What does it Cost to Make a Payment?

DAVID HUMPHREY*

Florida State University

MAGNUS WILLESSON

University of Gothenburg

TED LINDBLOM

University of Gothenburg

GÖRAN BERGENDAHL

University of Gothenburg

Abstract

We survey the limited data that exists concerning the cost of making/receiving a payment by banks, retailers, and other parties to a transaction. Since an electronic payment costs between one-third and one-half that of a paper-based instrument, a country may save 1% of its GDP annually as it shifts from a fully paper-based to a fully electronic-based payment system. Some gains have already been realized. Additional analysis indicates that bank costs of making a payment may have fallen by 45% in Europe as the share of electronic transactions in 12 countries rose from .43 to .79 over 1987-1999.

1 Introduction

There is surprisingly little public information about what it costs to make a payment. Of the limited information that is available, one consistent result is that an electronic payment seems to cost from one-half to one-third as much as its paper-based alternative. If these differences in costs could be operationalized in cost savings, the application to payment volumes suggests that a country may save 1% or more of its GDP annually by switching from all paper to all electronic payments.

* Contact author. Mailing address: Department of Finance, Florida State University, Tallahassee, FL, 32306-1042, U.S.A. Phone/Fax: (850)-668-0638, E-mail: dhumphr@garnet.acns.fsu.edu Comments by Santiago Carbó, Olaf Gresvik, Grete Øwre, Leo Van Hove, and a referee are acknowledged and appreciated.

Supporting these inferences, new payment cost function estimates suggest that aggregate nominal bank unit payment cost in 12 European countries may have fallen by some 45% during the 1990s. This time-series result is associated with three influences: a declining share of more expensive paper-based payments; a greater realization of improved electronic payment scale economies; and lower telecommunication costs due to technical change (and in some cases partial deregulation and expanded competition).

More detailed information on payment costs is obtained by looking at the expenses incurred by the various users/producers associated with a payment. To some consumers in Europe who are charged a per transaction fee differentiated by which payment instrument is used, the cost is the direct bank fee they pay along with any additional expense/inconvenience they incur. To other European consumers and consumers in the U.S., who typically only have to meet a bank minimum deposit balance requirement or obtain payment services tied to lower (higher) interest rates on deposits (loans), the marginal bank cost of making a payment is viewed as being zero. To retailers and businesses that receive payments in Europe or the U.S., the cost is the "all-in" expense of accepting different payment instruments and clearing them through a bank. To banks and other providers of payment instruments, the cost is the expense of processing the payment, internally settling the transaction, and (sometimes) a portion of branch office costs associated with maintaining payment/deposit accounts. Finally, if the bank receiving a payment differs from the one making the payment, there is the (very small) cost of an inter-bank settlement transaction through a central bank.

Overall, the total (social) cost of making a payment would be the sum of at least 4 different components: (1) consumer (payor) costs; (2) retailer/ business (payee) receiver costs; (3) bank processing costs; and (4), central bank settlement cost. As in the derivation of GNP, it is important not to double count but to sum only the value added at each of these 4 stages of production. Double counting exists since central bank costs (when priced) become bank expenses. In turn, bank expenses associated with receiving a payment and collecting it are a large portion of reported retailer expenses while bank costs associated with making a payment represent the majority of costs experienced by depositors (either through direct fees, "hidden" in balance requirements, or tied to and cross-subsidized by other bank deposit/loan services).[1]

The problem is not that payment cost information does not exist in some form somewhere, but that this information is considered confidential when it does exist and so is rarely publicly available. And, even when publicly available, these data typically refer to one point in time and can differ in their choice of what costs to include. It has to be remembered that information on something as basic as even the annual volume and value of different types of payment transactions at the country level – much less their cost at any level – only really became available about 15 years ago. This was through the efforts of some 11 developed country central banks co-ordinated by the Bank for International Settlements (BIS, various years). This volume/value information is now up-dated annually and has been expanded to 7 additional European countries by the European Monetary Institute and European Central Bank (EMI and ECB, various years).

In what follows, we survey some of the existing publicly available payment cost information in Section 2. This concerns payment costs generated at banks, incurred by payment receivers, as well as limited information on the social cost of representative paper and electronic payments.

[1] Bank fees paid by retailers or consumers are based on bank prices and may differ from bank real resource costs due to market power or pricing strategy. Bank costs would not be determined from the fees they charge but from the costs they incur. Retailer costs, of course, would be net of the bank fees they pay in order to identify the value added at that stage of the payment production process. Hence private and social costs can differ.

An attempt to identify changes in bank payment costs over time using a statistical cost model is outlined in Section 3. The model is applied to cross-country panel data on 12 European countries as well as to panel data for a single country. The results illustrate how much payment costs have fallen over the last decade as payments have shifted from paper-based to lower cost electronic payments.

The cost of making a payment will vary across different payment instruments due to differences in their production functions, most notably their scale economies and the technology used as well as their scale of operation. Section 4 illustrates these differences by surveying studies that have attempted to measure scale economy and technical change influences on payment costs at the bank or processor level. Finally, in Section 5, information is presented indicating that payment users are sensitive to the relative prices they may face when using different instruments. It follows that countries that have implemented direct pricing of payment services have experienced the largest shift to electronic payments and have saved the most. A brief summary is presented in Section 6.

2 Which cost to use: bank, retailer, or social?

2.1 Bank payment costs

For more than a decade, Norway has collected representative information on the bank cost of handling different payment instruments. This includes labor, building, materials, and computer expenses incurred in processing payments as well as allocated bank branch office expenses associated with payments.[2] Over 1988-2001, the approximate weighted average bank cost per transaction has fallen from € 1.93 to only € .73, a reduction of 62% over 13 years (7.2% a year in nominal terms).[3] This reduction in unit cost was the result of scale effects (since the total number of payment transactions in 2001 was 2.5 times the number in 1988) as well as the on-going shift to lower cost electronic payments.

