

## Payment Network Scale Economies, SEPA, and Cash Replacement

WILKO BOLT

Research Department, De Nederlandsche Bank, The Netherlands

DAVID HUMPHREY \*

Department of Finance, Florida State University, Tallahassee, FL, U.S.A.

### Abstract

The goal of SEPA (Single Euro Payments Area) is to facilitate the emergence of a competitive, intra-European market by making cross-border payments as easy as domestic transactions. With cross-border inter-operability for electronic payments, card transactions will increasingly replace cash and checks for all types of payments. Using different methods, we estimate card and other payment network scale economies for Europe. These indicate substantial cost efficiency gains if processing is consolidated across borders rather than “piggy-backed” onto existing national operations. Cost reductions likely to induce greater replacement of small value cash transactions are also illustrated.

## 1 Introduction

An important precondition for the emergence of a competitive, intra-European market is the ability of consumers and others to make cross-border transactions as easily as they now make domestic payments. To do this at lower cost requires the development of “...common instruments, standards, and infrastructures in order to foster substantial economies of scale” in electronic payment networks (European Central Bank, 2006). In this effort, electronic payments – specifically card transactions, credit transfers and direct debits – are the payment instruments of choice given the cost and other difficulties of using cash or checks. The promotion of a Single Euro Payments Area (SEPA) with cross-border inter-operability for electronic payments will, depending on realized cost efficiencies, make card transactions increasingly attractive to replace cash and checks for intra-European, as well as national payments. Our goal is to estimate card and other payment network scale economies for Europe and outline the potential cost efficiencies from consolidating card processing operations across-borders rather than “piggy-backing” onto current operations

---

\* Contact Author. David Humphrey, Department of Finance, Florida State University, Tallahassee, FL 32306-1042; (850)-644-7899. Email: [dhumphrey@cob.fsu.edu](mailto:dhumphrey@cob.fsu.edu) Comments by two anonymous referees, Santiago Carbo, Pierre Lafourcade, and Leo Van Hove are gratefully acknowledged, as well as data assistance by Magnus Willeson, Peter Keus, Philip Molyneux and Bent Vale.

at the national level (with likely little reduction in expense). Such cost efficiencies, if realized, will facilitate the on-going substitution of card payments for low value cash transactions, changing the future structure of consumer payments.

Most payments in Europe are currently domestic, with little cross-border activity, and electronic transactions are typically cheaper to produce and accept than paper-based instruments. For bill payments, this holds for both banks producing electronic giro payments and billers accepting them. At the point of sale, debit cards have largely replaced checks in many European countries (with France and the U.K. being the exceptions) and they continue to replace cash for medium value transactions. However, while electronic payments are cheaper for banks to produce than their paper-based substitutes (checks and cash over the counter), merchants often find it cheaper to accept cash, especially for smaller value transactions (Brits and Winder, 2005, Table 4.3; Garcia-Swartz, Hahn and Layne-Farrar, 2006, Tables 2 and 3). As a result, banks and other suppliers offered potentially lower cost stored value cards to replace cash for small value transactions.

At present, consumer adoption and use of stored value cards seems stalled at a relatively low level of market penetration (Van Hove, 2006). Although data are incomplete, stored value cards account for only € 1.2 billion in payments across 11 European countries in 2004. In contrast, the value of debit card transactions is estimated to be €1,146 billion while the value of cash withdrawals (a proxy for cash use) was €2,189 billion.<sup>1</sup> Overall, card payments comprised 34% of the total, cash withdrawals accounted for 66%, while stored value payments were only 0.04%. An important drawback of stored value cards is that consumers may have to carry two cards to replace cash – a debit card plus a stored value card – and the latter requires “filling” at terminals while the former does not. While convenience is enhanced if both technologies are on a single card and if merchants have a single terminal that can handle both types of transactions, banks often charge an extra fee to handle a stored value transaction.

There are two payment policy issues associated with the level of card and other transaction network scale economies. First, although the SEPA goal of making cross-border payments as easy as domestic transactions could be largely achieved by simply piggy-backing cross-border card transactions onto current national processing operations of Visa, MasterCard and other (more limited) card schemes, this is unlikely to result in any significant cost reduction. Cross-border transaction volume is currently low and these transactions require extra telecommunications expense so even if processing scale economies are large (as we show below) unit costs would not fall by much and could even rise. Lower cross-border, as well as domestic unit payment costs, however, would be achievable by consolidating processing operations across borders. A second policy issue is the replacement of cash for smaller value point of sale transactions where, currently, cash is cheaper and stored value cards have had only limited success. Here too, the realization of card scale economies that significantly reduce debit card unit costs would be important, especially to induce merchants to expand card terminal access (a precondition for consumer use).

In what follows, Section 2 presents estimates of payment network scale economies using public payment cost data for Norway over a short time period and also from recent

---

<sup>1</sup> Debit and credit card data are combined here since separate data are often unavailable and credit card values are typically very small compared to the U.S. Cash use is understated as it is mostly the value of ATM cash withdrawals, although a few countries report cash withdrawn over the counter at financial institutions (but only one country reported “cash-back” at the point of sale).

point estimates of payment costs in the Netherlands and Belgium. In Section 3, an econometric model of bank payment, ATM, and branch costs using panel data over 1987-2004 is specified for 11 European countries. Scale economy results are reported in Section 4 and concern both point of sale and bill payment transactions for each of the 11 countries. A distinction is made between simple sum and realized scale effects, the former corresponding to a proportional increase of all outputs (the usual assumption) and the latter a weighted sum based on the observed increase – a more accurate indicator of overall bank scale economies. In Section 5, scale economy estimates are paired with approximate current unit cost estimates of cash, debit card and stored value card payments to determine the possible time frame and context for debit card costs to fall sufficiently to likely replace a portion of small value cash and stored value transactions. Our conclusions are presented in Section 6.

## 2 Payment scale economies: Three countries

Detailed, publicly available, cost data on payments are hard to find. Only Norway has been collecting bank payment cost data over the last decade and making it available in summary form in annual reports. As the efficiency of a country's payment system has become an emerging policy issue, special surveys on retail payment costs were undertaken for the Netherlands and Belgium. Combined with econometric modeling using newly available data on national payment volumes and bank operating expenses, these different sources provide some initial estimates of payment scale economies.

### 2.1 Norway

The total estimated bank-incurred cost and number of EFTPOS transactions for Norway (due almost exclusively to debit cards) are shown in Table 1, along with average cost per transaction. The 47% reduction in average card costs between 1988 and 1994, and the 44% reduction between 1994 and 2001 are substantial and imply scale economies of 0.49 and 0.43 (computed from  $(\text{percent change in cost})/(\text{percent change in volume})$ ).<sup>2</sup> However, the total and average cost figures are in nominal terms and the cost of living index in Norway rose by 19% and 18% respectively, in these two periods so the true scale values are likely lower, indicating greater scale benefits. This is because scale economies should be a function of the number of transactions, holding input prices and changes in technology constant.

If we knew the share of labor expenses in total bank card costs and assumed (as is likely) that labor expenses rose in parallel with the rise in the cost of living index, we could then subtract the estimated rise in labor costs from total card costs and derive a more accurate scale estimate. In principle, this should be done for capital and material inputs as well. Although this is not possible with the data at hand, we can conclude that card scale economies in Norway are greater than those shown in Table 1 and the true scale statistics are less than 0.49 or 0.43. A complicating factor is that scale economies can vary as processing volumes become large, lowering average fixed costs and Norway with a

---

<sup>2</sup> As noted by a referee, an arc scale elasticity would usually be more appropriate given the length of time covered and the large changes in payment volumes. However, since point elasticities – even for long periods – are typically the way scale economies are measured and presented in the literature, the point elasticity measure is shown in Table 1.

population of 4.6 million can experience a different scale economy than larger European countries with greater volume.

	1988	1994	2001
EFTPOS Total Card Costs	NOK 51.85 mil	NOK 325.35 mil	NOK 822.75 mil
EFTPOS Card Transactions	6.1 mil	72.3 mil	329.1 mil
Average Cost per Transaction	NOK 8.50	NOK 4.50	NOK 2.50
Implied Card Scale Economy	1988-1994: 0.49	1994-2001: 0.43	

Sources: Gresvik and Øwre (2002) and Øwre (personal communication).

