Do Prices Reflect Costs?

A study of the price- and cost structure
of retail payment services in the Swedish
banking sector 2002

Gabriela Guibourg and Björn Segendorf

OCTOBER 2004

Do Prices Reflect Costs?

A study of the price- and cost structure of retail payment services in the Swedish banking sector 2002

Gabriela Guibourg and Björn Segendorf[*]

Abstract

We estimate private costs in the Swedish banking sector for the production of payment services and investigate to what extent the price structure reflects the estimated cost structure. We find that (i) banks tend to use two-part tariffs but (ii) variable costs are poorly reflected in transaction fees towards both consumers and corporate customers. (iii) there exist large cross subsidies between different payment services, foremost from acquiring card payments to cash distribution to the public, while payment services as a whole is not subsidized.

Keywords: retail payments, two-part tariffs, private costs, price structure, economies of scale

JEL-Classification: E58, G21, L11, L13, L14, L89

1 Introduction

In the public debate banks are often criticised for how they price their payment services. This critique seldom takes the banks' costs of producing the services into account and hence often neglects the role that prices have in providing information on production costs. The fact that the total cost of producing payment services in a country is substantial,[1] and that experiences from Norway[2] and Sweden[3] show that the demand for payment services is sensitive to prices, underline the importance of a well-functioning price system in the market for retail payments

With the price system's key role in providing cost information to consumers as a starting point in, this study aims to (i) estimate banks' private costs for producing different payment services, (ii) investigate to what extent the price structure reflects the cost structure. Our main findings are as follows. First, there are considerable differences in costs between payment instruments. As expected, paper based payments are more costly to produce than electronic payments, and debit card payments are less costly than credit card payments and cash withdrawals. Second, costs are poorly reflected in prices, in particular towards private customers. Third, the provision of payment services gives rise to an annual profit. However, cash distribution to the public is to a large extent financed through substantial cross subsidies from other payment instruments, foremost acquiring card payments. Hence, the banks could reduce total production costs and probably increase overall profitability by adopting more cost-based pricing strategies that would re-orient demand towards less costly payment services.

The costs of different payment instruments and cost recovery from the perspective of the banks have been investigated in a few earlier studies. A few other studies have empirically investigated the efficiency loss from the divergence between user costs and costs to society. The studies typically differ somewhat in the types of cost that are included.

Humphrey and Berger (1990) estimate private and social costs for nine payment instruments in the U.S. They examine, for each payment instrument, whether private average cost to users (i.e., prices) equals the corresponding average cost to society. The latter is defined as the real consumption of resources associated with the production of each payment. These costs were divided into production costs and user costs. In some cases float was added or subtracted to reflect private costs.[4] According to the estimations, there is a distortion explained by float, especially for checks.[5] Payments by credit- and charge cards give rise to the same type of distortion but of

---

[1] Humphrey and Pulley (1997) estimate the cost in the US as being as high as 3% of GDP.

[2] Humphrey et al. (1998).

[3] Nyberg and Guibourg (2003).

[4] Float is the time it takes for the payment to be completed from the point in time the payment amount is debited from the payer account to the point in time it is credited the payee's account. During this period the bank or banks involved in the transaction can invest the funds and earn interest.

[5] The time gap between the initiation of payment and the point in time where the payer's account is debited gives the payer an implicit credit and hence interest on the paid amount during the time gap. For large amounts the payer may even have a net income from checks and this may delay the transition from paper-based checks to corresponding electronic payments.

lower magnitude than for checks. The distortion was therefore deemed to be less harmful.

Wells (1996) replicated Humphrey and Berger's estimations on a new and richer cost data set for checks and ACH-payments[6] to see if float still could explain the relatively slow transition from paper-based to electronic payment instruments in the U.S. She used a weighted average of different cost estimates for each instrument; for example, the cost for a check payment was constructed as a weighted average of costs for different categories of check payments such as private-, corporate- and government payments. Her conclusion was that income from float was less than in Humphrey and Berger's study and that the results did not support their hypothesis of float being the cause of slow transition. The difference can be explained by a faster and more efficient check clearing process and lower interest rates in the latter study. Instead, Wells believed the inertia, if it existed, could be caused by economies of scale and network externalities.

Robinson and Flatraaker (1995a) study the relative cost efficiency of different payment instruments by analysing banks' costs in the Norwegian payment system. Cost recovery and financing is analysed in a second study, Robinson and Flatraaker (1995b). Both these studies were follow-ups on similar cost studies carried out by Norges Bank in 1988.[7] In their first study, Robinson and Flatraaker classify costs as either fixed or variable. Differences in variable cost are then used to compute potential cost savings from going from paper-based to electronically initiated payments. In the second study average cost is estimated for each instrument. It turns out that a larger part of the costs fall on the paying bank and that electronically initiated payments are more cost efficient than paper-based ones. They showed that instruments that had experienced a large change in volume between 1988 and 1995 also showed large changes in average costs, i.e. cost estimations were sensitive to the exploitation of economies of scale. Transaction fees covered 60% of average variable costs in 1995 compared with 20% in 1988.

An unconditional comparison with our results can be misleading. Their classification of costs as either variable or fixed differs to some extent from our classification, and almost 10 years have passed between the two studies, a time during which the market for retail payments has developed rapidly. However, if differences in the volume of transactions in similar payment categories are considered, the cost estimates obtained by Robinson and Flatraaker do not differ much from our estimates. For example, the share of costs for card payments in Sweden is twice as large as the corresponding share in Norway in 1994, while their share of the total number of transactions is almost 1,8 times larger. For electronic credit transfers, paper-based credit transfers, and direct debits, similar comparisons between the studies show almost no differences. However, there are two exceptions. In Sweden, costs associated with ATM withdrawals are almost 11 times higher than in Norway, while their share of the total number of transactions is only 1,4 times larger. Furthermore, the share of costs for credit transfers over the counter is 5 times larger in Norway than in Sweden, but these payments' share of the total number of transactions was 40 times larger in Norway at the time. Economies of

---

[6] Electronically initiated credit transfers.
[7] Fidjestöl et al. (1989a and 1989b).

scale may explain the deviation for credit transfers over the counter but not for ATM withdrawals.

Humphrey et al. (1998) used the Norwegian data to estimate own- and cross-price elasticities for different instruments and to estimate the potential cost savings from going from paper-based to electronic payments. They showed that the demand for payment services is elastic with respect to changes in relative prices, and that there is a substantial potential for cost savings from an automation of payment services.