Bank costs for five different payment instruments in Norway are shown in row 1 of Table 1.[4] Check costs are relatively high since they currently account for only 0.3% of all payment transactions in Norway. As checks are phased-out, their earlier processing scale economies will work in reverse and raise unit costs as volumes fall. Checks aside, the relative cost of a paper-based versus an electronic payment at the bank level can be judged by comparing the € 1.03 cost of a paper (mail) giro transaction with its electronic equivalent at approximately € .66 or with a debit card at € .34. For Norway, an electronic payment costs only 33% (debit card) to 64% (giro) as much as what a paper-based giro transaction costs.

At the bank level in Spain, the cost of either a giro or a debit card payment is about 28% of the cost of a check. Unlike Norway, the bank estimates for Spain do not include any branch office expense and so only reflect processing costs. The same is true for the U.S. in 1993 where

[2]Surveys of bank payment costs were made in 1988, 1994, and 2001. The 1988 results are in Norwegian but the other two are in English (see Flatraaker and Robinson, 1995; Robinson and Flatraaker, 1995; Gresvik and Øwre, 2002).

[3]These figures exclude the cost of cash withdrawals over the counter at a bank office, which is expensive (Gresvik and Øwre, 2002, Table 1). The exchange rate used was NOK 1 = € .137.

[4]Values in euros (U.S. dollars) are indicated by € ($) next to the year in column 1. As the exchange rate was € 1 = $ 1.027, there is no real need to express all values in only € or only $.

the relative cost of a check appears to be on par with that of an electronic payment. In 1993, checks accounted for 80% of all non-cash transactions while ACH comprised less than 2%. Although no hard estimates exist, the situation would be different today since use of checks has peaked (leading to a reversal of scale economies) while ACH costs appear to have fallen by over 80% at Federal Reserve processing offices (which would proxy bank costs).[5] Recent cross-country data on six ACH networks and four check systems suggest that the unit processing cost for an ACH payment averages 37% of that for checks (Khiaonarong, 2003).

Country/Year:		Check	Paper Giro	Credit Card	Electronic Giro/ ACH	Debit Card	Cash
			Bank Payment Costs				
Norway	2001 €	3.08	1.03	-	.62 - .69	.34	1.03-1.16
Spain	2001 €	.27	-	-	.08	.07	-
U.S.	1993 $	.15 - .43	-	-	.12 - .44	-	-
			Retailer Payment Costs				
Australia	2001 $	.28	-	.59 - 1.14	-	.10 - .23	.07
Germany	1999 €	.50 - .71[1]	-	-	.68	.87	.09 - .15
Netherlands	2002 €	-	-	3.40	-	.27	.15
Sweden	2001 €	-	-	1.54	-	.23	-
U.S.	2000 $	.36	-	.72[2]	.24[3]	.34[4]	.12
U.S.	1993 $	1.25	-	-	.23	-	-
			Social Cost				
U.S.	1993 $	2.78-3.09[5]	-	-	1.15 - 1.47	-	-

Table 1: Bank, retailer, and social cost of payment instruments (In Euros and U.S. dollars: € 1 = $ 1.027)

Notes: [1] All checks are truncated and collected electronically, lowering cost. [2] Covers both credit cards and off-line debit card cards. [3] Electronic Benefit Transfer, approximately the same as an electronic giro payment. [4] On-line debit card. [5] Excludes float benefit of $.09 per check.

Source: Bank payment cost: Norway (Gresvik and Øwre, 2002), NOK 1 = € .137; Spain (confidential source); U.S. (Wells, 1996). Retailer payment cost: Australia (Australian Retailers Association 2001), A$ 1 = $.565; Germany and the Netherlands (Van Hove, 2002); Sweden (Eklund and Larsson, 2001), SEK 1 = € .112; U.S. (Food Marketing Institute, 2001); U.S. (Wells, 1996). Social cost: U.S. (Wells, 1996).

[5]Estimates of ACH and check scale economies and technical change are presented below.

2.2 Retailer expense of accepting different payment instruments

Based on survey information, rows 4 to 9 in Table 1 illustrate what it costs retailers in different countries to accept different payment instruments, including cash.[6] These values are on a per transaction basis. They show that cash is typically the cheapest instrument to accept, followed by an on-line debit card, a check, or a giro/ACH payment, depending on the country. One consistent result is that a credit card is considerably more expensive for merchants to accept than any of the others. This is due to the relatively high merchant fee that is triggered with credit card use.

Additional retailer cost comparisons across countries are difficult to generalize from. For example, a debit card is cheaper than a check in Australia but not the U.S. in 2000.[7] Another example is Germany where check costs are apparently less than those for a debit card. This is explained by the fact that all checks in Germany are truncated and collected electronically. This makes a check in Germany equivalent to a paper-initiated electronic payment (which is not the usual case for checks).[8] Hence, while an electronic debit card payment can be cheaper to accept than a traditional check, if the check technology is advanced enough the check becomes equivalent to an electronic payment in practice as well as cost. If the costs shown for retailers in Table 1 were "normalized" so that they referred to a constant €100/$100 of retailer sales, the relative cost of cash would necessarily rise (since the average value per transaction is typically the lowest). To illustrate, the per transaction and per $100 of sales cost for different payment instruments in 2000 for the U.S. would be:

	Check	Credit Card	ACH	Debit Card	Cash
Cost per transaction:	$.36	$.72	$.24	$.34	$.12
Cost per $100 sales:	.80	1.80	1.00	.80	.90

Table 2: Retailer cost estimates for different payment instruments in 2000 for the US

Source: Food Marketing Institute (2001).

While cash is clearly the cheapest on a per transaction basis, it is on par with the cost per $100 in sales for a check, ACH, or a debit card. As seen, the main difference is that accepting a credit card – even though it is electronic – is more than twice as expensive for retailers than any other instrument except an ACH transaction (actually an Electronic Benefits Transfer). But this is mostly the result of a transfer payment to banks rather than a marked difference in real resource costs.

