Changing boundaries and structure of a technological system: lessons from UK retail banking

Davide Consoli*

Abstract. This article investigates the factors that have induced and shaped the process of industry evolution of banking in the United Kingdom and, in particular, the reorganization of the retail payments system. It will look at how the effects of technical progress within a changing regulatory framework have contributed to the flourishing of new consumer services, of increasingly specialized technologies and of new models of business organization. In relation to these issues, the paper develops an interpretative framework based on the rapidly expanding body of literature on technological systems. In so doing it argues also that the organization of the payment system has evolved towards a multilayered and increasingly heterogeneous industry in which competition has been fuelled at different levels by the growing diversity of the ecology of agents involved, as well as by the emerging patterns of interaction across them.


* ESRC Centre for Research on Innovation and Competition (CRIC), University of Manchester
Harold Hankins Building, Booth Street West; Manchester – M13 9QH (United Kingdom)
Tel. +44 161 275 7372; Fax. +44 161 275 7361. E-mail: davide.consoli@man.ac.uk
Preamble

The production and delivery of payment services is an information-intensive activity that has benefited enormously from the adoption and the development of Information and Communication Technologies (ICT) over the years. In *prima facie* such technologies entailed a progressively superior ability to harness information through the expansion of storage and transmission capacity. Further developments along this technological trajectory have brought about drastic structural and operational changes that impinged upon a number of complementary dimensions, including the technologies in place, the variety of agents in the industry, the extent and forms of competition and the patterns of consumption of retail banking services. Recent research on the topic (Llewellyn, 1996, 1999; Consoli 2005a, 2005b) has concluded that technological progress combined with the emergence of new forms of business organization paved the way to a number of new technological paradigms and trajectories (see Dosi, 1982) in retail banking. After three decades of turmoil, no one today could fail to notice that UK retail banking has changed radically and that its pattern of evolution delineates a process of transformation for the industry at large, but also for the organization of the array of underpinning activities.

The observation that technologies have played a crucial role in this process of transformation surely won’t come as a surprise to anyone. Nevertheless the extent of such an influence has been captured only on surface within the relevant literature. The literature on networks, for example, has been very affluent in highlighting the strategic consequences of joining a network of shared Automated Teller Machines (ATMs) and credit card schemes. But very little is said on how the changes in the technologies affect the environment. The business literature, also, has put forth a life-cycle approach to study the unfolding of the commercial fate of technologies, with no attention whatsoever at the changes that these stimulate in the competitive setting. In brief, following these approaches excludes a priori an understanding of the mutual interaction between firms and their environment. The objective of this paper is to address this grey area in the literature, by elaborating a conceptual framework that looks at banking as a technological system. We will cast this analysis in the context of the transition from paper-based to electronic-based transactions in UK banking. Understanding the way in which enhanced elaborating and storing capacity has revolutionized the banking industry requires an in-depth analysis of the transformations that have been necessary to accommodate general purpose technologies. In turn, these reflect the ways in which the growth of technological knowledge has paved the way to the observed patterns of specialization and the related customer service diversification. Finally, we posit that the emergent division of labour is a direct result of the redistribution of technological knowledge across various interdependent layers of the system.

The paper is structured as follows. Section one will spell out the key elements of interest in the process of structural change and innovation observed in UK retail banking. Section two will operate the conceptual shift and elaborate the notion of banking as a technological system. Under such a new perspective it will be possible in Section three to discuss the unfolding of innovation across layers, emerging as a direct effect of the evolving focus of competition in the banking industry. Here we will appreciate the forces that contributed in shaping the boundaries and the structure of the industry, including the emergence of specialized agents, the evolution of the various business models and organisational forms, the development of new products and services.
1. Introducing the play: banking without banks

Payment systems are at the core of the circulation of money in any capitalist economy. Their development rests upon the institution of a progressively complex nexus of interactions channeled through multiple layers of communication. Money, the vehicle of all commercial exchanges, is a carrier of information regarding the purchasing power of the holder. Ever since its ‘official’ appearance in the UK back in the XVII century, when precious metals were converted into promissory notes, the changing forms of money over time – from fiat money to the more recent use of electronic signals – embodied the changing array of social and institutional arrangements that regulate commercial exchanges (Moini, 2001). The creation of depository institutions is another key moment for the organization of retail payments as banks provide the channels within which such transactions are carried out. Banks ensure customers access over time and space to their wealth and in so doing they connect different agents to facilitate the transfer of value within an economy. In sum, the core business of modern retail banking revolves around the following activities (Llewellyn, 1996; 1999):

(1) Dealing with customers either directly or – as it is increasingly the case – indirectly by setting up dedicated channels for the provision of automated transactions;
(2) Processing and storing information;
(3) Monitoring and authorizing clearance and settlements.

Let us now try to put these notions in a more specific context. The British credit market has been traditionally crystallized by banks’ collusive practices for years. Regulatory changes like Competition and Credit Control in 1971 and the increasing pressure imposed by American banks in the domestic market compelled British financial institutions to embark on an unusual ride in search of new areas of profitability (Moran, 1984; Mullineux, 1987). In this perspective investing in retail banking seemed an attractive and relatively low-risk prospect. Payment systems had been traditionally organized around banks’ physical outlets, and, accordingly, the strategic choices of financial institutions were tied to the expansion of their branches on the national territory to capture as many customers as possible. In the wake of the opportunities opened up by information technologies payment services were forced out of the branches and, in a span of few years, electronic funds and information transfer emerged as the paradigm that modernized the organization of payments. As a consequence after the mid-1970s the automated procedures moved from the back- to the front- office, boosting the number as well as the variety of customer services. The example of the ATM is but a precursor of the changes that were about to take place (Child and Loveridge, 1990; Fincham et al, 1994; Buckle and Thompson, 1998; Bâtitz-Lazo and Wood, 2000).

