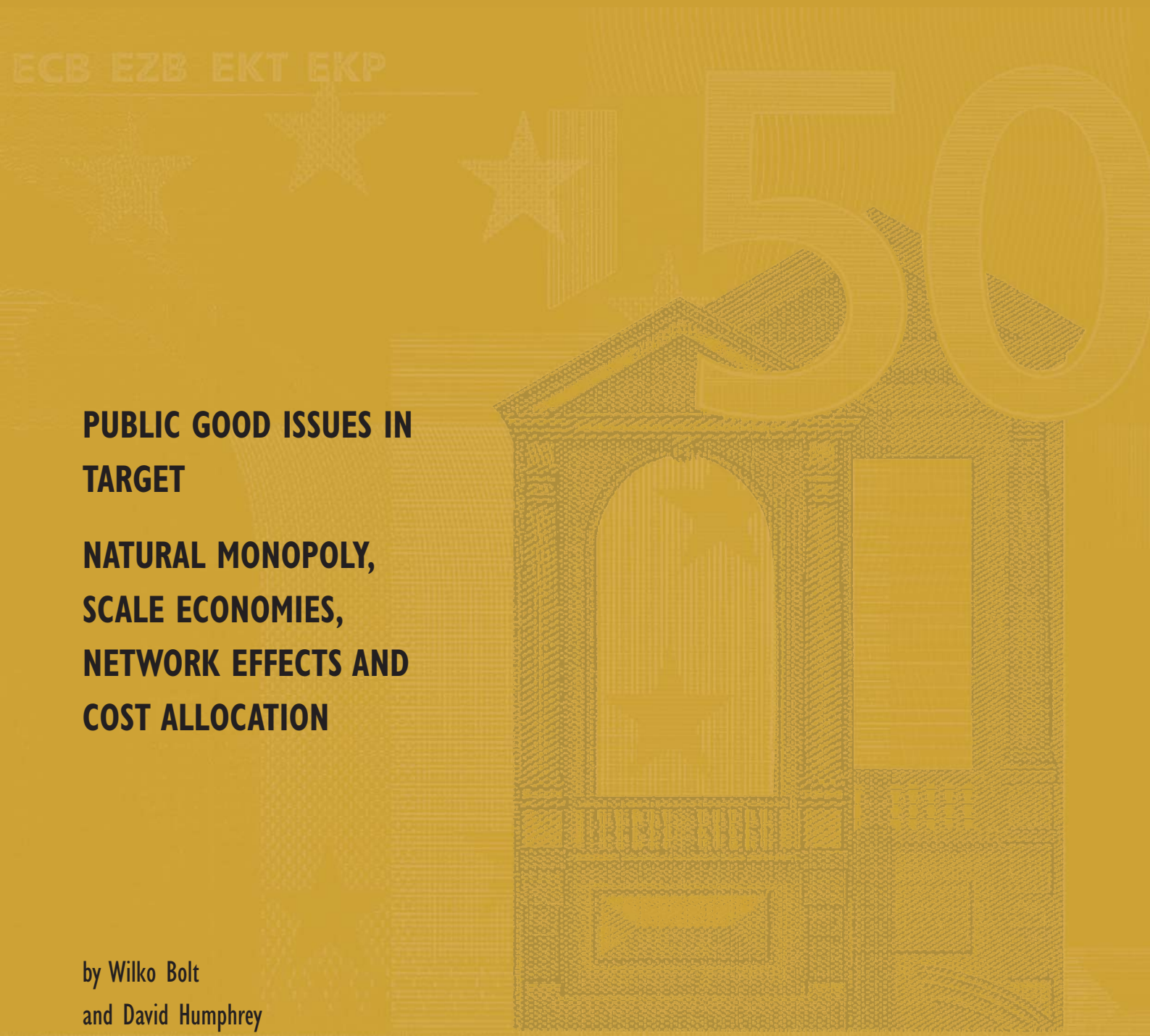
EUROPEAN CENTRAL BANK

WORKING PAPER SERIES
NO. 505 / JULY 2005



PUBLIC GOOD ISSUES IN TARGET

NATURAL MONOPOLY, SCALE ECONOMIES, NETWORK EFFECTS AND COST ALLOCATION

by Wilko Bolt
and David Humphrey

EUROPEAN CENTRAL BANK

WORKING PAPER SERIES
NO. 505 / JULY 2005

PUBLIC GOOD ISSUES IN TARGET

NATURAL MONOPOLY, SCALE ECONOMIES, NETWORK EFFECTS AND COST ALLOCATION

by Wilko Bolt [1]
and David Humphrey [2]

In 2005 all ECB publications will feature a motif taken from the €50 banknote.

1 Corresponding author: Research Division, De Nederlandsche Bank, Amsterdam, The Netherlands; e-mail: w.bolt@dnb.nl
2 Department of Finance, Florida State University, Tallahassee, FL, USA; e-mail: dhumphr@garnet.acns.fsu.edu

The Public Good Factor in TARGET2

This paper is part of the research conducted under a Special Study Group analysing various issues relevant for the design of TARGET2. TARGET2 is the second generation of the Eurosystem's Trans-European Automated Real-time Gross settlement Express Transfer system, which is planned to go live in 2007. (See http://www.ecb.int/paym/target/target2/html/index.en.html for further details on the TARGET2 project). The Special Study Group operated between spring 2003 and summer 2004. It was chaired by Philipp Hartmann, assisted by Thorsten Koeppl (both ECB). The Group was further composed of experts from the ECB (Dirk Bullmann, Peter Galos, Cornelia Holthausen, Dieter Reichwein and Kimmo Soramäki), researchers from national central banks (Paolo Angelini, Banca d'Italia, Morten Bech, Federal Reserve Bank of New York, Wilko Bolt, de Nederlandsche Bank, Harry Leinonen, Suomen Pankki, and Henri Pagès, Banque de France) and academic consultants (David Humphrey, Florida State University, Charles Kahn, University of Illinois at Urbana Champaign, and Jean-Charles Rochet, Université de Toulouse). Following the completion of the Group's work, the ECB Working Paper Series is issuing a selection of the papers it produced.