The latest contribution to the Norwegian series of cost studies is made by Gresvik and Owre (2003) in a descriptive study where they use an activity-based rule for allocating fixed costs. Comparing with the previous studies by Robinson and Flatraaker, Gresvik and Owre show that mainly because of the shift towards electronic payments, unit costs associated with the provision of payment services have decreased over time and the income that banks derive through direct pricing of these services has risen.

There are similarities but also differences between the current study and the previous ones. We focus on the signalling aspect of prices in the daily use of payment instruments and thus relate transaction fees to variable costs. In that respect our study differs from Humprey and Berger (1990) and Wells (1996) and is closer to Robinson and Flatraaker (1995a, 1995b) and Gresvik and Owre (2003). The main difference to the latter two studies is the signalling aspects of prices. When discussing welfare aspects we also make use of a relevant recent theoretical literature, e.g. Rochet and Tirole (2003), Wright (2004), and Bolt and Tieman (2003) to study the role of interchange fees in two-sided markets[8].

The payment process, framework and assumptions are presented in Section 2. Section 3 describes data and Section 4 contains analysis and results. Section 5 discusses our results.

2 The model

2.1 The payment process

Payments can be divided into two categories; cash and non-cash payments. The latter refers to methods of payment between accounts in financial institutions, typically banks. Only for this category it is meaningful to talk about a generic payment process, see Figure 1. It consists of the following main steps: (i) *authentication* of the involved parties, (ii) *validation* of the payment instrument, (iii) *verification* of the payer's ability to pay, (iv) *authorisation* of the transfer between the payer's bank and the payee's bank, (v) *information transmission* between the two banks and, possibly, a clearing organisation and (vi) *clearing and settlement.*[9] The process may look somewhat different depending on the payment instrument (credit transfer, card, cheque, direct debit) and channel/technology (on-line, off-line, paper based, etc.). Depending on the payment instrument used, the payment may be initiated by either the payer or the payee. Credit transfers are initiated by the payer and they may be initiated in either

---

[8] See Section 4.3.

[9] BIS (2000) defines a transaction process consisting of components (i)-(v). The payment process can then be viewed as an extended transaction process including clearing and settlement.

paper or electronic form. Card payments, direct debits and checks are initiated by the payee and are, with the exception of checks, generally initiated in electronic form. Two examples will help to illustrate these differences.

Example 1: A paper-based or electronic credit transfer is initiated by the payer's request to his bank to transfer some amount to the payee's account.[10] His bank performs steps (i)-(iv), debits the payer's account and transmits the payment data to a clearing house. This results in an interbank transfer and an information transmission to the payee's bank which credits the payee's account.

Example 2: For a debit card transaction, authentication, validation, verification and authorisation, i.e. steps (i)-(iv), are typically made on line at the point of sale (henceforth POS), e.g. by a PIN-code. The payee's request and the payer's authorisation are often routed over telecommunication lines and through a computer that works as a telephone switch, which contacts the card issuing and acquiring banks during the transaction process described in steps (i) – (iv). Then a clearing and settlement process takes place, an interbank transfer is made and subsequently the payer's and payee's accounts are debited and credited, respectively. For delayed debit cards (henceforth charge cards) and credit cards, the payer pays his financial institution later. A cheque payment is similar to a card payment but is paper-based and less automated, e.g., the payee has to present the cheque to his bank which sometimes is required to present it to the issuing bank. In the case of a direct debit, the request for payment is also initiated by the payee, usually through a preauthorised agreement with the payer. The payment request from the payee's bank to the payer's bank is presented in electronic form.

Figure 1. The generic payment process. The numbers refer to the example.
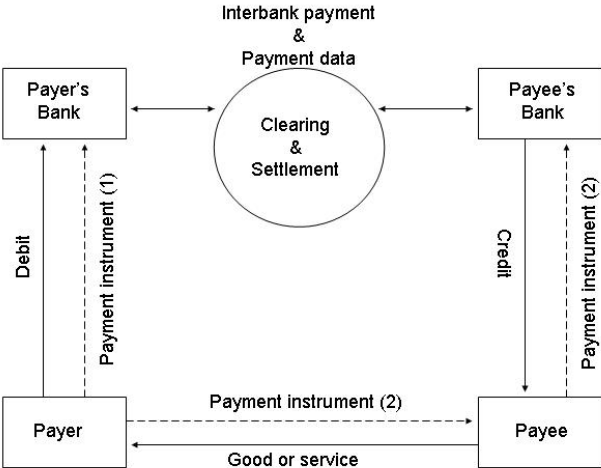
A transaction gives rise to costs in all involved banks. Typical for payments initiated by the payee is that the transaction requires the payment of a special fee between the

---

[10] This request can sometimes be made directly to the clearing house but this requires an agreement with the bank. The clearing house then has to perform steps (i)-(iv) in contact with the bank.

banks involved, the so called interchange fee,[11] and typically it is the payee's bank that pays the payer's bank. In the case of card payments the card issuer pays a fee to the acquirer and in the case of direct debits the payee's bank pays a fee to the payer's bank. Thus, the interchange usually flows in the opposite direction of the payment of the transaction itself. However, both the acquirer bank's and the card issuing bank's production costs need to be covered by customer fees. According to this explanation, the interchange fee may be seen as a way for the acquirer to cover the portion of the issuer's costs that are not covered by customer fees. Interchange also play an important role for the growth of two-sided markets, see Section 4.3. Sometimes cards are used to make cash withdrawals at ATM terminals. If the withdrawal is made with a card issued by a bank other than the bank that owns the ATM, the issuing bank reimburses the ATM owner for the amount of the withdrawal and, in addition, pays an interchange fee. Here the flow of the interchange payment goes in the same direction as the underlying payment itself. The interchange fee can thus be viewed as the price that the card-issuing bank pays to the ATM-owner for the service provided but the fee can also be explained by the presence of network effects. Bank customers' utility of holding a card that allows for ATM withdrawals increase in the number of ATM terminals. Expectations of interchange income create incentives for banks to increase the number of terminals and hence the utility of the issuing bank's customers.

2.2 Approach

The production of payment services is characterised by economies of scale and scope. Under these conditions marginal cost pricing leads to negative profits. A pricing structure that both provide cost information to consumers while allowing for full cost recovery is two-part tariffs. In the following we use two-part tariffs where the variable part is set equal to the relevant marginal cost as our benchmark. We thus focus on the short-run perspective where our benchmark allows banks to provide information to guide consumers in their day-to-day choice of payment services. The short-run focus reflects our key view that transaction fees play a key role for consumers' choice. In particular, most Swedish households already have access to most of the payment services we study below, in particular debit cards (including ATM functionality) and credit transfers by mail, internet and over the counter. Our decision to adopt a short-run focus also reflects our concern regarding the issues and problems associated with the distribution of fixed costs. We believe that adding this dimension may give rise to more problems than what is motivated by the benefits.