Given such an impressive transformation of the front-office dimension it is not surprising that the evolution of banking has been perceived as the progress of front-end technologies. During the 1970s electronic transactions were intended as a cheap(er) alternative to speed up common transactions and save on labour cost. Within any busy branch, the cash machines were just considered another piece of machinery which appeared on accounting books together with the likes of lighting, heating and cashiers’ registers (Howells and Hine, 1993). Less than twenty years after the introduction of the revolutionary cash machines, however, the banking system experienced the mushrooming of a great variety of alternative payment services, all incarnating and reinforcing the common purpose of ‘taking banking outside the banks’, that is, away from the traditional brick-and-mortar premises. This transformation stimulated by the institution of electronic-based transactions can be summarized in three
correlated dimensions of service provision, namely *automated machines*, *plastic cards* and *remote access banking*. The bottom part of Figure 1 provides a sketch of such changes in aggregate, which is also referred to as the “unbundling” of retail banking services.

FIGURE ONE ABOUT HERE

What is observed on the surface, however, is the reflection of a wider transformation in the organization of the payment system. The process of innovation involves the growth of knowledge, the institution of appropriate communication channels and, *de facto*, a change of behaviour (Nelson and Winter, 1982; Nelson, 1995; Metcalfe, 1998; Antonelli, 2001). New machinery is no more than the interface of a very articulated system that over time has grown, developed and changed both in the components as well as in the structure of the interactions. To be able to compete in their evolving landscape financial institutions had to learn how to cooperate in the process of building up the technical system that would have hosted their electronic clearing and settlement services. Accordingly, as new technologies emerged, new practices needed to be developed to govern the growing diversity of a system in which different types of specialization came to interact more intensely and frequently. Until fairly recently the conventional image of a bank was that of a vertically-integrated firm in which no distinction could be made between the manufacture and the delivery of financial services, as these were all carried out between the back-office and the front-office departments. Today these two complementary functions can be kept strongly separated (Llewellyn, 1999). The evolutionary mechanism at work allowed the propagation of the effects of increased specialization throughout the structure of the value chain, involving also the organization of modern financial institutions, which have progressed through different business models (Watkins, 1998; Consoli, 2005b).

The value chain, once entirely managed within the boundaries of vertically-integrated financial institutions, has been progressively broken down in a myriad of specialized activities and, correspondingly, of new types of firms and technologies. In the upstream market we observe the cross-entry of specialized firms, such as network providers and processing units, originally involved in previously unrelated industries. Boundaries have been crossed also in the downstream market where the ethos of wide accessibility (e.g. anywhere/anytime) to banking services has brought about the expansion of front-end devices such as new ATMs, EFTPOS, Internet banking, financial kiosks and, very recently, also mobile phones. As it happens, entrants from previously marginally-related industries are now qualified to compete in this disintegrated environment, thus forcing the existing players to react strategically by changing the organization and the nature of their business (Watkins, 1998; Llewellyn, 1999). In the early 1990s in the aftermath of a radical liberalization of the banking and finance sector established UK financial institutions have been increasingly challenged by an unprecedented surfacing of competition throughout the whole value chain of retail payments. Correspondingly, a number of non-traditional firms such as supermarkets and telecom providers are successfully capturing a significant share of the small payments market, as illustrated in the bottom part of Figure 1. The corollary of this pattern of progress is that the traditionally central role of financial institutions in the payment system has been effectively eroded. The consequences of this turmoil are far-reaching, to the point that, as recently suggested by David Llewellyn (1999) banking is irreversibly drifting away from banks as we have traditionally known them.

Briefly re-capping, the organization of retail payments in the UK has undergone significant changes during the last two decades: not only the volume of such transactions has grown, but also the ways in which these can be carried out have expanded thanks to the possibilities opened up by the development of Information and Communication Technologies. Today
financial institutions act as coordinators of a range of specialized suppliers. In so doing they assemble together the content, the format and the delivery mechanism of a piece of information that for the most part is processed and stored extra muros. It goes without saying that such an effort involves an increasing division of labour across growing forms of specialization and, subsequently, the coordination of different types of firms and, thus, of knowledge. We will refer to this process as the dynamic coordination in the evolution of the value chain of retail banking services.\(^1\) With this background in mind we can now move on to overview the main elements of our story: the development of retail banking access and how this fed back onto the organization of the payment system for both the firms involved and the structure of the industry.

2. A system perspective

The current understanding of competition in retail banking rests largely on the economics of networks (Cruickshank, 2000; Hunt, 2003; Saloner and Shepard, 1995; Pennings and Harianto, 1992; Shy, 2001), a framework which – we argue – may soon turn out to be inadequate. Indeed, despite having opened the way to a rising acknowledgement of the role of increasing returns to scale and of coordination across interacting firms, we argue that the network approach is essentially flawed by a twofold methodological bias. First, technologies are assumed to be given and constant, in clear contrast with the flow of service and process innovations that have characterized the banking industry in the last three decades. But ignoring the way in which technologies have changed over time conjures up a static environment which deprives the analysis of a plausible explanation for the processes that have excited the observed pattern of industrial and organizational change. Second, according to the theory the growth of a network corresponds bi-univocally to an increase in the number of participants, while the interlinkages across them are assumed qualitatively equal and, notably, unchanging. This analytical exercise may well suit the static industrial setting of British banking pre-1980s, when Building Societies and other competitors had not surfaced as a threat for the incumbents. But once we factor in the co-evolution between technologies and business models as a major inducement for the growth in variety in the ecology of the system, the network approach provides an arguably reductive picture. In particular, it falls short of accounting for the factors have elicited the emergence, both upstream and downstream, of a whole new nexus of firms in the now separated dimensions of production and supply of banking services.

Also the business literature has ridden the topic at large focussing mostly on the life-cycle of the technologies and very little about the ways in which the evolution of the technology feeds back onto the structure of the industry (Nolan and Gibson, 1974; Galliers and Sutherland, 1991; Friedman, 1994; Watkins, 1998). Although there is broad evidence that technologies – here narrowly understood as physical artifacts – undergo cyclical phases of development, the extent of this exercise is arguably limited, particularly in consideration of the outlined effects of competition in banking. We conjecture that such a misconstruction is due to a wide extent to the lack of appreciation of the variegated array of processes that the development of a technology brings about. Borrowing Richard Nelson’s (2001 with Sampat; 2002 with K. Nelson) terminology we are more inclined to consider physical technologies as the first building block of a wider process in which a complementary and central role needs to be played by the definition of coordinated patterns of behaviour. When the two processes

\(^1\) As put aptly by Langlois (2004: 2) “… coordination is the problem of making sure that goods, services, inputs, outputs, capabilities, etc., find themselves in the right place at the right time – and that they stay in the right place at the right time”.