Address
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address
Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone
+49 69 1344 0

Internet
http://www.ecb.int

Fax
+49 69 1344 6000

Telex
411 144 ecb d

ISSN 1561-0810 (print)
ISSN 1725-2806 (online)

CONTENTS

Abstract

This paper discusses various theoretic concepts which play a role in assessing the public benefits of Target, the large value RTGS payment network operated by the Eurosystem. These concepts touch upon natural monopoly, network externalities, competition and contestability, as well as economies of scale and scope. The existence of a natural monopoly provides a rationale for a temporary partial or full subsidy in order for Target to achieve the 'most efficient scale' or apply the most efficient technology to lower unit costs. Such a subsidy could be implemented through temporary 'penetration' pricing. Based on empirical results for the Federal Reserve's payment system (Fedwire), it is further argued that if Target decided to standardize its operating platforms and consolidate its processing sites into one or a few centers, it too could realize strong scale economy benefits and lower unit costs.

Key words: public good, natural monopoly, most efficient scale, partial subsidy

JEL codes: G20, H41, L10

NON-TECHNICAL SUMMARY

In a rapidly changing payment landscape, a lively policy debate was recently triggered about the cost effectiveness and the quantification of a possible 'public good factor' of the large-value payment system Target. This paper discusses the various public good issues related to the benefits that accrue to an interconnected Target payment network, the implications for market competition in large value payment services, how unit costs may change as payment volume expands on a network with standardized hardware and software, and the potential usefulness of knowing for pricing purposes how costs are associated with the different services provided by Target.

Broadly, the analysis is carried out along two parallel lines. First, economic theory is used to identify relevant concepts that play a role in assessing the public good character of large-value payment systems. Theoretically, we take the position that achieving a 'most efficient scale' could justify a full or partial subsidy on Target for a limited time period, until the point at which potential scale economies are fully realized given the size of the market. Second, learning from the Fedwire consolidation and standardization experience in the eighties and nineties, in the empirical part it is argued that if Target decided to consolidate its processing sites into one or few centers, complemented by the implementation of new standardized processing technologies, it could realize strong scale benefits and markedly lower unit costs. Temporary 'penetration' pricing in the form of subsidizing Target payments could achieve these benefits more rapidly and result, once the lower unit costs are realized, in full cost pricing to users while saving resources now expended by central banks but are unrecovered in current pricing arrangements.

The argument for a subsidy is based on the fact that if full cost pricing were implemented on the current Target system, scale economy benefits either would not be realized (as a portion of current users would find alternative payment arrangements/networks) or would not be realized as rapidly (extending the period where currently high unit costs are incurred). Once lower unit costs are achieved on Target, the subsidy would be removed and full cost pricing implemented so that users would face the full costs of their payment decisions. This matches user benefits with user costs, a result which more properly allocates resources in a society.

Possible scope economies on Target between central bank monetary operations responsibilities and providing payment services for the private sector could not be determined due to a lack of the necessary cost accounting data for these two broad categories of Target services. Information that does exist, however, suggests that scope benefits may be small, as they have been found to be in empirical studies of scope economies among payment services operated by the Federal Reserve.

I. Introduction

This paper discusses the various public good issues related to the benefits that accrue to an interconnected payment network like Target, the large-value RTGS payment network for the Eurozone. We also discuss the implications for market competition in large value payment services, how unit costs may change as payment volume expands on a network with standardized hardware and software, and the potential usefulness of knowing for pricing purposes how costs are associated with the different services provided by Target.

We take the position that achieving a 'most efficient scale' could justify a full or partial subsidy on Target for a limited time period, until the point at which potential scale economies are fully realized given the size of the market. Based on the consolidation experience of Fedwire in the U.S., if Target decided to consolidate its processing sites into one or a few centers, complemented by the implementation of new standardized processing technologies, it would likely realize strong scale benefits and markedly lower unit costs. Temporary 'penetration' pricing in the form of subsidizing Target payments could achieve these benefits more rapidly and result, once the lower unit costs are realized, in full cost pricing to users while saving resources now expended by central banks but are unrecovered in current pricing arrangements.

In what follows, section II discusses concepts that touch upon natural monopoly, network externalities, competition and contestability, as well as two-sided markets since these considerations determine the public good aspects of Target. Possible lessons from the telecommunications industry are also outlined. Section III notes the existing empirical evidence suggesting that large value payment operations, when properly structured, contain elements of a natural monopoly and can experience strong scale benefits. Although data are lacking, section IV outlines ways in which costs may be allocated and prices derived for a network such as Target which provides both monetary control services for central banks as well as funds transfers for the private banking sector. Section V concludes.

II. Payment system efficiency, natural monopoly, network effects, and competition

2.1 Efficiency, natural monopoly, scale and scope

Efficiency in payment systems is not easy to define. While prices and the social and private costs of making payments are important, other aspects such as speed, convenience, safety, and reliability are contributing factors to the efficiency of payment systems. There are basically three dimensions of economic efficiency: allocative efficiency (the extent to which total welfare of consumers and suppliers is being maximized); productive efficiency (the degree to which costs of production are being minimized); and dynamic efficiency (the extent to which suppliers are able to meet the changing needs and preferences of consumers over time in the most efficient way). Whereas allocative and productive efficiency are static and focus on how to split up existing costs, dynamic efficiency refers to how costs may change in the future. These efficiency concepts are useful when assessing the public good factor of Target.

Payment systems are characterized by strong economies of scale and are often said to be natural monopolies. A natural monopoly exists if, over the relevant range of demand,

the cost function is subadditive. This means that it costs less to produce two or more outputs together than to produce each one separately [3]. Cost subadditivity is important because it sheds light on the most efficient scale of an industry and can provide a rationale for a temporary partial or full subsidy on Target. We suggest that as long as the most efficient scale has not yet been reached, a 'penetration' pricing policy in the form of a subsidy (i.e., setting prices lower than marginal costs) might be employed to boost demand for Target payments and achieve a higher, more efficient scale. After scale benefits are realized, the subsidy would be removed and prices would reflect the full cost of production (and thereby align better user benefits with user costs, resulting in an efficient allocation of resources).