2.2.1 The hypothesis

The hypothesis that transaction fees equal marginal costs for individual payment services is tested for variable costs against corporate- and private consumer fees and a representative (market average) fee structure. We do this on data for individual banks and on the price- and cost structure of an average bank that is constructed using individual bank data.

---

[11] Checks are an exception to this rule. There is no interchange fee for them.

Data on fixed costs and fixed fees are used to calculate the average bank's profit and, for the sake of completeness, we account for it together with data on variable costs and fees.

2.2.2 Simplifying assumptions

We make two simplifications. First, float is not included in the calculations. In Sweden the payment process is typically completed within one day and the time during which float is generated is usually quite short, even though the point in time where the payment is credited the payee's account, and where it starts to generate interest, may differ. Interest rates have also been low during the period under study.[12] Secondly, the banks' fixed costs for offering accounts are not included in the study. This service is a prerequisite for all account-based payments. However, the variable cost for crediting and debiting accounts is included. This is a consequence of our focus on variable costs and the difficulty of distributing fixed cost among different payment services. These accounts may also be used for other purposes than payments, e.g. short-term savings, which add another difficulty. The banks' interest-margin revenue is excluded for the same reason.

Different payment services are seldom perfect substitutes. For example, debit- and charge cards differ in that a short-term credit is included in the charge card and the difference between a charge- and a credit card is the possibility to prolong the period of credit. However, when making cost- and price comparisons between different instruments one implicitly assumes a high degree of substitutability. This may not always be true and the results must therefore be interpreted with care.

3 Data

This study relies on price- and cost data that for each payment instrument and channel allows for computation of (i) price structure, i.e. variable and fixed fees, (ii) marginal costs, and (iii) fixed costs. Data was collected from the four largest Swedish banks, which together comprise 92 percent of the card and credit transfer markets and 96 percent of the cash distribution market.[13] The data is from 2002.

---

[12] The Swedish Financial Supervisory Authority, states in a report (Finansinspektionen (2002b)) that incoming credit transfers to private consumers starts to generate interest the first banking day after the account has been credited. Similarly, for outgoing payments interest is generated until the last banking day prior to the day the account was debited. In an example they calculate that a giro payment generates a float of SEK 0.32 of which the sender pays SEK 0.2. Per consumer this makes an annual cost of SEK 19 for both incoming and outgoing payments of which SEK 12 is for outgoing payments. We have, however, chosen not to use this information since (i) it only concerns private consumers, (ii) it concerns only credit transfers, (iii) it is unclear to what extent the example reflects actual interest rates and average amount transferred and (iv) whether the average number of float days is derived from actual payment patterns or is calculated in some other way. Moreover, the average number of float days seems to be inconsistent with an earlier report (Finansinspektionen (2002a)) and also the multilateral agreements regarding bank giro services.

[13] There is some data on credit transfers - both volumes and costs - that is missing, as one of the reporting banks did not provide complete information. According to information from the reporting bank, the missing data should not influence the results as regards unit costs. This is because the costs for the missing credit transfers are the same as for the reported ones.

3.1 Cost data

The banks were asked to provide data on variable costs that arise in the different stages of production of each payment service, i.e., transaction process, clearing, and settlement.[14] These costs include for example transaction fees to central switch networks, bilateral fees between the banks involved, and transaction fees to a central clearing house, an international clearing network or the settlement system. Also mail costs, costs associated with transactions initiated over the counter and credit costs associated with transactions initiated with charge or credit cards were provided. The data was collected for each process stage and every payment instrument and channel, i.e. debit, credit and charge cards, checks, credit transfers (electronic, paper-based and over the counter), direct debits and cash withdrawals (ATMs and over the counter). Also data on fixed costs – direct or associated – was collected for each process stage, instrument, and channel. We also asked for data on transaction volumes and the number of private and corporate customers for the different instruments and channels.

3.1.1 The payment instruments

For payment card transactions, the costs for the card-issuing bank and the acquiring bank were collected separately since the two sides represent different services. Here, acquirers pay an interchange fee to card issuers. As regards issuer costs, banks could not distinguish between costs that credit- and charge cards gave rise to and these two categories were reported as one. Variable costs for credit- and charge cards are higher than for debit cards as they include not only transaction process-, clearing-, and settlement costs but also credit costs. As regards acquirer costs, banks did not make any distinction between debit, credit and charge card. All acquiring costs were reported as pertaining to the same category, namely card payment at the POS.

Variable costs for credit transfers were collected separately for credit transfers (i) initiated over the counter, (ii) paper-based, and (iii) electronically initiated, e.g. through the internet. A distinction was also made between transfers transmitted through the Dataclearing system (henceforth DCL) and the BGC-system.[15] The DCL system is operated by the BGC but allows for less information to be embedded in the transfer than the BGC-system does, and is thus used for transfers with lower information requirements. We also distinguished between costs for sending and receiving a credit transfer, respectively.

As regards cash distribution, cost data for withdrawals over the counter at the bank office and withdrawals over the ATM network were collected separately. We also distinguished for ATM withdrawals done on terminals owned by the card issuing bank and terminals owned by other banks using the categories OC/OT (Own Card/Own Terminal), OC/FT (Own Card/Foreign Terminal) and FC/OT (Foreign Card/Own

---

[14] See Section 2.1 or, for a detailed discussion of the different stages. See also BIS (2000).

[15] In Sweden, Bankgirocentralen, BGC, is the main intermediary for retail payments between the banks. Typically both corporate and private customers use the BGC system for recurring payments as for example bill payments, i.e., for transactions between private customers and corporations or between corporations. They can also be used for payments between individuals or corporations and public authorities.

Terminal). For withdrawals on terminals owned by other banks the issuing bank pays an interchange fee and the variable cost is thus the dominating cost for this category of withdrawals.[16]

Fixed costs were divided into two categories; direct and indirect. The direct fixed cost consisted of fixed costs that could be directly associated with a payment instrument and/or channel. Indirect fixed cost could only be indirectly associated with a payment instrument or channel, e.g. a fixed cost for a computer-based accounting system that debits- and credits customers' accounts. Indirect fixed costs were allocated among the relevant instruments or channels according to the number of transactions. Fixed costs for each instruments and channel were then divided with the relevant volumes to obtain the fixed cost per transaction.