5
unfold synchronically, physical and social technologies contribute to the definition of the opportunities and the constraints that technical progress stimulates.

Given these premises, looking at banking as a system is an intuitively attractive step because it allows accounting for the process that led the structure of retail payments to become an ensemble of complementary types of specializations concerning technologies, business organization and consumption patterns. More generally, speaking of technological change in system language allows a multidimensional approach in which a variety of factors, internal and external to the firm, jointly contribute to the technological outcome. Advocating a system perspective makes it possible to spell out the patterns of interaction across three relevant dimensions: the external environment; the internal environment; and the interlinkages. In so doing we can concentrate on the continuous dialogue between the two complementary dimensions of technology: physical and social.

Let us first overview the conceptual elements of a system theory. Systems are made of components and of the connections across these: the organization of the latter defines a system’s structure and boundaries. Several scholars have pointed out with varying emphasis that large technical systems evolve according to rules that are contingent to the character of the institutions associated in a socio-economic environment (Hughes, 1983, 1987; Sahal, 1985; Freeman, 1995). In so doing the concept of a system is a useful benchmark to portray how economic action generates the stimuli to innovation and self-transformation. According to Carlsson and Stankiewicz, in particular, a technological system “consists of networks of agents interacting in a specific technology area under a particular institutional infrastructure for the purpose of generating, diffusing, and utilizing technology” (1995: 23). The concept of system here is employed with the caveat in mind that new physical technologies stimulate the implementation of new patterns of behaviour which, in turn, determine the ability of the components of a system to work together.

When a system encounters limitations due to the emergence of a reverse salient or to the re-definition of its purpose (see Rosenberg, 1976; Hughes, 1983) then it needs to evolve in that both the agents and – inevitably the interconnections across them – are forced to change. But the emergence of new varieties of behaviour entails also a re-configuration of both the relevant knowledge in minds of the individuals interacting in the system and of the process of communication across them (Metcalfe and Ramlogan, 2002). In the words of Boulding (1955), a re-configuration of the “structure” (see also Machlup, 1983 and Langlois, 2001). Seen this perspective, innovation is the process that alters the ensemble of activities that are carried out synchronically towards the purpose of a system. As economic interaction unfolds, such activities emerge, change and decline, and likewise their connections. In turn, the re-configuration of the whole rests on the degree of coordination that the agents in place across them. An important observation is that these features are not ex-ante properties but, rather, emergent characteristics of a system (Metcalfe, 2001). Indeed, the way in which they develop will determine the outcome of the system’s evolution. This is tantamount to say that no one can know in advance how the relative importance of a specific activity will change due to the unfolding of the patterns of technical, institutional and economic interaction.

2.1 Retail banking as a dynamic system

There are different ways to organize a technical system and to coordinate its subcomponents, and the latter cannot be detached from the contingency of the elements in place. In a dynamic system a change in the set of agents operating within entails a reconfiguration of the set of activities in place: as new ones emerge and old, unfit ones, are dismissed. This is certainly relevant to the case of the payment system, whose structure and boundaries have
been subject to radical transformations. To capture the main elements of this process we need to be able to correlate the changes that take place on surface, across access technologies, with the developments “within the wires”, that is, in the channels in which information is stored, organized and transmitted. More generally, we need to interpret the systematic development of three interconnected aspects: technologies, agents and industry structure (Nelson, 1996). Another brief historical excursus will assist this mission.

It is with the provision of customer services outside the physical banks’ premises in the mid-1970s that the paradigm of retail banking took a decisive diversion from its secular status-quo. As already outlined, until then the adoption of mainframe computers remained a rather private business for British financial institutions. As the prices of microchips fell and their capacity increased, banks begun to develop dedicated software for information processing and to play a stronger role in the design of their technologies. Soon, however, it became clear that the potentials of automation were rather restrained in the back-office of banks. I argued elsewhere that the ATM, in particular, has played a pivotal role in the evolution of the industry in consideration of the wide-encompassing cascade of interrelated developments that stimulated in various components of the British banking community (Consoli 2005b). Apart from the immediate and more localized developments, like the commercial application of cryptography to protect the data in the plastic cards, the ATM has represented the prototype for the automated interbank network to rationalize the intricate nexus of transactions across banks. Let us first focus on the how the UK ATM network has emerged and developed over the years using Figure 2 as a reference.

FIGURE TWO ABOUT HERE

In the early 1970s each bank used to manage the processing and clearing of payments internally and deployed own cash dispensers which could be used only by the bank’s customers (Fig. 2a). This was the phase in which technologies were developed for specific internal purposes and, thus, managed and designed on a localized basis. A few years later, in 1984, (Fig. 2b) we observe the outsourcing of interbank electronic clearing to a private company, CHAPS (Clearing House for Automated Payments) as a reflection of the emergent paradigm of managing transactions on shared platforms, a reasonably attractive undertaking for financial institutions who could foresee immediate cost-savings. At the time, however, the access channels were still proprietary and, thus, could be used only by own customers. Such restrictions became a drawback in terms of the bank’s ability to be attractive to customers and were partially overcome during the second half of the 1980s by means of strategic interbank alliances aimed at machine-sharing. By virtue of such agreements, (dotted line in Fig. 2c) customers of participating financial institutions were granted undifferentiated access to each others’ ATMs free of charge. Hence, in the figure we

---

2 Banks are represented by rectangles in Figure 3. In this stylized diagram we assume they carry out three basic functions: dealing with customers, processing information and clearing/settling transactions. This rationalization is referred to in the Figure by segmenting the banks/rectangles in three parts. As it happens in the figure, the parts are progressively reduced to indicate that some of these functions are outsourced.