**Table 1: Norway: Bank Card Costs, Volume and Scale Economy**

## 2.2 Netherlands and Belgium

Two recent detailed analyses of payments in the Netherlands (Brits and Winder, 2005) and Belgium (Quaden, 2005) have developed estimates of bank and merchant costs of providing and accepting different payment instruments. The average total cost and incremental variable cost of using cash, a debit card and a stored value card at the point of sale are presented in Table 2. As economies of scale can be expressed as the ratio of marginal to average cost, the cost figures in the table provide an implied estimate of scale economies for these three payment instruments.<sup>3</sup> The implied scale economy for cash ranges between 0.37 in the Netherlands to 0.25 in Belgium. The implied scale economy for debit cards in both countries is 0.39 while that for stored value cards is quite low at 0.04 to 0.16, indicating that average fixed cost here is quite high as transaction volume is low. These debit card scale economies are not very different from PIN and signature debit card scale values of 0.31 and 0.39 respectively, for card issuing banks in the U.S., derived from First Annapolis Consulting (2006). (See Cheney, 2006, for additional U.S. debit card cost information.) Similar payment scale estimates were also obtained for the Netherlands using confidential individual bank payment data (Bolt and Humphrey, 2007).

It is perhaps reassuring that the implied scale economy estimates for three countries (Norway, Netherlands and Belgium) are not too dissimilar. Even so, these three countries are small relative to Germany, France and the U.K. – the other countries we cover below – and our cross-country econometric scale estimates may well differ. As well, since the bank portion of payment costs likely includes some allocated expense from maintaining a branch network, as opposed to the back office processing of payments, scale economies associated only with payment processing (holding branches constant) are likely to be different. This is not an issue if these scale values were used to approximate future unit costs as payment volume expands within a given country (since branch and ATM networks may change as well). It would be something to consider, however, if the way that volume grows and unit cost falls is by consolidating back office processing centers among countries where expenses associated with front office payment expenses (for example, branch operations) would be little affected.

<sup>3</sup> Scale economies (SCE) equal  $(\partial OC/OC)/(\partial Q/Q)$  where  $OC$  is operating cost and  $Q$  is a measure of output. Thus,  $SCE = (\partial OC/\partial Q)/(OC/Q) = \text{marginal cost divided by average cost}$ , where incremental cost in Table 2 approximates marginal cost. As bank interest expenses should be unaffected by changes in payment volume, operating cost – not total cost – is the appropriate frame of reference.

	Cash	Debit Card	Stored Value
Netherlands, 2002			
Average Total Cost	€0.300	€0.489	€0.931
Incremental Cost	0.112	0.190	0.033
Implied Scale Economies	0.37	0.39	0.04
Belgium, 2003			
Average Total Cost	0.53	0.55	0.54
Incremental Cost	0.133	0.214	0.084
Implied Scale Economies	0.25	0.39	0.16

Sources: Brits and Winder (2005), Table 4.3; Quaden (2005), Table 3.

**Table 2: Total average cost, incremental cost, and implied scale economies: The Netherlands and Belgium**

### 3 Payment scale economies: A cross-country framework

Along with the payment survey information noted above, econometric analysis within a cost function framework can be used to estimate payment scale economies. Combining aggregate payment volume data with country-level banking cost information, we employ a cross-country panel data set to estimate payment scale economies for 11 European countries.

#### 3.1 Our approach

Determining scale economies for financial institutions – much less payment scale effects – has been difficult due to a lack of appropriate data. Instead of measuring the flow of banking payment, deposit account maintenance, security transaction activity and loan initiation and monitoring services directly, inferences on how costs may vary by size of bank and volume of service flow are typically obtained by relating total operating plus interest expenses across banks and over time to the value of their stock of loans, securities and (sometimes) deposits, or some other combination of on- or off-balance-sheet positions. In addition, information does not normally exist regarding the adoption of specific technical and other cost-saving innovations in banking and the default has been to assume that unknown technical change occurs linearly (or quadratically) with the passage of time and/or is somehow associated with (embodied in) the cost share or price of particular inputs.

We adopt a different approach which we feel can provide more solidly-based scale economy estimates. Specifically, we relate bank operating (not total) costs to measurable physical characteristics of banking output associated with payment processing and service delivery levels and mix. In this manner, we focus on those activities and expenses directly associated with the provision of payment services. Interest expenses paid to depositors and with a mark-up charged to borrowers are functionally separable from these activities. This approach allows us to determine how the level and mix of payment activities, along with the number of ATMs and bank branches, are directly associated with the size of a bank and its labor, capital, and materials operating cost from which scale economies may be

approximated.<sup>4</sup> In this regard, our approach represents an alternative and more specific way to identify the likely effect on costs from technical change in banking.<sup>5</sup> As point of sale and bill payment transactions are jointly processed in the deposit accounting function while aspects of service delivery are jointly produced via branches and ATMs, these two activities can be considered functionally separable.

A limitation of this approach is that our scale and technical change results are based on historical data. Our long time period covers payment technologies that initially were partially manual but end up being close to fully electronic (Leinonen, 2004, and personal communication August 13, 2006). As a result, evaluating our model at the mean of the data will likely understate cost reductions available with the very newest technology. However, evaluating the model at the beginning, middle, and end of our time period (as we do below) should capture the effect of a changing technology on operating cost. Another limitation is that costs statistically allocated to bank branch operations will cover a level and mix of non-payment services that will differ across countries, so branch scale effects should be viewed with caution (especially since a standardized branch measure had to be used).

### 3.2 Cross-country data

Payment volume (or value) data either do not exist for individual banks or are not publicly available (Norway is the only exception). However, annual aggregate payment information by country is available from European Central Bank and European Monetary Institute (“Blue Book”) as well as Bank for International Settlements (“Red Book”) publications. Thus, our data set consists of country-level panel data.

Important trends in these data are illustrated in Table 3 over 1987-2004 (18 years). As the euro did not exist for most of our time period, all value data are in U.S. dollars using purchasing power parity exchange rates (International Monetary Fund, 2006). The countries in Table 3 are ranked according to the value of their banking sector’s total operating cost in 2004 (Column 1 – the ranking is similar for 1987). The percent change in the ratio of operating cost (OC) to total assets (TA) between 1987 and 2004 is shown in Column 2 and is suggestive of how the cost of providing banking services has fallen over time (for example, an apparent 40% fall for Germany but no reduction for France).<sup>6</sup> For all

---

<sup>4</sup> The provision of deposit and loan services not directly associated with the number of payments or ATMs should be associated with the number of banking branches in a country. While this is similar to the common assumption made in the literature where the stock of the value of deposits, loans or assets is assumed to reflect the underlying flow of associated banking services, it is necessitated by a lack of data. The R-square value between the number of (standardized) branch offices and the value of total banking assets across our 11 countries and over time is 0.74.

<sup>5</sup> Our panel data set is a cross-section of 11 countries over 18 years. The standard way to identify scale effects from technical change is to presume that cross-section variation identifies scale while time-series variation identifies technical change. A problem is that payment volume tends to rise at a fairly constant rate over time while the simple passage of time itself is used as the (unknown) index of technical change. This collinearity problem is reduced when the time pattern of important technical changes is known (which rarely occurs).

<sup>6</sup> For an individual bank, the denominator of the OC/TA ratio can expand via purchased funds and generate a smaller rise in operating costs than if an asset expansion was funded by produced deposits using branches. As our data includes the operating cost of those banks producing the deposits that are sold as purchased funds to other banks, this potential bias does not exist. However, a rise in the share of assets held as securities as a bank grows (lowering the loan share) will lower the OC/TA ratio. This is a common occurrence over the business and interest rate cycle and typically reverses itself over time.

countries together, shown at the bottom of the table, the ratio of operating cost to total assets fell by a third (-34%).<sup>7</sup> At the same time, point of sale transactions rose by 140%, bill payments by 151%, ATMs by 434% but branches only by 9.8%. One thread tying these changes together is that ATMs, compared to branches, are a more cost-efficient way to deliver cash to depositors while debit cards reduce the need to provide cash in the first place. Both developments reduce bank costs.