3.1.2 Calculated costs

In some cases we had to complete the reported cost data. Banks reported the interchange costs they incurred as acquirers but not the interchange income they had as card issuers. This income was instead calculated by multiplying the volume of outgoing acquiring transactions by the average interchange fee for all reporting banks. Here we assumed that the four large banks formed a closed system, i.e., that total interchange costs equal total interchange revenue for the four banks. The same method was used to estimate the average interchange income for ATM transactions where card issuer and acquirer were different banks.

Another cost not reported by the banks was the forlorn interest revenue that arises on cash held to meet the public's demand for cash. If the bank could return this holding to the central bank, its central-bank account would be credited and it could invest the amount in some alternative interest-generating activity. To estimate this cost, we used available data on the total value of cash withdrawn in 2002 from the ATM system and over the counter at bank branches respectively. Divided by 365, this gives a daily average SEK 214 million that the average bank distributes to the public. We then added some extra 27.5% to meet the fluctuating demand. The forlorn interest revenue cost was then derived by multiplying the estimated cash holding with the average repo interest rate 2002 which was 4.07 percent. Dividing these costs with the number of cash withdrawals over the counter and at ATM terminals respectively gave us the unit forlorn interest revenue costs of SEK 0.12 for the average ATM-withdrawals and SEK 0.06 for the average withdrawal over the counter.

3.1.3 The average bank

For the purpose of estimating the production costs - and service fees - for the different payment services we constructed an "average bank" using a weighted average of the reporting banks' costs. As weights we used the banks' market shares for each payment instrument and channel, which we could calculate from the reported transaction volumes. The costs and volumes for the average bank are described in Table 1[17].

---

[16] The issuing bank may also have a fixed cost if the issuing and acquiring banks have agreed on a two-part tariff for access to the ATM-network.

[17] Data for individual banks are not accounted for reasons of secrecy.

Table 1: The average bank's fixed-, variable and unit costs per payment transaction, 2002.

Payment service		Costs (SEK)			Volumes
		Fixed/unit	Var/unit	Unit costs	(thousands)
Card	Debet	0,43	0,23	0,66	98834
issuers	Charge/credit	0,62	2,85	3,46	13419
Acquirers	Debet	0,09	1,09	1,18	98834
	Charge/credit	0,09	1,09	1,18	13419
Credit transfers (outgoing)	Paper based	0,25	1,76	2,01	51228
	OTC	1,89	4,72	6,62	644
	Electronic	0,41	0,80	1,21	66353
	Direct debits	0,25	-0,02	0,24	27405
Credit transfers (incoming)	Credit transfers	0,16	0,74	0,90	118225
	Direct debits	0,16	1,01	1,17	27405
Dataclearing	Internet	0,28	0,30	0,57	31473
	Received	0,05	0,18	0,23	17123
Cash Withdrawals	OC/OT	4,50	1,37	5,87	38301
	OC/FT	0,08	5,61	5,69	30841
	FC/OT	5,15	-3,18	1,97	30841
	OTC	10,98	0,06	11,04	11170
Cheque		18,05	1,97	20,02	932

3.2 Price data

Price information vis-à-vis private customers is readily available, e.g. through the banks' internet sites. In general, banks tend to use two-part tariffs, both against private consumers and corporate customers. For corporate customers the information is much less transparent and data often had to be collected separately through contacts with the banks. We used listed prices only and consequently did not acknowledge that lower prices sometimes are negotiated. Besides comparing private- and corporate fees with the estimated costs we also wanted to merge the two different fee structures into one.

First the average bank's fees, both towards private- and corporate customers, were estimated using the same method that was used to estimate the average bank's costs, i.e., by computing the banks' market share for each payment instrument and channel within each consumer group. The market shares were used to compute the average bank's prices for private and corporate customers separately. We then computed the average bank's share of transactions initiated by private- and corporate customers, respectively, using the two groups' relative size in terms of volume of transactions for each bank, instrument and channel. The private and corporate variable fees of the average bank could then be merged using these shares as weights.[18]

---

[18] Example: Suppose there are only two banks, A and B. Bank A has 70 percent of the debit card market (B's share is thus 30 percent) and 98 percent of A's debit card transactions are initiated by their private customers, and 2 percent by corporate customers. For Bank B, the corresponding shares are 95 and 5 percent respectively. Then, private debit card transactions constitutes 97 percent (=0.70*0.98+0.30*0.95) of the average banks total debit card transactions and corporate debit card transactions constitutes 3 percent (=0.70*0.02+0.30*0.05). The private and corporate fees were merged using these shares as weights, i.e. the variable fee that the average bank would implement on debit card transactions would equal bank 0.97*fee towards private customers + 0.03*fee towards corporate customers).

Fixed fees apply on a per customer basis and to facilitate comparisons we had to translate fixed fees into per transaction terms. The banks reported the number of private and corporate customers that they had for each payment service. This consumer- and price data was aggregated to apply to the composite bank in the usual way, i.e. to give the number of private and corporate customers of the average bank in each payment segment and the corresponding (annual per customer) fixed fees. Then we multiplied the numbers of customers with the relevant fixed fees and obtained the income that the average bank derives from fixed fees in each payment instrument and customer group. Adding the income from both customer groups and dividing this with the total number of transactions within each payment instrument gave us the fixed fee per transaction.[19]

Swedish banks also offer packages of services to private and corporate customers. This raises the issue of bundling and distribution of fixed fees between services in the same package. Private consumers, as an example, are often offered a combination of debit card and either internet bank or access to paper-based credit transfers. The debit card gives in turn access to ATMs and the internet bank to different types of credit transfers and other services. However, the difficulty to distribute joint fixed fees over different payment services does not pose a major problem here since, firstly, we have only used listed prices per payment service and, secondly, the analysis' focus is on variable fees and costs. Nevertheless it is likely that our calculated fixed fees are somewhat exaggerated because we don't allow for bundling.

As regards variable fees there is another complication. Acquiring banks implement a double transaction fee structure from merchant customers that accept charge- and credit card payments. First there is a flat per-transaction fee and secondly a turnover fee, i.e., a percentage of the transaction value. We calculated the average bank's flat fee and turnover percentage in the usual way, i.e. using the reporting banks' market shares as acquirers of charge- and credit cards as weights. We then needed to express the average bank's turnover percentage as an average turnover fee in terms of SEK. Using data on value and volumes of credit- and charge card transactions, we could calculate the average purchase amount, which we multiplied by the average bank's turnover percentage in order to find the average turnover fee. In 2002, the average purchase for both types of cards amounted to SEK 880 and the average turnover fee was 2.27 percent. Together with the flat fee of SEK 2.04 the average bank charged merchants on average SEK 22.01 (=2.04+0.0227*880) per transaction with charge- or credit card.

