Settling for Efficiency - A Framework for the European Securities Transaction Industry*

Baris Serifsoy Marco Weiß†

This version:
April, 25^{th}, 2005

Abstract

Despite a lot of re-structuring and many innovations in recent years, the securities transaction industry in the European Union is still a highly inefficient and inconsistently configured system for cross-border transactions. This paper analyzes the functions performed, the institutions involved and the parameters concerned that shape market and ownership structure in the industry. Of particular interest are microeconomic incentives of the main players that can be in contradiction to social welfare. We develop a framework and analyze three consistent systems for the securities transaction industry in the EU that offer superior efficiency than the current, inefficient arrangement. Some policy advice is given to select the 'best' system for the Single European Financial Market.

JEL Classification: *F 36, G 28, G 34, L22, P 51*
Keywords: *trading, clearing, settlement, straight-through processing, central counterparty, regulation, efficiency, EU, consistent systems*

*This is a substantially revised version of earlier work published in November 2003 as 'Efficient Systems for the Securities Transaction Industry - A Framework for the European Union'. We thank Thomas Kiehlborn, Reinhard H. Schmidt, Alexandre Soroko, Marcel Tyrell and Ingo Walter for their comments. We appreciate discussions with Stefan Mai and Andreas Wolf of Deutsche Börse Group.

†Address for correspondence: Wilhelm Merton-Chair for International Banking and Finance; Johann Wolfgang Goethe-University, Frankfurt/Main; http://www.finance.uni-frankfurt.de; Phone: ++49 (0)69 798-25115 or ++49(0)69 798-28268; e-mail: serifsoy@finance.uni-frankfurt.de and weiss@finance.uni-frankfurt.de

1 Introduction

Despite a lot of re-structuring and many innovations in recent years, the securities transaction industry in the European Union is still a highly inefficient and inconsistently configured system for cross-border transactions. Many EU politicians eagerly promote the completion of the Single European Market, but - few exceptions aside - the industry structure still resembles closely the former fragmented market structure of largely independent institutions operating along national lines. This causes higher costs in the handling of cross-border securities[1] which ultimately translates into higher costs of capital - a significant competitive disadvantage for European firms compared to companies in the USA.

Industry experts point at several aspects that impede the realization of an efficient securities transaction system. While trading is widely seen as efficient, clearing and settlement processes across different countries are still too costly. The fragmented industry structure, which does not allow for capturing the significant benefits from scale, scope and network effects, is paralyzed by several obstacles to consolidation. Besides political, cultural and legal barriers among the different countries, the motives of the market participants such as infrastructure providers and direct users sometimes contribute to the impediment of consolidation efforts and thus prevent a socially optimal solution.

What has become increasingly visible is the lack of a common communication standard among service providers. This could be a result of vertically integrated providers with incompatible information dissemination standards and post-trading routines. As a consequence, the typical cross-border trade requires substantial interaction among the pertaining different trading, clearing and settlement systems which can only be effectively dealt with by additional intermediaries such as (sub)custodians. This extends the length of the value chain and thereby increases the costs for the investors. More interaction requirement is also more risky due to the higher complexity of the trade and a higher likelihood of failures. Higher risks usually mean additional collateral requirements which is a further cost driver.

Not surprisingly, there are diverging opinions on how to cope with the current inefficiencies. Don Cruickshank, former CEO of the London Stock Exchange, favors a market structure that separates trading from post-trading activities while the latter should be organized as an utility comparable to the DTCC in the US:

> *"If the single market in financial services is to be delivered, then competition and regulatory policies must be allowed to work side by side. [...] We can see [that] some spring shoots of such an approach is in moves to allow exchanges to compete on a harmonised utility clearing*

[1] Lannoo and Levin (2001, p. 14 - 30) and Deutsche Börse Group and Clearstream International (2002, p. 15 - 29) present a cost analysis of cross-border transactions.

and settlement layer as the most effective way of reducing transaction costs in the securities industry as a whole, and maximising the potential for competition elsewhere in the securities value chain."[2]

This stands in stark contrast to Werner Seifert's view, CEO of Deutsche Börse, who claims that an efficient solution can be delivered by vertical integration of the activities and that the culprit are the myriads of different regulators in the EU:

"Many people claim that clearing and settlement should be done by a single, Europe-wide utility, like in the US, and that greedy private operators help themselves from the till by insisting that trading, clearing and settlement remain integrated. Not so! At Deutsche Börse we have looked at this very closely, and the overwhelming problem in the integration of European capital markets is driven by different regulations, even different applications of identical rules where they exist, - the whole messy business of EU regulation, with actual implementation left to the member states."[3]

What these two views reveal is that different and sometimes intertwined forces are at play in the European securities transaction industry. The motives of the different opponents can be biased by strategic deliberations and the desire to advance the industry structure to the own advantage. This paper sheds some light on these opposing claims by applying economic tools to identify the underlying economies in the industry and to comment on an efficient securities transaction system for the European Union. Our contribution is to provide a framework for the analysis of this industry which identifies and structures the different elements, interprets efficiency in a broader sense and offers policy advice by proposing consistently configured trading, clearing and settlement systems (TCS-Systems) that achieve high levels of efficiency from the perspective of a benevolent system designer. Applying the methodology of systems theory, we give answers to the following questions: (1)What are economic characteristics of the relevant activities and which institutions are involved in these activities? (2) Which strategic decisions can be conducted by the industry players and what are the consequences thereof? (3) What are consistently configured systems that provide superior efficiency to alternative set-ups?

Our framework presents three systems for the European securities transaction industry that are configured in a consistent way: The first assigns a prominent role towards regulation and allows to capture the economies in the industry to a great extent by integrating and consolidating the different activities in the different national markets. The second lets the market forces work and - although more fragmented -

[2]Cruickshank (2003).
[3]Seifert (2003, p. 82).

allows for high innovation and dynamics in the industry. The third consistent system is characterized by consolidation and integration as well, but instead of heavy regulation it is kept open to market forces by adjusting the necessary elements so that any ensuing monopoly remains contestable. We conclude that the third system has its advantages over the other two and policy makers in the European Union should strive to implement the various elements of it to a coherent whole to the benefit of the European capital market.

Related literature Public authorities and academics alike have taken an interest into the European securities transaction industry. Different regulatory bodies and committees established by the European Commission mainly focus on the identification of structural weaknesses of the industry and outline concrete recommendations to overcome these problems: In response to ECOFIN's request to give regulatory proposals for the European Securities Markets in 2001, the Committee of Wise Men, chaired by Alexandre Lamfalussy, demands a further restructuring and a scrutiny for the requirement of a regulatory framework in the clearing and settlement area. Furthermore, they point at competitive issues and general systemic risk aspects that may evolve in the context of monetary policy and the functioning of payment systems.[4] The Giovannini Group, a consultative group headed by Alberto Giovannini and appointed by the European Commission, analyzes the cross-border clearing and settlement arrangements in the EU and finds that international transactions are more complex and costly than domestic transactions due to fifteen barriers and that these inefficiencies represent a paramount barrier to integrated financial markets.[5]

The European Shadow Financial Regulatory Committee (ESFRC) disagrees to the claim of some market participants that a forced consolidation into a pan-European regulated utility would solve these problems. Instead they pledge for an ownership separation between trading and post-trading facilities in order to foster fair competition.[6] The Committee on Payment and Settlement Systems (CPSS) and the Technical Committee of the International Organization of Securities Commissions (IOSCO) jointly developed recommendations for securities settlement systems that aim to improve the safety and efficiency of these systems. In particular the report recommends minimum requirements that these systems should be obliged to fulfill and the best practices that they should strive for.[7] Based on these recommendations, the Committee of European Securities Regulators (CESR) and the ECB published 19 standards[8], which also incorporated comments on standardization, communication and messaging and business continuity[9]

[4] See Lamfalussy (2001).
[5] See Giovannini Group (2001) and Giovannini Group (2003) for details on these barriers in different areas.
[6] See European Shadow Financial Regulatory Committee (2001).
[7] See Committee on Payment and Settlement Systems and International Organization of Securities Commissions (2001).
[8] See Committee of European Securities Regulators and European Central Bank (2004).
[9] These issues were particularly raised by the G30, an international body composed of senior figures from the private and public sectors and academia.

as well as on standards for risk management controls[10]. Market participants were consulted on draft versions and many institutions used their chance to respond. Some organizations also published their own reports and white papers on the industry.[11]

Academic papers usually highlight particular aspects of the industry by focussing on a certain aspect in the securities transaction value chain while inevitably neglecting other potentially interrelated factors. Some contributions provide empirical research on the main activities: Malkamäki and Iftekhar (2000) investigate on potential economies of scale and scope at stock exchanges while Schmiedel, Malkamäki, and Tarkka (2002) focus on the same subject in the case of settlement systems. A related study was conducted on network effects at exchanges by Schmiedel and Iftekhar (2003). Additionally, formal models are presented for various topics such as on the economics of financial networks by Economides (1993), on vertical integration by Köppl and Monnet (2003) and Tapking and Yang (2004), on competition between central securities depositories (CSD) by Kauko (2003) and on competition between custodians and CSDs by Holthausen and Tapking (2003). Furthermore, moral hazard aspects were modelled to distinguish between net and gross settlement by Kahn, McAndrews, and Roberds (2003).