	Operating Cost (2004, Mil, PPP)	OC/TA	Point of Sale	Bill Payments	ATMs	Branches
France	\$82,850	0.02%	78%	185%	280%	1.4%
Germany	77,247	-40	501	115	601	14
U.K.	63,972	-52	117	214	160	-25
Italy	50,204	-29	121	117	809	133
Netherlands	34,157	-33	330	128	1,593	-50
Spain	32,120	-50	714	390	858	22
Belgium	12,070	-23	136	98	802	-48
Sweden	5,637	-38	685	8	70	-33
Denmark	4,112	-39	206	333	522	-38
Finland	2,783	-59	1,057	136	11	-46
Norway	2,160	-60	757	67	70	-38
All Countries Together:		-34%	140%	151%	434%	9.8%

**Table 3: Percent changes for 11 European countries between 1987 to 2004**

Generally, those countries with the largest reduction in the number of branches experienced the largest reductions in unit operating cost. Italy and Spain are exceptions here since they expanded the number of their offices (as branching restrictions were removed) but their unit operating cost still fell.<sup>8</sup> If Italy was excluded from the analysis, the number of branches in the remaining 10 countries would have fallen by 2% between 1987 and 2004 rather than rising by 9.8%. As we note below, large differences in the growth of payment transactions versus ATMs versus branches will influence our interpretation of scale economy results.

### 3.3 Estimation issues

There are two estimation issues that need to be addressed in determining payment scale economies. First, translog or Fourier cost function specifications used in these analyses are sensitive to the size of an independent variable's share in the cost of the dependent variable (operating cost). Variables with small cost shares are associated with unreliable results, a pattern well-documented in scope economy studies (Röller, 1990; Pulley and Humphrey, 1993) but the problem applies to scale analyses as well. This is of particular interest in our case since in the early part of our time period the volume of electronic giro payments in many countries and electronic card transactions in others were very small and generated

<sup>7</sup> This reflects the sum of OC across all 11 countries divided by the sum of TA for 1987 vs. the same ratio of sums for 2004. It is not a simple average of separate country ratios (which would be -38%).

<sup>8</sup> Both Italy and Spain are popularly considered to be over-branched compared to other European countries. This is misleading since their branches are smaller both in terms of deposits held and labor employed so service flow is not comparable to merely the number of offices in each country (see Section 3.5 below).

little cost. As well, toward the end of our time period, the volume of checks and paper giro transactions are zero or close to it in a number of countries.<sup>9</sup> Even a composite cost function (c.f., Pulley and Braunstein, 1992) which is better able to deal with this issue, since the data are not logged as in the translog or Fourier forms, experiences difficulties in our application.

Consequently, instead of specifying four separate payment categories where some transaction volumes are close to or at zero at different points of time, we sum card transactions with checks to represent point-of-sale transactions and aggregate paper and electronic giro transactions to reflect bill payments. This addresses the low cost share issue but at the expense of less specificity in the network scale economy estimates since the result is effectively a weighted average of the scale effects of the underlying composition of transactions. As we are primarily interested in card payment scale economies, the fact that debit cards accounted for 90% of non-cash point of sale transactions over 2002-04 suggests that the point of sale scale economy estimate for the latter part of our time period may be a good approximation for debit cards. Similarly, since 86% of bill payments over 2002-04 were electronic giro transactions, the bill payment scale effect for the latter portion of our time period should be close to that for electronic giro payments.

A second estimation issue concerns autocorrelation in the time-series component of our panel data set as our payment, ATM, and branch data are strongly trended. Indeed, three out of seven of our specified variables are non-stationary.<sup>10</sup> Since our variables are not cointegrated, we investigate the effect on our results from (a) adding time-specific indicator variables (one for each time period), as well as (b) transforming the data using a differencing parameter ( $\rho$ ) estimated from the residuals from equations (1) and (2) below.

### 3.4 Cost function specification

The variation of operating cost ( $OC$ ) across 11 European countries annually over 1987-2004 is used in a translog cost function to derive scale economies for point of sale transactions, bill payments, as well as for ATM and branch office networks. A translog function (1) and a more complex Fourier cost function (not shown) were both estimated for robustness. As the scale results were very similar and the figures of predicted unit operating cost (seen below for the translog) almost identical, only the translog results are reported here. The translog cost function is estimated jointly with cost share equations (2) for labor:<sup>11</sup>

$$\ln OC = \alpha_0 + \sum_{i=1}^4 \alpha_i \ln Q_i + 1/2 \sum_{i=1}^4 \sum_{j=1}^4 \alpha_{ij} \ln Q_i \ln Q_j + \sum_{i=1}^4 \sum_{k=1}^2 \delta_{ik} \ln Q_i \ln P_k + \sum_{k=1}^2 \beta_k \ln P_k + 1/2 \sum_{k=1}^2 \sum_{m=1}^2 \beta_{km} \ln P_k \ln P_m \quad (1)$$

<sup>9</sup> The Baumol, Panzar, and Willig (1982) average incremental cost measure (Chapter 4) derived from a translog multiproduct cost function also suffers from this problem as it is based on the evaluation of an estimated function at or close to the point where the production of an output is zero.

<sup>10</sup> Stationary variables are the logs of the number of card transactions, the number of ATMs and the two input prices, which are specified in the following section. Non-stationary variables are logs of operating costs, the number of giro transactions and the number of branch offices.

<sup>11</sup> The standard coefficient symmetry and linear homogeneity in input price restrictions are imposed in estimation.



$$S_k = \beta_k + \sum_{m=1}^2 \beta_{km} \ln P_m + \sum_{i=1}^4 \delta_{ik} \ln Q_i \quad (2)$$

where:

$OC$  = total operating cost, composed of all labor, materials, outsourcing and capital consumption costs (but no interest expenses);

$Q_i$  = four output characteristics ( $i = 1, \dots, 4$ ) composed of point of sale ( $POS$  = card and check) and bill payment ( $BP$  = electronic and paper giro) transactions along with the number of automated teller machines ( $ATM$ ) and size standardized branch offices ( $BR^{STD}$ ).<sup>12</sup> The standardization procedure is explained below;

$P_k$  = two input prices ( $k = 1, 2$ ) denoting the average labor cost per bank employee and an opportunity cost approximation to the price of bank physical capital and materials inputs represented by each country's market interest rate;<sup>13</sup> and

$S_k$  = the cost share for the labor input (the "mirror image" capital/materials input cost share is deleted to avoid singularity). It is expected that operating costs not directly associated with the type of payment or mode and level of service delivery will be represented in the intercept term.

### 3.5 Comparing branch offices across countries

It is clear that a single payment transaction in one country, whether by card, check, giro or cash withdrawal from an ATM, is equivalent to a single payment transaction in another country. While there is a great deal of size homogeneity among banking offices within a single country, this is not the case for banking offices across countries.<sup>14</sup> Although the average number of workers per branch office across our 11 countries was 15.9 it was only 6.7 in Spain but 26.2 in the U.K. Clearly, each branch office in Spain (or Belgium with 10.7 workers per branch) is providing a different level of banking service output than occurs in the U.K. (or Germany with 15.8 workers per office). An even greater dispersion exists for the value of assets generated per branch.<sup>15</sup> It is thus necessary to standardize branch office size according to some benchmark to make them more comparable across countries.<sup>16</sup> Our view is that the production function relationship reflected in the

<sup>12</sup> EFTPOS terminal availability is associated with the volume of electronic card payments – a variable we already use – thus, is not separately specified in the model.

<sup>13</sup> No information is available for banks' weighted average cost of capital (WACC) or a relevant risk premium on equity. Even though the nominal market interest rate should be lower than WACC, it is unlikely to deviate too much from it as banks have high leverage, a low beta value and often, a large share of low or zero interest deposits. The alternative of specifying a real interest rate had little effect on the payment scale economy values (changing them by 1 to 2 percentage points).