---

[19] Example: the annual fixed fee of the average bank for a debit card is SEK 275 for corporate customers and SEK 278 for private customers. The number of customers in each category is 32 124 and 996 891 respectively and the average bank's total income from fixed debit card fees amounts to over SEK 286 million (275*32 124+278*996 891). The total number of debit card transactions of the average bank is 395 million. However, besides POS transactions, debit cards are also used for cash withdrawals at ATMs[19]. There were 400 million such withdrawals making the total number of debit card transactions of the average bank 795 million. Hence, the fixed fee per transaction is SEK 0.36.

Table 2: The average bank's fixed and variable fees per transaction in 2002.

Payment service		Representative bank's fees per transaction			
		Average fixed fee	Average Variable fee	Consumers variable fee	Corporate customers variable fee
Card issuers	Debit	1,76	0,00	0,00	0,00
	Charge & credit	2,54	0,00	0,00	0,00
Acquirers	Debit	0,00	2,04	-	2,04
	Charge & credit	0,00	22,01	-	22,01
Credit transfers outgoing	Paper based	2,86	0,44	0,00	1,26
	Over the counter	0,00	41,93	41,93	41,93
	Electronic	4,02	0,17	0,00	0,34
	Direct debits	0,00	0,00	0,00	-
Credit transfers incoming	Credit transfers	0,00	0,00	0,00	0,00
	Direct debits	0,00	1,50	-	1,50
Dataclearing	Internet	0,00	0,00	0,00	0,00
	Received	0,00	0,00	0,00	0,00
Cash Withdrawals	ATM-OC/OT	1,65	0,00	0,00	0,00
	ATM-OC/FT	1,65	0,00	0,00	0,00
	ATM-FC/OT	-	0,00	0,00	0,00
	Over the counter	0,00	0,00	0,00	0,00
Cheque	Issuing	0,00	24,82	24,82	24,82

4 Empirical analysis

Recall the benchmark that transaction fees should equal marginal cost and that we use average variable cost as an approximation of marginal cost. Hence, below we examine whether (i) transaction fees equal average variable costs and (ii) relative prices reflect relative costs for some payment services that we consider to be close substitutes. Finally we discuss cross subsidies in the average bank and the profitability of different areas of payment services.

4.1 Testing P=AVC

As described in Section 3.2 and Table 2 banks tend to implement a two-part tariff structure. In Diagram 1 we compare variable fees and costs for each instrument and channel that was accounted for in Tables 1 and 2. The diagram shows high transaction fees for credit transfers made over the counter, acquiring services for charge- and credit card transactions and cheques. There are no transaction fees for card holders on card payments or on cash withdrawals. Regarding transaction fees on credit transfers (BGC) there is a difference between consumers and corporate customers, i.e., the latter meet transaction fees on both paper-based and electronically initiated credit transfers while the former do not. A more drastic interpretation of the situation would be to say that there seem to be a large disparity between variable costs and fees for all payment services with the exception of acquiring services for debit cards and direct debits.

Diagram 1: Variable costs and transaction fees in 2002
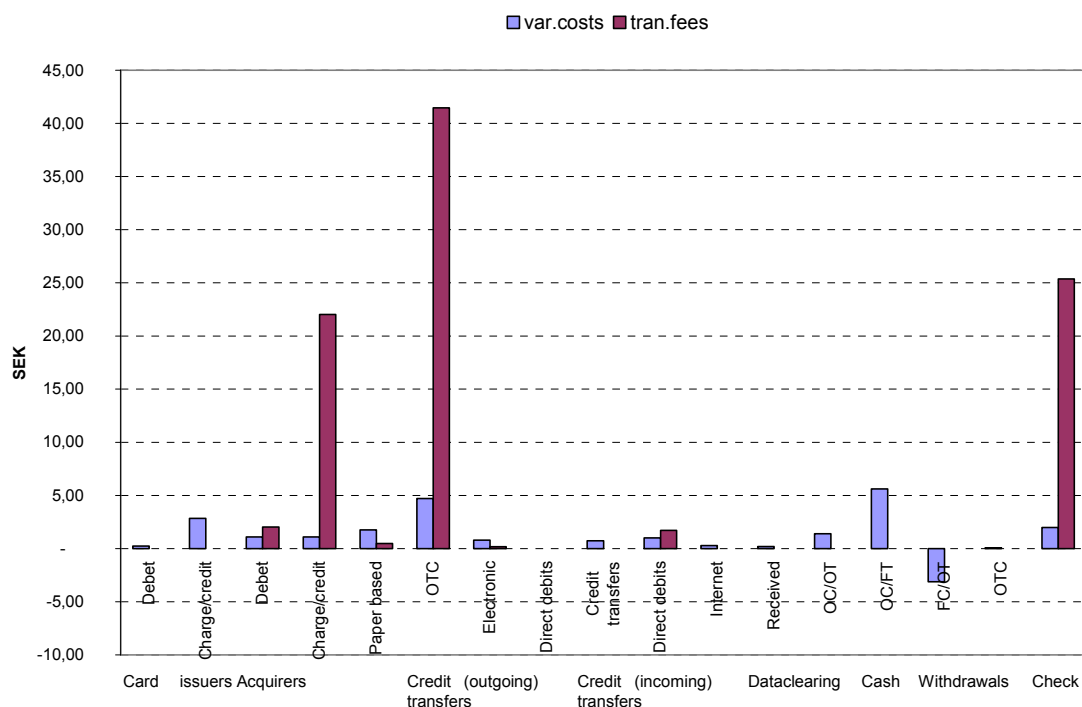
To see whether the differences are statistically significant we use the following test: for each payment instrument and bank we compute the difference between the transaction fee and the variable cost, i.e. the variable contribution for each instrument and bank. We compare variable costs with average transaction fees and with corporate- and private transaction fees separately. We then calculate the corresponding average contributions in the banking sector. This is what we call the individual bank approach since we make the comparison on the bank level. We test whether this average contribution is equal to zero at both 5 and 10 percent significance levels. We also repeat the tests using a slightly different approach, henceforth the average bank approach. Instead of computing the contribution for each bank and instrument and then the average of these individual banks' contributions, we compute average transaction fees and average variable costs separately. Comparison of the average transaction fee[20] with the average variable cost gives the average contribution of each instrument for the average bank. Notice that some services are not available to all categories of customers, e.g., acquirer services are only offered to firms. The number of services tested therefore differs between different customer categories. Table 3 summarises the results of the tests.

---

[20] As before, we compare total, private and corporate fees separately.

Table 3: Statistical inferences in the comparison between transaction fees and private variable costs*.