2.3 Putting it all together: social cost

The social cost of a payment would include payor costs, retailer costs, bank and central bank costs. For all but checks in the U.S. and checks and some giro payments in Europe, the social cost of a payment should equal the private cost. In times of high interest rates, however, the

[6] Values shown for the U.S. in 1993 are based on a constructed estimate rather than survey data.
[7] The U.S. 2000 check estimate in row 8 is from supermarket survey data. Unlike the 1993 U.S. check estimate in row 9, it has no billing component as transactions at supermarkets are at the point of sale which has lower costs.
[8] To reduce fraud and other problems, a small number of large value checks in Germany are physically presented so that the signature can be verified before payment is made.

benefits of float (a transfer payment) can place a wedge between social cost and private cost.[9] Social cost estimates are constructed estimates built piece by piece from different industry sources and educated guesses regarding the composition of payment instrument use by different users. Only three sets of estimates exist. A recent analysis compares checks with ACH and is shown in the last row of Table 1.[10]

Taking the average of the figures shown, the social cost of a check in the U.S. would be $2.94 while an electronic ACH transaction is $1.31, or 45% of the cost of a check. If we summed the average bank and retailer costs shown for the U.S. in 1993 in rows 3 and 9 in Table 1, we would get $1.54 for checks and $.51 for ACH. This sum is less than that for the average social cost and the difference may roughly approximate payor costs of $1.36 for checks and $.80 for ACH. Payor costs cover both consumer (point-of-sale, bill payment) and business check/ACH (payroll, bill payment) expenses as well as postage costs. Since the median value of a check payment back then was perhaps around $55, the average social cost of using a check ($2.94) accounts for about 5% of the transaction value.

The information presented in Table 1 is less homogeneous than one might expect to have to answer the simple question of "What does it cost to make a payment?". One reason for the heterogeneity is that countries process vastly different volumes of these instruments so scale effects, which are important to costs, will differ. Around 43 billion checks are processed in the U.S. today (Gerdes and Walton, 2002) but only about 8 billion in France plus the U.K. which are the two largest users of checks in Europe. A similar disparity exists for electronic giro and direct debit transactions which total under 7 billion in France plus the U.K. but more than 13 billion in Germany (even though the population of France plus the U.K. is 45% higher than that for Germany).[11] Thus generalizations about what it costs to make a payment are better made across instruments within a single country than across countries. However, there has been a general trend in all developed countries toward using more electronic payments. The effect of this trend on payment costs is addressed next; first for Europe as a whole and then for a single country.

3 Changes in bank payment costs over time

3.1 A cross-country analysis of changes in payment costs in Europe

As time-series bank level data on payment costs or transactions – when they exist – are not publicly available, determining how payment costs may have changed over time requires cross-country panel data using an "output characteristics" cost function. Specifically, we relate the annual operating cost of each of 12 European countries' banking sectors over 1987-1999 to the total annual number of check, paper giro, electronic giro, and card transactions in each country along with the number of ATMs and (standardized, size-adjusted) branch offices, controlling for

[9]Float occurs in a giro system when a payment is debited from a customer's account 1 or 2 days prior to good funds actually being transferred to a payee. This has been a common way for bank and postal giros to recover the cost of making a payment. For checks, the benefit flows in the opposite direction. Here consumers receive a product or service in exchange for a check which will take 1 to 2 days to collect and be debited from the consumer's account.

[10]An earlier effort covered more instruments but refers to 1987 (Humphrey and Berger, 1990). Both efforts contain detailed explanations on how the estimates were derived so others may improve the analysis. De Grauwe, Buyst, and Rinaldi (2000) compare the social costs of cards and cash for Iceland and Belgium.

[11]This disparity would be even greater if we counted checks in Germany as being an electronic payment, which they effectively are.

differences in the average price of labor and an indicator of the opportunity cost of physical capital and materials inputs.[12] Since the vast majority of banking expenses derive from processing and accounting for payments (including safekeeping) as well as delivering cash through ATMs and taking deposits and making loans at branch offices, the above specification should allow us to separate payment costs from service delivery expenses over time.[13]

For all 12 countries, the ratio of operating costs to total banking assets fell by 24% over 1987-1999.[14] This reduction in unit European bank operating expense is associated with a reduction in check and paper giro transactions (-10% and -79%, respectively), a rise in electronic giro and card-based payments (+192% and +671%, respectively), and a rise in ATMs (+318%) along with minuscule growth in the number of (unadjusted) branch offices (+0.3%). Overall, predicted bank unit payment costs fell by 45% as the share of electronic payments in total payment transactions of the 12 countries almost doubled, rising from .43 in 1987 to .79 in 1999.[15] In addition, while there was only 1 ATM for each of 3.5 branch offices in 1987, there were 1.2 ATMs per office by 1999.[16]

The relationship between bank operating cost, four payment instruments, two service delivery methods, and two input prices was estimated using a composite cost function (Pulley and Braunstein, 1992) as well as the more well-known translog and Fourier functional forms.[17] The predicted U.S. dollar value of operating cost of our four payment instruments for three different years, divided by the total volume of payment transactions, is shown in Figure 1. Unit operating cost, shown on the Y-axis, is arrayed against the log of the dollar value of total banking assets for each country on the X-axis. This is not an average cost curve for the four payment instruments. Predicted operating cost is determined by evaluating the estimated (composite) cost equation using observed transaction volumes for each of the four payment instruments while holding constant the number of ATMs, (adjusted) branch offices, and the two input prices at their sample mean values over 1987-1999. Thus Figure 1 reflects the sum of the mean cost of ATMs, branch offices, and the two input prices along with how the variation in payment transaction volumes affects operating cost over time and across countries. While the level shown on the Y-axis necessarily includes more than just payment costs, the slope of the curves indicates how these costs have, on a per transaction basis, varied over time.[18]

[12]From smallest to largest in terms of total banking assets, the 12 countries are: Norway, Finland, Denmark, Sweden, Belgium, Spain, Netherlands, Switzerland, Italy, U.K., France, and Germany. A maintained hypothesis in a pooled data set such as this is that all inputs (labor, physical capital) are variable since the levels of these inputs do in fact vary across countries. Operating cost reflects the payments to labor and the realized expense of physical capital.