<sup>14</sup> When new branches are established in a country they tend to expand in parallel with the growth of their local market. Once a branch reaches a given size, further bank growth occurs via establishing additional new branches (or through mergers). The large differences in branch office size across countries is probably due to differences in population density, earlier norms developed when (prior to ATMs) cash could only be obtained from a branch office and greater reliance in some countries (France and the U.K.) on checks versus cards or cash.

<sup>15</sup> The average total asset/branch ratio ranges from \$30 to \$193 million with an overall average of \$96 million.

<sup>16</sup> In partial support of the adjustment, the level of operating cost (or total assets) varies more closely with the standardized branch measure (R-squared = 0.76) than it does with the unadjusted number of branches (R-squared = 0.57).

“labor/capital” ratio – tying the labor input to branch payment processing, cash access, loan origination and monitoring – reflects better the service flow produced by branch offices than the unadjusted number of branches themselves. France, with an average of 16.1 workers per office over our 18-year period (close to our overall mean), was selected as the branch benchmark and other countries were adjusted accordingly.<sup>17</sup> This adjustment, while necessary given the heterogeneity of the cross-country branch data, is a weakness in our study.<sup>18</sup>

## 4 Payment scale economies: results

Based on the cost model presented above, we now show how predicted bank operating cost varies across countries, what the associated payment and service delivery scale economies are, as well as illustrate how predicted unit payment costs vary by transaction volume. Due to the non-proportional expansion of our set of four output characteristics, a weighted average of these scale economy values reflects better the overall scale effect for the banks and countries involved.

### 4.1 Network scale economies by country

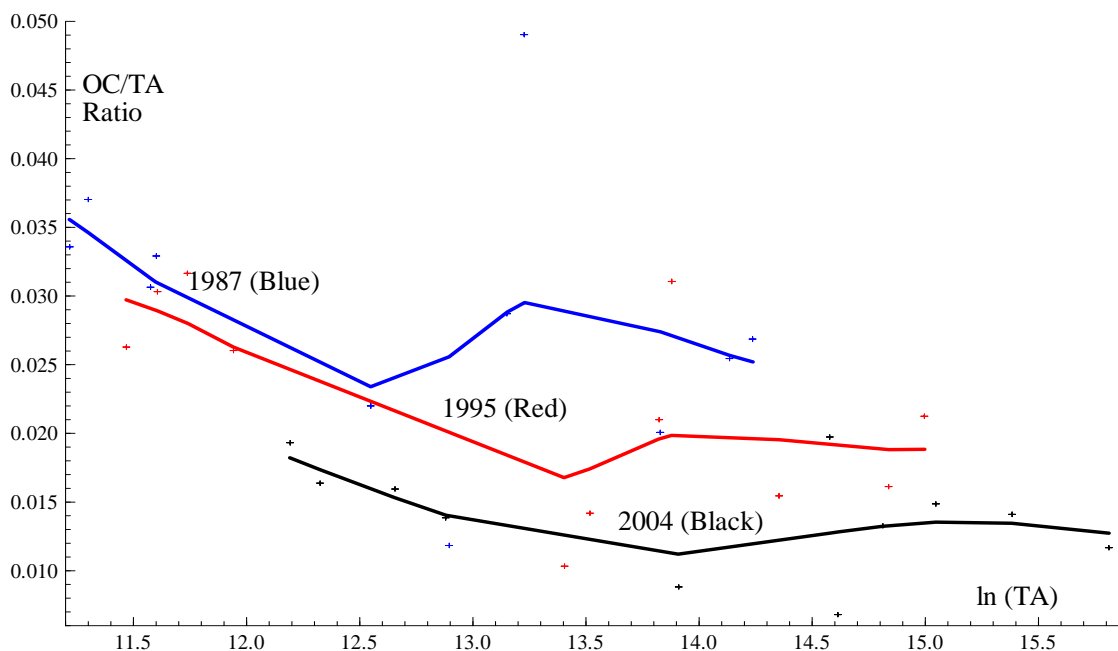
Figure 1 illustrates how the ratio of predicted operating cost to asset value (Y-axis) changes as the log of asset value rises (X-axis). Cubic splines were used to graph the average cost curves in Figure 1 and are suggestive of how each country’s banking sector predicted average operating cost varies with the asset size of this sector. The curves are based on a translog cost function and parameter estimates are in the Appendix.<sup>19</sup> Norway, Finland, Denmark and Sweden form the first part of each curve while the middle reflects Belgium, the Netherlands, Spain and the U.K. The last part of each curve covers Italy, France and Germany. In 1987 and also in 1995, Spain and the U.K., experienced a higher ratio of operating cost (OC) to total assets (TA) than did a number of countries that preceded it (Belgium and the Netherlands) and this accounts for the upward slope in the curve for these two years.

<sup>17</sup> Specifically, the number of each country’s banking offices ( $BR$ ) was adjusted as follows:

$BR^{STD} = BR[(L/BR)/16.1]$ , where  $L/BR$  is the observed labor/branch ratio in each country for each year and 16.1 workers per office is the standardized size of each office. This gives the number of standardized, size-adjusted branches ( $BR^{STD}$ ) which is used for each country in the estimations, not  $BR$ . For example, the average  $L/BR$  for the U.K. was 26.2 workers per office so dividing by the French benchmark gives  $26.2/16.1 = 1.63$  which increases by 63% the number of “standard” U.K. branches used in the analysis. In contrast, since Spain had an average  $L/BR$  of 6.7 workers per branch office, dividing by the French benchmark gives  $6.7/16.1 = 0.42$  which reduces the number of “standard” Spanish offices by close to 60%. This was done for each country for each year.

<sup>18</sup> There is no improvement if a different adjustment were adopted. In related work, results using a deposit/branch or total asset/branch ratio as an alternative standardization metric produced unrealistic scale economy values (Humphrey, Willeson, Bergendahl and Lindblom, 2006).

<sup>19</sup> Estimation of a Fourier cost function yielded a figure that was virtually identical to Figure 1.



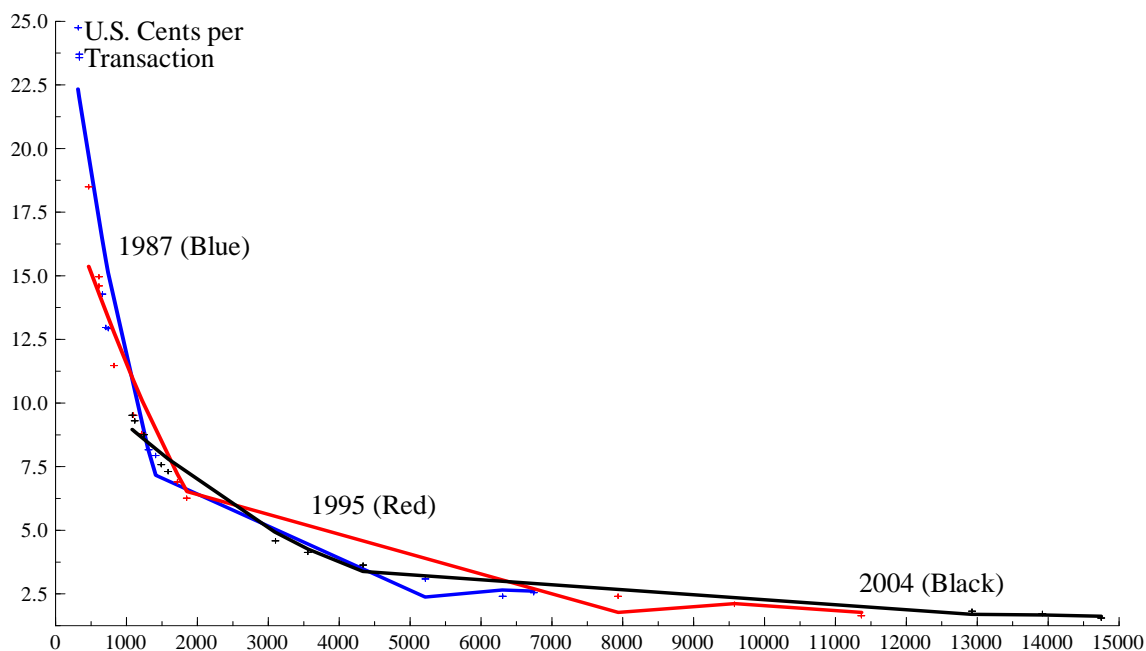
**Figure 1: Translog predicted unit operating cost (OC/TA) vs. ln(total asset value)**

Shifts in these curves between the three time periods shown illustrate the degree of cost reduction from changes in the composition of payments (small effect), the shift away from branches to ATMs (the major effect) and other technical change.<sup>20</sup> In predicting total operating cost (OC), only input prices were held constant at their mean values. As noted below, this is the closest we can get to a standard average cost curve for payment, ATM and branch networks.