TEST H0: (transaction fee - variable cost)		DECISION RULE					
		5% Significance level Accept if -3.18<t<3.18; REJECT= 5			10% Significance level Accept if -2.35<t<2.35; REJECT = 10		
		Individual Bank Approach			Average Bank Approach		
		All customer fees	Priv.cust. only	Corp.cust. only	All customer fees	Priv.cust. only	Corp.cust. only
		t-value	t-value	t-value	t-value	t-value	t-value
Card issuers	Debit	n.a.		n.a.	n.a.		n.a.
	Charge & credit				-10	-10	-10
Acquirers	Debit		-			-	
	Charge & credit	+5, +10	-	+5, +10	+5, +10	-	+5, +10
Credit transfers outgoing	Paper based		-5, -10			-5, -10	
	Over the counter	+5, +10	+5, +10	+5, +10	+5, +10	+5, +10	+5, +10
	Electronic		-5, -10				
	Direct debits			-			-
Credit transfers incoming	Credit transfers	-5, -10	-5, -10	-5, -10	-5, -10	-5, -10	-5, -10
	Direct debits		-			-	
Dataclearing	Internet	-10	-10	-10			
	Received	-5, -10	-5, -10	-5, -10			
Cash Withdrawals	ATM-OC/OT	-10	-10	-10			
	ATM-OC/FT	-5, -10	-5, -10	-5, -10	-5, -10	-5, -10	-5, -10
	ATM-FC/OT	+5, +10	+5, +10	+5, +10	+5, +10	+5, +10	+5, +10
	Over the counter	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Cheque	Issuing	+5, +10	+5, +10	+5, +10	+5, +10	+5, +10	+5, +10
	10 % level	9 of 17 (53%)	10 of 14 (71%)	9 of 16 (56%)	7 of 17 (41%)	7 of 14 (50%)	7 of 16 (44%)
	5 % level	7 of 17 (41%)	8 of 14 (57%)	7 of 16 (44%)	6 of 17 (35%)	6 of 14 (43%)	6 of 16 (38%)

* "n.a." has been used to indicate where the statistical formula could not be applied, e.g. either because of the variance was zero or that there was only one observation.

The comparison of transaction fees with variable costs shows that the hypothesis that transaction fees equal average variable costs can be rejected on both levels of significance and for both approaches for the following services: acquiring credit and charge cards, credit transfers OTC, incoming credit transfers (BGC), cheques and ATM withdrawals – OC/FT and FC/OT. The individual bank approach gives stronger results and rejects the hypothesis also in the case of received credit transfers (DCL) on both levels of significance and internet credit transfers and ATM withdrawals OC/OT on the higher significance level. On both levels of significance and in both the individual bank and the average bank approach, the efficiency condition is rejected for a larger number of instruments when only fees toward private customers are considered than when fees toward corporate customers are taken into account. Thus, pricing towards private customers appears to be less cost-reflecting than pricing of corporate customers' payment services.

4.2 Relative prices and costs

Even if variable fees do not reflect variable costs the price system may still provide important information on *relative* costs of different payment services through relative prices. Table 4 compares relative variable fees with relative variable cost and relative average costs for payments services that are judged to be close substitutes. Thus we have not made a comparison between BGC credit transfers or direct debits and DCL credit transfers since they often are used for different purposes. Nor are payments over the counter included in the discussion below. They are typically not viewed as close substitutes to corresponding paper-based or electronic payment services. The volumes are also relatively small. As before, relative prices are divided into prices facing consumers, corporate customers, and the average payer, i.e. a weighted average of private consumers and firms.

Table 4. The average bank's relative prices and costs.[*]

		Relative transaction fees			Relative costs	
		Private	Corp.	Average	Variable	Average
Card/Card	Debit/Charge	"0/0"	"0/0"	"0/0"	0,08	0,19
Card/Cash	Debit/ATM-total	"0/0"	"0/0"	"0/0"	0,18	0,14
	Charge/ATM-total	0	0	0	2,23	0,75
Credit transf.	Paper-Giro/Electr.-Giro	"0/0"	3,70	2,64	2,20	1,66
	Paper-Giro/Direct-debit (out)	"0/0"	-	-	-88,00	8,49
	Electr-Giro/Direct-debit (out)	"0/0"	-	-	-40,00	5,12

* Relative prices and costs are computed in accordance to the expression in the second column. A relative value lower (greater) than 1 indicates that the "numerator" is cheaper (more expensive) to buy or produce than the "denominator".

Table 4 confirms the preliminary conclusions from the data presented in section 4.1 regarding relative variable costs. From the viewpoint of the card-issuing bank, debit card transactions are cheaper to produce than both charge- and credit card transactions since the latter also include the cost of extending credit to the user. For them debit card transactions are also cheaper to produce than corresponding cash withdrawal while cash is less costly than charge- and credit cards.[21] Paper-based credit transfers are more expensive to produce than their electronic counterpart. Direct debits are considerably cheaper to produce than both paper-based and electronic credit transfers.[22]

Average relative costs show the same qualitative pattern as relative variable costs even if less strongly. The cost advantage of direct debits over both paper-based and electronic giro is smaller because the number of direct debit transactions is relatively small compared to both paper-based and electronically initiated credit transfers. Thus economies of scale are not well exploited and unit fixed costs are consequently relatively large. Also, not only debit cards but also charge cards are cheaper to produce than corresponding cash withdrawals in terms of average costs. This is also explained by relatively large fixed unit costs for cash withdrawals.

Table 4 shows that almost no information about these relative costs for close substitutes is passed on to consumers through the price mechanism. Especially in the case of consumers, relative prices convey no information at all. They face zero variable fees for all kinds of card payments. In the POS market, relative prices give the consumer no information to support a cost-efficient choice between debit- and charge cards and debit cards and ATM cash withdrawals. In both situations, the debit card has a large cost advantage not reflected in private variable fees. The same problem arises in the market for remote payments, where zero variable fees to consumers do not reflect the large cost advantages that electronic credit transfers have over paper-based credit transfers. Of the situations captured in Table 4 there is only one in which a consumer may face a transaction fee, namely if a charge- or credit card is used to withdraw cash from an ATM. This fact is not displayed in the table since the number of

---

[21] Here cards are compared with an average ATM withdrawal. However, recall that ATM costs varies with the issuing- and ATM-owning banks, see Diagram 1 and Table 2.
[22] An outgoing direct debit generates an interchange income or negative cost for the sending bank. Thus the negative sign in the variable cost ratios

such transactions are very small compared to withdrawals with debit cards and thus not reflected in the average transaction fee for ATM withdrawals. However, if the consumer considers making a withdrawal using his/hers charge or credit card, the relative transaction fee signals a cost advantage for the use of the charge- or credit card at the point of sale instead of paying with cash. This price signal contradicts actual relative variable costs. In terms of the efficiency of the price system, this is not a real problem as such withdrawals are rare.