[13]The tasks of making and monitoring loans, using security positions for liquidity, and managing off-balance sheet loan commitment and derivative activities account for a relatively small portion of total bank operating costs and should be reflected in an intercept term or captured in the cost associated with the branching variable.

[14]This is a ratio of the sum of operating cost over all 12 countries divided by the sum of the value of total banking assets (not a simple average of operating cost/total asset ratios across countries).

[15]The number of EFTPOS card terminals rose from 147,000 to over 3,500,000 over this period and obviously made possible the rapid expansion in card transactions at the point of sale.

[16]Specifically, the ATM/branch ratio was 49,098/172,354 in 1987 but rose to 205,071/172,943 in 1999.

[17]Unlike the translog and Fourier forms, the composite specification does not log the payment or service delivery variables but does log the input prices. As a result, the predicted costs can be more accurate when the number of check and paper-based giro transactions are absolutely small, as occurs in some countries. With logs of input prices, homogeneity of degree one in input prices can be imposed.

[18]Additional cost information is available from predicted changes in marginal cost as paper-based payments shift to electronics. Such analysis should also consider the effect of possible "stranded" or sunk costs as the volume of paper-based payments has contracted rapidly in some countries over time.

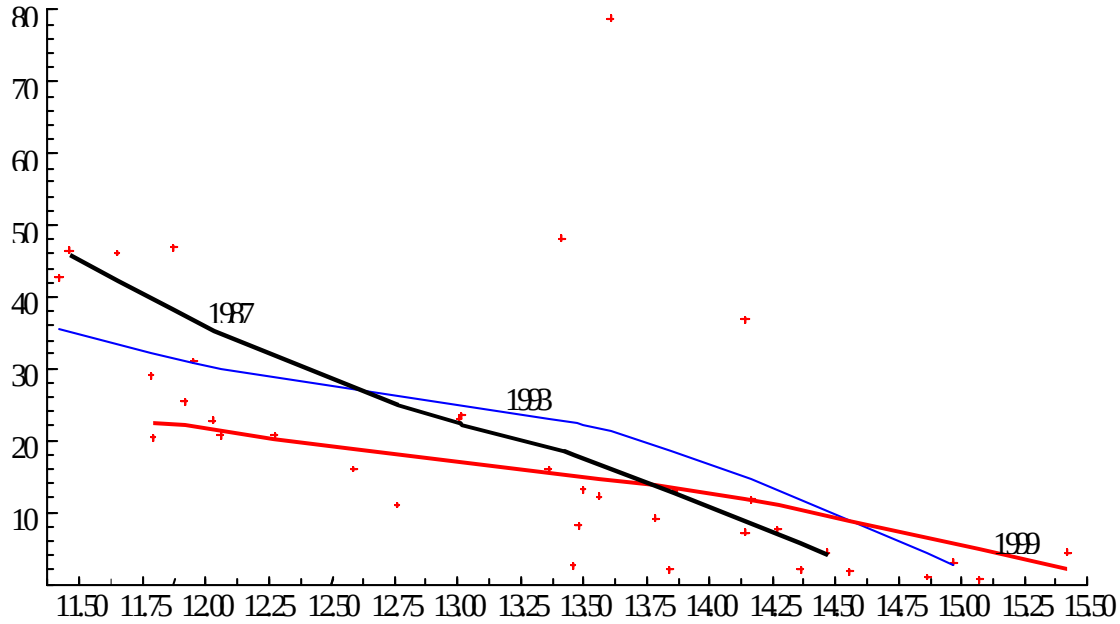

Figure 1: Predicted unit payment cost for Europe by log of total assets: 1987, 1993, 1999 in dollars

Note: Predicted values are from a composite cost function evaluated by allowing the volumes of four payment instruments (check, paper giro, electronic giro, and card transactions) to vary while holding the number of ATMs, branch offices, and input prices constant at their sample mean values.

Looking at the slope of the three payment cost curves in Figure 1, it is clear that there is a strong scale effect in processing payments. The implied scale economy is on the order of .20 for the three cost function estimations (composite, Fourier, and translog) which means that a 10% rise in payment volume is associated across countries with only a 2% rise in allocated operating cost, so unit costs would fall. The shift in the three payment cost curves suggests that technical change mostly lowered costs in countries with smaller banking systems while raising it somewhat in nominal terms in the larger countries.

More information is obtained when electronic giro and card-based payments are separated from checks and paper-based giro payments. The predicted labor, physical capital, and materials operating cost per transaction for electronic payments in Figure 2A is seen to have fallen markedly, mostly in the smaller countries (Norway, Finland, Denmark, Sweden, Belgium, Spain, and the Netherlands). The other five larger countries also experienced a reduction which looks small by comparison but was on the order of 50% over the 1987-1999 period. As scale economies for electronic payments are large and the volume of these payments have greatly expanded over the period (192% for electronic giros and 671% for card payments), large reductions in predicted unit costs would be expected.

The predicted cost per transaction for both types of paper-based payments in Figure 2B shows a different result. Here there was a large rise in the nominal cost of processing paper payments that affected all countries but was focused on those larger countries that use the most checks (the U.K., Italy, and France). Although within each country paper-based scale economies may indeed exist, as paper-based payments continue to be replaced by electronic payments the

lower volumes will be seen to raise unit processing costs rather than reduce them.[19] Hence the apparent scale diseconomies in processing paper-based payments seen in the cross-country analysis.