There are however also contributions that apply a more holistic approach and thus are closer related to the framework developed in our paper. One of the first academic contributions is by Giddy, Saunders, and Walter (1996): They analyze four alternative models for the European clearing and settlement market mainly from the perspective of the users of these services. Differences between their models exist in the way that linkages between the CSDs are structured. Our approach is similar to theirs in respect to evaluating trading, clearing and settlement along three dimensions and deriving distinct systems for a future industry setting. Unlike their approach, we take microeconomic incentives of the key industry players in more detail into account and base the analysis of possible systems on sounder foundations regarding these aspects.

The paper by Milne (2002) establishes analogies between utility network industries such as the telecommunication sector and the securities settlement market. These markets, he argues, have similarities in possessing a natural monopoly that has to be regulated. Milne identifies the book-entry function and the transmission of corporate actions as the two functions of the value chain that need to be regulated via access pricing and the establishment of common communication standards. He concludes that this minimal regulatory effort should suffice to create a level playing field on all other stages such as clearing, settlement and custody, rendering further public interventions unnecessary. While we agree on this notion, we extent his policy advice by presenting and applying a more comprehensive framework that -

[10] The European Association of Central Counterparty Clearing Houses (EACH) was one of the initiators of this point.

[11] See for example the White Paper by Deutsche Börse Group (2005) which describes the post-trade market. While trying to influence policy-making, such efforts are often rather biased.

together with the methodology of system theory - allows a broader policy advice.

The paper closest to ours in spirit and result is by van Cayseele (2004) who agrees with our reasoning proposed in an earlier version of this paper.[12] Van Cayseele starts from the same premises, asking what an economically optimal outcome for the clearing and settlement industry would be. Relying on the concepts of 'essential facilities' and of 'two-sided markets'[13], he concludes that the advantage of a market solution relying on contestability of the industry is to be preferred against a single regulated monopoly with the potential costs of government failure.

Structure of the paper Section 2 briefly describes the three efficiency concepts deployed in this paper. We concentrate on the activities of trading, clearing and settlement along the value chain and the providing institutions.[14] We explain the underlying economies of the activities at each level of the value chain and the interdependencies across the whole chain in section 3. The role of regulatory institutions is briefly discussed as well. Section 4 regards possible strategies and associated actions to highlight the microeconomic incentives of the infrastructure providers. Three different decisions have to be made - where to set the boundaries, whether to adopt an industry-wide standard and how to assign ownership rights. We show consistent TCS-systems in section 5 that are efficient from the viewpoint of social welfare. Its individual components are complementary and thus reinforcing each other. Potential drawbacks and implications for social welfare are discussed. Section 6 concludes.

2 Three concepts for evaluation

Economic rents are created through 'good' investment decisions by the various constituencies and allocated to these institutions through 'good' distribution rules. This interplay between ex ante incentives and ex post distribution determines the efficiency of possible TCS-systems. We analyze efficiency along the three lines of static, dynamic and systemic efficiency to (1) evaluate the generated economic rents and (2) estimate the resulting overall efficiency. We shortly describe each concept and potential trade-offs between them.

Static efficiency A certain activity is performed in a *statically efficient* way if there is no solution that would allow a less costly implementation. It is under this notion that the commonly used concept of cost efficiency is considered. Parameters influencing the static efficiency are the costs of production

[12] See Serifsoy and Weiß (2003).

[13] See also Rochet and Tirole (2003).

[14] Custody functions follow the settlement process. These ensue the distribution of coupon payments, the implementation of corporate actions and the lending of securities besides the trading-induced transfer of ownership. We will subsume these transaction-induced custody aspects under the settlement activity and ignore the other services in custody. Taking into account all aspects of custody would add to the complexity, while providing only limited value-added for our purposes.

which in turn are influenced by the underlying technologies and the economies arising from these. In the securities transaction industry, network externalities[15] are prevalent. They lead in many areas, like the trading of a single derivative instrument or the settlement of a particular stock, to an efficient market structure that is a natural monopoly.

Static efficiency thus generally increases if the number of companies conducting business along the securities transaction value chain decreases due to the underlying economies. The costs of any regulation that has to be set up to keep the remaining companies and their rent-extracting potential in check, however, lead to a lowering of static efficiency.

Dynamic efficiency Activities are performed in a *dynamically efficient* way if today's structures and investments do not hamper the performance of these activities in the future. By investing in a certain technology or by institutionalizing a certain industry structure, the ability to change and to adapt becomes affected. Particularly the dominance by a network provider or the sponsor of a network - as Economides calls it - may have detrimental effects on the innovativeness of the market.[16] Industry structures and processes that do not allow for innovation and for quality improvement in future thus are not efficient under the notion of dynamic efficiency. Competition in the market usually helps to alleviate problems such as low innovativeness, poor quality of the goods and services whereas the absence of competition may lead to complacency and to less innovation as it is common in a monopolistic environment. For estimating dynamic efficiency key parameters are (1) the industry structure that determines the difficulty of entering the TCS-industry, (2) the rate of technological innovation and (3) the propensity of all institutions to invest and the resulting sum of all investments.

Systemic efficiency Our third evaluation concept, denoted *systemic efficiency*, provides insight on systemic risk issues that are inherent on various stages of the value chain and takes into account the stability of the TCS-industry when faced with adverse systemic events. We define systemic efficiency as the degree of robustness of the activities in the securities transaction industry to systemic risks that are borne from strong adverse systemic events.[17] A systemic event occurs when a 'bad event' for one or more market participant(s) has subsequent negative repercussions on other market participants. Such an event may vary in severity, ranging from a delay in payment or delivery of the securities in question to a full-blown failure of a party to meet the agreed-upon obligations. Potential contagion effects have to be taken seriously most notably in cases of strong negative systemic events like a failure of an institution. Systemic

[15] For a detailed discussion on network externalities confer Shapiro and Varian (1999, p. 173 - 225).
[16] See Economides (1993, p. 92).
[17] The terminology used is adapted from de Bandt and Hartmann (2000) albeit the authors discuss this issue in much greater depth. See de Bandt and Hartmann (2000, p. 10 - 17) for further details.

risk issues are treated with great care by public and private entities. Both ex ante (crisis prevention) and ex post (crisis management) measures have to be introduced in order to deal with systemic risks. 'Good' regulation has to ensure this.

Interdependencies Note that the three concepts of efficiency are interdependent: (1) The statically efficient solution of a monopoly conveys only minor incentives to innovate whereas a few players in an oligopoly can interact in heavy competition and try to develop better products and processes thereby increasing dynamic efficiency. They also compete for monopoly rents, that are non-existent in a perfect-competition-environment where the users of the infrastructure reap the main part of economic rents. The potential profit that can be gained is therefore a big enough incentive to undertake the large technological investments needed up-front. (2) Static efficiency can decline when measures are taken to increase systemic efficiency: The provision of collateral e.g. increases the stability of the industry against adverse shocks but levies opportunity costs on the market participants. The existence of economic rents also facilitates the build-up of a financial buffer that allows these companies to be more stable in times of systemic crises. (3) Perfect competition would contribute potentially more in terms of innovativeness thereby increasing dynamic efficiency but the systemic efficiency could be damaged since a more fragmented structure may impose more work on regulators to keep the overall system sound. However, competition also fosters innovations in risk management tools which are beneficial to systemic efficiency.

3 The securities transaction industry

To analyze the securities transaction industry we define the *securities transaction value chain* and the institutions that are involved in these activities. The value chain has three main activities: Securities need to be *traded*, the results of the trade have to be confirmed and calculated by a *clearing* process and the delivery of money and paper to the parties to a trade has to be *settled* by institutions that provide this infrastructure, e.g. exchanges, clearing houses and central securities depositories (CSD). Two other classes of institutions are involved in the value chain, namely the users and regulators. The users are the clients of the infrastructure provider and can be further broken down into banks and brokers as direct users[18] and investors and issuers as indirect users. The regulators monitor the processes in the industry to ensure a sound and efficient transaction environment. Their appropriate role is discussed at the end of this section.

[18] Banks and brokers are the main institutions using the infrastructure as the immediate users. They play a pivotal role in the securities transaction value chain for institutional as well as retail investors on the one side and for companies with their underwriting business on the other side. By internalizing security transactions and by acting as subcustodians, banks are to a certain degree also direct competitors of the infrastructure providers. A model on the competitive relationship between CSDs and custodians can be found in Holthausen and Tapking (2003).

3.1 Economies of the securities transaction value chain

Network effects There exist strong positive network externalities on each of the three stages: In trading, network effects can be both observed on the investors' as well as on the issuers' side. For the former, becoming a member of an already large network of investors that trade on the same platform increases both her own and the other's utility by providing additional liquidity to the market.[19] The latter group benefits from larger networks as these can absorb the issuers' need for capital more easily.

There are viable positive network externalities on the user side for both clearing and settlement. A concentration on few transaction systems allows for a higher portion of clearing and settlement instructions to be processed internally. This increases the utility of all users because costly links to other networks become less necessary. In the extreme, a single clearing and settlement network would be faster and less costly in comparison to processes which require the interaction with several clearing and settlement systems.