3 The UK payment system is run by the Bank of England. The main UK payment systems are CHAPS, BACS and CREST. CHAPS is the dominant system for Sterling and Euro payments. BACS is the UK’s automated clearinghouse which provides retail payment services that include direct debit, credit and standing order. CREST is the system that settles the most payments in a given day. These are now jointly coordinated by APACS (Association for Payment Clearing Services) set up in 1985. In addition to those listed, APACS is also responsible for Cheque and Credit Clearing, Currency Clearings, LINK, Mastercard, Switch/Maestro, Visa. See http://www.apacs.org.uk
see that customers of Banks 1 and 2 could use shared ATMs 1 and 2, while Bank 3’s customers were restricted to proprietary machines, ATM 3.

The emergence of a unique ATM network is a relatively recent accomplishment in the United Kingdom, made possible by a collective interbank agreement which has instituted LINK Interchange Network Ltd as the transaction management company in charge for switching and settling shared ATM transactions. LINK was formed in 1985 and operated initially as a circuit for smaller building societies and medium-sized banks which were precluded strategic ATM sharing by bigger players. The growing role of such smaller institutions within the British financial system has acted as an inducement mechanism for the creation of a unique shared network that is now critical to the functioning of the UK payments infrastructure (Consoli, 2005a, 2005b). LINK is now owned by the 22 leading British financial institutions and among its 51 members, other than Banks and Building Societies, there are non-financial institutions such as Card companies and independent ATM deployers.4

Technically speaking LINK is a central switching hub across participating institutions hosts’ to which it connects through dedicated telecommunication links. LINK is also in charge for the settlement process and the management of information services. From the point of view of the system structure, the creation of a centralized access platform responds to a specific design strategy: achieve the overall coordination of both the technologies and the governance mechanisms of the ATM network. This particular class of solutions has been widely observed and discussed, with a varying extent, in various industries and is often referred to as the business of system integration, following the original definition by Rothwell (1992) (see e.g. Brusoni et al, 2001; Prencipe, 2001). Running the system through a centralized hub maximized the convenience of customers, who can use virtually any cash machine, also implied that the major financial institutions lose competitiveness due to territorial supremacy. This induced the pursuit of alternative strategies to capture and/or to keep customers with new products and services. Therefore the retail banking industry experienced two counteracting forces: the stabilizing implementation of shared platforms for automated transactions and the staggering differentiation of products and services which became the real competitive battlefield. Indeed, product differentiation has not been restricted to ATMs, which have progressively evolved towards the model of the multi-service point of sale enriched by various ancillary services (e.g. mobile phones’ top-up), but entailed also the creation of entirely new transaction channels. Inevitably, this stimulated a shift of the gravitational centre of retail payments.

Figure 3 provides a stylized sketch of this process. To put things in perspective, cash dispensers, ATM machines and every other device – including PCs and mobile phones – that allows carrying out payments are simply access points to the system, the expansion of which are indicated in the figure by means of the black arrows. Accordingly, their management entails a bundle of technologies and agreements across a variety of agents, from both inside and (increasingly) outside the ‘traditional’ sphere of the banking activity. Together with the shared ATM we observe the emergence of three ‘blocks’ of business types, corresponding to alternative ways of accessing the payment system:

- the Points of Sale (POS), whose access points are the terminals installed in retailer’s premises;

4 See http://www.link.co.uk
- **Home technologies** such as the telephone or Internet, whose access points are the PCs of customers which can connect either directly to their bank’s website or to the retailers’ website for online purchases;

- **Financial kiosks**, which are a hybrid form of access to the banking system as they operate both as Internet points as well as ATM with expanded functions.

**FIGURE THREE ABOUT HERE**

The implementation of this wide array of payment options, in turn, required a long and complicated process of social and economic interaction, that provides a good example to appreciate the increasing complexity of technological change when variety of the interacting components (and not only their number) drives the evolution of an industry. The story behind the emergence of the EFTPOS (Electronic Funds Transfer at the Point of Sale) network is an interesting case in point. The automation of funds transfer and data interchange took place while the principal incumbents of the UK banking system were losing their market dominance. At that time, however, indiscriminate investments in ITCs eschewed a detailed assessment in terms of unit costs, simply because financial institutions had not developed a method to account for them yet (Mullineux, 1987; Howells and Hine, 1993). As the EFTPOS seemed to embody the long-craved cashless society, the creation of the network was initially undertaken as a joint programme by the major British clearing banks. This process required consistent investments and a close collaboration within the banking community as well as, for the first time, between banks and retailers. The whole process however collapsed in both fronts. The attitude of some major British banker shifted from being cooperative to being competitive, culminating with the adoption of different cards schemes. In so doing the cost of the entire operation boosted and, more subtly, also the nature of EFTPOS changed from being a cost-reducing banking “product” to a utility for electronic transactions. Moreover, its effectiveness depended heavily on the collaboration of customers and retailers, with whom the rules of deployment and use of terminals had to be negotiated, at times on a case-by-case basis. This was the second pitfall. The solutions were often a compromise between financial institutions’ priorities (i.e. maintain security in the system through data encryption) and retailers’ suspicious attitude (i.e. who would have borne the high cost of deploying sophisticated terminals?) (Mullineux, 1987; Howell and Hine, 1993). In this case, UK banks became twofold victims of their own strategy: by setting competing card schemes on the EFTPOS network, they needed to capture retailers quickly to guarantee a wide availability of the services on the streets but, at the same time, in order to accommodate retailer’s requests they enabled a variety of mixed schemes, both on- and offline with strong consequences on the front of costs. If anything, the build-up of the EFTPOS network has been biased towards retailers’ needs. At this point, however, banks were irreversibly tied to the machine that was changing the nature of their core business – more and more based on electronic money transmission – but also grinding their traditional predominance in the market of money.