In addition, scope economies may also exist among the various payments services supplied on a large-value payment network. In particular, Target currently provides both government-related services and payment services to private banks of financial institutions. It is likely that the joint costs of providing government-related and banking industry services on Target are lower than the 'stand-alone' costs of providing these services separately. Interestingly, as is shown in Tirole (1989), the concept of cost subadditivity for multiproduct firms also allows a meaningful definition of economics of scope in terms of these stand-alone costs.

2.2 Network effects, competition, and two-sided markets

Payment systems exhibit strong network externalities. These externatilities arise when a good or service is more valuable to a user the more users that adopt the same good or service. On the demand side, the size of a network is a key factor for its total value to users. However, in attracting more and more users an existing network may increase its advantage over competing networks, ultimately pushing smaller networks out of the market (called 'tipping'). Here end-users would be effectively locked-in by the single existing network.

On the supply side, compatibility, interoperability and standardization are necessary ingredients for expanding the size of a network (McAndrews, 1997). In payment systems, compatibility can be achieved by adherence to technical standards, infrastructural arrangements and interbank cooperation. Standardization and compatibility make sure that economies of scale are fully exhausted. There are, however, potential dangers. For example, in the early stages of setting up a network firms might have been coordinating on what turns out later to be a less efficient technology. Having attained a large installed base, it would be difficult and expensive to then upgrade to newer technology. As well, a dominant network provider may have an incentive to supply incompatible services in order to improve its market position. In these instances, allocative and dynamic efficiency would be reduced. Overall, network externalities and economies of scale may induce a monopolistic market structure since duplication of fixed costs is often not socially desirable.

The potential negative effects from a monopolistic market structure for large value payments can be mitigated if markets are contestable. A necessary condition for contestability to occur is the removal of significant barriers to entry to and exit from the market. In the current context of competition between Target and the private large value payment network Euro1, allowing a price subsidy for Target makes the market less

[3] See Tirole (1989) for a thorough discussion about economies of scale and cost subadditivity.

contestable and might increase economic barriers for Euro1 or others to enter in the same market.

Electronic payment systems typically have large set-up costs because of their relative capital intensity, making actual network competition hardly a viable option. Thus the ability for a potential entrant to access the existing network becomes an important issue. This type of competition--often dubbed 'platform' competition--triggers the complex issue of access pricing. An incumbent network may levy fees on new entrants for accessing its network. If these fees are too high (e.g., a fee higher than its own end-user price), then competition cannot be of any benefit to end-users as their costs would be higher with the new entrant. At the same time, access fees cannot be set too low, in order to prevent inefficient competitors to enter the market causing productive inefficiencies. The practice of setting high access fees is a form of 'raising rival costs' and may arise in the context of Target and Euro1 as Euro1 now accesses the Target network for final settlement.

The market for large-value payments is in effect a two-sided market where payees and payors (initiators and receivers) 'consume' the good (the payment). For a payment to be executed in Target, both sides need to be connected to Target. In these markets, 'platforms' coordinate the demands of these two distinct segments accounting for the interactions between them on the demand side when devising pricing and investment policies. A key aspect involves the optimal price structure which involves the division of fees between the two sides of the market to get them both on board. Optimal prices for the different segments of end-users must balance the demand among these segments--and the need for an optimal pricing structure as well as an optimal pricing level distinguishes a one-sided from a two-sided market. As Evans (2003) points out, in two-sided markets the demand for the product may completely vanish if the pricing structure is incorrect.

More precisely, Rochet and Tirole (2004) define a two-sided market as a market in which the total volume of transactions varies with the price for one side of the market while keeping the total price constant. They show that the distribution of the total fee over the two segments depends on the demand elasticities of the product.[4] Although in their framework the optimal total fee obeys the well-known inverse-elasticity rule for monopolistic price setting, it may indeed be the case that the optimal price for one side of the market will be set below its marginal cost. These two-sided pricing principles may also be important for large-value payment markets. In setting the prices for a large value payment, Target needs to get both segments on board. Currently, only one side of the market (initiators of payments) is charged for executing payments on Target. Although, on average, this pricing scheme will not result in major cost differentials between the two types of Target users (as each user is likely to be initiator and receiver about half of the time), it can affect the total realized volume within Target.

2.3 Lessons from the telecommunications industry

Payment systems require telecommunication networks. Therefore, the study of telecommunication networks can provide useful insights about the functioning and design of payment system networks. Mason and Valletti (2001) describe several main

[4] Surprisingly, at an interior solution, the most elastic side is charged the highest fee in the monopoly situation, see Rochet and Tirole (2003). Bolt and Tieman (2003) show that the reverse may be true for corner solutions.

defining characteristics of communication networks that may also apply to payment systems. The first characteristic is the specific cost structure of networks, where one typically incurs large fixed cost attached to setting up the network but faces a low marginal cost of operating the network. Strong complementarity between the different network components is the second factor, so that there are large gains to connecting two networks, but can also create a bottleneck problem. The third factor is demand externalities which, in short, states that networks are more valuable if there are more people using them. The fourth characteristic refers to social obligations, in the sense that these networks are often viewed as providing essential services. On top of these, we may add that both industries are two-sided where we distinguish payees and payors in payment systems, and senders and receivers in telecommunication networks. The combination of all these characteristics pose a real challenge to the analysis of pricing, social benefits and regulation of these networks.

The experience of the telecommunication firm AT&T in the US offers some parallels to determining the public good benefits/costs associated with Target. AT&T previously operated as a legal monopoly in providing both local and long-distance phone service. This allowed it to achieve lower unit costs due to the existence of marked scale economies in the provision of both local and long-distance phone service. AT&T was treated as if it were a natural monopoly since the public benefits of having a single producer results in the lowest cost of production. To ensure that these low costs are passed on to consumers, prices were regulated to cover all cost plus a normal return on invested capital.