Economies of scale The providers of trading, clearing and settlement facilities can reap significant economies of scale as the setup costs for a transaction platform have a substantial portion of fixed costs so that average costs fall with increasing transaction volume. This view is confirmed by empirical investigations. Malkamäki's analysis on the processing of trades at stock exchanges shows scale economies for increasing trading volume.[20] Another contribution by Schmiedel, Malkamäki, and Tarkka (2002) measured significant economies of scale for settlement systems: Platforms with high transaction volumes will be able to offer lower transaction costs to users than low-volume competitors. For non-automated transaction systems, i.e. floor-based trading, this effect is not as pronounced as for automated trading systems since the ratio of fixed costs to variable costs is higher for the computerized system.

Economies of scale are also present in counterparty risk management. Especially if a central couterparty (CCP) is used in the clearing stage, the users of the facility can save resources on the management and control of counterparty risks. By pooling risk management facilities at the CCP, costs can be eliminated by risk management specialization effects. Additionally, if netting mechanisms are used, the users will enjoy reduced capital provision requirements and therefore lower opportunity costs.[21] As a consequence, scale economies may be even more pronounced in clearing than in other stages.

[19] The pivotal role of liquidity stems from the potentially large costs that can arise from illiquidity during trading. According to Deutsche Börse Group and Clearstream International (2002, p. 17 - 22) their proportion of total trading costs is substantial. Liquidity can be characterized along four dimensions, namely width, depth, immediacy and resilience. See Harris (1991, p. 3). For an excellent model of such two-sided markets see Rochet and Tirole (2003).

[20] However, he confines his findings for very large stock exchanges. See Malkamäki (1999) for further details.

[21] See Van Cauwenberge (2003, p. 94).

Economies of scope All three activities exhibit potential economies of scope. Providers are able to process different types of securities on the same platform while incurring only relatively low incremental costs. Clearing facilities that process different classes of securities such as stocks, bonds, and derivatives have additional leeway for scope economies as they are able to implement innovative risk management procedures such as cross-collateralizing along different securities classes. This would lead to an overall decrease in capital provision requirements to the users and would consequently save costs.

The upshot of the results above suggests a tendency towards a strong concentration of the activities or even a natural monopoly on each stage of the value chain due to the underlying and mutually reinforcing economies. According to our definition in section 2 a concentration in the industry translates into high static efficiency.

Contestability of the market A concentrated market in turn lowers dynamic efficiency as the latter falls with decreasing levels of competition. The existence of substantial network externalities, scale and scope economies create a barrier to entry and offer established platforms some protection from competitors.[22] Nevertheless, high levels of dynamic efficiency can be achieved if the market for securities transactions remains contestable, i.e. if competitive and innovative infrastructure providers can gain market share at the expense of established competitors.

The diversion of transaction volume is the more likely the more of the following aspects coincide: (1) The competitor demands lower fees.Domowitz and Steil (1999, p. 8 - 9) give several examples of this behavior. (2) A competing provider offers a significantly better service for users based on a better technology which manifests itself in faster, more reliable or more convenient transaction handling. (3) The competitor offers new products or services which have not been supplied and therefore 'monopolized' by an established provider yet. (4) Clearing and settlement institutions reduce their capital provision requirements by introducing netting processes. However, this point may be a matter of regulatory concern as competing institutions might want to apply less stringent risk management to successfully underbid the fee structure of competitors.

Systemic risk issues Systemic efficiency is particularly relevant in the clearing and settlement of securities and appendant funds. There are several sources of and alleviation efforts to systemic risks. Clearing and settlement institutions have developed risk management tools that attempt to reduce both ex ante and ex post the various types of settlement risk. Since the various types of systemic risks such as

[22]This view is shared by Economides (1993, p. 92 - 93). He states that as a consequence of the reinforcing nature, a financial "network exhibits positive critical mass". A further consequence of networks "is that history matters [...] because of significant switching costs" which protect established players in the market.

counterparty, custody and cross-border risks have been elaborated in detail by other contributors[23] we focus on the relationship between systemic efficiency and market concentration. The relationship is not straight-forward. Both a fragmented and a consolidated industry structure have to deal with trade-offs.

A concentrated industry structure can exploit economies in centralizing risk management efforts as it is more cost efficient to have one party collect the information and to monitor the other parties instead of having all parties monitoring each other. Thus, a central risk manager will be more cost efficient and more sophisticated. However, the central risk manager may bank too strongly on its dominant position and believe that public entities would bail him out in case of a failure. Moral hazard may materialize in the form of reduced monitoring efforts.[24] Therefore regulatory effort - albeit rather easy as only one institution has to be controlled - might be necessary.[25] Additionally, a higher degree of consolidation leads to less complexity in the interaction between the providers and thus reduces the probability of failures in communication and asset transfers.

A fragmented industry structure on the other hand may provide systemic efficiency that is superior to a concentrated market. More industry players will usually lead to higher levels of competition. A possible parameter of competition can be the provision of sound and stable transaction systems among providers which boosts systemic efficiency. Another positive aspect of fragmentation is the existence of redundancies which - if communication protocols between different transaction systems are compatible - can be used to re-route transactions from a failed to an intact system. Multiple transaction systems may thus increase the robustness of the industry although potential contagion effects between the providers may weaken this advantage.

3.2 Vertical interdependencies in the securities transaction value chain

In several European countries, the dominant trading institution often also exercises control in the activities further downstream, i.e. the exchange is vertically integrated into the domestic clearing and settlement activities. This setting is mainly driven by efficiency motives as it enables to process straight-through the whole transaction in a faster, more cost efficient and more reliable manner. Straight-through-processing (STP) at a single institution offers significant economies to both users and providers in comparison to the processing between separate entities as: (1) It lowers communication costs between the respective activities thereby improving static efficiency. (2) Innovations concerning the processing of transactions are easier to implement since coordination efforts with other providers along the value chain are not necessary. This shortens the implementation period and therefore increases dynamic efficiency. (3) It

[23] See Giovannini Group (2001, p. 18 - 19) and de Bandt and Hartmann (2000).
[24] See Diamond (1984) on the use of such delegated monitoring as a rationale for the existence of banks.
[25] Confer de Bandt and Hartmann (2000, p. 17) for further details.

makes transaction failures less frequent since the data transmission process is optimized in-house, for example by implementing a proprietary communication standard. This represents an improvement in systemic efficiency.

However, with the decline in IT-infrastructure costs in recent years, the arguments for vertical integration are not as strong as they have been some years ago as transmission costs to outside institutions are now significantly less costly and not necessarily higher than in-house transmission costs.

Furthermore, as trading habits of investors gradually shift from a domestic towards a more international approach, the national 'silos', as the vertically integrated entities are also called, not only no longer represent the investor's scope of transaction activities but even hamper frictionless processing of cross-border transactions in Europe. This is due to their incompatible proprietary communication standards which each national silo had developed to communicate along its own controlled value chain - a legacy that makes communication between silos a highly complex and inefficient task.

Differing communication standards between vertical silos also de facto impede the contestability of the downstream activities, i.e. clearing and settlement markets. They represent an effective entry-barrier against other providers that strive to enter the market of an established silo. They are unable to do so because once a trade is made on the established trading platform, competitors are restricted to offer their services for the downstream functions due to the existing proprietary communication standard of the established provider. Therefore, clearing and settlement activities of the established provider are protected by its trading activity and are thus barely subject to contestability. This may result in dynamic inefficiencies.

3.3 Regulation

Efficient trading, clearing and settlement of securities is important for the functioning of the whole economy. Companies need to get access to finance and private households need a vehicle by which they can save their financial surplus. This assigns financial markets in general and the TCS-industry in particular a pivotal role. The well-being of other industries and many people depends on it. Adverse effects spill over into other parts of the economy implying negative externalities. Therefore, regulation of the TCS-industry is a means to avoid or to mitigate these external effects.

These spill-over effects are very material in the settlement stage of the securities transaction value chain when the payment system is involved. A failure of one party to meet its obligations might lead to contagion effects that have negative effects on the liquidity of the banking system and threatening the economy by this transmission channel. The central bank as lender of last resort therefore has an

incentive to deal with these regulatory issues. It therefore needs to be - and also is - one of the key regulating institutions since central bank money is frequently involved in settling the cash side of securities transactions. Other regulatory bodies are concerned with different aspects: For example, the performance of each activity for all users - which are of a considerably heterogeneous degree - in a fair manner needs to be ensured, i.e. access to the infrastructure must be open and in an undiscriminating way. The European Commission with various reports - the Lamfalussy Report, the two reports of the Giovannini Group and the Investment Services Directive for example - is committed to this task in the Single European Market. National agencies implement actions to put these aspects into practice today. Regulators often join their forces to set standards after consulting the relevant industry players, e.g. the joint working group set up by the ECB and the Committee of European Securities Regulators (CESR). A single European Financial Services Authority might one day take over this job. In all cases, the right balance between static, dynamic and systemic efficiency must be chosen by regulators.

