as plant, which make the firm vulnerable to the risk of hold-up by raw material suppliers. For example, copper smelters are built for particular types of ore, effectively tying the smelter to a single mine. This rules out spot-market contracting, forcing smelters to enter very long-term contracts or vertically integrate into mining. As with knowledge complementarities, it is transaction costs related to hold-up which explain the use of alternative modes. Clearly the research agenda of business historians should include these more general applications of the transaction cost framework to the growth of firms. Below, we suggest further items on the research agenda to advance the application of the transaction cost model within the alternative institutional framework.

## V

American business historians have been better served by secondary studies of changing channels of distribution than their British, European or Australian counterparts who have few detailed case studies upon which to draw. One major exception in British business history is James Jeffreys' two books on British distribution channels.<sup>41</sup> Jeffreys' *Retail Trading in Britain* identified the shift in distribution away from wholesalers to direct manufacturer-retailer and manufacturer-consumer transacting modes by the large producer after 1870. While this shift was analysed only briefly, Jeffreys recognised that transaction-specific investments in buildings, branding and advertising, the frequency or volume of sales and opportunism in terms of the 'whims and fancies' of wholesalers were key variables in the shift.<sup>42</sup> Without detailed secondary sources, one item on the research agenda into alternative transactional modes is to integrate distribution channels, including agents, franchises, dealerships and independent wholesalers, into firm and industry business

<sup>40</sup> D. HUTCHINSON, *Corporation Diversification in the Australian Manufacturing Sector*, (unpublished PhD., University of New South Wales, 1986), Chapter 6; D. HUTCHINSON, "Diversification Performance and Scope Economies: British Tobacco (Australia) Ltd." (unpublished manuscript).

<sup>41</sup> J. JEFFREYS, *The Distribution of Consumer Goods* (Cambridge, 1950) and *Retail Trading in Britain 1850-1950* (Cambridge, 1954).

<sup>42</sup> *ibid.* pp. 9, 12, 47-48.

histories. A recent example is Roy Church's work on independent distributors and dealers as distribution channels for the British auto industry before 1939.<sup>43</sup> While largely a description of distribution channels, Church highlights changes in the modes of transacting as a neglected factor in the growth of British auto firms, identifying the weakening of independent distributors and the rise of a franchise dealer network as topics awaiting detailed analysis.

Business historians interested in alternative transactional modes also need to discover why some large firms integrated into retailing while others only vertically integrated into wholesaling. Before 1939 it was more likely that manufacturing firms took over wholesaling than selling directly to final consumers through their own sale offices. There is a need for comparative studies contrasting industries using intermediate arrangements between independent firms with those industries where vertical integration was characterised by the manufacturers' take-over of successive intermediate distribution channels.

Inter-country comparisons show that the degree of vertical integration was not simply a matter of the nature of the product. For example, British meatpackers, margarine makers and drug manufacturers integrated into retailing while their American counterparts did not.<sup>44</sup> The degree of vertical integration and the use of different transactional modes for the same product in different countries are closely related, but unexplored, research areas.

Finally, there has been a tendency to identify conflict between opportunistic parties as characterising intermediate arrangements. While conflict pervades some contracts between independent firms, most intermediate modes are based on co-operation. Co-operation reflects the relational aspects of agreements between contracting parties, which depend on social as well as economic factors. Trust and commitment between contracting parties are powerful forces reducing transaction costs, allowing contractual agreements to be modified for unforeseen events. A shared value system between partners also reduces the costs of monitoring contractual compliance. Social relations, then, are important in sustaining intermediate modes, and guaranteeing that the joint welfare of the two parties is greater than that which could be attained separately. The need to assess social factors, including religion, family ties and class, in co-operative modes promises a fruitful new approach to inter-firm contracting.

<sup>43</sup> R. CHURCH, "The Marketing of Automobiles in Britain and the United States Before 1939", in A. OKOCHI and K. SHOMOKAWA, eds. *Development of Mass Marketing*, (Tokyo, 1981).

<sup>44</sup> This problem was pointed out to the authors by A. D. CHANDLER.