At the time, AT&T was charging prices lower than costs for local service but higher than costs on long-distance. When combined, all costs were covered but the pricing structure led to challenges by potential new entrants who wished to enter the long distance market but could only do so if AT&T was no longer deemed a natural monopoly. It was determined that the cost of the joint provision of local and long distance phone services by AT&T alone was not notably lower than if these two services would be provided separately by different firms. As scope economies did not apparently importantly exist between local and long distance, even though both services experience very large scale economies, AT&T was broken up and new firms entered the long distance market. The relevance of the AT&T experience and other telecom businesses to Target is that Target produces two services--government payment services as well transaction services used by banks. The same scope economy issue faced in the telecommunication industry also applies to Target.

III. Scale economies on large-value payment networks

3.1. Estimates of scale effects

Scale economies exist on payment networks as unit cost per transaction typically falls as volume is increased. This is seen in a recent analysis showing how unit cost falls as transaction volume expands across 11 Target processing sites in 2000 (European Central Bank, 2001, page 9). The figure shown there reflects strong scale effects when processed volume varies between 1 to 5 million transactions annually but unit costs appear to be flat for volumes greater than 5 million. While 22 million transactions were processed at the largest site on Target in 2000, two other analyses have found that unit

cost continues to fall, bringing scale benefits, when annual volume exceeds the experience on Target.

An empirical study of payment costs on 21 public and private large value payment networks across 20 countries for 2001 found that the unit cost per transaction averaged $1.30 and ranged between $.29 to $3.91(Khiaonarong, 2003). This also tended to vary by region, as shown in Table 1.[5]

Table 1: Average Unit Cost Per Transaction on 21 Large Value Payment Networks (in U.S. dollars, 2001)

Region:	Number of Networks	Average Unit Cost
European Union	10	$1.54
East Asia-Pacific	9	1.15
North America	2	0.84
Total	21	$1.30

Source: Adapted from Khiaonarong (2003) Table 6, page 29.

Regressing unit or average cost (AC) on transaction volume (Vol) for these 21 large value networks gave the following results:

(1) ln AC = 3.83 - 0.58 ln Vol

where both parameters were significantly different from zero at the 99% level of confidence. Other variables added to this simple regression gave significant scale effect parameters ranging between -0.49 to -0.67. Focusing on the result in (1), this says that if transaction volume on a large value payment network were to double then average cost per transaction on average falls by 58%.[6] Taking the average cost per transaction of $1.30 across the 21 networks, a doubling of transaction volume suggests that average cost could fall by perhaps $0.75 to a level of around $0.55 if no prices rose during this period (i.e., if inflation was zero). [7] At present only the very largest Target processing sites seem to incur a unit cost of around $1.30 per transaction (others are higher). If transaction growth was 10% a year, as it was on Target over 1999-2002, it would take about 7.5 years to double transaction volume.[8] Equation (1) tells us the percent change

[5] Unit costs in local currency values were translated into U.S. dollars at market exchange rates. Had purchasing power rates been used these average values would be somewhat higher. In the range of unit costs shown in the text, one (clearly non-comparable) network in Khiaonarong (2003) was ignored.

[6] In a cross-section study such as this, the common maintained hypothesis is that a payment network with relatively low volume and relatively high unit cost would "look like" the average of payment networks with higher volumes and lower unit costs if its volume were to expand.

[7] Since (1) can be re-expressed as ln TC = 3.83 + .42 ln Vol, the cross-section relationship for 21 large value payment networks in 2001 yields a scale economy value of .42. The re-expression relies on the fact that ln AC = lnTC - ln Vol.

[8] Greater accuracy in estimating scale effects would be obtained if (1) had been specified as a log quadratic with $(\ln Vol)^2$ as an additional explanatory variable to capture better how scale benefits tend to be reduced as larger and larger volumes are attained. As the data used in Khiaonarong (2003) are confidential, this could not be done.

in average cost as volume expands across the different networks, networks with a wider range of annual transaction volumes than currently exists on Target.

A more comprehensive analysis of large value payment network scale effects concerns an effort to estimate the relative importance of three major determinants of the reduction in unit cost on Fedwire over 1979 to 1996 (Hancock, Humphrey, and Wilcox, 1999).[9] Between 1979 and 1996, Fedwire average cost per transaction fell by about 24% in nominal terms (but 62% in real terms).

The three major determinants of the reduction in Fedwire costs were: processing center consolidation, scale economies from expanded volume, and technological change which lowered processing and telecommunication costs directly. Data processing and telecommunication costs accounted for upwards of 86% of Fedwire expenses (the remaining costs were distributed among labor, building, and materials inputs). Data processing input prices (which reflected on-going technical improvements) fell by 51% over 1979-1996. Telecommunication input prices (which also reflected technical change and deregulation) rose by 38% over the same period while the prices of labor, buildings, and materials rose by 100% to 246%. The reduction in the price of data processing inputs relative to telecommunication inputs was associated with a rise in the quantity of data processing inputs relative to those for telecommunication, indicating that these two inputs are strong substitutes for one another.[10]

Simply put, large value payments can be processed in a distributed manner where there are many processing sites and relatively little use is made of telecommunications or processing can be centralized and greater use is made of telecommunications. The change in relative prices over 1979-1996 clearly favored centralized over distributed processing. Responding to price incentives, and reflecting a need to control Fedwire costs due to legislation that required the Federal Reserve to fully price its payment services, the 12 separate Fedwire processing sites in 1979 were consolidated into a single site in 1996. The cost reduction from consolidating processing operations was only partially offset by incurring greater telecommunication costs. In 1979, 36 customer service offices which handled wire transfer requests had telecommunication links with 12 processing sites. By 1996, there were 12 customer service offices (one in each Federal Reserve District) that communicated with a single processing center.