Regulation lowers static efficiency since it is costly to set up a bureaucracy - or a publicly owned entity in the extreme - to achieve the performance of the three activities in the value chain. The 'outsourcing' of regulation to the providers and to the users of the infrastructure could be a cost-efficient alternative. Possible means for this outsourcing lie in the self-regulation by the infrastructure provider. Whenever they can compete on quality as many exchanges do with different market segments and the attached regulatory conditions, regulation need not be of the costly public variant. As a second means, the infrastructure can also be user-governed. In this setting, the club of users writes its own rules. Whenever little entry in this club is required, this can again be better than publicly provided regulation.

The current system with competing regulatory regimes in Europe - with sometimes overlapping competencies, sometimes unattended areas - is seen nearly unanimously by the industry as a major barrier to business since a level playing field is not provided. We come back to the question of regulation in the context of the three efficient systems in section 5.

4 Strategic conduct - The provider's action set

In this section we analyze the three key parameters in our framework that the providers of the infrastructure, the users and the regulating institutions can use to interact strategically to shape the future of the securities transaction industry. The focus is on the providers of the infrastructure, possible actions and reactions of the other two classes of institutions are taken into account where necessary.[26] The three

[26] A more elaborate description of the three classes of institutions can be found in an earlier version of this paper. Confer Serifsoy and Weiß (2003).

parameters in the action set that we look at are (1) the *boundary* decision of the infrastructure providers, (2) the decision whether to adopt an open standard or to develop a proprietary *communication* tool and (3) the *governance* of the infrastructure providers.

4.1 Boundary decision

The institutions providing the infrastructure for trading, clearing and settling the different financial instruments face the problem whether to integrate different activities under the institutional roof of one firm or whether to concentrate on just one function or one specific financial instrument. We describe two distinct business models - the vertically integrated silo and the vertically focused firm.

The vertically integrated silo The first business model is the vertical silo - the combination of trading, clearing and settlement under the roof of one firm. It is applied for example by Deutsche Börse AG. The advantage of such a model is that it allows to reap the benefits that derive from the economies of scope between the three functions as described in section 3.2. Communication is easier when the three functions are performed in close proximity within the same organization. Specific forms of data exchanges between the three stages of the value chain and straight-through processing allow for the emergence of economic rents.[27]

However, one of the adverse effects such a business model has, which might be a prevalent microeconomic motive behind this strategy, is the leverage of a (natural) monopoly on one stage of the value chain upstream or downstream to other stages. Particularly, a vertical silo may cross-subsidize its trading costs - and thereby attracting customers from other platforms - through its monopoly profits on the clearing and settlement stage or vice versa.[28] By this strategy, an institution following the business model of vertical integration effectively strengthens its competitive position. Furthermore, the vertical silo forecloses the market for competitors: By restricting access for them in one activity, users can be forced also to 'buy' the solution for another activity from the same institution. If there is no choice for them but to deal with the same provider, a monopoly rent can be extracted from the users further increasing the economic rents generated in this model due to the specificity inherent in it.

Therefore, the interesting question arises whether Deutsche Börse can deliver its promises given that it controls downstream activities and de facto can foreclose the market due to its monopoly on the following stage in the value chain. When faced with the specific and co-specialized investments have to undertake, banks and brokers could be reluctant to join in this venture.

[27] See Williamson (1985) for the role of specificity in explaining vertical integration.

[28] The detrimental effect of vertical integration on horizontal consolidation between different infrastructure providers on the same stage of the securities transaction value chain is formalized by Köppl and Monnet (2003) in a mechanism design model taking into account the asymmetry of information between the different players.

The vertically focused firm The other business model that has promising features is that of more focused infrastructure providers and the use of market mechanisms between them. A prominent example for this industry setting was used in England, where three independent institutions, namely London Stock Exchange, London Clearing House and Crest, provided infrastructure services only for one stage of the value chain, respectively.[29] If a good solution for the data transfer between the three activities of the securities transaction value chain is implemented, this model has appeal because it does not give too much power into the hands of a single entity that can control access to its infrastructure - an infrastructure that exhibits strong network effects. For each activity, the users can choose the best institution that provides it in the most efficient way. Competition forces less efficient institutions out of business and sets high-powered incentives for the surviving. As such firms do not have to worry about any interdependencies between their different lines of business, they are more eager to adopt new and better technologies and processes. Any cannibalization of value propositions within the same firm cannot happen. These institutions increase therefore static and dynamic efficiency. A possible drawback which might be taken into account by the regulators is less systemic stability that too high-powered incentives might induce. How the problem of establishing a market mechanism for the intermediate goods - the information transfer from one stage to the next in the value chain - can be dealt with is the topic of the next section.

4.2 Communication standards and accessibility

The interaction between the stages of the value chain is of crucial importance to the way business is performed in the TCS-industry. It necessitates the infrastructure providers to make decisions both on the information transfer mode, i.e. the type of communication standard, and the degree of accessibility of their activities to competitors.

Proprietary versus open communication standards *Proprietary standards* infer that the information format of the transactions cannot be interpreted without co-specialized investments so that competitors are discriminated whereas an *open standard* enables competitors to process the information and allows users to switch the providers more easily.[30]

The decision whether to adopt an open standard or to set up a proprietary system is intertwined with the vertical boundary decision. In the case of a vertically focused infrastructure provider, the case is trivial. Such an institution has to rely on the market for the performance of upstream and downstream

[29] The London Clearing House has recently teamed up with Euronext's Clearnet, whereas Crest has merged with Euroclear in 2002.

[30] The battle for an unique communication standard and the different competing approaches are described in Weitzel, Martin, and König (2003).

activities, the communication protocol has to be in an open and understandable format.[31] The case is different for companies following the business model of a vertical silo: Such companies can develop a solution that allows them to keep the information that has to be passed along the securities transaction value chain private. By doing so, it can develop an idiosyncratic data exchange format that allows them to generate an economic rent due to the specific nature.

However, an economic rent could also be generated by foreclosing the market for an upstream or downstream activity. Users are forced to rely on the same institution and buy the bundled product in a one-stop shop. They have to invest in co-specialized computer systems that allow them to handle this proprietary data format. With an open communication standard between the different stages of the value chain a deconstruction becomes possible. There would be a choice for customers to deal with the best and most efficient institution.

Analyzing open and proprietary communication standards on the basis of our three efficiency concepts there are two major advantages for proprietary in comparison to open communication standards: (1) They can be more specific in relation to a certain financial instrument or a certain institution than open communication standards and thus be more statically efficient. (2) A higher dynamic efficiency can be reached since a proprietary communication standard allows for the complete appropriation of the economic rent that is generated by innovations. Additionally, the benefits of innovations made with open standards could be enjoyed by every participant without being obliged to invest into this innovation. Thus, underinvestment problems may arise with open standards.

However, proprietary communication has also two major drawbacks when compared to open standards: (1) Proprietary standards provide more incentives for strategic behavior to infrastructure providers which can be to the detriment of the users. Thus, market foreclosure strategies and mutual reinforcements of monopolies on different stages of the value chain can have a negative impact on users. Both static and dynamic efficiency may be impeded. A communication standard that is open to all market participants prevents or at least alleviates this strategic behavior: It is easier for competitors of the infrastructure provider or for the users themselves by means of internalization to work around such a foreclosure. (2) Proprietary standards raise more regulatory concerns if a regulator wants to ensure the proper functioning of the market and access for other institutions. Therefore, static efficiency may be lowered due to the increased regulation costs. Additionally, systemic efficiency may be low if regulators do not control the proprietary standard properly.

[31] One could also imagine the case that a proprietary standard is used. In this case, the outcome would be a hybrid solution along the lines of Williamson (1985): Long-term contracts or strategic alliances are necessary to account for the hold-up problem since specific investments have to be made.

Restricted versus open accessibility Accessibility is the flip-side of communications and refers to interactions between competitors across different stages of the value chain, i.e. between trading and clearing or between clearing and settlement. *Open access* describes the ability of institutions to provide their services on one stage of a transaction although other stages of the value chain may be performed by competitors. This stands in contrast to transactions where access is *restricted* by a provider. Restriction of access is possible whenever a provider is able to leverage its dominant position on other stages of the value chain. This may for example occur when a dominant trading facility prevents other providers of downstream activities to receive the transaction and automatically route it to their own clearing facility instead. Another example for dominance can be found in the opposite direction, when a settlement provider has the monopoly on a certain security and refuses to accept transactions that are traded or cleared from anybody but its own upstream activity provider. Therefore restricted accessibility can strongly impede fair competition among providers in the TCS-industry.[32]

Accessibility as well as the communication standard decision primarily depend on the industry structure[33], the allocation of power between the three classes of institutions, and the governance and ownership of the providers of the infrastructure. Using (or being forced by a regulatory institution to use) a common means of communication technology effectively opens markets. The power that is conveyed by open markets to users allows them to search for the best price and quality. This in turn eventually forces a redistribution of economic rents away from the incumbent providers who would otherwise hang on to an inefficient allocation of resources from a welfare perspective. It is to these governance aspects that we turn next.