Corporate customers face, just as private consumers, zero variable fees for card payments and a variable fee for ATM withdrawals with charge- or credit cards and, thus have the same economic incentives when choosing between these substitutes. In contrast to private consumers however, they face variable fees for electronic and paper-based credit transfers that give them correct signals compared to relative variable costs. These comparisons only capture situations where the firm act as payer, i.e. the initiation of direct debits is not considered. Regarding acquiring services the relative variable fee suggest a strong cost advantage for the debit card that is not reflected in the actual relative variable- and unit costs. To sum up, we find that variable costs are reflected in relative variable prices only to a low extent.

4.3 Welfare aspects

According to our results, and given that user demand is price elastic, a pricing strategy based on variable costs in line with our benchmark would lead to a larger use of debit cards, less use of credit cards and cash, larger user of electronic credit transfers and direct debits, and less use of paper based credit transfers. Such changes in the choice of payment instruments would lower costs for the banks involved. However, this would not automatically lead to efficiency gains for the whole society. In order to be able to discuss the welfare aspects of changes in payment patterns, we would need to consider two further issues.

1. We need to consider costs that private consumers and corporate customers incur from the consumption of payment services in addition to the fee, e.g. cost of time needed to initiate or receive the payment, to handle cash, cheques, etc. We would also have to investigate production costs in the market for infrastructure services, i.e., clearing services from e.g. VISA, MasterCard/Europay, and BGC, switch services from e.g. CEKAB, and settlement services from the Riksbank and one of the major banks where certain card transactions are settled.

2. Price equal to marginal cost gives an efficient allocation of resources if there are no market imperfections. In the market for retail payments there are two market imperfections that affect the welfare analysis: economies of scale in the production and the presence network effects. Economies of scale seem, as already said, be handled by two-part tariffs. In what follows, we therefore focus on network externalities and briefly discuss how their presence may require a deviation from marginal cost pricing and the impact on our results.

Network effects is present in all payment services and would in general require pricing below marginal cost since the positive effect of an additional participant should be compensated for in the price. Often this can be done by a reduction of the fixed fee. These subsidies may be financed by other bank activities or by the society by taxes on other markets depending on which method gives rise to the smallest dead-weight loss. However, in practice it may not be enough to subsidise fixed fees. Sometimes network effects are closely connected with the use of a payment instrument. As an example, for a merchant the value of accepting cards does not stem from the number of card holders but from the number of *active* card holders.[23] The market for card payments is an example of so-called two-sided markets that have been given special attention in the recent economic literature.[24] According to this literature, interchange fees can be explained by the need to exploit network externalities in two-sided markets. In this context, the interchange fee is seen as a part of the pricing structure aimed to create appropriate incentives for the two sides of the market. If the demand of the two sides of the market have different price elasticities, a lower price should be charged from the group with the higher price elasticity, in order to increase the size of this side of the market. This would explain why transaction fees for card payments are taken mostly from the merchant side.

Network externalities are exploited to a larger or smaller extent in each segment of the market, often reflecting the maturity of the product. In Sweden, credit transfers is a mature product. Thus, it is reasonable to believe that network externalities are already exploited to a large extent since almost everyone, or at least every household in Sweden, has a transaction account and access to this service. Hence, there is probably no need for a subsidy there. The same could be said of the ATM market. For these services, the private costs that have been estimated in this study may serve as an approximation for the corresponding social costs. Regarding the two-sided markets, i.e., the market for card services and direct debits, there is probably potential for further exploitation of network effects. Hence, from a social point of view, interchange is likely to be well motivated in these markets. However, the current interchange fees may be too high or too low compared to the optimal interchange. Here, a welfare analysis would greatly benefit from an estimation of optimal interchange fees. Hence, for these services it is difficult to say how the private costs of production relate to the corresponding social cost.

4.4 The profit of the average bank

From Tables 1 and 2 it follows that all payment services except for cash distribution gives rise to net revenues for the bank. Credit transfers as a whole give a net revenue of SEK 161 million. However, most net revenue is generated by card payments, SEK 460 million of which SEK 280 million comes from acquiring charge- and credit card transactions which constitute only about 12 percent of the total volume of card transactions. Issuing cards gives in total net revenue of SEK 96 million. The only net

---

[23] The Swedish e-money scheme CashCard is an example of this. It was introduced by three of the major banks and at most all four of them participated in the scheme. The e-money functionality was incorporated in ordinary debit cards which greatly contributed to the number of CashCard holders. However, the number CashCard holders using the card actively was relatively low, as was the turnover in the system. The potential network effects could never be exploited and the scheme was terminated in 2004.

[24] See for example Rochet and Tirole (2003) and Bolt and Tieman (2003).

cost comes from cash withdrawals, either due to large volumes (ATM) or high loss per unit (over the counter). The average bank's annual net cost associated with cash withdrawals amounts to SEK 466 million (including checks) of which SEK 347 million come from ATM-services. In total, fixed and variable fees cover 69 percent of the average bank's cost for producing payment services while variable fees cover 114 percent of variable costs. However, these numbers depend to a large extent on the revenues from acquiring services for charge- and credit card transactions.

The average bank's total interchange fee (income) as acquirer of card payments (card issuer), amounts to SEK 78.6 million per year. The corresponding income (cost) for outgoing (incoming) direct debits is SEK 26.0 million and SEK 158.5 million for ATM withdrawals respectively. However, since we have assumed that the four large banks constitute a closed system, interchange costs and revenue will cancel out for the average bank. As regards forlorn interest from cash, there is a private annual cost for the ATM-owing bank of SEK 10 million and of SEK 1 million for cash withdrawals over the counter.