The overall cost effect of consolidation, normal volume growth, and technical change on Fedwire unit cost is illustrated in Figure 1. The jagged line shows how actual Fedwire average cost per transaction (noted on the Y-axis) varied over 75 quarterly observations during 1979-1997 (numbered 1 to 75 on the X-axis). Over this period, annual Fedwire volume expanded from 35 to 83 million. The straight line is from a simple linear regression of average cost against time while the smooth curved line is a cubic spline fitted to the same data.

[9] Unit payment costs combine processing costs and telecommunication expenses. Processing costs are composed of labor, building, and computer expenses to electronically transfer funds from one party to another while telecommunication and other associated expenses are incurred in sending and receiving payment messages from the funds transfer initiating and receiving parties.
[10] Quantities were derived by dividing nominal expenditures by the relevant technology or quality adjusted price indices.

Average costs on Fedwire fell from about $0.48 per transaction in 1979 to around $0.35 in 1996, a reduction of 26% in nominal terms. This cost reduction in nominal terms occurred even as the U.S. GDP deflator rose by 81% over this period and the cost-of-living index expanded by 116%. Thus it is clear that strong scale economies exist for large value payment networks and that Target, Euro1, and other networks could expect to experience roughly similar scale effects with increases in their transaction volume over time.

Annual transaction volume on Target is around 65 million and its largest single processing site handled about 22 million transactions in 2000. This single site figure is considerably less than the single site volume of 115 million annual transactions on Fedwire or the 64 million on CHIPS. To obtain the benefits from scale economies, Target essentially has two choices: it can realize these economies over time consistent with its historical rate of volume growth of 10% a year or it can choose to consolidate its processing operations by folding some or most of its smaller volume sites into one of the currently larger sites. In the latter case, greater scale economies will be realized more rapidly. Alternatively, Target could adopt standardized and more efficient processing technology at each of its existing sites and this too could markedly reduce unit costs (Leinonen, 2002). However, also other considerations--such as operational ease and security--need to be considered along with unit cost.[11] Of course, neither arrangement would disrupt the legal account relationship between banks and their individual central bank for settlement purposes.

The average cost per transaction on Target in 2000 was around 1.80 euro while the average cost at the two central banks with the greatest volume averaged around 1.15 euros (European Central Bank, 2001).[12] Although these figures are only approximate and preliminary, they contrast sharply with the current $0.34 cost per transaction on Fedwire and CHIPS which have centralized their processing operations.[13] CHIPS processes about three times the volume as the largest volume site on Target (64 million transactions compared to 22 million on Target) while Fedwire, which incurs greater telecommunication costs than does CHIPS, processes over five times as much (115 million compared to 22 million).[14] Since annual volume on Target is around 65 million,

[11] While the experience of Fedwire and CHIPS indicate that strong scale economies exist within a centralized processing framework, we have no similar experience to draw upon to assess the likely unit costs associated with implementing new and lower cost technology--along with related telecommunication and other expenses--within a distributed processing framework. Such a detailed and comparative study would be necessary before the ECB (and member central banks) could appropriately decide on the most cost effective configuration for Target.

[12] At early 2004 exchange rates, 1.15 euros translates into around $1.30, which was the average cost over 21 networks shown in Table 1.

[13] CHIPS and Fedwire incur about the same unit costs ($0.34 per transaction) and both split this cost equally between the paying and receiving banks. Thus fees are about $0.17 for each party (Mohr, 2003). CHIPS is jointly owned by its commercial bank users and is operated as a "payment utility" that recovers all of its costs and funds transfers over Fedwire and CHIPS are in practice strong substitutes for one another. While this suggests that the type of ownership of a large value network per se has little effect on user choice, it does support the view that competition can be important for achieving low costs per transaction.

[14] The low unit cost experienced on CHIPS is made possible by having a centralized processing facility with strong scale economies plus low telecommunication expenses (since all CHIPS transactions are made by entities located in New York City, rather than geographically distributed across the country as they are on Fedwire).

consolidating processing on Target would close to triple the volume that now exists at the largest processing site and yield considerably greater scale benefits than relying on the normal rate of volume growth of about 10% a year. By lowering overall unit costs and transaction prices, consolidation of processing centers on Target would enhance allocative and productive efficiency.

3.2 Target and the public benefits of a natural monopoly

Given the strong scale economies that appear to exist on large value networks, an argument can be made that Target be considered a natural monopoly. To provide an incentive for potential users to expand transaction volume on Target and realize its full potential for scale economies, Target need not be required--at least initially--to cover all or most of its current costs. Instead, prices could be set to equal the lower unit cost expected to be incurred once most of its scale benefits (or benefits from improved processing technology--c.f., Leinonen, 2002) have been achieved. The social trade-off would then basically be a trade-off between static and dynamic efficiency. In particular, between:

(1) The cost of a transition period which equals the discounted value of the expenses not recovered while Target expands its transaction volume at a consolidated processing site up to the limit established by the size of a market; and

(2) The benefits of a mature system which equals the discounted value of the future cost savings from lowering expenses from their current level.

This comparison puts a premium on keeping the transition period as short as possible and illustrates the importance of determining when transaction volume growth has reached the limit associated with the size of the market. It is at this point that the justification for subsidizing unit cost only to achieve unrealized scale economies no longer applies or is very weak in economic terms.[15]

At the present time, Target does not fully charge for the payment services it provides to the banking industry. Thus continuing this arrangement on Target, or only charging a low future expected price equivalent to unit cost once scale economies have been fully realized in the market (or new technology is in place), would be unlikely to result in any extra growth in payment volume over the "normal" growth it already experiences. Target has, in its present configuration of some 16 separate processing sites, already effectively reached the limit of the size of its market and future scale benefits will depend on the normal expansion of this market over time. Put differently, if Target were to continue with its current configuration of distributed processing and existing nonstandard technology, a scale economy argument for subsidizing unit cost does not apply as the scale benefits associated with the size of the market have already been realized (although unit costs are higher than they would otherwise be with distributed processing and standardized technology).[16]

[15] However, Target could continue to be subsidized if the provision of its payment services were determined to have strong public benefits in terms of ensuring financial market stability. This, effectively, has been the judgement the Bank of Japan made for its large value payment network. This consideration is developed more fully in Angelini and Maresca (2004) and Pagès and Humphrey (2004).