4.3 Ownership structure and governance

Ownership of a good is an incentive device: When residual decision rights are aligned with the rights to residual income, decisions are made in such a way as to maximize this share. The maximization of it optimizes social welfare whenever these decisions can be made independent of others. It therefore matters who ultimately has control over a certain good or resource. This is also the case for a firm - a much more complex 'good' and a whole bundle of resources. For our analysis, we take into account the ownership structure of a provider of the infrastructure in the securities transaction value chain to check whether economically sound decisions will be made by this institution. Three distinct forms of ownership can be identified: (1) A *for-profit firm* that operates to maximize the profit that is distributable to its

[32] See also Milne (2002) who proposes to regulate access to the book transfer (which would fall under the notion of settlement in our paper) as he identifies it as the natural monopoly within the clearing and settlement industry.

[33] There is a strong interdependency and complementarity between accessibility, communication standard and vertical boundary decision of the providers: We observe the tendency that a vertically integrated firm often employs a proprietary standard with restricted access while a vertically focused firm prefers open standards and open accessibility.

shareholders as dividends, (2) a *non-profit mutual* that operates to maximize the utility of its members and (3) a *publicly-owned entity* that provides a good or service that would not be provided efficiently by a private firm due to its public good nature and the underlying external effects.

Public ownership of an infrastructure provider can be a means to deliver a service that must be provided by a natural monopolist. The public policymaker - acting in the interest of society as a whole - is not interested in narrow profit motives but rather tries to provide this service in the efficient quantity. This gain is however very likely to be offset by inefficiencies that public bodies bring with them. Without a profit motive the resulting incentives in the publicly owned 'firm' are weakened and inefficiencies are re-introduced.

In recent years, many publicly-owned monopolies in diverse industries were therefore privatized and for-profit firms were established instead. In this form of ownership - the standard capitalist form most commonly analyzed in economic theory - the residual decision rights are aligned with the residual claims and better incentives are thereby conveyed. The public interest of the provision of the right quantity for the correct price in these network industries is better served by a regulator that has less decision power (and less potential for meddling) than an outright publicly owned enterprise. The shift in control and power away from a public authority towards the private agents that use and provide the infrastructure increases overall welfare by nurturing better decisions because the resulting economic rents are exploitable by these decision makers.

The third possible ownership arrangement is that of mutual ownership. In the mutual form, the users of the infrastructure provide the necessary investments themselves so that the statically efficient quantity is produced for a price that is lower than the monopoly price. The direct users are members in the providing institution and take into account the supplementing function that the infrastructure has for their core business in which they ultimately want to generate economic rents. The amount of economic rent that is generated in such a mutually-owned institution is therefore lower compared to a for-profit firm while inefficiencies are reduced in comparison to a publicly-owned and over-regulated authority. Such a structure has its drawbacks however. The membership of the mutual can procrastinate and new entrants can be discouraged from using the same infrastructure. If the members are too heterogeneous, the governance of 'one member - one vote' instead of 'one share - one vote' can cause decisions to be distorted by the divergent interests and the larger players can be held up by the smaller institutions.[34] In recent years there has been a wave of demutualizations, especially among exchanges, due to these problems.[35]

[34] Hart and Moore (1996) present a model in which the heterogeneity of users makes a mutual structure less preferable.
[35] See also Domowitz and Steil (1999, p. 14 - 16) on this issue.

Why are ownership and governance aspects important? The governance of the infrastructure providers is of particular importance for the efficiency with which securities transactions are performed. The Council of Institutional Investors - representing 130 pension funds holding 3 trillion dollars in assets in the USA - criticizes that of the three constituencies of the New York Stock Exchange - members (intermediaries like broker-dealers and specialists), listed companies and investors - only the members are allowed to vote and to choose the board. This structure has a potential negative effect on the self-regulation of the exchange since that is biased towards the members' interests.[36]

The governance structure also influences the ability of a firm to innovate and to be efficient in a more dynamic sense. Too much power in the wrong hands hinders the necessary innovation in the face of disruptive technologies. The introduction of electronic trading systems for example was heavily fought by floor-based brokers that have an important voice in the governance of exchanges. When these are not only the users of the infrastructure but can also exert power through a mutual ownership arrangement, they can block such innovations that would make them worse off but lead to big gains for other users.

4.4 Interdependencies between the parameters and its configurations

The three described action parameters are not independent of each other. We want to highlight some interdependencies here as a precursor to the analysis of the three systems compared in section 5.4.

Boundary decision and communication standards Open standards are a means to credibly commit an institution performing a certain function in the securities transaction value chain not to pursue a foreclosure strategy by vertical integration. The leveraging of a natural monopoly on one stage of the value chain for a certain financial instrument cannot be used to force customers to use the infrastructure of the same institution on the previous or next stage in the value chain as well. The choice for customers and the threat of market entry by upstarts do not allow institutions to use their power to extract more than their 'fair' share of economic rents generated by the activities of the securities transaction value chain. This makes a strategy of vertical integration less attractive. On the contrary, in such a setting it would be necessary for the integrated institution to compete with many focused firms that know their activity by heart. Any advantage in terms of higher economic rents these focused institutions can gain can only come from better service which leads them to pursue a strategy that puts a premium on innovation. Even if the vertically integrated firm also pursued aggressive R&D activities, it would be faced with the dilemma of cannibalizing its own success whenever it came across an innovation on one stage of the securities transaction value chain that would force it to restructure the relationship between the

[36] See Economist (2003, p. 59) for this example.

integrated stages. The need to meddle with transfer prices weakens incentives for middle managers or even leads to outright sabotage of the new product or process by the managers of the less innovative stage.

Ownership structure and communication standard Economic rents can be generated by for-profit firms by suitably using the ideas of industrial organization theory to structure the industry to make it more difficult to enter the market. One such tool is a proprietary standard probably in combination with a strategy of vertical integration. In the other two ownership forms we described - public and mutual ownership - these incentives to foreclose the market by opting for a proprietary communication standard are not that prominent since the appropriation possibilities of any rents generated are less good for the owners of such institutions. In the case of a publicly provided infrastructure of the natural monopoly functions this institution will settle on its own (proprietary) standard but fair access is usually granted by the provider of the public good. In the case of a mutual ownership structure the tendency is for an open communication standard because the users themselves will gain from less diversity between different providers since they then have to invest only in one system to cope with data from numerous institutions on the other stages of the securities transaction value chain. However, incentives to develop the common standard and to take account of better possibilities in data exchange through broadband connections and better encryption and decompression algorithms are needed. One possibility is the use of open source-like structures: Franck and Jungwirth (2002) see the advantage of such structures in donations that are made by interested institutions without a crowding out of valuable investments in the case of an emerging standard. Cooperatives are then a preferable institutional setting in the case establishing a standard without the effects of competition that would lead to a fragmentation or to a lock-in in an inefficient system.[37]

Boundary decision and ownership structure Mutually owned institutions have their drawbacks in terms of slow decision making and weakened incentives due to the lack of the profit motive. Vertical integration augments this disadvantage by making the institution even more complex. The users of the infrastructure for the securities transaction value chain are therefore more likely to set up several cooperatives each one highly focused along the value chain that have probably different members and to rely on open standards for the exchange of information between them. The users themselves restrict the activities of a cooperative to the absolute necessary.

The solution of public ownership is more likely to be vertically integrated but a sensible and eco-

[37] For a theoretical underpinning of cooperatives and their investment incentives in emerging standards see the work of Rey and Tirole (2001).

nomically minded policy maker would again opt for a deconstruction of the value chain and a private provision for the activities where this is the best option. Unregulated private institutions run in the best interest of their shareholders, i.e. without ignoring incentives for managers to engage in empire-building activities, are very likely to pursue a strategy to shape the industry in their favor and to erect entry barriers whenever possible. As mentioned before, a foreclosure strategy of leveraging a monopoly position from one activity to the next makes perfect sense for such institutions. Privately owned companies are therefore likely to increase their scale and scope by actively integrating along the value chain when no countervailing forces prevail.

We have now outlined the constituents of our framework and discussed them in detail. In the following section we put these parameters together to concentrate on the systemic relationships between them that make some combinations better from a socially efficient viewpoint than others.

5 Three proposals for an efficient TCS-system

Systems in general consist of various modules. Between these modules or elements there can be a complementary relationship. Complementarity between any two such elements implies that the simultaneous increase in both elements leads to an overall superior performance. In the case that such a complementary relation between the elements of a system exists, the right configuration of these modules matters. Only if they are adjusted in a coherent fashion, the system in question will be internally consistent. Such a consistent system will perform better than any system in which deviations from the coherent configuration in one or a few parameters occur.[38]

This section presents three *idealized systems* where the configurations of selected elements, especially the parameters of the action set described above, are set thus that the complementary relationship is taken account for and the overall system is efficient from the viewpoint of a benevolent system designer. Small deviations from the configurations suggested lead to an overall less efficient solution.

5.1 System 1 - Regulated monopoly

Description The system of a regulated monopoly has two distinct features implied by its name: (1) there is no competition in the provision of the activities on those levels of the securities transaction value

[38] Mathematically, complementarity relates to a positive cross-derivative: The first-order returns for the increase in one element are still more enhanced if the second element is also increased. Consistency is the characteristic that any pairwise combination between any two elements has a non-negative cross-derivative - also referred to as a supermodular relationship - between them. See Milgrom and Roberts (1990) for an application of system theory in the context of modern manufacturing. See Topkis (1998) for a mathematical approach.

chain that are consolidated and (2) the role of the regulating institutions is very pronounced in these stages. Usually the roles of regulators and providers are combined and the infrastructure is publicly owned.