Figure 2 shows how an approximation to unit payment cost varies by the total number of payment transactions. The curves are cubic splines of the ratio of predicted payment costs (for both point of sale and bill payments together) divided by the total number of transactions (Y-axis) and arrayed against total transactions (X-axis). Importantly, the curves in Figure 2 are not average cost curves as the level of these curves (in U.S. cents per transaction) is too high.<sup>21</sup> However, the slopes shown are a fair reflection of how payment unit costs change with payment volume. As seen, these slopes (and their associated scale economies) vary with payment volume but are quite similar for the three years shown, suggesting that a changing payment mix or technology seemingly adds little to the operating cost reductions shown over time in Figure 1.

<sup>20</sup> We do not have quality adjusted input prices for bank use of computers and telecommunication facilities. Thus, cost reductions here will be incorporated in and shown as decreases in predicted payment operating cost in Figure 1. When quality adjusted input prices exist and are specified, technical change can be separated into the portion derived from the cost function and the portion embodied in input prices. In either case, operating cost falls.

<sup>21</sup> This because predicted payment operating costs are obtained by evaluating the estimated cost function by holding input prices and the number of ATMs and branches constant at their mean value. Although constant, these mean values and their associated costs add to the level of the payment costs being predicted. The inability to obtain average costs for a subset of outputs in a multi-output cost function was noted in Baumol, Panzar and Willig (1982). The exception is for Figure 1 where total operating cost is being predicted, not the portion of this cost associated with payments (or ATMs or branches) separately.



**Figure 2: Translog predicted unit payment costs vs. payment transaction volume (11 countries, PPP exchange rates, POS and bill payments)**

Our payment-related scale economy (SCE) or cost elasticity estimates using the translog form are presented in Table 4.<sup>22</sup> The scale economy for the  $i^{th}$  payment service ( $SCE_i$ ) from the cost function (1) is  $\partial \ln OC / \partial \ln Q_i = \alpha_i + 1/2 \sum_{j=1}^4 \alpha_{ij} \ln Q_j + \sum_{k=1}^2 \delta_{ik} P_k$ , where:  $i, j$  = point of sale transactions ( $POS$ ) and bill payments ( $BP$ ), number of ATMs ( $ATM$ ), and number of branches ( $BR$ ). In Column 1, countries are ranked according to their total non-cash payment volume in 2004. Note that in 2004, non-cash payment volume in our 11 European countries was 59 billion transactions while in the U.S., it was 85 billion. On a per person basis, the U.S. makes 74% more non-cash transactions per year than individuals in Europe, indicating a greater degree of cash replacement.

The average scale economy for all payments for each country using the translog cost function is shown in Column 2. The average across countries is 0.27 and indicates that substantial scale effects would be expected as payment volume rises. The point of sale and bill payment SCEs appear quite low and at 0.11 and 0.16 respectively, are considerably lower than independent estimates for Norway, the Netherlands or Belgium using different data sources. The independent estimates include some branch expenses while the econometric analysis holds branches (and their associated expenses) constant in order to focus on only payment processing costs and this likely accounts for some of the difference.

<sup>22</sup>Scale economies are  $SCE_i = \partial \ln OC / \partial \ln Q_i$  and values  $< (>) 1.0$  represent economies (diseconomies).

	Payment Volume* 2004	Average Payment SCE	Point of Sale SCE	Bill Payment SCE	ATM SCE	Branch SCE	Total SCE	Realized SCE
Germany	14,748	0.23	0.06	0.17	0.22	0.59	1.05	0.31
France	13,926	0.30	0.08	0.22	0.31	0.36	.97	0.47
U.K.	12,919	0.35	0.11	0.24	0.36	0.27	.99	0.54
Spain	4,335	0.30	0.10	0.20	0.23	0.48	1.01	0.45
Netherlands	3,563	0.17	0.09	0.09	0.24	0.65	1.06	0.24
Italy	3,094	0.21	0.05	0.16	0.17	0.62	.99	0.30
Belgium	1,594	0.20	0.10	0.10	0.26	0.59	1.05	0.26
Sweden	1,488	0.33	0.18	0.15	0.39	0.37	1.09	0.21
Finland	1,244	0.35	0.19	0.16	0.40	0.34	1.09	0.20
Norway	1,117	0.34	0.19	0.15	0.40	0.34	1.08	0.23
Denmark	1,081	0.24	0.12	0.12	0.28	0.52	1.04	0.37
Average	5,374	0.27	0.11	0.16	0.30	0.47	1.04	0.40 <sup>#</sup>

Total 11 country non-cash transactions 59.1 billion in 2004 (U.S. 84.5 bil)

Note: \*Millions of non-cash transactions. All data rounded. <sup>#</sup> Calculated separately.

#### Table 4: Non-cash payment volumes and translog scale economies by country

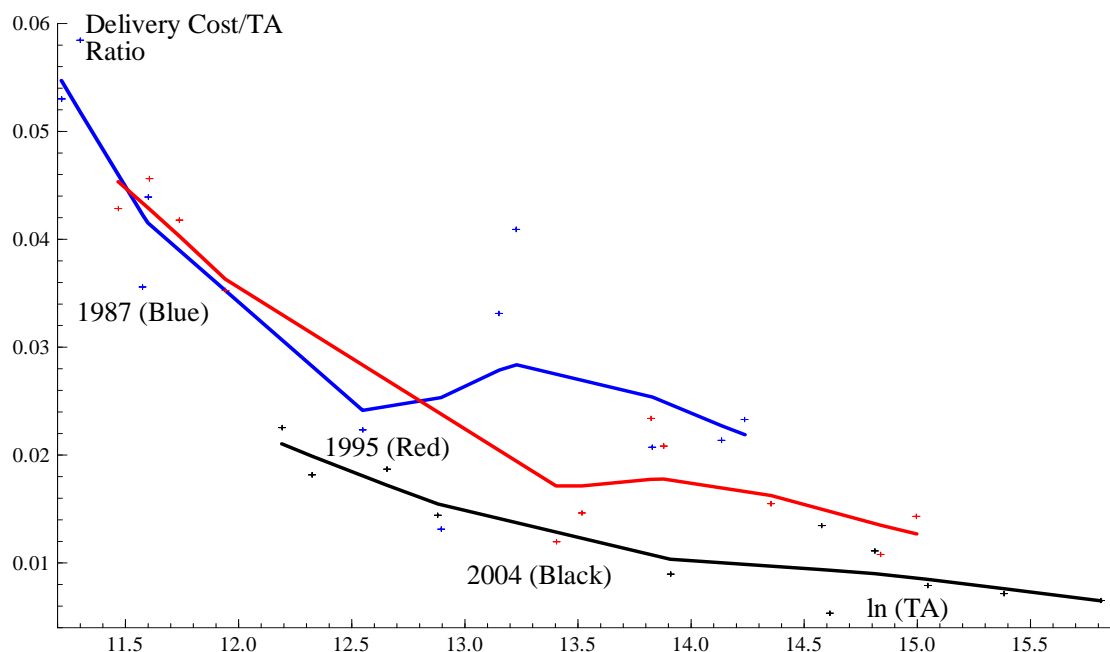
To gauge robustness, the estimation process was repeated using data from earlier versus later time periods, segmenting countries by smaller versus larger payment volumes and adding 11 country-specific indicator variables to the model. The results indicated that the average payment SCE results were not markedly different by earlier or later time periods nor for countries having smaller versus larger payment volumes compared to using all 11 countries together for the entire time period.<sup>23</sup> Also, adding 18 time-specific indicator variables yielded little reduction in the positive autocorrelation evidenced in the Durbin-Watson statistic for the panel data and had no effect on the number of significant parameters (as 17 out of the 20 that were common were significant).<sup>24</sup> Finally, a differencing parameter ( $\rho = 0.85$ ) was estimated from the residuals from equations (1) and (2) and used to transform the data. This generated a D-W value close to 2.00 but the cost function concavity condition was not met and branch scale economies were negative, suggesting that operating costs fall absolutely for countries with larger branch networks.<sup>25</sup>