[16] Put in terms of the infant industry argument often used to justify tariffs or quotas adopted to protect a domestic industry until a scale of operation and lower costs are achieved to allow it to compete with foreign imports, Target--in its current configuration--would not be considered to be

As Target has already reached the limit of its market share under its current configuration, a justification for continued subsidization on scale economy grounds alone can only reasonably be made if:

(a) The scale benefit of having only one supplier for all large value payments is so strong that the provision of large value payments is deemed to be a regulated monopoly (similar to arguments that have in the past supported regulated monopolies for electricity and telephone services); or

(b) Target decides to consolidate its processing sites into one or a few centers to realize greater scale benefits (and lower costs). At least initially, these consolidated sites would not yet have realized the limit imposed by the size of the market since volume growth there would exceed normal, historical rates during a transition period.

There is no international agreement on the issue in (a). In the U.S., with two suppliers of large value payments (CHIPS and Fedwire), the regulated monopoly argument has been rejected in favor of competition versus greater scale benefits. In Japan this argument has been accepted and implemented with the full subsidization of their single large value network. Thus (a) is a judgement call on the part of policy makers who could point to arguments on both sides of this issue--but not rely on scale economies alone to make the case one way or another.

The situation in (b) could be used to justify full or partial subsidization on Target for a limited time period. Policy makers could decide to consolidate processing sites to achieve more rapidly scale economies compared to what the normal expansion of the payments market would bring. Here subsidization could be justified using the above criteria that compares the discounted cost of a subsidized transition period with the discounted cost savings obtained when a mature, consolidated system is achieved or when new standardized and lower cost technology is in place in a distributed processing system.

IV. Cost recovery and pricing on Target

4.1 Pricing policy and cost recovery with scale economies

Marginal cost pricing--where the extra benefit to the user of a payment service is equated with the extra cost to the supplier--achieves allocative efficiency and avoids cross-subsidization.[17] Despite its theoretic soundness, marginal cost pricing is often difficult to implement since marginal costs are hard to accurately measure and, as well, will not fully recover costs when scale economies exist. While average cost pricing implemented with a single fee can recover all costs, it would likely discourage users to utilize fully the scale benefits that are available.

A reasonable departure from marginal cost pricing is 'two-part pricing' consisting of a fixed price to recover the average fixed cost of producing payment services (reflecting processing economies of scale) and a variable price or transaction-based fee to recover

---

an "infant industry" since it has already "grown up" and achieved the limit of scale benefits associated with the current size of the market under its current configuration.

[17] However, in two-sided markets the 'user's pay principle' ceases to hold for the individual market sides, rather one should compare total price with joint marginal costs (Wright, 2003).

average variable costs. Here the benefits of high volume operations are passed on to those high volume users that generate the economies of scale and the average price of the payment service decreases with the number of units demanded. Two-part tariffs correspond to a quantity discount and also assures that the degree of cross-subsidization between different end-users is limited. A two-part pricing approach would resolve most of the scale problems and is indeed a standard pricing approach where fixed costs are large but where all costs may need to be recovered.[18]

4.2 Pricing policy and cost recovery with scope economies

Target currently provides both government-related services and payment services to private banks or financial institutions. Government-related services involve: (1) the maintenance of reserve accounts with the banking system for final settlement of net positions of retail or large value payment transactions made on other payment networks; (2) monetary control payment transactions in financial markets; and (3) other government payment functions. Bank-related services involve the processing of large value RTGS payment transactions. These two services are jointly produced as they likely share some of the costs of maintaining accounts, computer processing of payments, and telecommunication facilities between banks and the central bank. This suggests the existence of scope economies on Target. Here the sum of the cost of a stand-alone payment network providing only central bank and government-related payment services plus the cost of a second stand-alone network that only provides payment services to the banking industry would likely be more expensive than a single network that provides both services jointly.

When scope economies exist, the allocation of joint production costs need not rely on a simple accounting rule-of-thumb to distribute the joint cost of the two main service categories on Target. Indeed, the cost that is allocated to the provision of payment services for the banking industry can be the additional cost incurred for processing, telecommunication, and site operations after the stand-alone cost associated with providing government-related services is determined. All costs would be recovered as the government would cover the costs it incurs (the stand-alone government service cost) and the banking sector would cover the cost it incurs (the additional cost).[19]

The relative importance of the stand-alone costs of providing only government-related payment services on Target is roughly indicated by how large the share of government-related payment volume is. This share averages less than 1% across 16 Target processing sites with even lower values at sites with the largest total volumes. Even if the 1% government volume is associated with 10 times its share of processing equipment, telecommunication connections, labor input, and site facility cost--so that 10% of the cost on Target could be allocated to stand-alone government costs, this would still leave 90% of total costs to be recovered from private sector users. In sum, the scope economy argument for cost allocation could be applied but its impact is likely to be relatively small.

[18] Holthausen and Rochet (2002) consider efficient two-part pricing of large-value payments in the presence of unobservable heterogeneity about banks' future payment volumes.
[19] This approach would need to be refined with the assistance of technical experts who have a detailed knowledge of the internal cost structure and operations of Target.

We know of no study that has attempted to determine the possible scope economies among the various payment services offered on a large value network. However, there have been a few empirical studies that have attempted to determine scope economies among a broader range of payment and banking services. Importantly, when a proper functional form is used, these investigations have found little support for the existence of significant scope economies either within the commercial banking industry or between different payment services provided by the Federal Reserve.[20] This suggests that, aside from general managerial overhead and some likely sharing of infrastructure costs, scope effects are more likely to be small than very large for the central banks that comprise Target. A more definitive answer here will have to wait until the requisite cost accounting data on Target becomes available.