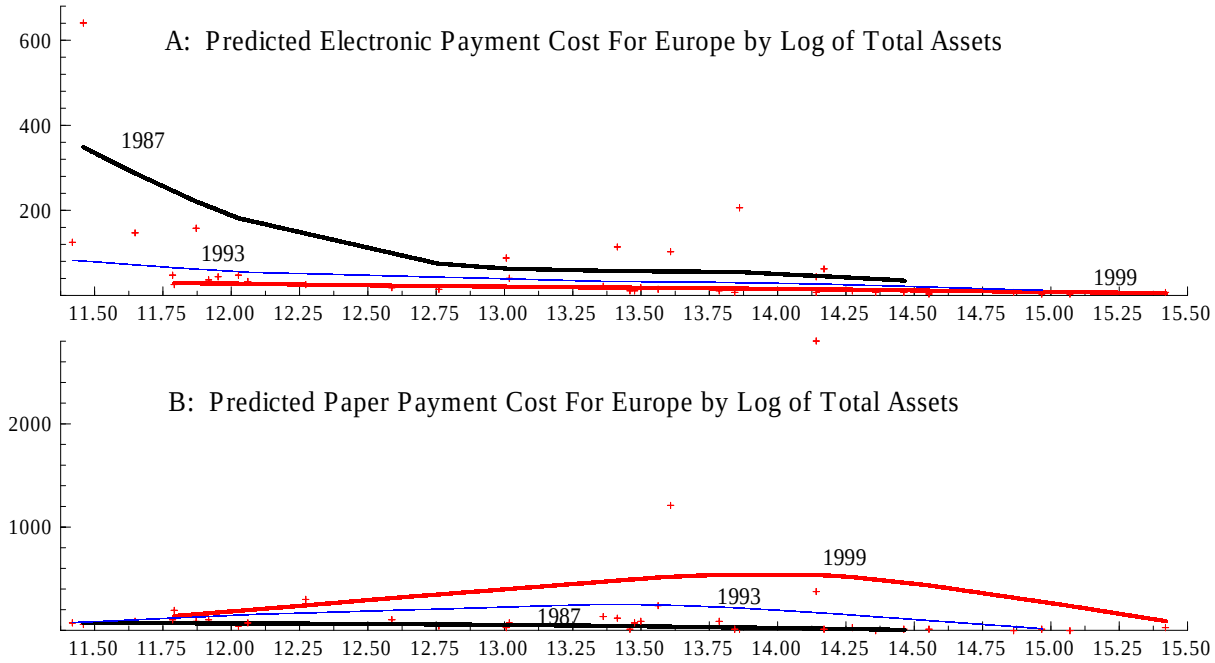

Figure 2: Predicted unit electronic & paper cost for Europe by log of total assets: 1987, 1993, 1999 in dollars

Note: Predicted values determined in the same manner as in Figure 1 except that in part A (B) the volume of check and paper giro (electronic giro and card) transactions were additionally held constant at their mean values in the cost evaluation.

Due to the nature of the multiple product cost function, our small sample size, and the fact that the cross-country observations on payment volumes have no countries (except Spain) that process zero amounts of any payment instrument, it is not possible to accurately determine the average cost of processing different electronic or paper-based payments for our 12 European countries. More success is obtained using a larger sample size and a more homogeneous data set of a single country.

3.2 Changes in payment costs in one country: Spain

A similar cost function approach has been applied to a panel of 1,541 observations of savings and commercial banks in Spain over 1992-2000 (Carbo, Humphrey, and Lopez, 2002). The overall effect of shifting to electronic from paper-based payments, along with the move to increasingly deliver cash services through ATMs rather than expanding branch offices, was

[19] As well, fixed investments in check and paper-based giro processing could well be written off more rapidly over time, raising observed expenses, if the reduction in payment volumes proceeds faster than normal depreciation rates.

estimated to perhaps have saved the Spanish banking sector € 5 billion over a nine-year period.[20] Unit operating cost for the banking sector fell by 35% while predicted unit payment transaction expenses fell by 47%. The reduction in payment cost is associated with a 85% rise in giro payments, a 78% rise in card payments, and a 18% reduction in check payments. The share of check transactions in total non-cash payments fell from .19 in 1992 to only .10 in 2000. Giro transactions accounted for a .56 share in 2000 while the card share was .34.[21]