<sup>23</sup>The translog average payment SCE is 0.27 over the entire period and is 0.24 using data for only the first half (1987-1995) and 0.35 for the second half (1996-2004). Restricting the estimation to the 5 countries with the smallest payment volumes gives an average payment SCE of 0.30 while for the 6 largest countries it is 0.36. Estimation with 11 country-specific dummies raises the average payment SCE to 0.34 so country identification explains some of the reduction in operating cost, leaving less to be associated with our four outputs.

<sup>24</sup> Estimation with 18 time-specific dummies gives an average payment SCE of 0.20. Here, all the time dummies are positive indicating that the passage of time raises (not lowers) operating cost. This anomalous result likely occurs because we already apparently identify the banking outputs associated with cost curve shifts over time in our originally specified variables.

<sup>25</sup> Setting  $\rho = 0.75$  met the concavity condition but branch scale economies were still negative. Although the time sequence of observations is fixed for each country, the ordering of countries in the panel data is arbitrary and changes here did lead to some reduction in positive autocorrelation (but of course, had no effect on parameter or scale results).

Based on these results, a fixed effects estimation framework would not offer any improvement in estimation. More importantly, fully adjusting for positive autocorrelation yields anomalous results.<sup>26</sup> Consequently, we focus on the scale economy estimates in Table 4 which applies a translog cost model using unadjusted levels data for the entire time period for all 11 countries together (in panel econometrics jargon, a “total pooled regression”). To correct for heteroskedasticity, Robust-White standard errors were calculated and are shown in the Appendix with the estimated parameters.



**Figure 3: Translog predicted delivery cost/asset ratio vs. ln(total asset value)**

As noted above for Figure 2, the payment scale effects shown there are not the source of the cost reductions between time periods illustrated in Figure 1. This source appears to be associated with the shift away from expensive branch offices toward ATMs for the delivery of cash acquisition and other banking services (probably along with lower computer and telecommunication input expenses and other technical changes we can not directly identify). This is seen from Figure 3 where the ratio of predicted operating cost associated with ATMs and branches to the value of banking assets for each country (Y-axis) is arrayed against the log of asset value (X-axis). For the same reasons noted in the discussion of Figure 2, the cost curves shown in Figure 3 are not average cost curves: their level is too high but their slopes and shifts over time are indicative of scale effects and cost reductions in delivering banking services to depositors.

## 4.2 Realized scale economies

The average payment SCE of 0.27 in Table 4 is the simple sum of the point of sale (0.11) plus bill payment (0.16) scale economies. Equally weighting and adding these two

<sup>26</sup> Indeed, a grid search using values of  $\rho = 0.25, 0.50, 0.75$  and  $1.0$  to transform the data indicated that the average payment SCE (Column 2) was fairly stable up to the point where  $\rho = 0.50$  but completely fell apart when  $\rho = 1.0$ .

payment SCEs together is reasonable in this case since the average percent changes in point of sale transactions (140%) and bill payments (151%) shown at the bottom of Table 3 over 1987-2004 are very similar. However, the percent changes in ATMs and branches are quite different so a simple sum of these two elasticities (0.30 and 0.47 respectively) is not appropriate.

Individually, all four of the specified characteristics of banking output – two classes of payments, ATMs and branches – experience scale economies. It is common to report the simple sum of all individual SCEs as representing the overall scale effect. This is shown in the next to last column in Table 4 under the heading “Total SCE” and seems to suggest that overall operating cost economies reflect constant returns to scale or even slight diseconomies. The simple sum is informative when all outputs expand proportionally, which is probably acceptable when banking “outputs” are specified as summary measures of types of loans, security holdings and/or deposits since these balance sheet categories tend to expand at similar rates. As shown in Table 3, our set of output characteristics expanded at markedly different rates and a weighted SCE would be more informative of the scale effect actually realized.

If all banking outputs have grown by the same percentage amount, then the scale economies actually realized would be:  $RealizedSCE = 1.0(0.11) + 1.0(0.16) + 1.0(0.30) + 1.0(0.47) = 1.04$  from  $\sum_{i=1}^4 SCE_i$  for  $i = POS, BP, ATM, BR$  using the average scale economies at the bottom of Table 4. If the number of branch offices did not rise on average by 9.8% overall (as shown in Table 3) but instead was constant while the other three outputs all expanded at the same rate, then the weight of branches in the overall cross-country scale measure should be zero since the overall effect of branches on operating cost should also be zero. This would give:  $RealizedSCE = 1.0(0.11) + 1.0(0.16) + 1.0(0.30) + 0.0(0.47) = 0.57$ . Finally, if only ATMs had expanded and nothing else, then the realized scale economy should be 0.30 – the scale measure of ATMs alone. In general, the output experiencing the greatest growth should have the largest weight for its scale economy value while other outputs experiencing less growth should have smaller weights.<sup>27</sup> A reasonable measure of realized scale economies should be able to generate the same numerical values for the previous three simple examples, handle situations where the growth rates of the various outputs are not all the same and weight the output with the largest growth the highest.<sup>28</sup>

One way of meeting these conditions (there may be others) would be to define the scale economy weights as ratios of each output’s growth rate divided by the growth rate of the output that experienced the greatest expansion, whether this be 10%, 100% or in our case 434% for ATMs giving:  $RealizedSCE = (1.40/4.34)(0.11) + (1.51/4.34)(0.16) + (4.34/4.34)(0.30) + (0.098/4.34)(0.47) = 0.40$ .<sup>29</sup> Realized operating cost scale economies are less than

<sup>27</sup> If growth of one of the output characteristics, say branches, was negative its realized operating cost would rise as the number of branches falls due to a reversal of the scale effect. This would reduce the value of the *RealizedSCE*.

<sup>28</sup> A straightforward weighting scheme could be:  $SCE_{tot} = \sum w_i SCE_i$  with  $w_i = s_i^C g_i^T / (\sum s_i^C g_i^T)$  where  $s_i^C$  denotes the (average) cost share of output  $i$ ,  $s_i^T$  the (average) transaction share and  $g_i^T$  the average growth rate of transactions. Unfortunately, separate cost accounting data on  $s_i^C$  rarely exists and is not available from estimation of a multiproduct cost function which specifies interactions among the variables.

<sup>29</sup> If all the outputs had expanded by 434%, all the numerators would be 4.34 yielding weights of 1.0 and we obtain the Total SCE value of 1.04. If the first three outputs expanded by 434% but branches were constant

half the value of the simple sum of derivatives and overall operating cost scale economies are apparently quite large.<sup>30</sup> Other things being equal, banks that reduce their branch operations the most should realize the greatest overall scale benefit as payments and ATMs expand.