V. Conclusion

This paper discusses various theoretic concepts which contribute to assessing the public benefits of Target. These concepts touch upon natural monopoly, network externalities, competition, as well as two-sided markets and economies of scale and scope. The results of an earlier econometric analysis of the effects on unit cost from standardization and consolidation of the Federal Reserve's large value payment network (Fedwire) largely comprise the empirical part of this paper. Based on this analysis, it is argued that if Target decided to standardize its operating platforms and consolidate its processing sites into one or a few centers, it too could realize strong scale economy benefits and lower unit costs.

The concept of a natural monopoly provides a rationale for a temporary partial or full subsidy in order for Target to achieve the 'most efficient scale' and lower unit costs. Such a subsidy can be implemented through temporary 'penetration' pricing (i.e., pricing at less than full current cost). When the potential scale benefits are fully realized, given the size of the market, the subsidy would be removed and full cost pricing implemented. After unit costs have been lowered and reflected in the prices charged, users would face the full costs of their payment decisions. This would permit them to better match benefits with actual costs resulting in a more proper and efficient allocation of resources than occurs today on Target.

Possible scope economies on Target between central bank monetary control responsibilities and providing payment services for the private sector could not be determined due to a lack of the necessary cost accounting data for these two broad categories of Target services. Information that does exist, however, suggests that scope benefits may be small, as they have been found to be in empirical studies of scope economies among payment services operated by the Federal Reserve.

[20] Little to no scope effects were found between Federal Reserve ACH and wire transfer operations (Adams, Bauer, and Sickles, 2002) and only weak scope effects were identified between U.S. commercial bank deposits and loans (Pulley and Humphrey, 1993) where most had expected to see a strong relationship.

Bibliography

Adams, R., P. Bauer and R. Sickles, (2002), “Scope and Scale Economies in Federal Reserve Payment Processing”, Working Paper 02-13, Research Department, Federal Reserve Bank of Cleveland, November.

Angelini, P., and G. Maresca (2004), "Do Intraday Liquidity-Externalities Warrant Subsidization of Target?", European Central Bank.

Bolt, W. and A.F. Tieman, (2003), "Pricing Debit Card Payment Services: An IO Approach", IMF Working Paper 03/202.

European Central Bank, (2001), "Overview of Target Cost Figures for 1999 and 2000", June 1, Confidential document.

Evans, D., (2003), "The Antitrust Economics of Multi-Sided Platform Markets", Yale Journal on Regulation, 20: 325-382.

Hancock, D., D. Humphrey, and J. Wilcox, (1999), "Cost Reductions in Electronic Payments: The Roles of Consolidation, Economies of Scale, and Technical Change", Journal of Banking and Finance, 23: 391-421.

Holthausen, C., and J-C. Rochet, (2002), "Efficient Pricing of Large-Value Interbank Payment Systems, ECB Working Paper 184.

Khiaonarong, T., (2003), "Payment Systems Efficiency, Policy Approaches, and the Role of the Central Bank", Bank of Finland Discussion Papers, 1/2003, Helsinki.

Leinonen, H., (2002), "Network-Based Payments and e-Settlement--A Long-Term Perspective", Bank of Finland Bulletin, 4: 29-35.

Mason, R., and T. Valletti, (2001), “ Competition in communication networks: pricing and regulation”, Oxford Review of Economic Policy, 17: 389-415.

McAndrews, J., (1997), “Network Issues and Payment Systems”, Federal Reserve Bank of Philadelphia Business Review, Nov.-Dec.: 15-24.

Mohr, J., (2003), “Personal Communication on the Average Cost of a Large Value Payment Transaction on CHIPS", e-mail, August 15.

Pagès, H., and D. Humphrey (2004), "Settlement Finality as a Public Good in Large-Value Payment Systems", European Central Bank.

Pulley, L., and D. Humphrey, (1993), "The Role of Fixed Costs and Cost Complementarities in Determining Scope Economies and the Cost of Narrow Banking Proposals", Journal of Business, 66: 437-462.

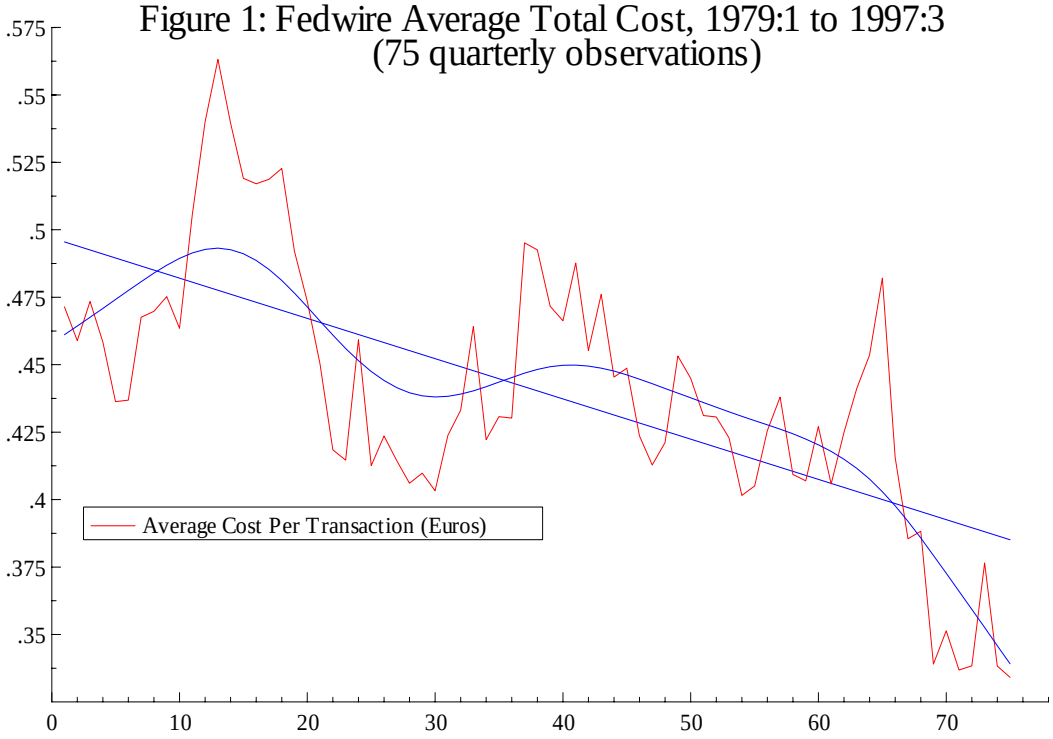
Rochet, J-C. and J. Tirole, (2003), "Platform Competition in Two-Sided Markets", Journal of the European Economic Association, 1: 990-1029.

