

COMPETITION IN BANKING: SWITCHING COSTS AND THE LIMITS OF ANTITRUST ENFORCEMENT*

Donatella Porrini^α, Giovanni B. Ramello^β

1. Introduction

A recurring theme in the analysis of competition in the banking sector is the problem of stability, and the regulatory constraints that are consequently imposed on economic agents operating in this particular market.

Generally speaking, antitrust intervention in the banking is heavily influenced by considerations of stability, because although competitive processes are inherently selective, and presuppose the possible exit from the market of inefficient competitors, this is precisely the eventuality that economic policy decisions seek to avert. Therefore, as discussed more in detail in the paragraphs below, the regulation has historically given precedence to the stability objective, relegating competition to second place. This is borne out by the many structural and operational constraints imposed on the authorities and laws that ought to safeguard competition, and the elevation of administrative barriers to entry.

Now, under a law and economics perspective, regulatory intervention in the market is justified as a means of counteracting the emergence of inefficiencies, and so we can apply this same justification to the banking sector, where a specific inefficiency arises from the macroeconomic and systemic repercussions of the normal workings of the competitive process. The central problem, in this case, is entrepreneurial risk, which must necessarily exist in any competitive market, and plays a decisive role in ordaining the entry and exit of competitors. However in the specific case of banks, price competition tends to encourage overly speculative behaviours, which essentially entail

acceptance of excessive risk, with a resultant volatility that could potentially harm depositors, and ultimately compromise the stability of the economic system as a whole.

From the perspective of economic policy, this eventuality translates into a natural friction between stability and competition, that cannot always be overcome without penalising one of the two, and which in the particular sector under study is resolved at the expense of competition.

The consequence of this approach is that banking market becomes extremely rigid on the supply side and structurally not equipped for a competitive orientation, and banks come to occupy a privileged position vis-à-vis governments that--to a greater or lesser extent, depending on the countries and the situations--enables them to sidestep the antitrust authorities.

In such a scenario, the trade-off between stability and competition cannot be totally resolved through traditional antitrust actions, which are sometimes at odds with the stability objective and hampered by the constraints of the previously defined regulatory framework. On the other hand, the supply side approach which characterises a great deal of the scientific literature on competition policy appears unable to pursue one objective without penalising the other.

It is precisely these considerations, found in a significant portion of the literature, that provide the starting base for the hypothesis of this work--described more in depth in the second section of the paper--and namely the proposal of a novel demand side perspective, i.e. one which focuses on the central role of consumers in the competitive process. If intervention on the supply side is hampered *a priori* by the regulatory framework, it is nevertheless possible to implement pro-competition actions on the demand side, for example by enhancing the ability of consumers to change from one provider to the other without impacting on the market structure. In operational terms, the proposed approach is to leverage consumer mobility in order to stimulate the currently weakened competition between firms. This would make it possible to pursue the traditional antitrust objectives of efficiency and welfare maximisation, without necessarily impacting on stability.

In the sector under study there exists a simple solution for implementing such a strategy, which is to reduce the switching costs that currently restrict the mobility of consumers between different banks, with an obvious impact on competition. These policy implications will be presented in the closing sections.

2. Relationship between competition policy and the stability objective

If stability is a crucial objective common to all banking and financial systems, control of competition policy is a fundamental tool for pursuing such a goal. In fact, an unrestricted competitive mechanism, with the attendant risk of entry-exit of firms from the market, clearly leaves open the possibility of bank failures: the risk that inefficient firms might fail is in fact a necessary condition for the existence of a competitive market. However, in the particular case of banking, because of the ties generally linking this sector to the rest of industry, failure of one firm is liable to trigger a contagion or domino-effect, causing other banks and financial institutions to fail in their turn, and culminating in the collapse of the entire market, with very serious repercussions on the economic system as a whole, at both the national and, ever more frequently, international level¹.

Confirming the above is the fact that, in most industrialised countries, the stability objective was first enshrined in regulations issued in the aftermath of the 1930s as a reaction to the Great Depression, i.e. the most serious instance of market failure in modern history². Therefore, even though the national regulations for the safeguard of stability developed independently, under separate institutional frameworks and market regulation authorities, all systems nevertheless exhibit common traits. In fact, the various national frameworks embody certain shared assumptions that can be summarised as follows: 1) the risk of failure within the banking system, with the attendant danger of systemic repercussions on the economy as a whole, requires special treatment; 2) depositors, in their role as producers of savings, should be protected and 3) competition in this specific sector has the effect of increasing risk and must therefore be controlled.

Associated with this is the idea that guaranteeing market power to banks will help attenuate the risks to which they are exposed, thereby furthering the goal of stability, as expressed in the so-called theory of "charter value"³.

The stability objective thus becomes a real hindrance to competition policy and antitrust enforcement, due to the many special measures and exceptions that it causes to be imposed (though here again with national variants), for example with respect to mergers, supervision, etc.

In recent years the trend has been toward a partial restoration of market competition, while at the same time seeking to safeguard--or at least avoid significantly compromising--the stability objective⁴.

It thus remains an open question whether competition and stability are necessarily in opposition, or whether it is instead possible to pursue the one without necessarily compromising the other. The conventional result which emerges from the literature is that resolution of the conflict through regulation is generally the optimal solution. However the new industrial economics, as we discuss below, provides a means for reconciling this position with (at least) a partial safeguard of competition.

3. Regulation as a constraint on competition policy

Generally speaking, regulation is the practical mean by which competition is restricted and the stability objective pursued. But it can also sometimes be deployed to support competition, which places them in a sort of no-man's-land that effectively renews the dilemma of stability versus competition⁵. On the one hand, regulation restricts competition *ex-ante*, creating a safety net to guard against bank failures, as well as providing for discretionary *ex-post* interventions, through the monetary authorities, to rescue banks that run into difficulty. But on the other hand, regulation also seeks to uphold competition through *ad hoc* market interventions in specific situations, for the most part in the case of mergers. This is a source of further ambiguity, because in certain countries the activities of the central bank will then overlap with those of the antitrust authority, generating what is in essence a conflict, as will be clarified below⁶.

Each national situation is therefore characterised by its own set of regulatory measures, such as prohibitions on listing in the stock-exchange, limits on diversification into non-banking products, barriers to entry of foreign banks, restrictions on branching, controls on interest rate levels, and capital requirements.