The horizontally consolidated and possibly even vertically integrated structure can take several forms depending on the scale of integration in each activity. In the USA for example, the Depositories Trust & Clearing Corporation is the monopoly for clearing and settlement activities and trading occurs on several exchanges. Many national markets in Europe were structured as a vertical silo with all three activities integrated into one entity. Where are the advantages of such a system?

The consolidation in each activity allows to reap economies of scale and scope along the securities transaction value chain for the providers. Users on the other hand enjoy the strong benefits of a single network. The public ownership and the lack of competition lead to low incentives to innovation activity. The threat of entry is subdued since the underlying economies as well as the publicly sanctioned role as the sole provider entrench this institution.

Vertical integration even enhances this entrenchment but leads also to the possibility of straight-through processing and an efficient use and dissemination of information from one stage of the value chain to the next. The low rate of innovation and the resulting stability in the industry makes it feasible to write detailed plans. The low innovation activity is also consistent with a low investment propensity of all players and low total investments. The users of the infrastructure are willing to undertake the necessary co-specialized investments. The standardization process is organized by the regulator which uses its powerful position to enforce and set the standard means of communication. The sum of total investments is low since no company can compete with such a vertically integrated, publicly owned institution that uses the underlying economies of scale and scope.

Figure 1 provides an illustration of the three possible industry outcomes of a regulated monopoly as well as the position of the power center between the three classes of institutions.

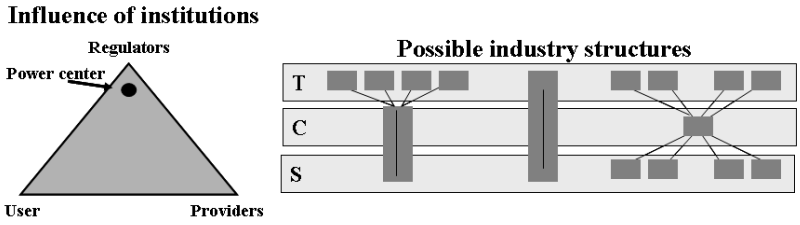

Figure 1: Regulated Monopoly

Efficiency analysis *Static efficiency* in these settings is relatively high due to the strong exploitation of network, scale and scope economies. The significant market power of the providers has its counterweight in the public ownership structure so that the inefficiencies of a monopoly do not prevail. However, the incentive structure within a big public agency brings also some costs in terms of lost efficiency. It depends on the actual processes and organizational structures of this body whether the combined effect is still positive.

Dynamic efficiency is rather low due to the lack of competition. The pressure for product or service innovations will remain limited to the detriment of the users. The overall investment activity is too low and potential competitors are deterred from entering the industry.

The analysis of *systemic efficiency* in these settings has a two-sided result. Consolidation enables the centralization of risk management at the infrastructure providers which can be more efficient than a decentralized risk management solution. However, as mentioned above, moral hazard aspects such as the reduction of risk management efforts due to a too-big-to-fail-feeling may endanger systemic efficiency in these settings.

The regulating institutions are the center of power in the system of a regulated monopoly. Ideally, this should reduce potential moral hazard issues in risk management, ensure fair transaction prices for the users, i.e. the users should benefit from exploited scale economies. However, regulation itself is costly so that increasing systemic efficiency will lead to a loss in dynamic efficiency. This system is most notably interesting when static efficiency aspects outweigh dynamic considerations. This may be the case when disruptive innovations are expected to be rather rare in the future and the processes in the industry are settled and stable.

5.2 System 2 - Competitive fragmentation

Description In contrast to system 1 the market structure of system 2 is rather scattered, i.e. it features polypolistic characteristics including several providers for trading, clearing and settlement. A high level of competition on all stages of the securities transaction value chain leads to a high rate of technological innovation. The fragmented industry structure necessitates the use of open standards since otherwise the users would have to undertake co-specialized investments to several providers. Open standards and good access possibilities allow for new institutions to enter the market easily whenever they see fit. This is consistent with the high rate of innovation that is increased by such new entrants. The tendency to invest is high since it can be the basis of Schumpeterian rents. The resulting overall investment level is therefore possibly too high when too many uncoordinated investments are undertaken. Over-investment

and a resulting bubble can lead to cycles of investments that exacerbate the economic cycle and the ups and downs of financial markets.

The ownership of these firms rests in private hands since that is the most efficient incentive tool to sustain the needed rate of technological progress to keep this system stable. The role of the regulators is very subdued: Any exaggerated activity by them would lead to a lowering of investment incentives for the private companies that then would have to fear a meddling of the regulators. The only activity they should engage in is to ensure the open access. Self-regulation by the competing providers is a means for them to differentiate themselves from competitors and attract more users and a better competitive advantage. The epicenter of power lies with the privately owned providers or with the users depending on the providers' ownership structure. The tendency for vertically integrating the securities transaction value chain is low: Cross-subsidies from one stage to the next are not possible due to the fierce competition on each stage and the reluctance for change in such a vertically integrated institution that is faced with cannibalizing its own success whenever new processes or products occur makes it a suboptimal solution.

Figure 2 provides an illustration of the industry setting as well as the position of the power center.

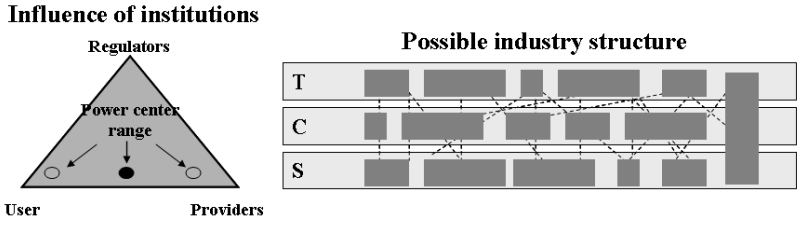

Figure 2: Competitive Fragmentation

Efficiency analysis *Static efficiency* is low in this setting. The relatively small size of the providers does not allow for the harvest of scale and scope economies. Also, positive network effects for the users remain unexploited due to the large number of smaller networks in this fragmented industry setup. Consolidation efforts exist, but a constant stream of new industry entrants armed with new, innovative products and services prevent the creation of one dominant monopoly. A system with several relatively small market participants prevails.

Dynamic efficiency on the other hand is high due to fierce competition and low entry barriers of the market. Open communication standards ensure that providers with better services will be able to offer their service to users without being strongly hampered by established providers. Even users such as banks can effectively threat providers to internalize transactions should they not be satisfied with

existing products and services. An exact configuration of the elements is necessary to keep this system stable between two countervailing forces. On the one hand, an industry setting with open standards may not provide enough incentives for the providers to develop the standardized technology further due to free-riding inducements. On the other hand, uncoordinated investments can lead to value destruction when too much is invested in boom times of a cycle. The system has therefore to strike a delicate balance between this under- and overinvestment problem.

Systemic efficiency is rather high. Although this setting has no (public) regulator but is mainly self-regulated, the robustness in the provision of securities transaction possibilities is nevertheless quite high. The driving factor is the competition among providers, in this case, the competition for the most stable and secure transaction system. Thus, infrastructure providers have an incentive to compete on quality and create a safe TCS-environment for their clients. However, an important precondition for this scenario is the transparency of the providers' risk management efforts to the users. If it is difficult for the latter to evaluate the quality of risk management, the providers may have the adverse incentive to boost profitability by cutting down on costly risk management procedures and endanger systemic efficiency. This race-to-the-bottom effect may be prevented by a user-dominated governance structure.

Another positive aspect of the competitive fragmentation on systemic efficiency is that the fragmented market structure which is characterized by high levels of infrastructure redundancies and open standards enables a relatively easy re-routing of transactions in emergency cases. Ample substitution possibilities for the users and low switching costs due to open communication standards ensure systemic robustness in times of the failure of one institution. However, depending on the nature of an adverse systemic event, contagion between the different transaction systems may occur and thus neutralize the positive redundancy aspects.

A system of competitive fragmentation is particularly interesting in a dynamically changing environment when returns on innovations are high and static efficiency considerations are dominated by dynamic efficiency aspects.

5.3 System 3 - Contestable monopolies

Description There are two crucial characteristics to the third system we propose: (1) the market for infrastructure providers is more or less consolidated and (2) communication between the industry participants - both horizontally and vertically - is performed via open standards. New entrants into the market are able to communicate with the others and are granted access to established providers. The efficient size with respect to scale and scope economies limits the number of direct competitors on each

horizontal stage and natural monopolies prevail. Also, the users benefit from the merits of a large single network. Vertical integration is rather detrimental in such a system since the monopoly positions on different stages could be used strategically by the providers to re-enforce rent extraction possibilities in other activities to the disadvantage of the users. Two possibilities exist by which such behavior can be ruled out: For one, a public regulator can ensure open access and limit any vertical integration attempts. For the other, the users themselves can mutually own the necessary infrastructure and restrict such behavior by the management of the provider in question. Depending on which concept is used to restrict the infrastructure provider from capturing too large a slice of the economic rents, the epicenter of power is somewhere between the regulators and the users.