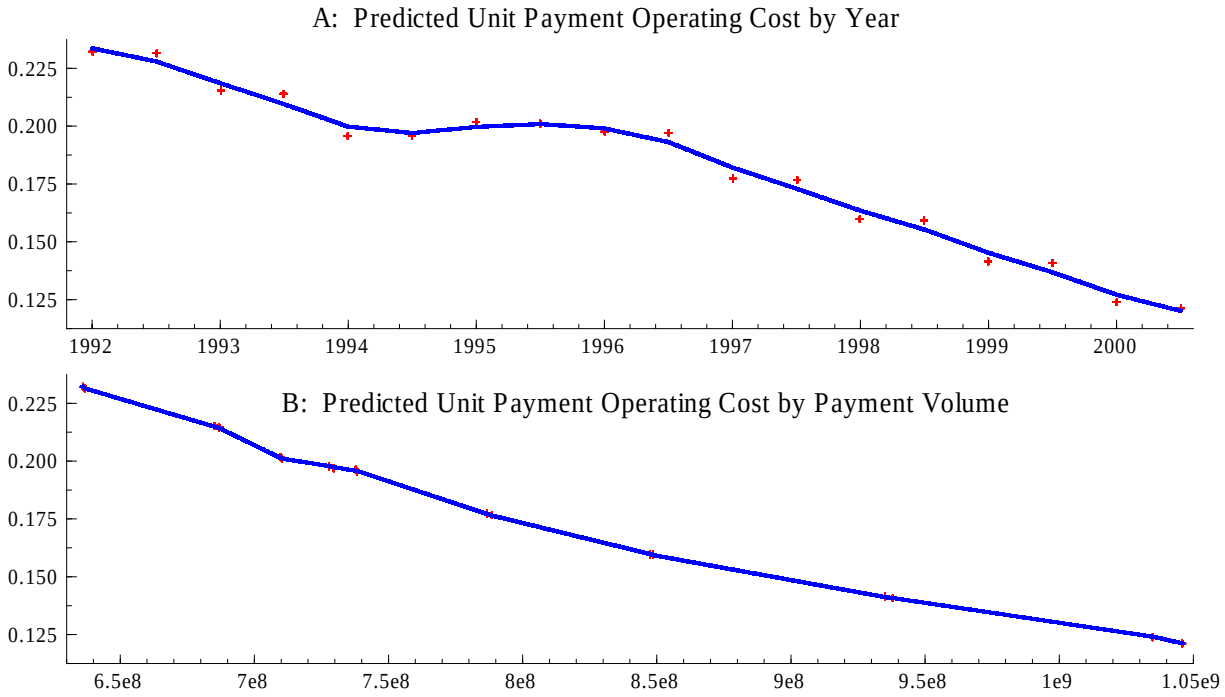

Figure 3: Predicted unit payment operating cost for Spain by year & payment volume: 1992-2000 in Euros

Note: Predicted values determined in the same manner as Figure 1 using a composite function for Spain.

Figure 3A illustrates the reduction in unit payment costs by year over 1992-2000 while 3B shows that payment costs experience substantial scale economies when predicted unit cost is arrayed by bank size (the log of individual bank total asset value). As before, the values in Figure 3 are not actual average payment costs but rather show how the predicted values of the estimated operating cost function (divided by total payment volume) varies when evaluated using observed and changing levels of check, giro, and card transaction volumes but holds costs associated with ATMs, branch offices, and input prices constant at their sample mean values. Doing the same for

[20]While ATM and branch information is available on a per-bank basis, payment volume data exists only at the national level in Spain and all other countries (although Norway can separate payment volumes by instrument between all commercial banks and all savings banks at the national level).

[21]There is no separate information on paper-based versus electronic giro transactions for Spain but anecdotal information indicates that probably all reported giro payments during this period were electronic.

our three separate payment instruments suggests that unit check costs rose by 58% over 1992-2000 while giro and card unit costs fell by 31% and 27%, respectively.[22]

It is possible to at least approximate the likely actual average cost of processing different payment instruments using an ad hoc scaling procedure (not explained here). Applying this procedure to three cost function estimations, Table 3 shows the approximate values of the estimated average bank cost for three payment instruments.[23] The results support earlier conclusions that, with the exception of credit cards, an electronic payment costs only one-third to one-half as much as an equivalent paper-based transaction whether at the point of sale, for a bill payment, or for employee disbursements.

	Cost Function:			Average of Actual Bank Unit Cost Estimates
	Composite	Fourier	Translog	
CHECK	€ .25	€ .27	€ .15	€ .27
GIRO	.07	.10	.05	.08
CARD	.10	.15	.08	.07

Table 3: Estimated average bank cost of check, Giro, and card payments in Spain

4 Payment scale economies and technical change

As data on private sector check, debit card, and credit card processing costs and transaction volumes are not publicly available, almost all payment scale economy and technical change information has been based on Federal Reserve operations. This is not a severe limitation since the production function for check and ACH processing is almost identical not only between banks and the Federal Reserve within the U.S., but also among different countries processing the same types of payments (i.e., checks and electronic payments or direct debits for bill payments).

4.1 Scale economies

Scale economies exist for check operations but the average cost curve is essentially L-shaped. Earlier scale estimates were close to constant costs. More recent estimates over 1993-1997 suggest a cost elasticity of between .67 to .80, depending on the cost specification estimated, for the weighted average of 47 processing offices (Bauer, 2002).[24] This combines significant scale benefits for small check processing offices with typically constant costs for the very largest operations.

The mechanical check processing and sorting component of the production process experiences scale economies up to the point where an additional reader/sorter machine is

[22] As before, the rise in check expense is explained by the fact that the scale economies associated with checks have a reverse effect on unit costs when processing volumes are being reduced.

[23] This allocation does not include a portion of branch office expense associated with maintaining deposits as occurred for some countries in Table 1. The last column in the table represents an average of internal estimates from a confidential industry source.

[24] Although no numerical scale estimates are presented, a non-parametric data envelopment analysis of all Federal Reserve check processing offices over 1980-1999 that specified check volume and delivery endpoints as "outputs" could not reject variable returns to scale for all offices as a group (Gilbert, Wheelock, and Wilson, 2002).

required to handle higher volumes and there are strong scale benefits with the computer component of the processing operation. However, the land/air physical transportation collection and presentment component is sensitive to the spatial distribution of delivery endpoints (paying banks or check clearing houses). If additional delivery endpoints are associated with a higher volume of checks, as when market share expands, then economies on the processing side can be reduced by constant costs or diseconomies on the collection side.[25]

Scale economies in electronic payments are the result of a tradeoff between the relative cost of centralized versus distributed processing. If the cost of communication links between the point where an electronic payment is initiated and where it is to be processed are low enough, then processing will be centralized. Here the cost savings from processing electronic payments using large, scale efficient computer facilities more than offsets the higher communication costs associated with transmitting payment information back and forth over greater distances to a centralized location. If communication costs are relatively high, then processing will be distributed among more numerous and closer locations with each processing center handling a smaller volume of transactions.[26]

Overall scale economies at U.S. automated clearing houses are estimated to be .48, so a 10% rise in ACH processing volume is associated with only a 4.8% rise in total processing and communication costs (Bauer and Ferrier, 1996). The average cost curve for ACH payments is downward sloping rather than L-shaped as it is for checks. Since an ACH transaction is similar in many respects to an electronic giro payment in Europe, as well as debit and credit card transactions today, the scale economies here should be similar to those for an ACH (although we know of no published studies on this topic). Apparently, there are little or no scope economies among ACH and wire transfer operations (Adams, Bauer, and Sickles, 2002) and the same would likely also hold for check and ACH.