However there is a particular distinction when it comes to the divide between depositors and investors. For the former, the most important thing is stability, whereas the latter are mainly interested in transparency and complete information. Therefore, the banking and financial markets have different needs which call for distinct regulatory frameworks. In this connection, we can identify two main lines of approach adopted by the various countries. At one end of the spectrum are nations whose regulatory

frameworks make a distinction between banking and financial market operators, such as the United States, where the Banking Law and Securities and Exchange Act were enacted almost simultaneously, but as completely separate legislations. And at the other end are countries such as Italy, where banks function both as collectors of savings in the form of deposits and as financial intermediaries for investors, characterised by a unified banking law strongly conditioned by the stability objective, which takes precedence over transparency.

But even within the banking market, regulatory actions can be further sub-classified as a function of the relevant market, identified through application of three criteria, of which two are borrowed from antitrust practice while one is specific to the sector under study: 1) the geographical market, which in this case is delimited by national and regional borders; 2) the product market, expressing the substitutability of products as perceived by consumers and, finally, 3) the nature of the market players, which distinguishes between saving banks, and private, cooperative or state-owned banks⁷.

With regard to the second criterion, the European Commission provides clear guidelines for identifying markets, at least for those relating to traditional banking products⁸:

- 1) Retail banking, which comprises the various groups of individual products and services that banks offer to households (current accounts, savings accounts, bonds, pension funds, short term and long term loans, mortgages)⁹
- 2) Corporate banking, which comprises services aimed at businesses (domestic corporate banking, public sector banking, international credits to public companies)¹⁰;
- 3) The financial market sector, which comprises services relating to government securities and capital markets (trading equities, bonds and derivatives, foreign exchange and money markets)¹¹

These three product categories are seen as giving rise to distinct markets, due to differences in both the composition of demand and the nature of the offerings, which are provided through different channels. And, according to antitrust practice, each individual product category can then be further broken down into a series of separate product markets--known as submarkets--corresponding for example to different categories of clients, distinguished according to the characteristics of their demand¹².

Naturally, in defining relevant markets it is also necessary to carefully consider the ongoing evolution of the Internet and electronic commerce, which may introduce new classifications. In fact, the future prospect of electronic banking, with no need for physical branch banks, would entail a relaxation of the geographical constraint and a general increase in competition. However this prospective development is not considered in the present work, nor in the cited literature, because it is still poorly defined, and in any case negligible compared with the other activities. In the current scenario, the uptake of e-banking is still hampered by a variety of constraints, arising from the need to obtain information about clients, the substantial advertising investments required to establish a new product, and the difficulty of stipulating contracts on-line¹³.

The three banking markets listed above have different attributes from a regulatory and competition-policy perspective. This work proposes to look more in depth at the retail market, i.e. the specific market of products offered by banking institutions to households, examining its characteristics from the competitive standpoint. The analysis will illustrate how, from a supply side perspective, the regulatory framework is a major obstacle to the development of policies for competition in this market. These limitations are discussed more in detail in the paragraphs that follow.

4. Competition in banking law and practice

Competition policy in banking is different from competition policy in any other market. This peculiarity of banking policy can be traced back to the special charters granted to banks, and has been perpetuated to this day by a continuing friction between the need to guarantee stability on the one hand, and to sustain entrepreneurial opportunity and productive efficiency on the other.

Bank competition policy arose simultaneously with the birth of the banking market itself, and therefore long predates the advent of national antitrust laws. However, after the 1930s such policies were everywhere affected by a growing tension between the competition objective and the need to protect the national economic system as a whole from the risk of bank failure.

As a result of this, the national banking laws placed very little emphasis on the safeguard of competition. In the US, this historical bias was corrected after World War

II, with the integration of antitrust and banking law, which marked a move toward upholding competition in banking after a period of anti-competitive restrictions imposed by the government¹⁴. In fact, even though the US introduced a general competition law very early on (the Sherman Act was enacted in 1890 and the Clayton Act in 1914), a specific competition law for banking dates back only to the 1960s (the Bank Holding Act was enacted in 1957, the Bank Merger Act in 1960).

Similarly, in the case of Europe: “For a long time, banking in many countries was exempted from the reach of competition law and subjected to regulation only. As a consequence of deregulation, this is different today. For example on the matter of mergers, banks in the EU are fully subjected to EU competition law. In this way, competition policy and regulatory interests can become intertwined and enter into conflict” (Canoy et al., 2001, p. 31).

However, since the 1980s, there has been a general reassessment of public policy in banking, with pro-competition policies emerging in the wake of structural, behavioural and technical developments, and changes to the banking law. In recent times, the banking market has felt the effects of novel trends, such as the relaxation of geographical constraints and the lifting of restrictions on the scope of activity, an unprecedented movement toward bank mergers, a sharp decline in the number of banking institutions, and the advent of secondary markets and new systems of payment--all of which presage a fundamental transformation in the industry.

Beginning in the 1980s, the European Commission and the national governments initiated a so-called "competitive deregulation" process, embodied in the Banking Directives (first 77/780 and second 89/646) as well as the Own Funds (89/299 and 91/633) and Solvency Ratio Directives (89/647 and 94/7). This opened the way to the abolition of restrictions within the European common market, under the principle of mutual recognition. At the same time, however, the Basle Committee on Banking Supervision began formulating a series of capital adequacy requirements, starting from the agreement of 1988.

Despite the recent thrust toward increased competitiveness, competition policy in banking remains atypical due to the issue of financial market stability. This subordination of competition to stability continues to dictate the adoption of supervision mechanisms for ascertaining and limiting the risks to which banks are exposed.

Generally speaking, in the banking sector national antitrust authorities are confronted not so much with cases of abuse of dominant position, but rather with cases of cartels and especially mergers, resulting from the concentration process that has in recent years characterised, and continues to strongly characterise, this sector¹⁵. Now, the specific case of mergers and acquisitions is an interesting example of the peculiar manner in which competition, and the authorities charged with upholding it, are dealt with in the banking sector.