In fact, the involvement of banks in the supply of retail payment services has been further eroded during the 1990s. After the implementation of the competing card schemes, large

---

5 Barclays launched an unexpected Visa-branded card scheme, Connect, in 1987 which would operate outside of the Unified Card Scheme that had been discussed throughout the 1980s. This broke de facto the negotiations with the other major clearing banks and stimulated a strong competition on the newly emerging market of the debit card. Lloyd followed Barclays with a Visa debit card. The other banks Midland, NatWest and the Royal Bank of Scotland instead formed an alternative card scheme, Switch in 1988. See the full story of the competing schemes in Howells and Hine, 1993, pp. 30-32.
retailers realized that a great volume of money transmission was processed through their hardware. The EFTPOS terminals in fact transmit the transactions to the head offices through a proprietary network that passes these on to the bank for clearing. Such retailers were the ones that effectively enjoyed significant cost savings from the switch from paper-based to automated transmission. Moreover, supermarkets and large retailers soon realized that they had all the technology and the bargaining power in place to develop their own card schemes and retail services packages, thus becoming effectively competitors for financial institutions. Today there is a wide spectrum of different types of banking organizations, from the traditional fully integrated bank to the ‘virtual’ Internet-only banks through the vast and heterogeneous population of organizations that supply financial services and compete with the former either over the whole product range or, more likely, in a niche. This is represented in a stylized form in Figure 3 with the appearance of Processing Units (PU) as autonomous elements next to banks.6

The intersection of various types of specialization opens up new markets, or new niches within existing markets, and an ancillary activity like information processing becomes a self-standing business run by specialists in the management of information-intensive activities. This has excited the outsourcing of banking “core” activities at the other end of the value chain, in the upstream market, in recent years. Financial institutions have been traditionally hesitant to outsource their core business, due to the risk of losing customers to potential competitors, but rising volumes of information have compelled the transfer of some processing activities to specialized companies. While no British banking organization has opted for total outsourcing as yet, the partial outsourcing of IT-intensive processes managed by means of different types of contracts is becoming increasingly common. Today outsourcing is used as a means of reducing costs and/or achieving strategic aims. Its impact can spread across many business activities including information technology (e.g., applications development, programming, and coding), specific operations (e.g., some aspects of finance and accounting, back-office activities, processing and administration), and contract functions (e.g., call centres). Clearly, the potential for cost savings through outsourcing are linked to the existence of scale economies in a particular area of activity, or significant lower cost labour in another country. Recent examples of partial outsourcing in the UK are: Barclays Bank’s recent contract with Siemens concerning Business Process Outsourcing (BPO); Barclays and Lloyds’ partnership with Unisys to outsource their check clearing system; Bank of Scotland – recently merged with Halifax – outsourcing deal with IBM; Barclays Bank’s agreement with Teradata – a division of NCR Corporation – to merge its legacy databases onto a single platform in order to reduce licensing and supporting costs associated with maintaining multiple database platforms7.

The outlined process stimulates some reflections. As with the famous pins example of Adam Smith – re-elaborated in different contexts by Stigler (1951), Richardson (1972), Pavitt (1998), Metcalfe (2003), Langlois (2002) – the emergence of new opportunities stimulates the disaggregation of productive activities into refined and more specialized tasks. At the same time, this entails the need to coordinate the units in charge of a small portion of the overall set of activities within the value chain. In so doing the division of labour feeds back into the boundaries and the structure of the originating system, eventually changing it. In retail banking the growing involvement of new agents has taken its toll from financial institutions that have lost the power to impose their preferred technologies and organization of services to the industry. We learn from this historical excursus that the banking industry

---

6 R’s PU stands for Retailers’ Processing Units.

7 Sources: http://www.sbs.siemens.co.uk/; http://www.finextra.com/; http://www.itweek.co.uk/
has evolved from being a relatively enclosed community restricted to professionals to an intricate texture of commercial and strategic relationships shaped by collaboration but also by contrasting interests and the bargaining ability of the actors involved. We can thus pinpoint the moment in which retail banking lost its mono-dimensional network character (as it is intended in that literature) and acquired the traits of a system (as it is intended in the literature outlined here). This is coherent also with the observation that the technologies we are considering have been progressively embedded in a twofold process of development: the physical dimension, embodied in the engineering principles that make mechanical and electrical parts work together, and the social process that patterns human interactions (see Nelson and Sampat, 2001; Nelson, R. and Nelson, K., 2002). To capture the full extent of the process of structural change in banking we will now spell out the interplay among technologies, organizational forms and the changing needs of customers are the barebones of the process that has shaped (and still is) the dynamics of competition in the banking system.

3. Innovation in Retail Banking Services: Structural change across layers

The interpretative framework that we intend to develop here requires the effort of refining the meaning of competition. This process is commonly referred to a context in which firms strive for market shares either through product and/or process innovations or, in the orthodox tradition, by price/quantity adjustments. The type of competition we refer to is complementary to the former approach in that we think of firms as repositories of knowledge in the tradition of Penrose (1959) (see also Langlois and Robertson, 1993; Winter, 1993; Kogut and Zander, 1996; Spender, 1996), and of competition among them as the contraposition of different types of knowledge and, contextually, of different ways to organize labour and manage the set of activities through which the knowledge base finds practical applications. As said before, the evolution in a system is stimulated by the need to re-define the purpose of the set of internal activities, or the way in which these are organized. When preferences and production possibilities are not given but endogenous, competition drives specialization through the proliferation of new activities. Specialization through the division of labour is the mechanism through which the reconfiguration of knowledge in a business unit finds new practical applications with respect to the changing goal. But to compensate for the disaggregative force of the division of labour, competition need also to include a mechanism of selection and coordination. In a context where knowledge is distributed unevenly, the interaction of different types of agents involves also the effort to make their patterns of behaviour consistent with each other and with the purpose of the system. Therefore, competition is a relational concept: a change in the circumstances in place in a system entails the replenishment of the relevant economic activities. Arguably, this process concerns technologies, products, capabilities, preferences, organizational structures and routines.

Turning back to the banking business, the outlined process of transformation has impinged upon many levels of the structure of the industry as summarized in Figure 4, where we sketch the various layers and facets of competition.