Open standards and guaranteed access allow new entrants to enter the market and thereby limit the rent appropriation potential of the incumbent further. With better products or processes they are in a good position to challenge the incumbent and to gain the upper hand eventually. The rate of technological progress and innovation is therefore higher than in system 1 of a regulated monopoly. The investment propensity and also total investment are higher since the incumbent has to keep up with the innovative progress or risks to become obsolete and to lose his position to an upstart. Again, self-regulation can be a means of competition with the better quality and stability gaining an advantage. This allows to reduce the public regulation and the costs associated with it.

Figure 3 shows three idealized industry settings in analogy to system 1. The firms with dashed lines illustrate potential new entrants into the industry.

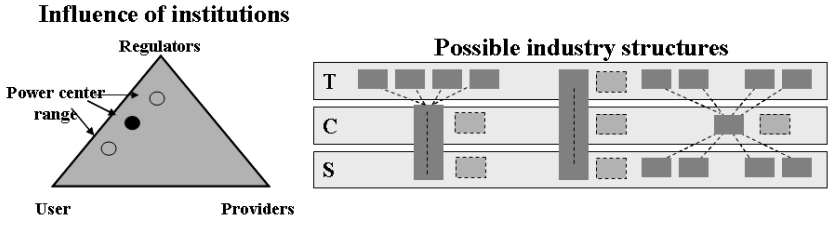

Figure 3: Contestable Monopolies

Efficiency analysis *Static efficiency* in a system of contestable monopolies is enhanced in this setting due to the high level of consolidation in the three stages of the value chain.The existence of a quasi-monopolistic infrastructure enables full exploitation of existing economies of scale and scope and network effects. Static efficiency gains are passed on to the users in this system as each level on the value chain is contestable to market entries due to open communication standards. Furthermore, the costs of regulation

can be kept to a minimum and is not distorting investment incentives for the providers.

Dynamic efficiency is also high and is achieved by the open standards architecture which results in contestability in each activity and low barriers to entry. This prevents the existing monopolist from appropriating a too large portion of the monopoly rent as potential entrants with better service offerings pose an effective threat. Nevertheless, some barriers to entry still exist such as liquidity in a certain security on the trading level, and allow a monopolist to reap rents from his dominant market position. These rents further ensure that the incumbent has a strong incentive to maintain this position and react to the incentives provided thereby. As mentioned in section 4.2, there are also some drawbacks to open communication standards with regard to dynamic efficiency aspects such as potential free-riding behavior in the development of innovations.

Systemic efficiency is high in this system. The setting benefits from its two main characteristics of being rather consolidated and having open communication standards. The former aspect enables the industry to centralize its risk management at one institution while the latter ensures competition for the most stable transaction system so that the quasi-monopolist is forced to maintain a high quality of risk management in order not to lose users to other providers. Additionally, open communication standards enable a wider proliferation of knowledge on the transmission of transaction data. As communication technology becomes generic knowledge, it is likely that market participants react faster and better on systemic emergency events. However, contestability of the market may also bring along adverse aspects such as the aforementioned race-to-the-bottom incentives for the provider. An effective no-bail-out commitment by governments or central banks may prevent the monopolist from assuming himself to be too-big-to-fail.

5.4 Comparison of the systems

The three systems described above are all consistent systems that maximize social welfare in the sense that an incremental deviation from the configuration of its elements would not lead to further improvements. As such they are better than the current, inefficient industry structure employed in the European securities transaction industry which both fails to capture the static efficiency benefits from a full-blown consolidation and the dynamic efficiency gains from competitive fragmentation. In this section we compare the three systems derived from our framework and evaluate their relative merits and drawbacks to come to a conclusion which system policy makers in the EU should strive to implement. In particular we pose the following questions:

- How robust are these three systems when small deviations from the optimal configuration occur

and how likely is a system to deteriorate into an inconsistent system that is inefficient given the micro-motives of the different institutions?

- Which of these three system dominates the other two systems when social welfare is to be maximized, i.e. which system is the global optimum?

Stability of the systems and the threat of inefficient systems The system of a *regulated monopoly* is very stable and not in danger of falling apart easily once its elements are configured in a consistent way. By its ownership of the infrastructure providing institutions or by the power it devolved to its regulating institutions the government commits credibly to stay in this system. New entrants cannot upset this system and the incumbent monopolist has only weak incentives to engage in innovative activities. The stability itself puts a positive feedback into the system since long-range planning and routinization become possible that lower the cost imposed by regulation.

The system is little prone to deteriorate into an inefficient system: In many European countries dominant regulated monopolies along the securities transaction value chain ensured that the underlying economies of scale and scope could be reaped at the domestic level without incurring too big a social welfare loss due to efficient regulation. By striking the right balance between these costs, securities trading, clearing and settlement in national markets is highly efficient, at least from a static and systemic point of view. The past has shown that such a system needs a very big shock - like the integration of formerly apart financial markets into the single European one - to overcome its inertia.

The system of *competitive fragmentation* is not very stable and small deviations from the consistent configuration can lead to a deterioration into an inefficient setting. If, for example, too much uncoordinated investments are undertaken, the problem of over-investing arises. If a bubble builds and subsequently bursts due to any sufficient macroeconomic shock, it can force these investors to sell many assets below their value. Many providers become insolvent and are forced to leave the industry. A consolidation process is started by the institutions that are in a better position.

These institutions start to consolidate horizontally to achieve greater economies of scale and increase the degree of static efficiency. They also integrate vertically to safeguard this horizontal expansion and to leverage the resulting market power. Since all institutions concentrate on getting financially sound again, the rate of technological innovation drops, new entry looks less attractive and the whole system can transire to one of the other systems or falls into inefficiency if no regulation is introduced, to keep the market contestable or if the ownership is reorganized to a more mutual structure (which is the less likely possibility). The surviving institutions can extract too much of the economic rent and their monopoly

power is not compensated by a regulating institution.

The system of *contestable monopolies* is equally hard to sustain. A monopolist on one stage of the securities transaction value chain for example, might be tempted to integrate forwards or backwards. Such a merger of two dominant monopolies might look good at first sight: By integrating the two institutions the communication between them can be streamlined and straight-through processing might be facilitated to the benefit of the users. Open access is guaranteed by the acquiring institution and formally assured by the small regulator. However, the realization in practice might look different and many potential entrants are deterred by the more entrenched position of the merged institution. This protection induces the incumbent divert its efforts towards rent-seeking and engage in investments in the 'open' communication standard that slightly favors its own business. If regulation is not adjusted accordingly towards a stronger regime, such an institution can lower the overall amount of economic rent that is generated thereby decreasing social welfare. At the same time it can gain an economic rent that is bigger than it would be in a consistently configured system at the expense of a relatively larger loss in the economic rents that the users enjoy.

Like the system of competitive fragmentation the system of contestable monopolies is prone to deteriorate when even small deviations from the consistent configuration occur. The micro-motives of the infrastructure providing institutions will generally lead to a situation in which a monopoly prevails that is entrenched by vertical integration, a proprietary communication standard and an ownership structure that places too little weight on the benefit of the users and society as a whole. Such a mixture of different configurations will not maximize overall welfare.

Evaluation of the systems and global optimum So far we have not discussed if one of the idealized system is better than the others. Calculating an exact figure for social welfare in each of the three systems is nearly impossible: Too many parameters would need to be measured and too many errors would be made in measuring the efficiency of the involved institutions. We therefore restrict ourselves to indications only. Which of the three systems might be the global optimum that dominates the others? The system of a *regulated monopoly* produces at an efficient level so that economies of scale and scope can be reaped. However, it fares poorly when dynamic aspects of efficiency are taken into account. No investment incentives are set and the cost of regulation or public ownership further decreases the overall welfare generated by this system.