As we know, merger and acquisition operations unite two formerly independent firms into a single entity, and are generally forbidden if they result in a dominant position that can in some way impair and/or restrict competition¹⁶. Or, alternatively, the operation may be authorised on condition that certain correctives are applied to neutralise its anti-competitive effects. However, in banking all these decisions, and the manner in which they are taken, are once again influenced by the implicit or explicit consideration of a trade-off between stability and competition.

There is, first of all, a widespread belief that larger sized banks, with the resultant attenuation of competition, can help protect against systemic instability, and this view to some extent prejudices the decisions. This position is also consistent with the theory of scale and scope economies, according to which organisational fixed costs imply economies of scale while the joint provision of deposit and credit activities can produce important scope economies¹⁷.

A further indication of the special status granted to this market is the peculiar attribution of competencies over competition to institutions different from the antitrust authorities. Despite differences between countries attributable to historical legacies not directly linked to the efficiency-stability paradigm, and to varying preferences accorded to the authorities responsible for bank competition and financial stability, we can nonetheless say that the banking market everywhere enjoys a privileged status.

Looking for example at the situation in certain European nations, it should first of all be borne in mind that, in the European Union, national institutions are competent to decide on mergers below 'community dimension', whereas the Merger Task Force of the European Commission (DG Competition) handles those which exceed this dimension¹⁸.

In France, the body responsible for reviewing bank mergers is the 'Comité des établissements de crédit et des entreprises d'investissement', one of the Committees in

charge of prudential supervision in the financial sector, headed by the Governor of the Banque de France. The criteria applied by the Comité are not those defined in general competition law, but rather those defined in banking law, where the stability objective strongly prevails over competition.

In Germany, the Federal Cartel Office is entitled to issue a first opinion on bank mergers, based principally on the application of competition law. However, before a merger can be blocked it is also necessary to have the approval of the Federal Supervisory Office, which examines the case from the perspective of banking law, and therefore also taking into account considerations of stability. Moreover, if the opinions of the Cartel and Supervisory offices are not in agreement, bank mergers can be submitted for political review to the Federal Minister for Economics, who issues a decision based upon macroeconomic and common welfare considerations.

In United Kingdom, bank mergers are reviewed through the application of competition law, in the same way as any industry. Only in cases of particular importance a report from the Office of Fair trading is also requested. So, although the Financial Services Authority and the Bank of England are consulted during the process, considerations of stability are comparatively less influential¹⁹.

In Italy, the primacy of the stability objective over competition has led to responsibility for competition policy being transferred from the antitrust authority (Autorità Garante della Concorrenza e del Mercato) to the banking market supervision authority, the Banca d'Italia, which also handles antitrust issues and therefore any conflicts arising from the application of the two. This is true across the board for mergers, while in general antitrust practice there is an occasional overlapping of competencies that is a feature peculiar to the Italian financial system²⁰. In fact, although Law no. 287 of 1990 which instituted the Autorità does include credit institutions under the scope of application of antitrust discipline, treating them in the same way as any other business, it also assigns enforcement of this discipline to the Banca d'Italia²¹.

In the case of bank mergers which exceed the community dimension, these are examined by the Merger Task Force of the European Commission, which applies the EC Merger Regulation, as is for any industry. However, in its process for examining mergers, the Commission requests all the necessary information from the competent national authorities, so that in countries such as France and Italy central banks can still

raise concerns over stability. Having said that, in its activities as a whole the Commission generally tends to emphasise competition aspects.

Ultimately, despite the fact that the community approach to mergers of community dimension is strongly competition-oriented in theory, there is no specific community-wide supervision authority, so that European nations retain a certain degree of discretion in practice, through the involvement of their national supervisors who look at mergers from a prudential perspective²².

Therefore, in the decision-making on mergers at both the national and community level, we once again detect a friction between stability and competition, and here again banking emerges as a peculiar market heavily influenced by monetary and financial policies, and generally by the views of central banks and other member-state authorities charged with financial stability.

5. Demand as a pro-competitive device: considerations drawn from the new industrial economics

The analysis so far has repeatedly highlighted the conflict between regulation considered necessary for assuring the stability of national and international economic systems, and antitrust laws, which instead appear necessary for assuring market efficiency. The recurring theme, therefore, is that regulation is the inefficient but necessary solution by which competition is sacrificed to serve the ‘greater good’ of economic stability. In the paragraphs that follow we shall propose an alternative means of resolving this dilemma, based on the new industrial economics.

The central idea of this alternative solution is encapsulated in the following general principle, which has been emerging in the industrial economics literature: when competition is limited for whatever reason on the supply side, it can be at least partially stimulated on the demand side, provided that sufficient consumer mobility can be induced²³. In certain industries, for example, competition may be restricted or impeded due to structural reasons. Within such contexts, antitrust enforcement is largely irrelevant because it is barred from removing the cause of inefficiency, and in any case can only be sporadically applied. However in certain situations this rigidity can be sidestepped, by approaching the analysis from the consumer's perspective, if there are elements that, for whatever reason, work on demand to confer *de facto* market power to

the firms. In such circumstances, these same elements can be leveraged to pursue a pro-competitive policy. However such actions lie outside the province of conventional antitrust practice, which generally looks only at the supply side, so that a novel approach is called for.

One situation where the above approach could be applied is the case of lock-ins, which prevent consumer mobility from one product to the other. This is an issue that has been widely debated in the technology sector, where incompatibility between different technical components can attribute market power to a particular firm. The precedent-setting case on this matter was that of photocopy machines and their spare parts²⁴, so that the literature in question generally makes reference to markets for durable goods and their related markets, which are termed ‘aftermarkets’²⁵. It is still possible, in some cases, for antitrust authorities to take direct action against firms that pursue allegedly abusive strategies and, in such circumstances, the remedies will be the traditional ones. But even when such conventional remedies are impracticable, competition can still be promoted through policies that enable consumers to choose alternative products.

Looking now specifically at the banking sector, we find certain significant affinities with the situation described above: this is in effect a market where repetitive consumption over time presents analogies with the previous representation, with the opening of an account or acquisition of a service constituting the primary market, while the continuation of this relationship in subsequent years generates the aftermarkets. In other words, for reasons connected with the cost of switching, consumers face a problem of compatibility between their initial purchase and successive purchases, and this locks them into the original provider, conferring market power to that firm.