FIGURE FOUR ABOUT HERE

Note that as the industry evolves, various layers of competition emerge along the unfolding of new technologies, new activities and new agents. The first layer is competition among

---

8 And in principle there is no reason why this rationale could not be applied to the dynamics of consumption, but this is beyond the objective of this paper (Langlois and Cosgel, 1998; Loasby, 2001; Consoli, 2004).

9 Fransman (2001) put forth the idea of ‘layers’ of competition in the evolution of the Telecommunication industry, namely competition among microtechnologies, networks, services and firms. Differently from him,
banks. The progressive demise of the oligopolistic setting in the British financial industry after the 1960s has revived the innovative efforts of banks which pointed towards the expansion of their retail activity. The increasing number of transactions moved the attention to the information capacity management and paved the way to the “creative” adoption of a number of emerging technologies (Antonelli, 2004). In turn, the adoption of Information and Communication Technologies (ICT) has induced the definition of new procedures, the emergence of a new division of labour and, eventually, changes in the nature and the quality of the provision of customer services.

In this perspective, the adoption of a technology moulds the second layer of competition, among activities within firms. The adoption of microprocessors provided a key impulse to establish new operational practices in the management of transactions and information. In this process firms try to align emergent opportunities and constraints and to adapt the organization of their activity around the new technology. This, in turn, implies that diffusion stimulates a change in adaptive firms and that, subsequently, the way in which the population of adopters changes determines the direction of the inventive efforts and, thus, of technological change (Metcalfe, 1981, Metcalfe & Cameron, 1988; Antonelli, 2004). This perspective opens the way to a multidimensional analysis of diffusion as a sequence of partially overlapping technological substitutions among complementary processes over time.

This brings us to the third layer of competition among services, that is, among different ways of providing access to the payment system. In presence of a bundle of coexisting (and competing) sub-systems, banks are no longer the focal point of the whole retail payment process. The transaction to electronic-based transactions has also entailed the adoption of new payment instruments. Cash and most notably cheques have been substituted by plastic cards and the Direct Debit. Major sources of differentiation have subsequently emerged within the paradigm of Credit and Debit cards, with the proliferation of additional services as well as of access mode to a changing retail landscape. In turn, technologies are progressively more embodied in multicomponent products and cooperation among firms has become the norm rather than the exception. In the context of banking this has been translated in a bifurcation: retail transactions have become a multilateral business characterized on the one hand by an increasing degree of specialization to provide services on standardized distribution channels and, on the other hand, by a stronger emphasis on marketing in search of differentiation to attract new customers and to retain existing ones in a starkly competitive environment.

Transformations of the retail financial market have altered the three-sided relationship among retailers, financial institutions and customers. According to the traditional scheme, retailers acted as acceptors of payment instruments issued by banks. This status quo was established during the first phase of the front-office revolution, where the retail chain supply was still fragmented and the provision of services through credit payment systems was pretty much in the hands of banks. Further technical advances and the subsequent emergence of debit cards and EFTPOS in the UK have challenged the conventional payment system, bringing on a new structure in which the financial institution can be partially, if not totally, excluded from retail transactions (Howells and Hine, 1993; Alexander and Colgate, 1998). This is certainly true in the case of those UK retailers who entered the plastic card payment market supply as well as offering low-cost accounts and developing their own facilities. The
prospective gains of such an undertaking are two-fold: the revenues due to the widespread physical location of retailers together with the elimination of the costs – for both customers and retailers – typically entailed by the use of cheques and bank-issued cards. Hence, two ties have been broken contemporarily: those linking the financial institutions with the customer and with the retailers. The systemic aspects of this transfer of roles are clear. The nexus of arrangements upon which retail transactions are built has been enriched by putting the act of purchasing at the core. The evolving structure altered the relative position of the players: banks are no longer the gateway to the intricate bundle of payments, as each agent in the new, fragmented value chain can offer access with its own facilities.

The consequences of this process are remarkable also under another point of view. The recent spurt of alternatives for micro-payments impinges on the circulation of money, whose physical use is allegedly vanishing (Moini, 2001). As transactions are increasingly carried out in the form of electronic signals, the connotation of money as a commodity is overshadowed by a wider perspective which – perhaps atypically – calls upon the Hayekian theme of coordination. *Money* is now understood as a standard that carries a combination of quantitative and qualitative information about the availability of some purchasing power. Thus, money is a convention that has been instituted to reduce the uncertainty of value exchange by creating a scale for the measurement and the widespread acceptation of some type of economic transactions. The concept of money as an instituted standard across interacting agents matches the aforementioned theme of the importance of coordination across a myriad of specialized economic activities. Conversely, *payment instruments* are the gateway to the ganglia underpinning the changing structure of the payment systems. In this fundamental sense, the role of financial services as system integrators emerges even more forcefully for they impinge upon the speed of value exchanges. The tangibility of money, as much as the intangibility of services, is a non-issue. Similarly, the evolution of electronic payment instruments is but the tip of the iceberg. The endeavour which emerges from this analysis puts the evolution of the conventions that regulate social relations and economic exchange at the centre. Innovation in retail banking, thus, is essentially rooted in the dynamics of social interaction, the starting point and the finish line of our historical overview.

3.1 The changing sources of increasing returns

Several studies on the dynamics of large technical systems have provided important insights on the role that increasing returns play in shaping the evolution of important industries such as the railway, the electric power. Arguably, there are two dimensions in which increasing returns are winnored to economic development. One is the quantitative aspect of growth in utilization capacity as the system expands. The other is a qualitative – perhaps less frequently discussed – feature related to the structural changes that the growth of the system generates. A system does never grow only in size, but also in structure. Indeed, growth in variety in a system brings about more heterogeneity in the form of new activities, new forms of knowledge and new modes of communication. This translates into relevant changes in the mode of operation of the agents involved: as the organization of the system evolves, both the objectives and the means to pursue them evolve. As seen above, the evolution of the banking system has been characterized by the progressive emergence of new areas of competition. This is coherent with the outlined idea that the growth of a system involves three correlated dimensions: physical expansion, density and variety – both of the agents operating within and of the nature of their interconnections. The introduction of a number of large communication networks in banking has reshaped the dynamic of the costs involved in the management of the payment system. True, as the business literature has it the whole process
would have not been possible without the various technological transitions and incremental innovations, but the viability of heavy investments in large network technologies depends ultimately on the moving source of competitive advantage that characterizes dynamic industries such as retail banking. With the caveat that automation in banking brings about cost advantages in the form of saving from labour costs, the plot thickens as process and product innovations feed back onto the firms’ organization, the areas of competition and, eventually, the industry structure. Let us now explore more in detail how the shifts in technologies, organization and competition outlined above can be read in terms of shifting sources of increasing returns.