The system of *competitive fragmentation* scores especially high when these aspects of dynamic efficiency are important. However, due to the small scale of the providing institutions too little of the underlying economies are utilized. The system suffers also from the coordination problem between the

		Regulated Monopoly	Competitive Fragmentation	Contestable Monopolies
Action set	Boundary decision	strong horizontal consolidation, vertical integration, STP-possibility, no need for transfer prices	low horizontal consolidation, VI detrimental: cross-subsidies and cannibalization effects	strong horizontal consolidation, VI detrimental: incentive to appropriate open standard
	Standard and accessibility	proprietary standard, closed system	open standard, open access	open standard, regulated access
	Ownership and governance	public ownership, not for profit, heavy public regulation	private ownership, for profit, self-regulation as parameter of competition	private ownership (possibly as mutual), for profit; 'regulation light' or self-regulation
Constituencies	Users	low level of decision making power (DMP), low investment propensity	low - medium level of DMP, (too) high investment propensity	high level of DMP if mutual ownership, else lower; low - medium investment propensity
	Providers	very few players (possibly only one), low – medium level of DMP, low investment propensity	many players, high level of DMP, (too) high investment propensity	very few players, Low – medium level of DMP, low – medium investment propensity
	Regulators	high level of DMP	very low level of DMP	medium level of DMP if providers are not mutually owned, else lower
Comparative analysis	Static efficiency	high due to realized economies, lessened by regulatory costs	Low due to unrealized economies	high due to realized economies, low regulatory costs
	Dynamic efficiency	very low due to lack of competition, low technological progress, high entry barriers	Very high due to competition, high technological progress, low entry barriers	high due to threat from potential competitors, Medium technological progress, Low – medium entry barriers
	Systemic efficiency	high due to centralized risk management, possible moral hazard problems	High due to redundancies and open standards, Lessened by low coordination among players (over-investment)	high due to centralized risk management; moral hazard issues unlikely due to entry threat by potential competitors
	System stability	rather robust against changes in parameters	robust as long as high degree of innovation is existent, bubble-prone	rather fragile; precise configuration necessary; regulation must ensure accessibility for potential competitors

Table 1: Comparison of the three idealized TCS-systems

many institutions so that too many duplicate and incompatible investments are undertaken. The system of *contestable monopolies* does not have these drawbacks once configured in a consistent way: The small number of institutions deploys the underlying economies of scale and scope and the limited role of the regulator ensures that these costs are kept to a minimum as well. The market stays open for new entrants so that improvements due to innovations need not to be forgone.

This 'guesstimate' leads to the conclusion that the third system of contestable monopolies is the best and should be implemented in the TCS industry in the European Union. A caveat however must be applied since this system is of rather instable nature and prone to slide down towards an inefficient system of unregulated monopoly if not adjusted in a consistent way. Table 1 summarizes the merits and drawbacks of the three systems.

6 The transformation of the securities transaction industry

The introduction of the Single European Market was a strong catalyst that upset the system of a regulated monopoly that many European countries had in place. Many features of the established system were suddenly and simultaneously changed. By simply opening the markets and leaving everything else unchanged, the result however is inefficient. Too many inconsistent configurations of important elements are in place: Too many regulators increase the costs and thereby decrease social welfare. Publicly owned or heavily regulated institutions do not have the incentives to make the right decisions. And previously

vertically integrated institutions can bar others from using parts of their infrastructure. Divergent objectives of the many regulators or even unhealthy competition between them decreases efficiency even further.

The response by many regulators was therefore to withdraw a bit and let the market mechanism work. The system in the securities industry in the European Union in the 1990s therefore had some characteristics of the system of competitive fragmentation: The rate of innovations like automated trading and the demutualization of exchanges increased dramatically and many new entrants tried to do business in the industry. The total amount of new investments was high and duplication of investments occurred in the process of battling for the dominant position in a segment of the market. The users were fully aware of the costs imposed by the incompatible communication standards between the national institutions and tried to shape the industry to their liking.

Now that the investment boom is over and the rate of technological progress has receded a bit, the securities industry in the European Union is again at a crossroads. The rate of consolidation - horizontally as well as vertically - remains high and many unsuccessful ventures were forced to close down and leave the industry. New entrants that could keep up the pressure to innovate cannot be seen. The surviving providers try to entrench their monopoly position by vertical integration and proprietary communication standards.

It is an open question how and if their rent appropriation possibilities will be countered either by tougher regulation that would put the European securities industry back to a system of a regulated monopoly - although now at the European level - or whether the users of the infrastructure can ensure together with a cut-down regulatory institution that the system of contestable monopolies can be reached which is our policy advice. It is very crucial that a consistent configuration of key parameters is achieved to avoid a system with a quasi-unregulated monopoly which might be preferred by infrastructure providers but certainly not by the users and society at large.

References

Committee of European Securities Regulators and European Central Bank (2004): "Standards for Securities Clearing and Settlement in the European Union," *Report*.

Committee on Payment and Settlement Systems and International Organization of Securities Commissions (2001): "Recommendations For Securities Settlement Systems," *Consultative Report*.

Cruickshank, D. (2003): "The Impact of the EU Financial Services Action Plan on the Regulation of the EU Securities Market," *Speech at the "Conference on the Impact of the EU Financial Services Action Plan on the Regulation of the EU Securities Market", 6^{th} March 2003, Guildhall, London*.

de Bandt, O., and P. Hartmann (2000): "Systemic Risk: A Survey," *ECB Working Paper Series*, 35, 1 – 77.

Deutsche Börse Group (2005): "The European Post-Trade Market - An Introduction," White paper, Deutsche Börse AG, Frankfurt/Main.

Deutsche Börse Group and Clearstream International (2002): "Cross-Border Equity Trading, Clearing and Settlement in Europe," .

Diamond, D. (1984): "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393 – 414.

Domowitz, I., and B. Steil (1999): "Automation, Trading Costs, and the Structure of the Trading Services Industry," *Brookings-Wharton Papers on Financial Services*.

Economides, N. (1993): "Network Economics with Application to Finance," *Financial Markets, Institutions and Instruments*, 2, 89 – 97.

Economist (2003): "Plus ça change," *The Economist*, August 16^{th}, 2003.

European Shadow Financial Regulatory Committee (2001): "Statement No. 12: Re-Plumbing European Securities Markets," .

Franck, E., and C. Jungwirth (2002): "Reconciling investors and donators - The governance structure of open source," *Universität Zürich, Working Paper No. 8*.

Giddy, I., A. Saunders, and I. Walter (1996): "Alternative Models for Clearance and Settlement: The Case of the Single European Capital Market," *Journal of Money, Credit and Banking*, 28(4), 986 – 1000.

Giovannini Group (2001): "Cross-border clearing and settlement arrangements in the European Union," Discussion paper, European Commission.

——— (2003): "Second report on EU clearing and settlement arrangements," Discussion paper, European Commission.

Harris, L. E. (1991): *Liquidity, Trading Rules, and Electronic Trading Systems*. New York University Salomon Center, New York.

Hart, O., and J. Moore (1996): "The Governance of Exchanges: Member's Cooperatives versus Outside Ownership," *Oxford Review of Economic Policy*, 12(4), 53 – 69.

Holthausen, C., and J. Tapking (2003): "Raising Rival's Costs in the Securities Settlement Industry," *European Central Bank Working Paper*.

Kahn, C. M., J. McAndrews, and W. Roberds (2003): "Settlement Risk under Gross and Net Settlement," *Journal of Money, Credit and Banking*, 35(4), 591 – 608.

Kauko, K. (2003): "Interlinking securities settlement systems: A strategic commitment," *Bank of Finland Discussion Papers*.

Köppl, T., and C. Monnet (2003): "Guess what: It's the settlements!," *European Central Bank Directorate General Research*.

Lamfalussy, A. (2001): *Final Report of the Committee of Wise Men on the Regulation of European Securities Markets*. Brussels.

Lannoo, K., and M. Levin (2001): "The securities settlement industry in the EU - Structure, costs and the way forward," *CEPS Research Report*.

Malkamäki, M. (1999): "Are there Economies of Scale in Stock Exchange Activities?," *Bank of Finland Discussion Papers*.

Malkamäki, M., and H. Iftekhar (2000): "Are Expansions Cost Effective for Stock Exchanges? A Global Perspective," *Bank of Finland Discussion Papers*.

Milgrom, P., and J. Roberts (1990): "The Economics of Modern Manufacturing: Technology, Strategy and Organisation," *American Economic Review*, 57, 511 – 528.

Milne, A. (2002): "Competition and the rationalisation of European securities clearing and settlement," *Working Paper City University Business School London*.

Rey, P., and J. Tirole (2001): "Financing and Access in Cooperatives," *Working Paper*.

Rochet, J.-C., and J. Tirole (2003): "Platform competition in two-sided markets," *Journal of the European Econommic Association*, 1(4), 990 – 1029.

Schmiedel, H., and H. Iftekhar (2003): "Do networks in the stock exchange industry pay off? European evidence," *Bank of Finland Discussion Papers*.

Schmiedel, H., M. Malkamäki, and J. Tarkka (2002): "Economies of scale and technological development in securities depository and settlement systems," *Bank of Finland Discussion Papers*.

Seifert, W. (2003): "On Europe plc and America Inc.," *The Finance Foundation Review*, 2, 75–82.

Serifsoy, B., and M. Weiss (2003): "Efficient Systems for the Securities Transaction Industry - A Framework for the European Union," *WP Series Finance and Accounting, University Frankfurt*, 120.

Shapiro, C., and H. R. Varian (1999): *Information Rules*. Harvard Business School Press.

Tapking, J., and J. Yang (2004): "Horizontal and vertical integration in securities trading and settlement," *Working Paper*.

Topkis, D. M. (1998): *Supermodularity and Complementarity*. Princeton University Press.

Van Cauwenberge, S. (2003): "New Structure for Clearing and Settlement Systems in the EU," *Financial Stability Review*, pp. 83 – 103.

van Cayseele, P. (2004): "Competition and the organisation of the clearing and settlement industry," *Working Paper*.

Weitzel, T., S. V. Martin, and W. König (2003): "Straight Through Processing auf XML-Basis im Wertpapiergeschäft," *Wirtschaftsinformatik*, 4.

Williamson, O. E. (1985): *The Economic Institutions of Capitalism*. The Free Press.