In light of this, therefore, the existence of competition at an initial time t does not guarantee that competition will continue to exist in subsequent periods, with the paradoxical outcome that even a firm with no apparent market power when competing for new clients might nevertheless exert market power over its acquired consumers, who in a sense constitute a specific relevant market²⁶.

One likely explanation for the persistency of consumption is hence the existence of switching costs, whose origins can either be structural, i.e. relating to certain exogenous attributes of the sector, or strategic, i.e. endogenously determined by the firms in order to create market power. But irrespective of their cause, the important point is that these

costs lock consumers into their initial purchase, thereby conferring a certain degree of market power to the original provider.

In the case of exogenously determined switching costs, the level of market competition can be elevated through additional regulatory intervention, and this--at least in Europe--falls outside the remit of the antitrust legislation and authorities. In the case of switching costs that are endogenously determined, i.e. for the explicit purpose of acquiring market power, antitrust enforcement could theoretically be applied, though this proves difficult in practice because the judgement will inevitably have an arbitrary component, requiring application of a 'rule of reason' on a case by case basis²⁷. And such a judgement becomes even more difficult where there are *ex-ante* reasons for restricting antitrust actions, as is the case in the banking sector.

6. Demand-side restrictions on competition: switching costs

Promoting competition on the demand side should not be seen as a universal alternative applicable to every market structure, but rather as a remedy geared to specific markets where certain peculiar features characterise the relationship between consumers and firms: namely the existence of switching costs, which are exploited to secure market power over particular segments of demand, thereby restricting competition. The remaining sections shall first of all describe the nature of such costs and their workings in the market, before proceeding to identify the specific types of switching costs found in the banking industry. This with a view to formulating policy indications that are specific to that context, but extensible (and have in part already been applied, though not systematically) to other markets as well.

6.1. The nature of switching costs

Switching costs can emerge in markets that are characterised by repeated consumption. Within such markets, consumers who initially purchase a good or service from a firm will remain 'loyal' to that firm in order to avoid incurring these costs again at a later date. This has the effect of weakening the substitutability between (even identical) products, once the initial act of consumption has been made.

So even if price competition exists at an initial time t , when firms are attracting new consumers, this will tend to diminish in following periods, when consumers must

repeat the purchase, due to the effect of switching costs²⁸. We can therefore say, in this scenario, that at a time $t+n$ (where $n=1,2,..$) the existence of switching costs can transform the relationship between firm and consumer in the relevant market. This in itself clearly complicates the antitrust analysis, because the implied dynamic perspective creates more scope for ambiguity, unless--as we saw earlier--concrete reasons exist for attributing the emergence of switching costs to clearly abusive behaviours.

There is a vast body of literature in economic theory devoted to switching costs, in which their impact on competition is examined from a variety of starting assumptions, such as whether consumers do or do not have perfect information²⁹. Now, although no univocal conclusion exists, as a general principle (setting aside special cases in which switching costs do not compromise competition) we can say that such costs generally confer a certain amount of market power to firms, precisely by virtue of the exclusive tie which they create (or which is purposely created) between consumers and the firm. This is indirectly confirmed, for example, by the structural persistence of above-marginal cost pricing, even where there is apparent competition for the capture of new consumers, and by the emergence of multi-product firms not attributable to scope economies, but rather to forms of scale economies on the demand side, deriving precisely from a consumers' desire to amortise switching costs³⁰.

The balance of collective welfare which results is usually negative, because "switching costs generally raise prices and create deadweight losses of the usual kind [...] and may also discourage new entry and so further reduce the market competitiveness" (Klemperer, 1995, p. 536)³¹.

The conventional representation of switching costs is as a sunk investment (effective or perceived) made by consumers, which creates a certain inertia in the choice between apparently identical alternatives³².

Such an investment can be 'real', for example connected with issues of technical compatibility, as in the case of a decoder for a particular pay-TV system, which would need to be replaced if the consumer switched to a new provider. Or it can be 'perceived', for example connected with the effort expended by the user to research a particular purchase, which would have to be repeated to select an alternative, or the time invested in becoming proficient with a new system, which creates a path-dependence, as in the learning curves confronted by computer users. It can also take the form of transaction costs, in situations where migrating from one product or service to another incurs added

intermediation charges. And finally, there are also switching costs associated with the emotional sphere, when the user's familiarity with a particular product or service engenders a sort of affective dependency. This is in fact one of the mechanisms at the root of brand loyalty, and of course the effect can also be enhanced by artificial means, through commercial initiatives such as cumulative discount schemes (for example airline 'frequent flyer' programmes), or coupons promotions.

It should be emphasised that the existence of switching costs does not in itself imply behaviours in violation of the antitrust laws. Nevertheless, as mentioned previously, from the standpoint of firms switching costs do help create market power, and thus have a beneficial effect on profits, especially if price discrimination can be practiced between old and new consumers. The monopoly pockets thus created will naturally have the side-effect of rationing demand. The cost of switching might prompt certain individuals not to consume at all, with a distorting effect on the allocation of resources. And further distorting the market structure is the emergence of multi-product firms. In fact, if switching costs encourage consumers to source a variety of products or services from the same firm, by the same token they encourage firms to diversify into a variety of products and services, even inefficiently, exploiting the fact that consumers have already been 'captured' by their original purchase.

This mechanism can in its turn constitute a form of market foreclosure against potential competitors who are able to efficiently produce only a single product or service³³.

However, although it is fairly straightforward to describe the anti-competitive effect of switching costs, it is rather more difficult to identify behaviours that are clearly anti-competitive, and therefore punishable by the authorities. In other words, because switching costs can arise for so many different reasons, it can sometimes be difficult to determine their causes and judge whether they are explicitly anti-competitive. As a general principle, switching costs can be likened to a product differentiation strategy, so that although the underlying rent-seeking logic is quite clear, a univocal antitrust interpretation is more difficult: in the same way that a promotional campaign causes two objectively identical products to become differentiated in the eyes of consumers (we speak in this case of perceived differentiation), switching costs introduce an *ex-post* differentiation between products that were substitutable *ex-ante*.

Ultimately, the crux of the antitrust enforcement is the nature of the costs, which can be simultaneously exogenous, i.e. arising from structural features of the market and/or technology in question, or endogenous, i.e. arising from strategic policy decisions taken by firms to create market power, in much the same way as with product differentiation. It is only in rare cases that unequivocally abusive behaviours will emerge, although switching costs always create some amount of market power, and must thus have a correspondingly anti-competitive effect.