4.2 Technical change

There have been many cost reducing improvements in check processing operations over time. Cost savings have been experienced in the past with: (a) the use of magnetic ink on checks for identifying the customer account and paying bank to be debited; (b) the standardization of the dimensions and thickness of a check;[27] (c) the requirement that bank endorsement information be placed where it can be easily read (to facilitate return item processing); (d) the development of ever faster check reader/sorter machines (made possible by the use of magnetic ink and standard check dimensions);[28] and (e) recent efforts to truncate checks at the point of sale or bank of first deposit while transmitting the payment information needed for collection via the ACH or some other electronic presentment arrangement (as Germany now does). Some efforts to estimate the effect of technical change in check operations have largely occurred after improvements (a) to (d) were in place and prior to any significant use of (e). As a result, these studies have not found

[25]Check processing managers, of course, will say they can handle a slight rise in check volume at virtually zero marginal cost and this is true. A more relevant indicator of scale effects, however, is associated with a significant increase in long-run volume rather than a de minimis expansion which merely absorbs current short-run excess capacity.

[26]While the same tradeoff exists for checks, the higher costs of physical transport (rather than electronic communication) typically require extensive distributed processing as opposed to centralization.

[27]Although rare, in the past one could write a perfectly legal check on a block of wood which then would go through the collection process – with expensive special handling – and be presented for payment at a paying bank.

[28]Today, the limit to mechanical processing of checks has effectively been reached. While reader/sorter speeds could be increased, it would just rip the checks and jam the machines.

evidence of statistically significant technical change in check operations (Bauer and Hancock, 1995; Bauer and Ferrier, 1996). Looking at a longer period, however, suggests that median productivity in check processing ranged from 1.4% to 1.9% a year over 1980-1999, with the largest increases occurring in the mid-1980s (Gilbert, Wheelock, and Wilson, 2002).

Using cost accounting information, the unit cost of an ACH payment at the Federal Reserve fell from $.87 in 1990 to $.18 in 2000, a reduction of almost 80% (Berger, 2003). Estimated ACH marginal costs appear to have declined by upwards to 81% during this period while marginal costs for wire transfers may have fallen by 45% (Adams, Bauer, and Sickles, 2002).[29] Technical change in ACH operations is estimated to have reduced unit costs by about 10% a year since 1989 (Bauer and Hancock 1995; Bauer and Ferrier, 1996). Due to a combination of scale economies from higher volumes, technical advances in computers, deregulation of the communications industry, and changes in organizational structure, very large reductions were realized in the unit cost of ACH payments, a payment method very similar to a giro transaction in Europe.

5 Incentives to use the lowest cost payment instrument

The check float wedge between the social and private cost of checks for payors has in some countries hindered the movement away from checks to lower cost electronic payments.[30] Another barrier lies in how payment services are priced by banks. While retailers and businesses usually pay a price per transaction they make or receive through their bank, consumers–when given a choice – typically choose to hold a required minimum balance and/or pay fixed monthly account fees rather than pay a fee per transaction (Hannan, 2002). Although this balance requirement or account fee often reflects the average expected use of different payment instruments across users, from the individual consumer's perspective each payment made is viewed as having a zero marginal cost.[31]

A zero marginal cost for payment instrument use, paired with float benefits for checks and frequent flyer miles (or other rebates) for credit card use, serve to hinder the adoption of other types of electronic payments with lower costs.[32] Analysis of the consumer and business response to changes in bank payment prices over a decade in Norway suggest that the price mechanism can strongly affect the composition of payments in a country. After payments were priced and differentiated to cover something close to unit variable cost, the share of check and paper-based giro transactions in all non-cash payments fell from 90% in 1987 to 40% in 1996 (Robinson and Flatraaker, 1995; Flatraaker and Robinson, 1995). The share of paper-based instruments has

[29]These estimates are from a separable quadratic cost function similar to a composite function noted in Section 3 above. A translog estimation suggested a 75% reduction in ACH marginal cost with an 82% fall for wire transfers.
[30]Canada has eliminated the check float wedge for consumers (which was small in any event) as well as for businesses. Canadian banks give payees same - day funds availability for deposited checks and charge firms for the check float they create. The central bank also backdates interbank check settlements by one day to effectively eliminate interbank float. By shifting the cost of float to payors, and backdating check settlements, Canada has removed an incentive to use checks and thus promoted electronic payments.
[31]Business depositors, who have much higher transaction volumes, are typically charged on a per transaction basis. Payment is made through a debit to their account or by holding a compensating balance with the bank that is set to generate the same level of revenue as would occur with direct fees when evaluated by a market interest rate.
[32]When retailers assess fees on use of (expensive) credit cards, this misallocation is reduced (Lindblom, 2001). A similar argument has been made regarding the use of (subsidized) cash versus an electronic purse for smaller value transactions (Van Hove, 2002).

continued to fall since then and was less than 20% in 2001 (Norges Bank, 2002).[33] With this response to pricing, it is no surprise to find empirically that: (a) checks and debit cards, along with checks and ATM withdrawals (reflecting cash use), are significant net substitutes; and (b), that the own price elasticity of demand for checks is elastic while that for debit cards is inelastic (Humphrey, Kim, and Vale, 2001).[34]

6 Summary and conclusions

The cost of making a payment can differ across countries depending on:

(1) The type of payment instrument being used (checks and paper giro transactions versus an electronic giro and debit cards);

(2) The level of use (which affects the realization of existing scale economies);

(3) The degree of electronic technology applied to paper-based payments (e.g., in Germany, a check is really a paper-initiated electronic payment);

(4) The extent that deposit account maintenance and associated branch expenses are included with payment processing costs; and,

(5) Pricing methodology (credit card merchant fees versus full cost recovery from users).