## 5 Cash replacement and SEPA

The average value of a cash transaction at the point of sale in the Netherlands (€9) and Belgium (€18) is much smaller than for debit cards (€44 and €50). As seen earlier in Table 2, at these transaction values the average Netherlands combined bank plus merchant total cost of using cash (€0.30) is cheaper than a debit card (€0.49) by €0.19 but only €0.02 lower in Belgium.<sup>31</sup> Cash is also cheaper than debit cards for both banks and merchants separately in the Netherlands, as illustrated in Table 5 (rows 2 and 3).<sup>32</sup>

As these cost differences are not reflected in explicit per transaction prices, users have no real price incentive to use one instrument over the other and relative use is driven by apparent consumer preferences and trade-offs, such as the universal acceptability of cash, the availability of a card terminal or the convenience of using a debit card that does not require “refilling” at an ATM or other terminal. While underlying cost does not directly influence consumer use, banks and merchants who pay these expenses are influenced by them along with the revealed preferences of users.<sup>33</sup> The apparent consumer preference for a debit card over cash for medium value transactions can be accommodated for low value transactions if the scale economies associated with debit cards were more fully realized to lower current bank expenses. If successful, this could permit banks to reduce fees charged merchants for accepting a debit card, make card terminals more widely available and equalize better the relative costs of cash and cards for low value transactions (where cash now appears to have an advantage). This may also lead banks to replace stored value cards so that a single technology could be used for both medium and smaller value transactions.

Using scale economies computed for the Netherlands and Belgium, as well as from our cross-country estimation (shown in the lower half of Table 5), it is possible to illustrate the required debit card volume growth and time it may take at current rates of expansion to reduce the average total cost of debit card use (€0.49) down to a level close to that of cash (€0.30), a potential reduction of €0.19 per transaction. Using the scale economies (SCEs)

(a zero in the numerator), we get 0.57 while if only ATMs expanded, we obtain 0.30. The Realized SCE using scale estimates from a Fourier cost function is 0.43.

<sup>30</sup> Our scale economy estimates are derivatives which are appropriate for small changes in output while our growth figures are large and cover 18 years. Thus, the Realized SCE figures shown by country in Table 4 are approximations.

<sup>31</sup> For Belgium, merchant cost of cash is €0.27 per transaction while a debit card costs €0.28. At the bank level, cash costs €0.26 while a debit card costs €0.27. Hence, the €0.02 overall difference.

<sup>32</sup> Estimates for the U.S. focus on marginal costs and suggest that the bank plus merchant cost of a cash payment is \$0.37 (\$0.07 plus \$0.30, respectively) while that of a PIN debit is \$0.84 (\$0.27 plus \$0.57). Thus the bank plus merchant difference in marginal cost for a low value transaction averaging \$11.52 favors cash by \$0.47 (Garcia-Swartz, Hahn and Layne-Farrar, 2006, Table 3). This is a considerably larger spread than that for the Netherlands or Belgium.

<sup>33</sup> Banks introduced ATMs to reduce the costs of supplying cash to depositors over the counter and the benefits were twofold. A cash withdrawal was cheaper at an ATM than at a branch office and in addition, ATMs could be substituted for more expensive branch offices and still provide the same (or greater) level of cash access for depositors.



shown in Table 5, the approximate transaction volume expansion ranges from 44% to 64%.<sup>34</sup> As the 2003-4 annual average card transaction growth rate in the Netherlands was 8%, this could take 6 to 8 years to be realized. However, since the year-to-year growth rates over 2000-2004 were, respectively, 19.0%, 12.1% 8.2% and 7.8%, the use of an 8% growth rate is likely too optimistic. Debit card volume growth has been increasing at a decreasing rate over this period, an expected result when adoption of a “new” payment technology has seemingly passed its inflection point on a logistic S-curve (illustrated in Bolt, Humphrey and Uittenbogaard, 2007). If domestic debit card volume growth in the near future was only 6% annually, the illustrative reduction in debit card unit costs would take from seven to 11 years. Indeed, if S-curve card volume growth projections are correct, domestic volume growth in the Netherlands will be well below 6% in the near future.

Netherlands, 2002	Cash	Debit Card	Stored Value
Average Total Cost:	€0.30	€0.49	€0.93
Banking Sector Unit Cost:	0.14	0.25	0.78
Retail Sector Unit Cost:	0.16	0.24	0.15
Scale Economy Value:	SCE	Volume Growth	Years
Netherlands/Belgium SCE	0.39	64%	8.0
Average Payment SCE	0.27	53%	6.6
Average Point of Sale SCE	0.11	44%	5.5

Source: Brits and Winder (2005), Table 4.5, data are rounded.

**Table 5: Illustration of debit card cost reduction**

A better alternative and one consistent with the goals of SEPA, would be to consolidate card processing centers among European countries to achieve larger scale benefits. To illustrate, consolidating debit card volume in Belgium (671 million in 2004) with that of the Netherlands (1,247 million) would expand volume at a consolidated processing center by 54%, suggesting that unit processing cost may decline by 33% if the Netherlands/Belgium scale economy of 0.39 in Table 5 is the appropriate metric.<sup>35</sup> If the average payment or average point of sale SCEs in Table 5 is used, the change in unit processing costs could be -39% or -48%, respectively. These reductions in unit processing costs would, however, not translate directly into reductions in total unit costs since additional telecommunication expenses would need to be factored in (as well as the amortized cost of the expanded investment in processing equipment). Consolidation of existing processing arrangements is clearly more difficult to arrange than merely piggy-backing the currently small amount of cross-border card volume onto current national processing center operations.

<sup>34</sup> The relationship between percent changes in transaction volume ( $\partial \ln Q$ ) and percent reductions in average payment operating cost ( $\partial \ln AC$ ) cannot be estimated statistically since separate cost accounting data on  $AC$  does not exist. However, scale economies  $\partial \ln OC / \partial \ln Q = SCE$  from an operating cost ( $OC$ ) function  $\ln OC = f(\ln Q, \ln P)$  of output ( $Q$ ) and input prices ( $P$ ) can be rearranged as  $\partial \ln AC / \partial \ln Q = SCE - 1$ . Setting  $\partial \ln AC = -39\%$  (from a €0.19 reduction in  $AC$  divided by the current  $AC$  level €0.49), the implied growth in transaction volume shown in Table 5 is  $\partial \ln Q = -39\% / (SCE - 1)$  for different values of  $SCE$ .

<sup>35</sup> As noted in the previous footnote,  $\partial \ln AC / \partial \ln Q = SCE - 1$ . Thus  $\partial \ln AC = \partial \ln Q (SCE - 1) = 54\% (0.39 - 1) = -33\%$ .

It is likely that the advent of SEPA will spur consolidations in a competitive, intra-European payment processing market. In order to maintain or increase market share, European payment processors will be looking for partners and alliances. One such alliance has already been formed. Interpay, the Dutch processor of debit card and giro payments, just completed a merger with its German equivalent Transaktionsinstitut (TAI) at the end of 2006. The merged company – called Equens – will double its yearly processed volume from around 3.5 billion transactions each to 7 billion in total, making it one of the largest payment processing centers in Europe. In a recent press release (January 23, 2007), Equens announced that the volume growth from their merger will generate large scale benefits and by 2010, they could “realise cost savings of approximately 25 percent, which include a reduction of around 400 full-time positions and plan to increase the yearly processed payments transactions to 10 billion”.<sup>36</sup>

In light of these market dynamics, some small players have already stated their intention to shift processing to other centers. This happened in Luxembourg, where it was decided to shut down the national retail system LIPS-net from 2007 onward and to process all its domestic payments via the Euro Banking Association’s (EBA) Step2 platform for euro interbank payments. As well, Spain and Italy announced their intentions to start using Step2 by the second quarter of 2007. In this environment, it would be instructive to review the experience of the U.S. – which operates large card and automated clearing house processing centers – to get a better idea of the cost and benefits of a possible cross-border consolidation, much in the way that the consolidation of Fedwire operations in the U.S. suggested that similar cost benefits may be obtained from cross-border consolidation of Target (the European large value payment network).

## 6 Summary and conclusions

The goal of a Single Euro Payment Area (SEPA) is to make cross-border payments as easy as domestic transactions as this will support the emergence of a competitive intra-European market. While this could be largely achieved by simply piggy-backing cross-border card transactions onto existing national processing operations in Europe, this is unlikely to result in any significant cost reduction. Lower cross-border as well as domestic unit payment costs, however, would be achievable by consolidating network processing operations across borders. In addition, markedly lower debit card unit costs would likely promote greater merchant adoption of card terminals leading to greater replacement of cash for small value point of sale transactions where currently, cash is often cheaper for merchants and the ability of stored value cards to substitute for cash has had only limited success.