6.2. Switching costs and the banking industry

In the discussion so far we have already mentioned the peculiar structure of the banking market, characterised by considerable rigidity on the supply side--largely a result of the regulation--and by high switching costs--associated with repeated consumption and various endogenous and exogenous traits of this sector, such as multi-product supply, market power over consumers, etc.³⁴. For example, the discrepancy between the interest rates offered to savers and those earned by banks on their investments, which in a perfectly competitive market ought to coincide (minus the operating intermediation costs), is a sign of market power which can be accounted for by the existence of switching costs³⁵.

A growing body of literature supports the key role played by switching costs in the banking sector, although few studies have as yet attempted to define the magnitude of this effect, or to examine possible policy indications from an antitrust perspective³⁶. In any case, the costs in question are sufficiently high to discourage movement of consumers from one bank to another, even where there are large differences in charges and interest rates, making them a significant constituent of banks profits³⁷.

Switching costs appear in a variety of forms and can also be added together, thereby reinforcing their overall effect. The following paragraphs provide a brief, and by no means exhaustive, taxonomy specific to the banking sector. One category of switching costs are those relating to direct-debit payments from a customer's bank account for utilities and other bills, for example phone, gas, electricity, credit cards, and so forth. And of course a similar consideration applies to credit transfer operations into the account, such as the payment of salaries or other remunerations, dividends, and so forth. On closing the account the customer would have to transfer all the aforementioned credit and debit operations to his new account, thereby incurring significant transaction

costs³⁸. Further compounding these are costs relating to uncertainty, due to the possibility of oversights or errors occurring during the transfer³⁹.

The transfer times themselves are highly variable, and generally take longer than theoretically necessary, suggesting possible strategic behaviours pursued by the originator banks expressly to increase the cost of switching⁴⁰.

Surveys of Italian operators have found that banks explicitly take such costs into account, and generally advise their staff to activate as many services as possible when a customer opens an account⁴¹. A similar conclusion was reached in the UK by the Competition Commission (2001, sect. 1.6 and 1.8) which found that personal current accounts “the core product in personal banking [...] also serve as ‘gateway’ through which suppliers can sell other financial product, such as credit cards and personal loans [...]”, thus raising the overall switching costs⁴².

A similar trend exists in the US, where competition in banking is increasingly centred not so much on individual products as on the provision of bundles of services⁴³.

Then there are additional switching costs associated with the information that consumers need to collect when deciding to open an account, and with the fixed fee normally charged for opening the account: both these elements fall under the category of start-up costs, i.e. those costs that are unavoidably incurred when entering into a relationship with a bank, and which must be repeated whenever the individual starts again with a different supplier⁴⁴.

Finally there is a category of switching costs associated with the bank's direct knowledge of the customer and, conversely, the latter's familiarity with the bank. This makes many operations quicker and easier, because the bank staff are directly acquainted with customer and his history of solvency/insolvency and so forth. In the transfer to a new bank, all this accumulated background information will be lost. The bank customer, on his part, is likewise facilitated by familiarity with the procedures and staff, making it more convenient to continue with the existing bank rather than have to get used to a new one⁴⁵.

It is important to remember that, taken individually, each of the switching costs listed so far can be relatively small, and not bring much market power by itself. Nevertheless, through the cumulative effect of various costs, it is possible to achieve considerable market power in practice.

Some authors have pointed out that this mechanism is offset by an implicit trade-off in the exploitation of switching costs, because the exercise of market power must be weighed against competitiveness in attracting new consumers. However in this particular sector, as mentioned previously, such costs appear to have a significant impact on the profits of banks⁴⁶.

There is some fragmented evidence in support of this in the literature. For example, a study of the Norwegian mortgage market was found that 25% of the marginal profit of banks (i.e. the profit arising from each additional borrower) could essentially be ascribed to the lock-in effect, with the bank-client relationship lasting on average 13.5 years⁴⁷. This span of time is consistent with the values found in both the European and US markets⁴⁸. This confirms, at least in part, that the market power deriving from switching costs is being exploited, and does have an impact on competitive mechanisms and market outcomes. And many other signs further corroborate this assertion. For example, the phenomenon of diminishing interest rates on bank accounts--whereby higher rates are offered on opening the account and then gradually decreased--would seem consistent with the hypothesis that, as the relationship is prolonged in time, and the switching costs consequently increase, a lower remuneration is needed to retain the consumer.

A corollary to the above assertions is the empirical fact that the principal reason for which consumers switch banks is when they move house--in other words a drastic event that incurs very high overall transaction costs, which far exceed the switching costs themselves⁴⁹.

All this is not to suggest, however, that the banking industry is pervaded by systematically abusive behaviours. Such an interpretation is unjustified because the behaviours are to some extent structurally determined by the particular conditions in which competition is played out, which are largely defined *ex-lege*. In a sense, switching costs are built into the peculiar relationship between customer and bank, so that the market configuration makes them inevitable in the current context. What is more, the related observation--that firms strive to increase these costs to secure new pockets of monopoly--is also open to ambiguous interpretations, because this is in effect a competitive strategy in the particular market structure.

Irrespective of the market, any practice that shifts competition onto non-price elements has the aim of attenuating price competition, and in this sense must have a corresponding anti-competitive effect. In the banking industry such practices are a

rational and competitive course of actions for firms operating in a peculiar market structure. They constitute rent-seeking behaviours only to the same extent as does advertising or product differentiation, and so are not punishable on this count by the antitrust authority.

We note also that, when antitrust enforcement falls outside the price competition paradigm, and unless there is clear circumstantial evidence of infractions, it faces increasing difficulties because the analytical tools available are to some extent obsolete, and referred chiefly to static price-based competitive paradigms which are not very useful in complex contexts⁵⁰.

7. Policy indications

The conventional analysis framework has accustomed us to look at competition from the supply side and to approach its stimulation from the perspective of firms. In this paper we have instead shown how, in the banking market, competition on the supply side is institutionally restricted by the imperative of assuring the stability of the banking system, and the economic system as a whole.