Rochet, J-C. and J. Tirole, (2004), "Defining Two-Sided Markets", Mimeo, Toulouse.

Tirole, J., (1989), The Theory of Industrial Organization, MIT Press, Second Edition.

Wright, J., (2004), "One-Sided Logic in Two-Sided Markets", Review of Network Economics, 3: 42-63.

Figure 1: Fedwire Average Total Cost, 1979:1 to 1997:3
(75 quarterly observations)



European Central Bank working paper series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (http://www.ecb.int)

466 "Regulated and services' prices and inflation persistence" by P. Lünnemann and T. Y. Mathä, April 2005.

467 "Socio-economic development and fiscal policy: lessons from the cohesion countries for the new member states" by A. N. Mehrotra and T. A. Peltonen, April 2005.

468 "Endogeneities of optimum currency areas: what brings countries sharing a single currency closer together?" by P. De Grauwe and F. P. Mongelli, April 2005.

469 "Money and prices in models of bounded rationality in high inflation economies" by A. Marcet and J. P. Nicolini, April 2005.

470 "Structural filters for monetary analysis: the inflationary movements of money in the euro area" by A. Bruggeman, G. Camba-Méndez, B. Fischer and J. Sousa, April 2005.

471 "Real wages and local unemployment in the euro area" by A. Sanz de Galdeano and J. Turunen, April 2005.

472 "Yield curve prediction for the strategic investor" by C. Bernadell, J. Coche and K. Nyholm, April 2005.

473 "Fiscal consolidations in the Central and Eastern European countries" by A. Afonso, C. Nickel and P. Rother, April 2005.

474 "Calvo pricing and imperfect common knowledge: a forward looking model of rational inflation inertia" by K. P. Nimark, April 2005.

475 "Monetary policy analysis with potentially misspecified models" by M. Del Negro and F. Schorfheide, April 2005.

476 "Monetary policy with judgment: forecast targeting" by L. E. O. Svensson, April 2005.

477 "Parameter misspecification and robust monetary policy rules" by C. E. Walsh, April 2005.

478 "The conquest of U.S. inflation: learning and robustness to model uncertainty" by T. Cogley and T. J. Sargent, April 2005.

479 "The performance and robustness of interest-rate rules in models of the euro area" by R. Adalid, G. Coenen, P. McAdam and S. Siviero, April 2005.

480 "Insurance policies for monetary policy in the euro area" by K. Küster and V. Wieland, April 2005.

481 "Output and inflation responses to credit shocks: are there threshold effects in the euro area?" by A. Calza and J. Sousa, April 2005.

482 "Forecasting macroeconomic variables for the new member states of the European Union" by A. Banerjee, M. Marcellino and I. Masten, May 2005.

483 "Money supply and the implementation of interest rate targets" by A. Schabert, May 2005.

484 "Fiscal federalism and public inputs provision: vertical externalities matter" by D. Martínez-López, May 2005.

485 "Corporate investment and cash flow sensitivity: what drives the relationship?" by P. Mizen and P. Vermeulen, May 2005.

486 "What drives productivity growth in the new EU member states? The case of Poland" by M. Kolasa, May 2005.

487 "Computing second-order-accurate solutions for rational expectation models using linear solution methods" by G. Lombardo and A. Sutherland, May 2005.

488 "Communication and decision-making by central bank committees: different strategies, same effectiveness?" by M. Ehrmann and M. Fratzscher, May 2005.

489 "Persistence and nominal inertia in a generalized Taylor economy: how longer contracts dominate shorter contracts" by H. Dixon and E. Kara, May 2005.

490 "Unions, wage setting and monetary policy uncertainty" by H. P. Grüner, B. Hayo and C. Hefeker, June 2005.

491 "On the fit and forecasting performance of New-Keynesian models" by M. Del Negro, F. Schorfheide, F. Smets and R. Wouters, June 2005.

492 "Experimental evidence on the persistence of output and inflation" by K. Adam, June 2005.

493 "Optimal research in financial markets with heterogeneous private information: a rational expectations model" by K. Tinn, June 2005.

494 "Cross-country efficiency of secondary education provision: a semi-parametric analysis with non-discretionary inputs" by A. Afonso and M. St. Aubyn, June 2005.

495 "Measuring inflation persistence: a structural time series approach" by M. Dossche and G. Everaert, June 2005.

496 "Estimates of the open economy New Keynesian Phillips curve for euro area countries" by F. Rumler, June 2005.

497 "Early-warning tools to forecast general government deficit in the euro area: the role of intra-annual fiscal indicators" by J. J. Pérez, June 2005.

498 "Financial integration and entrepreneurial activity: evidence from foreign bank entry in emerging markets" by M. Giannetti and S. Ongena, June 2005.

499 "A trend-cycle(-season) filter" by M. Mohr, July 2005.

500 "Fleshing out the monetary transmission mechanism: output composition and the role of financial frictions" by A. Meier and G. J. Müller, July 2005.

501 "Measuring comovements by regression quantiles" by L. Cappiello, B. Gérard, and S. Manganelli, July 2005.

502 "Fiscal and monetary rules for a currency union" by A. Ferrero, July 2005

503 "World trade and global integration in production processes: a re-assessment of import demand equations" by R. Barrell and S. Dées, July 2005.

504 "Monetary policy predictability in the euro area: an international comparison" by B.-R. Wilhelmsen and A. Zaghini, July 2005.

505 "Public good issues in TARGET: natural monopoly, scale economies, network effects and cost allocation" by W. Bolt and D. Humphrey, July 2005.