As a general rule of thumb, an electronic payment costs only from one-third to one-half as much as a paper-based payment. If a country moves from a wholly paper-based payment system to close to an all electronic system, it may save 1% or more of its GDP annually once the transition costs are absorbed. While scale economies exist for paper-based payments, they are small compared to electronic payment methods. Some countries, notably Norway, have encouraged banks to directly price the use of different payment instruments based roughly on their different production costs. This speeds up the transition to cheaper electronic payments, improving social welfare, given that users' non-price considerations (convenience, safety, accuracy) are largely met.

For 12 European countries, the share of electronic payments in total non-cash transactions rose from .43 in 1987 to .79 in 1999. Econometric analysis suggests that the shift to electronic payments over 1987-1999 may have reduced the cost of making a payment by 45%. This shift, along with the large expansion of ATMs in place of expensive branch offices, is the main reason why the ratio of bank operating cost to total asset value in Europe has fallen by 24% over the same period. A few European countries have already largely shifted to electronic payments and are realizing the cost savings involved. Others, along with the U.S., are at an earlier stage with

[33]Moving from 100% paper - based payments to 100% electronic is estimated to save 0.6% of GDP annually in Norway, and this only refers to savings in bank costs.

[34]Pricing can speed up the shift to electronic payments but is difficult to implement. Past practice almost everywhere has been to recoup payment costs indirectly via float (early debiting/late crediting of accounts), use of minimum balance requirements, and return item fees all of which hide the true price consumers currently pay. Consequently, shifting to an explicit price is viewed as a price increase rather than just a different way to cover costs. Compounding this problem, antitrust considerations inhibit or prevent the coordinated implementation of differential payment pricing so explicit pricing by the first bank to do so can lead to a lower market share unless others quickly follow.

less than 40% of non-cash transactions in electronic form. These countries may be interested in encouraging their banks to price more directly payment instrument use to realize sooner the gains associated with most electronic payments.

7 References

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.

Australian Retailers Association (2001) "Submission to the Reserve Bank of Australia Credit Card Schemes in Australia," in Reserve Bank of Australia, *Reform of Credit Card Schemes in Australia III, Submissions Received*, Volume 1, Sydney: H2-1-24.

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

Bauer, P. (2002) "Personal Communication on New Check Scale Economy Estimates," December 24.

Bauer, P. and G. Ferrier (1996) "Scale Economies, Cost Efficiencies, and Technological Change in Federal Reserve Payments Processing," *Journal of Money, Credit and Banking*, 28: 1004-39.

Bauer, P. and D. Hancock (1995) "Scale Economies and Technical Change in the Federal Reserve Automated Clearinghouse Payment Processing," *Federal Reserve Bank of Cleveland Economic Review*, 31: 14-29.

Berger, A. (2003) "The Economic Effects of Technological Progress: Evidence from the Banking Industry," *Journal of Money, Credit and Banking*, 35: 141-76.

Carbó, S., D. Humphrey and R. Lopez (2002) "Effects of ATMs and Electronic Payments on Banking Costs," Working Paper, Department of Finance, Florida State University, September.

De Grauwe, P., E. Buyst and L. Rinaldi (2000) "The Costs of Cash and Cards Compared: The Cases of Iceland and Belgium," mimeo, University of Leuven.

Eklund, H. and S. Larsson (2001) *Kortbetalningar – En studie av den svenska marknaden*, Rapport från Svensk Handel och SHR, Stockholm, October.

European Central Bank (ECB) (2002) *Payment and Securities Settlement Systems in the European Union, Addendum Incorporating 2000 Figures*, Frankfurt, July.

European Monetary Institute (EMI), *Payment Systems in the European Union*, Frankfurt, various years.

Flatraaker, D. and P. Robinson (1995) "Income, Costs, and Pricing in the Payment System," *Norges Bank Economic Bulletin*, 66: 321-32.

Food Marketing Institute (2001) *It All Adds Up–An Activity Based Cost Study of Retail Payments*, Washington, D.C.

Gerdes, G. and J. Walton (2002) "The Use of Checks and Other Noncash Payment Instruments in the United States," *Federal Reserve Bulletin*, 88: 360-74.

Gilbert, A., D. Wheelock and P. Wilson (2002) "New Evidence on the Fed's Productivity in Providing Payments Services," Working Paper 2002-020A, Research Department, Federal Reserve Bank of St. Louis, September.

Gresvik, O. and G. Øwre (2002) "Banks' Costs and Income in the Payment System in 2001," *Norges Bank Economic Bulletin*, 73: 125-33.

Hannan, T. (2002) "Retail Fees of Depository Institutions, 1997-2001," *Federal Reserve Bulletin*, 88: 405-13.

Humphrey, D. and Allen Berger (1990) "Market Failure and Resource Use: Economic Incentives to use Different Payment Instruments," in D. Humphrey (ed.), *The U.S. Payments System: Efficiency, Risk, and the Role of the Federal Reserve*, Kluwer Academic Publishing, Norwell, MA: 45-86.

Humphrey, D., M. Kim and B. Vale (2001) "Realizing the Gains from Electronic Payments: Costs, Pricing, and Payment Choice," *Journal of Money, Credit and Banking*, 33: 216-34.

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

Lindblom, T. (2001) "Card Payments – Pricing and Competition," in Gardener and Versluijs (eds.), *Bank Strategies and Challenges in the New Europe*, Palgrave, New York: 87-105.

Norges Bank (2002), *Annual Report on Payment Systems, 2001*, Oslo, July.

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.

Robinson, P. and D. Flatraaker (1995) "Costs in the Payment System," *Norges Bank Economic Bulletin*, 66: 207-16.

Van Hove, L. (2002) "Cost-Based Pricing of Payment Instruments: The State of the Debate," Working Paper, Free University of Brussels, December.

Wells, K. (1996) "Are Checks Overused?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 20: 2-12.