This creates a significant problem for the application of antitrust law, because it interferes substantially with regulatory activity. For example, if we accept the antitrust approach which interprets competition laws as tools "for prohibiting bigness or facilitating ease of entry for small businesses" (Howenkamp, 1994, p. 275), the natural result is a profound ambiguity. The obvious solution would in fact be, as Justice Brandeis said in a celebrated US case, to punish practices that 'destroy' competition⁵¹. However in our particular case this route is precluded by definition, and what is more the practices in question are (ambiguously) competitive in the banking sector, due to the context created by the regulation.

This last observation once again raises the conventional antitrust dilemma of per se rules versus rules of reason. Application of the latter criterion, with the attendant connotations of political motivation and expediency, will tend to make decisions appear less absolute and objective, more open to question⁵². With, in this case, the aggravating factor that the underlying assumptions of the regulation have the practical effect of blocking any form of antitrust application.

For these reasons, there seems to be very little scope, overall, for increasing competition through traditional antitrust interventions directed at behaviours on the supply side.

But by turning to a different antitrust tradition--namely that which pursues the objective of maximising consumer welfare--it is possible to map out an alternative strategy of demand-side interventions, as has in fact already been done in some other markets, though not in any systematic manner.

The idea is to precisely pinpoint and remove the elements that inhibit competition, not by altering the competitive scenario as such, but simply by operating in the sphere of consumers. In the case in point, if switching costs (coupled with the effects of national regulations) are one of the main sources of market power for firms, a competitive equilibrium may be approached by taking explicit actions to reduce such costs⁵³.

Such a strategy has already been successfully deployed to increase competition in certain other sectors. In the telephone industry, for example, liberalisation of the so-called 'last mile' has enabled many countries to introduce competition into a sector that appeared locked into monopoly positions arising from historical inertia. Something similar has occurred in the sector of car liability insurance, through the introduction of a more effective system based on the portability of the insured party's risk status, and efforts to combat fraudulent practices. So taking precisely this insurance market as a model, one could envisage applying a similar system to the banking industry as well. The relationship between a customer and bank, much like that between insurance companies and their clients, is one of long standing that involves collection of a large data set, and is generally characterised by high transaction costs.

But notwithstanding these similarities, the sector of car liability insurance has seen a gradual reduction in its switching costs: contracts are generally annual and renewable, and customers are able to switch between different companies without loss of data, thanks to a standardised procedure that guarantees the accuracy of the information received by the new company, and retains any benefits (such as no-claims bonuses) matured with the original insurer.

A similar strategy could be deployed within the banking sector, in order to promote increased competition without compromising the overall market structure, and so preserving stability.

Some cautious moves in this direction have recently been attempted, though as yet not backed by an explicit theoretical position or clear economic policy. In a recent case of a proposed merger between two British banks, Lloyds TSB and Abbey National, the Competition Commission (2001, sect. 1.11) notes for example that “Steps are being taken to improve the process of switching personal current accounts between banks. This is important in creating conditions for the market to become more competitive [...]. The current project to automate the necessary exchanges of information must be completed and effectively implemented without delay and the process must be speeded up.” In the same market, a slightly earlier study underlined the need to ensure greater transparency and, most importantly, greater representativeness to consumers⁵⁴.

So an added advantage of the proposed solution is that, because it still passes through regulatory channels, unified control of the market can be retained. The market structure is ultimately not altered in any way and can continue to be defined in agreement with the stability objective. The increased mobility of consumers between the various banks on the market will simply have the primary effect of reducing the market power that each firm exerts over its locked consumers (arising from the switching costs). However the resultant increase in competition and efficiency will only minimally impact upon the likelihood of failure and subsequent contagion, because both entry and exit remain regulated, and the banks are diversified into other submarkets such as financial products sold to large firms, wholesale banking, and so on.

8. Conclusive remarks

This contribution has addressed the central dilemma of the banking sector, that of stability versus competition, and attempted to identify a novel solution. In fact, the prevailing idea, in the abundant literature on the subject, is that a trade-off exists between these two objectives, and that competition must largely be sacrificed to the greater good of economic stability. The result is an inevitable antagonism between regulatory activities and those which promote competition.

The thesis presented in this work does not attempt to refute this antagonism, but more simply leverages unexplored market resources in order to obtain an alternative competition strategy that sidesteps the problem. The proposed solution involves increasing the mobility of consumers, without any significant alteration to the regulated

supply. More specifically, the existence of exogenous and endogenous switching costs makes it possible to pursue pro-competition policies that do not require imposition of penalties or changes to the market structure, but focus instead on eliminating the lock-in mechanisms, and hence the market power and inefficiencies associated with them. Although the described solution is undoubtedly a second-best alternative, it nevertheless makes it possible to pursue the goal of consumer welfare maximisation, which is one of the stated aims of antitrust policies.

An analysis of the specific competitive attributes of the banking market - equally found in other markets - also indicates that there is an entire area as yet unexplored by antitrust practice, which could be exploited to at least partially pursue the efficiency objective, in contexts where competition appears to be structurally ruled out. In particular, such a solution could provide a new and alternative route for stimulating competition in sectors that are regulated, or whose market structure cannot be drastically altered.