The current composition of payment instrument use at the point of sale in European countries is primarily determined by consumer preferences and trade-offs such as the universal acceptability of cash and the convenience of debit cards which do not need to be “refilled” at terminals or banking offices (unlike cash or a stored value card). Explicit pricing per payment transaction, used broadly in Norway, is not the norm and explicit payment prices that do exist in Europe are typically piecemeal, need not reflect the

---

<sup>36</sup> In a recent interview (in *Dialogue*, Q4 2006, p.10), Ben Haasdijk – former chairman of Equens – stated “scale was the main driver for the merger that gave birth to Equens” and that “we will be passing on benefits of scale we achieve”.

underlying relative costs of different instruments and in some cases, induce users to use the more expensive instrument (c.f., Sveriges Riksbank, 2004).

Information on card network scale economies for three countries (Norway, the Netherlands, and Belgium) ranged from 0.43 to 0.39 and was compared to economies derived from a cost function using an 11 European country panel of payment and banking data over 18 years. Estimated scale economies for point of sale transactions (primarily debit cards) averaged 0.11 while bill payment (paper and electronic giro) transactions averaged 0.16. For all payments together, the scale economy was 0.27 (the simple sum of point of sale and bill payment economies). These scale estimates were used to illustrate the possible context and time frame for debit card costs to fall to a level similar to that for cash, potentially enabling debit cards to replace cash for small value cash and stored value transactions, as well as lowering unit costs for higher value transactions. For greater accuracy, expanded telecommunication expenses, as well as transition expenses would need to be included rather than relying solely on the scale effects estimated here.

This study provides preliminary scale economy information that may be helpful in outlining possible benefits from a Single Euro Payments Area arising from the consolidation of electronic payment processing centers across the euro zone. If this approach were pursued, the experience of the U.S. which consolidates card processing across states may serve as a useful example concerning the realized costs and benefits, as well as likely implementation issues of cross-border consolidation. For SEPA to become a success, however, both consumers and merchants will need to accept the harmonized pan-European payment instruments and this may necessitate some direct pricing of payment instruments in order to induce users to adopt more fully the most cost effective methods. The interplay between scale benefits, emerging payment market structure and pricing requirements is complex and has not been previously addressed.

## 7 References

Baumol, W., J. Panzar and R. Willig (1982). *Contestable Markets and the Theory of Industrial Structure*. Harcourt Brace Jovanovich: San Diego.

Bank for International Settlements (various years). *Statistics on Payment Systems in the Group of Ten Countries*. Basle: Switzerland.

Bolt, W., D. Humphrey and R. Uittenbogaard (2007) "Transaction Pricing and the Adoption of Electronic Payments: A Cross-Country Comparison," *International Journal of Central Banking*, forthcoming.

Bolt, W. and D. Humphrey (2007) "Scale Economies from Individual Bank Payment Data," Working Paper, Florida State University.

Brits, H. and C. Winder (2005) "Payments Are No Free Lunch," De Nederlandsche Bank, Occasional Studies, Vol. 3/Nr. 2 (preliminary version published as: De Nederlandsche Bank (2004), *The Costs of Payments: Survey on the Costs Involved in POS Payment Products*, Working Group on Costs of POS Payment Products, March).

- Cheney, J. (2006) "Supply – and Demand – Side Developments Influencing Growth in the Debit Market," Payment Cards Center Discussion Paper, Federal Reserve Bank of Philadelphia, October.
- European Central Bank (2006) *Towards a Single Euro Payments Area*, Fourth Progress Report, Frankfurt, Germany.
- European Central Bank (various years) *Payment and Securities Settlement Systems in the European Union*, Frankfurt, Germany.
- European Monetary Institute (various years) *Payment Systems in the European Union*, Frankfurt, Germany.
- First Annapolis Consulting (2006) "2005 Issuer Cost of Payments Study" (data used with permission from First Data Corporation).
- Garcia-Swartz, D., R. Hahn and A. Layne-Farrar (2006) "The Move Toward a Cashless Society: A Closer Look at Payment Instrument Economics," *Review of Network Economics*, 5: 175-98.
- Gresvik, O. and G. Øwre (2002) "Banks' Costs and Income in the Payment System in 2001," *Norges Bank Economic Bulletin*, 73: 125-33.
- Humphrey, D., M. Willeson, G. Bergendahl and T. Lindblom (2006) "Benefits from a Changing Payment Technology in European Banking," *Journal of Banking and Finance*, 30: 1631-52.
- International Monetary Fund (2006) *World Economic Outlook*, Washington, D.C. Purchasing Power Parity Exchange Rates, available at <http://www.imf.org/external/pubs/ft/weo/2006/01/data/dbginim.cfm>
- Leinonen, H. (2004) "Payment System Developments – Their Dependence on Competition Cooperation and Authority Actions," presentation September 2-3, Unitas Congress Centre, Helsinki
- Pulley, L. and Y. Braunstein (1992) "A Composite Cost Function for Multiproduct Firms with an Application to Economies of Scope in Banking," *Review of Economics and Statistics*, 74: 221-230.
- 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.
- Quaden, G. (2005) *Costs, Advantages and Drawbacks of the Various Means of Payment*, National Bank of Belgium, English Summary, December: 41-7.
- Röller, L.-H. (1990) "Proper Quadratic Cost Functions with an Application to the Bell System," *Review of Economics and Statistics*, 72: 202-210.

Sveriges Riksbank (2004) “Prices and Costs in the Swedish Payment System,” *Financial Stability Report*: 65-73.

Van Hove, L. (2006) “Why Electronic Purses Should be Promoted,” *Banking and Information Technology* (Universität Regensburg, Germany), Nr. 2, June: 20-31.

## 8 Appendix: Parameter estimates for the translog cost function

Variable	Parameter	Estimation	t-statistic
Constant	$a_0$	2.1562	4.54
$\ln Q_{CARD}$	$a_1$	0.2117	3.44
$\ln Q_{GIRO}$	$a_2$	-0.4530	-5.77
$\ln Q_{ATM}$	$a_3$	0.2192	2.30
$\ln Q_{BR}$	$a_4$	1.2095	13.49
$1/2(\ln Q_{CARD})^2$	$a_{11}$	0.0063	0.34
$1/2(\ln Q_{GIRO})^2$	$a_{22}$	0.0084	0.24
$1/2(\ln Q_{ATM})^2$	$a_{33}$	0.0668	5.04
$1/2(\ln Q_{BR})^2$	$a_{44}$	0.3614	12.63
$\ln Q_{CARD} \ln Q_{GIRO}$	$a_{12}$	0.0267	1.57
$\ln Q_{CARD} \ln Q_{ATM}$	$a_{13}$	0.0476	3.52
$\ln Q_{CARD} \ln Q_{BR}$	$a_{14}$	-0.1076	-7.17
$\ln Q_{GIRO} \ln Q_{ATM}$	$a_{23}$	0.0684	4.06
$\ln Q_{GIRO} \ln Q_{BR}$	$a_{24}$	-0.0695	-3.72
$\ln Q_{ATM} \ln Q_{BR}$	$a_{34}$	-0.1998	-9.91
$\ln Q_{CARD} \ln P_L$	$d_{11}$	-0.0450	-6.64
$\ln Q_{GIRO} \ln P_L$	$d_{21}$	-0.0819	-8.45
$\ln Q_{ATM} \ln P_L$	$d_{31}$	-0.1007	-11.22
$\ln Q_{BR} \ln P_L$	$d_{41}$	0.2120	18.19
$\ln P_L$	$b_1$	1.2916	25.95
$1/2(\ln P_L)^2$	$b_{11}$	0.1826	37.94

There were 198 observations and the Log likelihood was 546.821. Standard errors were computed from a heteroscedastic-consistent matrix (Robust-White) and the cost function was concave. The R-square and Durbin-Watson statistics were 0.97 and 0.43, respectively, for the cost equation (1) and were 0.42 and 0.44, respectively, for the cost share equation (2).