**Economies of scale:** The effective provision of payment services rests upon the extension of the network and, accordingly, of the degree of access to this. In the first phase of automation culminated in the bulky cash dispensers first and the ATMs in the mid-1970s, banks replicated the winning strategy adopted in their first decisive move towards the consolidation of their domestic retail market: the larger – in terms of territorial extension – is the presence on the territory, the more likely it is to capture and/or retain customers. Such a rationale applied to the traditional brick-and-mortar premises until the end of the 1960s and even more so to the adoption of ATMs. The decision to embark in such high fixed-costs strategies was driven by the prospective cost advantages of investing in additional operating units and equipment on the territory to increase the load of their existing processing capacity. In this phase, vertical integration is the emergent form of business organization. As discussed above, the first ATMs debuted in the British streets in the form of small competing clusters of machines operated separately by each proprietary bank. The payment industry then settled in an oligopoly with undifferentiated products in which cost advantages stemmed from enlarged territorial presence accomplished through strategic machine-sharing.

**Economies of scope:** When the fixed costs of incremental product innovation fell, banks moved towards a model of differentiated supply of a bundle of complementary services accessible through the same shared platform. In this phase economies of scale matter in the transmission function (e.g. automated processing and clearing) but economies of scope dominate the distribution of services with a growing variety of close substitutes enriched by ancillary functions. In presence of economies of scope, the growth of the system consists in the growing heterogeneity of services, that is, satellite activities “plug in” the central hub of the automated clearing house as illustrated in Figure 3. The rationale of this strategy is to realize a volume of traffic that ensures capacity utilization and decreases the unit cost per service: when the customer uses the ATM to carry out more than one transaction (e.g. cash withdrawal and information) it is possible to realize further saving on labour costs. Economies of scope are concentrated in access technologies and can generate cost advantages so far as imitation is limited through specialization. From the end of the 1980s the clearing banks experienced the growing competitive pressure exerted by Building Societies and other financial organizations after the Building Societies Act and the Financial Services Act in 1986. The combination of these two threats – the former from within the enclosed circle of the banking oligopoly and the latter from the outside – stimulated specialization of access technologies through horizontal alliances with specialized external providers.

**Economies of system:** when external collaborations enter the picture, the ability of the financial institution need to change radically. While in presence of scale and scope banks had to deal with a structurally stable network in which size and range of products matter, when the influence exerted by specialized suppliers increases, the industry becomes a system, characterized by new species of agents. With such an industry configuration, financial
Institutions can realize cost advantages depending on the ability to implement organizational changes to coordinate the range of heterogeneous specialized suppliers to which core functions such as processing have been outsourced. More in general, economies of system emerge as a dominant characteristic of large technical systems in which control strategies require technical and organizational innovation (Davies, 1994; 1996). Recent literature on retail banking confirms indirectly this trend. Brady and Targett (1995) argue that the growing number of legacy systems and the management of a growing volume of customer information had put banks in front of the necessity to rationalize high fixed ICTs costs. Similarly, Chiesa and Manzini (1998) conclude that investments in ICT have become a strategic necessity. As they put it, in a dynamic competitive environment, such as banking, can be met only through continuity and coherence of skills and knowledge accumulation rather than product stability.

3.2 Discussion

In front of the growing variety of the agents operating within the retail payment system the interaction across agents, institutions and technologies defines the pattern of dynamic coordination that allows keeping their patterns of change consistent with each other. Put differently, organization is at the core of economic development and all the components of the system need to achieve and maintain a high degree of complementarity. This is the mechanism that we call dynamic coordination: as the system evolves and its constituent parts change, the new configuration will make them work together. But since new knowledge is likely to be dispersed, it is also not fully capable of living up to its potential if it is not properly coordinated. This, in turn, implies an explicit effort in communication based on codification to achieve a correlating understanding of the issue at hand (Metcalfe and Ramlogan, 2002). When every unit in a system is in charge of a particular activity and all such activities can be coordinated the growth of knowledge can be understood as a collective (Antonelli, 2001) or as a distributed (Andersen et al, 2000; Metcalfe and Ramlogan, 2002) process. Unfortunately such an idyllic result cannot be always taken for granted. The ways in which idiosyncratic forms of knowledge, which reside in humans’ minds, achieve a correlated understanding is subject to a good deal of noise. The engine of economic growth and development is the ability to correlate/de-correlate technological knowledge in presence of radical or incremental mutations of the competitive environment. The extent to which this can be done successfully in our picture is arguably endogenous to the system and is shaped by the interactions at work. In a system characterized by growing variety due to specialization, it is natural to think that the emergent mechanism of coordination has to change over time.