References

- Allen F. and Gale D. (2003), 'Competition and Financial Stability', *Journal of Money, Credit, and Banking*, forthcoming.
- Amel D. and Starr-McCluer M. (2001), 'Market Definition in Banking', *Working Paper Board of Governors of the Federal Reserve Board*, Washington D.C.
- Biehl A.R. (2002), 'The Extent of the Market for Retail Banking Deposits', *Antitrust Bulletin*, Spring, 91-106.
- Bruzzo G., Polo M. (2000), 'Fonti di potere di mercato nel settore bancario: teoria e implicazioni per l'analisi antitrust', in Polo M. (ed.), *Industria bancaria e concorrenza*, Bologna, Il Mulino.
- Canoy M., van Dijk M., Lemmen J., de Mooij R., Weigand J. (2001), 'Competitional and Stability in Banking', *CPB Document*, n. 15, December.
- Carletti E., Hartmann P. (2002), 'Competition and Stability: What's Special about Banking', in Mizen P. (ed), *History, Exchange Rates and Financial Market: Essays in Honour of Charles Goodhart*, vol. 2, Edward Elgard, Cheltenham.
- Ciocca P. (1998), *Concorrenza e concentrazione nel sistema bancario italiano*, Banca d'Italia, Roma.
- Competition Commission (2001), *Lloyds TSB Group plc and Abbey National plc*, at <www.Competition-Commission.org.uk>
- Cruickshank D. (2000), *Competition in UK Banking: A Report to the Chancellor of the Exchequer*, The Stationery Office, London.
- De Bandt O., Hartmann P. (2002), 'Systemic Risk in Banking: a Survey', in Goodhart C. and Illing G. (eds), *Financial Crises, Contagion, and the Lender of Last Resort*, Oxford University Press, London.
- Di Giorgio G., Di Noia C. (2001), 'L'impatto della tecnologia sulla regolamentazione finanziaria: il caso italiano', in Masciandaro D. and G. Bracchi (ed.), *Dalla banca all'euro@ank: nuovi mercati e nuove regole*, vol. 2, Edibank, Roma, 145-171.
- Ghezzi F. and Magnani P. (1999), *Banche e concorrenza – Riflessioni sull'esperienza statunitense, comunitaria e italiana*, EGEA, Milano.
- Goodhart C. (1987), *The Central Bank and the Financial System*, Macmillan, London.
- Group of Ten (2001), *Consolidation in the Financial Sector*, Summary Report, at <www.imf.org>
- Hovenkamp H. (1994), *Federal Antitrust Policy. The Law of Competition and Its Practice*, West Publishing Co., St. Paul, Minn.
- Kiser E.K. (2002), 'Predicting Household Switching Behavior and Switching Costs at Depository Institutions', *Review of Industrial Organisation*, 20, 349-365.
- Klemperer P. (1995), 'Competition When Consumers Have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics and International Trade', *Review of Economics Studies*, 62, 515-539.

- Kim M., Kliger D., Vale B. (2001), 'Estimating Switching Costs: the Case of Banking', *Journal of Financial Intermediation*, 12, 25-56.
- Masciandaro D. (1993), 'Indipendenza della banca centrale, vigilanza bancaria e stabilità monetaria', *Moneta Credito*, 184, 527-538.
- Minervini G., Onado M. (2000), 'Efficienza dei sistemi finanziari e tutela del risparmio: disciplina o deregolamentazione?', in Tesouro G. and D'Alberti M. (eds) *Regolazione e Concorrenza*, Il Mulino, Bologna.
- Mishkin F.S. (1999), 'Financial Consolidation: Dangers and Opportunities', *Journal of Banking and Finance*, 23, 675-691.
- OECD (1998), *Enhancing the Role of Competition in Bank Regulation*, DAF/CLP/ (98)16, OECD, Paris.
- OECD (2000), *Mergers in Financial Services*, DAF/CLP/ (2000)17, OECD, Paris.
- Shy O. (2001), *The Economics of Network Industries*, Cambridge University Press.
- Shy O. (2002), 'A Quick-and-easy Method for Estimating Switching Costs', *International Journal of Industrial Organization*, 20, 71-87.
- Panzar J.C., Willig R.D. (1981), 'Economies of Scope', *American Economic Review*, 71, 268-272.
- Polo M. (ed) (2000), *Industria bancaria e concorrenza*, Il Mulino, Bologna.
- Porrini D. (1994), 'Economie di scopo e sistema bancario italiano: alcuni spunti teorici ed un tentativo di verifica empirica', *Economia Internazionale*, 47, 3-22.
- Porrini D. (1996), 'La diversificazione produttiva nelle banche: definizione, applicazione e stima delle 'economie di scopo'', *Il Risparmio*, 4-5, 789-816.
- Ramello G.B. (2003), 'Copyright and Antitrust Issues', in Gordon W. e Watt R (eds) *The Economics of Copyright: Developments in Research and Analysis*, Elgar, Cheltenham-Northampton.
- Sharpe S.A. (1997), 'The Effect of Consumer Switching Costs on Prices: A Theory and its Application to the Bank Deposit Market', *Review of Industrial Organisation*, 12, 79-94.
- Shull B. (1996), 'The Origins of Antitrust in Banking: an Historical Perspective', *Antitrust Bulletin*, 41, 255-288
- Waterson M. (2002), 'The Role of Consumers in Competition and Competition Policy', *International Journal of Industrial Organization*, 21, 129-150.

Notes

* Forthcoming in Josselin J. e Marciano A. (eds), *Law and the State. A Political Economy Approach*, Elgar. We are greatly indebted for comments and suggestions to the participants to the 4th Corsica Law and Economics Workshop, to Francesco Denozza, Federico Ghezzi, Donato Masciandaro. The usual disclaimer applies.

^a Università degli Studi di Milano

^β Università Carlo Cattaneo - LIUC, Castellanza

¹ De Bandt, Hartmann (2002).

² Minervini, Onado (2000).

³ A typical argument found in the literature is that a highly concentrated banking sector will be generally less susceptible to crisis. Naturally this thesis has both its supporters (Allen and Gale, 2003) and detractors (Mishkin, 1999). The relevant point, for the purposes of this paper, is that the general predominance of the former view has prompted specific economic policy measures which effectively restrict competition. For an overview see Carletti and Hartmann (2002).

⁴ The Italian experience provides an interesting example. Historically, the stability objective took precedence over competition, as one can see from the Italian banking law of 1936, enacted in the aftermath of the Great Depression of the 1930s. It is only recently, in the 1990s, that the competition objective has been formalised in Italy through the consolidation Act on Banking (Testo Unico) and the law instituting the antitrust authority. With these two legislations the country began placing greater emphasis on competition, relying on instruments such as risk ratios and deposit insurance to counter the risk of instability (Ciocca, 1998).

⁵ For a description of the peculiarity of banking systems from a stability perspective see Goodhart (1987).

⁶ OECD (1998 and 2000).

⁷ Amel Starr-McCluer (2001). For a more in-depth discussion of the first two, see also Hovenkamp (1994).

⁸ Dec. 25/11/1993, Fortis/Cger, in EECM.C. Rep., p.1241. Naturally, the present-day dynamics are in part changing the market, for example with the advent of e-banking, so that the listed taxonomy may well need to be extended, as is discussed in the following paragraphs.

⁹ Dec. 13/6/1994, Banco Santander/Banesto, in EECM.C. Rep., p.1471.

¹⁰ Dec. 10/12/1997, Merita/Nordbanken, in CELEX, n. 397M1029.

¹¹ Dec. 10/12/1997, Merita/Nordbanken, in CELEX, n. 397M1029.

¹² For the EU, see. EC Commission, XXIII Report of the Commission on Competition Policy, COM(94) 161 def., 1994; for the US see 2A Antitrust Law par. 533(2nd ed. 1995). For a general discussion of submarkets as a relevant market see Hovenkamp (1994).

¹³ Group of Ten (2001).

¹⁴ Shull (1996).

¹⁵ Carletti, Hartmann (2002).

¹⁶ Hovenkamp (1994):

¹⁷ Porrini (1994, 1996).

¹⁸ Since 1997 an amendment of the EC Merger Regulation, income figures are used as a measure. So that the "Community dimension" is reached when:

The aggregate world-wide income of the merging banks exceeds 5000 million euro.

The aggregate community-wide income of each of the merging banks exceeds 250 million euro.

¹⁹ See OECD (1998 and 2000).

²⁰ Di Giorgio, Di Noia (2001).

²¹ This approach is perhaps a legacy of the pro-competitive function carried out by the Banca d'Italia before to the enactment of the "latecomer" law n. 287 of 1990. See Ghezzi, Magnani (1999).

²² Carletti and Hartmann (2002).

²³ Waterson (2002).

²⁴ Eastman Kodak. Co. v. Image Technical Services, Inc., US, 112 S.Ct. 2072, 1992

²⁵ Hovenkamp (1994).

²⁶ This should naturally be taken into account when defining the relevant market where, as we have said, market power plays a central role.

- ²⁷ It is also possible for switching costs to be partly endogenous and partly exogenous. The highly diverse nature of switching costs, which can originate from a multiplicity of sources, as discussed later in the text, makes distinguishing between these two components by no means a trivial matter.
- ²⁸ Unless consumers have perfect information and are able to exactly calculate the costs for all periods, in which case competition would be preserved. But the hypothesis is nevertheless debatable, and the situation not likely in practice.
- ²⁹ Klemperer (1995).
- ³⁰ These are the so-called ‘purchasing economies of scope’ which make it more worthwhile for consumers to purchase n different products from the same firm. See Panzar and Willig (1981). See also Klemperer (1995); Shy (2002); Kiser (2002).
- ³¹ Note, however, that the overall competitive picture can be quite complex, with incumbent or dominant firms that impose higher switching costs on consumers, and newcomer or fringe firms which instead limit these costs as a strategy for attracting new consumers. See Shy (2002).
- ³² Bruzzone and Polo (2000).
- ³³ Klemperer (1995)
- ³⁴ Kiser (2002); Shy (2002).
- ³⁵ Sharpe (1997).
- ³⁶ Bruzzone and Polo (2000); Kim, Kliger and Vale (2001); Kiser (2002); Shy (2002); Waterson (2002).
- ³⁷ In the former case “it has been widely observed that consumers refrain from switching between banks even when they are fully informed of large differences in bank services fees. The main (perhaps the only) explanation for this consumer behaviour is the existence of switching costs that are encountered by consumers each time they terminate services with one bank and switch to a competing bank” (Shy, 2001, p. 188). The British Competition Commission (sect. 1.8) confirms that “consumers tend to see switching between banks as a difficult and unrewarding process, and the rate of switching is very low”.
- ³⁸ In general, the existence of transaction costs in the banking sector is also taken into account for the purposes of defining the relevant market: “Since the Supreme Court ruling in the Philadelphia National Bank case of 1963 [US v. Philadelphia National Bank, et al., 374 U.S. 321] [...] [t]ransaction costs were seen to limit a consumer’s choice to a set of financial institutions that were within some relatively small distance of his or her home or work.” (Biehl, 2002, p. 91). As a result, the relevant market could be defined in geographical terms.
- Now, the interpretation of switching costs as transaction costs that significantly alter the substitutability of products in the eyes of consumers could further call into question, at least in certain cases, the definition of the relevant market.
- ³⁹ Shy (2001).
- ⁴⁰ In the English market the average delay for the four leading banks in May 2001 was 11-13 days, but a process of slow deterioration could extend this period to up to six weeks (Competition Commission 2001).
- ⁴¹ From private communication to the authors in April 2003.
- ⁴² On this point, see also the Competition Commission (2001, sect. 2.36 and after).
- ⁴³ Biehl (2002).
- ⁴⁴ It is possible for different switching costs to coexist at the same time. Consider for example the process of applying for a credit card: as a rule there are certain *start up costs*, largely deriving from the initial commission for opening the card account, in addition to switching costs relating to all the necessary paperwork, and others still arising from the possible inconvenience, when changing bank, of having to temporarily return the card and wait for a new one to be issued.
- ⁴⁵ Bruzzone, Polo (2000); Shy (2001).
- ⁴⁶ Bruzzone, Polo (2000); Shy (2001).
- ⁴⁷ Kim, Kliger and Vale (2001).
- ⁴⁸ Shy (2001). Other sector surveys of the British market have found anomalous levels of profitability compared with various sectors. See Cruickshank (2000).
- ⁴⁹ There appears to be a consensus among authors on this matter. Among these see Sharpe (1997) and Kiser (2002).
- ⁵⁰ In this sense, Ramello (2003).
- ⁵¹ Board of Trade v. U.S., 246 U.S. 231, 238, 38 S.Ct. 242, 244 (1918).
- ⁵² If for example a bank dynamites a competitor, the abusive behaviour is clearly punishable by a ‘per se rule’, because the violation is manifest and, moreover, extends beyond the bounds of competition protection. But if a bank raises its switching costs (for example through bundling, which has the effect of tightening the bonds between the firm and consumers), deciding whether such a practice is anti-

competitive relies essentially on a 'rule of reason' that is inevitably characterised by an ideological component, or at least by an explicit scientific stance, and these are also susceptible to change over the years following new developments in economic theory, as evinced by over a century of US antitrust activity (Howenkamp, 1994; Ramello, 2003).

⁵³ Waterson (2002).

⁵⁴ Cruickshank (2001).