This last observation is our connecting link to the role of economies of system as a source of increasing returns for the banking industry. Various scholars have addressed these issues arguing that in several industries cooperation emerges as a defining feature of the process of competition (Penrose, 1959; Richardson, 1972; Antonelli & Foray, 1992; Metcalfe, 1992; Loasby, 1999). The accomplishment of coordination needs adjustments at several levels ranging from the arrangements of procedures to the implementation of new artefacts. Coordination as such requires complementarity at various levels: among members within one business unit through the institution of organisational rules that open the way to increased variety; among producers through the process of competition and the resulting selective forces which reduces variety; and ultimately, among the subcomponents of the economic system through the interplay between demand and supply to generate developmental feedback necessary for the replenishment of variety. The key aspect of the question is that firms are repositories of knowledge, which is not freely available and neither
can be traded as such (Nelson and Winter, 1982; Antonelli, 2001; Metcalfe 2002). In presence of a variety of knowledge it is likely that a variety of formal and informal coordination processes will regulate the various parties involved, thus making coordination a contingent process (Antonelli, 1999). This theme is not new at all for those who are familiar with the seminal contribution of Richardson (1972). In his own words:

“It is convenient to think of an industry as carrying out an indefinitely large numbers of activities related to the discovery and estimation of future wants, to research, to development and design, to the execution and co-ordination of processes of physical transformation (...). And we have to recognize that these activities have to be carried out by organizations with appropriate capabilities, or, in other words, with appropriate knowledge, experience and skills. (...) What concerns us here is the fact that organizations will tend to specialize in activities for which their capabilities offer some comparative advantages. (...) But the organization of industry has also to adapt itself to the fact that activities may be complementary.

Richardson, 1972, p.888

Thus the boundaries of the firm itself come to be subject to a great deal of reassessment. In these cases, in alternative to internal accumulation, the augmentation of capabilities rests on the possibility to outsource (Langlois & Robertson, 1993). On the one hand, specialization exerts a centrifugal effect resulting in increasing dispersion as knowledge grows. At the same time, however, the ability to harness the general principles of knowledge into economic activities elicits the coordination between new and existing knowledge. Here, centripetal forces are at work to uphold what Hayek called the “spontaneous interaction of a number of people, each possessing only bits of knowledge” (1945: 527). In this perspective competition is at the same time an engine of change and of stabilization for the system.

The micro-macro interlinkages within this pattern of industrial and economic evolution are crucial to our story. At micro level the demand and the supply of financial services has been subject to a process of growth and development. The differentiation of the provision and the demand of several types of services entailed significant changes in the organization of financial institutions as well as of customers, inducing cooperation among competitors on the one hand and substitution effects on the other hand (Consoli, 2005b).

**Concluding remarks**

This paper provided the opportunity to reflect on the necessary steps to be taken to proceed with a re-styling of the analysis of innovation in retail banking. It is clear that the introduction of new technologies combined with a changing regulatory framework broke down the barriers that characterized for years the British financial structure. Within the textures of this process we have highlighted the existence of various interrelated layers that contributed to the building-up of the modern retail payment system. In so doing we noted in particular that various forms of social institutions such as consumers’ changing habits, new regulatory settings, the collaboration of retailers and the increasing specialization of various components, all guided the rate and the direction of the shift towards the electronic transactions paradigm in the UK. As such, the development of the system carries the traits of all the myriad of interactions taking place over time across its changing components. This is why the emergence of this technical system has required more than new pieces of machinery to work together through standardizes interfaces. In a banking industry characterized by growing product diversification, vertical disintegration and cross-industry competition, the ultimate task has become to put in place various patterns of interaction and the mechanisms to coordinate them. The ability to choose within such a growing menu an appropriate mix of contractual agreements, investments in technology and in human capital is a strategic
labyrinth complicated by the fact that the growth in the number and the variety of agents in
the system brings about new modes of organizing the production and delivery process.

With this background in mind we have then provided historical and empirical evidence of
the corollary that the degree of involvement of banks within the payment systems has been
eroded. In light of the growing variety of technologies, customer services, business models
and coordination mechanisms we noted that the concept of technological system embodies
better what banking has become today and, conversely, why the inherited theoretical
approaches from the economics of networks and the business literature are somewhat
outdated. The paper places the evolution of retail banking in the literature of business
organization at large, with a hint of where it could be interesting to look at in the future. One
characteristic that seems to emerge with more clarity from our analysis is a progressive bias
towards modularity in the organization and in the technologies of retail banking. This
conclusion should be the starting point of further research to investigate what kind (or, more
precisely, what degree) of modularity is emerging in retail banking (See Wegberg, 2004). In
so doing it would be possible to spell out the dynamic relation between emergent
technologies (such as middleware technologies) and new forms of business organizations
(such as the growing variety of outsourcing contracts).

References:

relationships: a conceptual framework”, International Journal of Retail & Distribution
Management 26 (2), 225-237.

Economic Processes”, in J.S. Metcalfe and I. Miles (eds.), Innovation Systems in the
Service Economy, Kluwer Academic Publishers, Norwell, MA, USA.


University Press.

adoption”, Dipartimento di Economia “S. Cognetti de Martiis”, University of Turin,

Innovation and New Technology 2, 37-47.

commercial banking”, In: Lenk, K., Traunmueller, R., Schmid B. F., Pavlikova, L.


grail or poison chalice?”, Technology Analysis and Strategic Management 7(4), 387-
406.


Figure one

Traditional Banks

Financial Services

Branches

Cash Machines

ATM

EFTPOS

Remote

New ATM

Kiosks

Credit cards

Debit Cards

Online Banking

Phone Banking

Fringe of competitors including: building societies, supermarkets, virtual banks
Fig. 2: The Evolution of the ATM network

2.a – Cash machines, no network, in-house processing and clearing (1970s)

2.b – ATM, no network, in-house processing, outsourced clearing (post 1984)

2.c – Limited ATM sharing, strategic alliances, in-house processing, outsourced clearing (mid-1980s until mid-1990s)

2.d – Total ATM sharing, unique network, in-house processing, outsourced clearing (post 1996)
Fig. 3: The Evolution of the UK retail banking system
Figure 4

Within the industry: competition among firms

- New expertise (e.g. accommodating new technologies)
- "Creative" adoption of technology in firms
- Advancements in technology (e.g. Magnetic strip)

Within firms: coevolution of organization and technologies

1970s

1980s

1990s

New Regulation

Production of services

Consumer knowledge (e.g. learning how to use new services)

Within services: competition among modes of provision

Standardization e.g. shared platforms

Specialization and outsourcing

Within the new industry

Competition within networks: e.g. Internet vs. EFTPOS

Competition within niches: e.g. Retailers vs. Banks