Figure 2: The average's banks net result in the provision of payment services
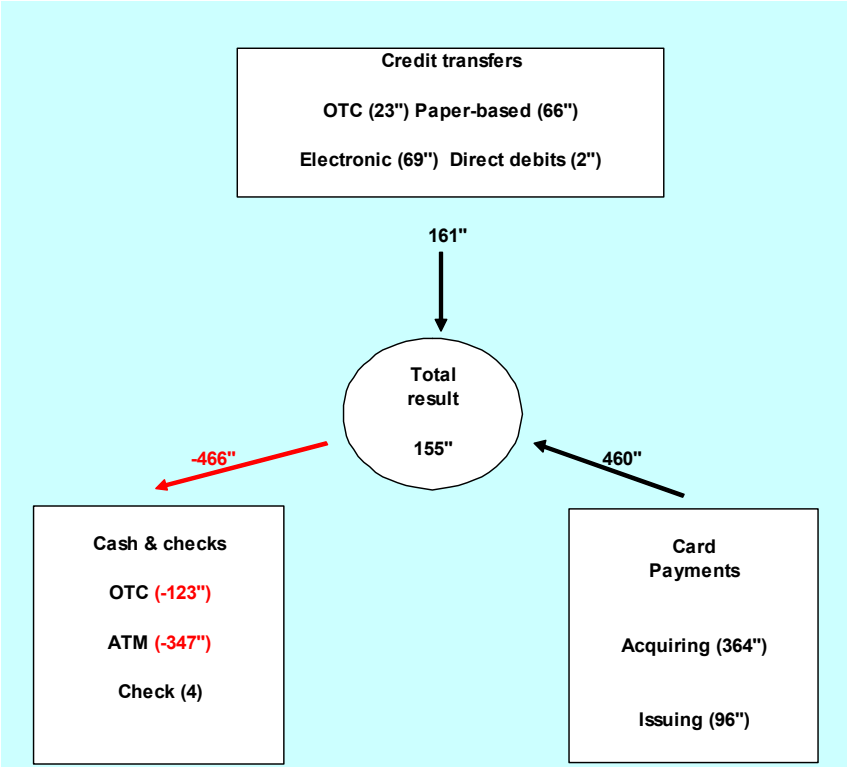
5 Discussion

Our study shows that the banks to a large extent use a two-part pricing structure when pricing payment services. However, a positive variable fee is, with one exception, taken only from corporate customers. With few exceptions, consumers do not face any variable fees and therefore receive insufficient price signals in their choice of payment instruments. If they did receive correct price signals, they would have economic incentives to shift to direct debits and electronic credit transfers from paper-based ones, and to debit cards from charge- and credit cards.

Both fixed and variable fees apply for corporate customers but the fee structure is not transparent. Corporate customers probably find it difficult to get a clear picture of relative prices and possibly of the costs they incur themselves by their choice of payment instrument. Lack of transparency also limits price competition. Like private customers, corporate customers do not face any variable fee for their use of card payments. Thus, they do not get correct incentives in their choice between on one hand debit cards and on the other charge- or credit cards.

The results discussed above may, at least to some extent, explain the differences in utilisation of the retail payment system between Sweden and the other Nordic countries, i.e. Denmark, Finland, and Norway, where it is more common to impose transaction fees on ATM withdrawals.[25] When making country comparisons of retail payment systems the usage of cash in an economy is usually measured by relating the stock of notes and coins in circulation to GDP, and the usage of card is measured by card payments per capita. Table 5 shows that Sweden has a relatively high use of cash (first column), a low card usage (third and fourth columns) and relatively few card terminals per capita which is fully consistent with our results.

Table 5. Retail payment statistics 2001.[26]

	BNP/M0	Electr. Payments	card transact. /capita	Cards per capita	Terminals per 1000 inhabitants
Denmark	2,90%	n.a.	87	0,69	n.a.
Finland	1,84%	88%	76	1,31	12,9
Norway	2.75%	87%	99	1,43	15,9
Sweden	4,48%	89%	45	0,85	9,9

Source: Blue Book 2003 and Norges Bank

All in all, the payment market gives rise to an annual net revenue of in total SEK 620 million to the major Swedish banks. However, net costs occur in the distribution of cash while acquiring of card payments leaves a large positive contribution. Thus, the pricing structure implemented requires cross-subsidising between different payment instruments.

If the banks were to move towards a more cost-based price setting strategy, this would result in the introduction of transaction fees for ATM withdrawals. Private consumers would also face transaction fees on paper-based and electronically initiated credit transfers. Fees for acquiring services would have to decrease. Given the demand's high responsiveness to prices that the evidence suggests, it is possible that a cost-based pricing strategy would give rise to substantial changes in the use of payment services and thus contribute to lower variable costs for the banks. Based on the data on variable

---

[25] In 2002 in Norway, customers paid a fee per withdrawal of NOK 3.76 if they used the own bank's ATMs outside business hours, of NOK 3.89 if they used another bank's ATMs during business hours and of NOK 4.79 if they used another bank's ATMs outside business hours. In Denmark, customers paid a small fee for cash withdrawals at other bank's ATMs and for cash withdrawals at the issuing bank outside bank opening hours. In Finland a small fee applies only for withdrawals at other bank's ATMs.

[26] The year 2001 was chosen because it is the last year with data on M0 in Finland – they adopted the Euro in 2002. Data from 2002 shows the same qualitative pattern. Also, to our knowledge there has only been small changes in price- and cost structures between 2001 and 2002, which enables us to discuss qualitative results based on observations from 2001.

costs and volumes in Table 1, the bank sector could lower their variable costs with SEK 563 million annually if bank customers replaced charge- and credit cards and cash withdrawals with debit card payments. Similarly, a move from paper-based to electronically initiated credit transfers could result in a reduction in variable cost of SEK 197 annually. These calculations do not take into account the impact that a change in consumer behaviour may have on fixed costs in the long run.

An interesting aspect of the banks' pricing strategies is their rationale. Banks are most likely aware of the potential cost-saving effect of more cost-based pricing strategies. One possible explanation to why cost-based pricing has not been adopted is that banks are locked into a prisoners'-dilemma type of situation. Here, every bank would gain from a joint shift to cost-based pricing. At the same time it may be very costly in terms of lost customers and market shares to be the first bank to switch to this more direct and transparent pricing strategy. Then no bank dares to initiate such a process and the banks are locked into the high-cost equilibrium. An alternative explanation is that banks view payment services as complementary with each other and perhaps other business activities. Banks then maximise profits by establishing customer relations by selling cheap payment services and use this relationship to sell more profitable complementary services, e.g., saving products, mortgage loans, financial services, etc. It may also be a combination of the two explanations.

Regarding welfare aspects, one would have to estimate the social cost for the different payment instruments. In addition to the data accounted for here, this requires knowledge of costs in relevant parts of the financial infrastructure, e.g. clearing houses, switches, etc. and additional costs to end users, i.e. private consumers, merchants and other type of firms. This could be one important area for future research. Another interesting area of research is to empirically investigate optimal interchange fees for card payments and direct debits.

Literature

BIS (2000), "Clearing and Settlement Arrangements for Retail Payments in Selected Countries", CPSS, 2000.

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

Fidjestøl, A., D. Flatraaker, and C.J. Vogt (1989a), "Kostnader i betalninsformidlingen" (Costs in the payment system), *Penger og Kreditt* no. 3/89, Norges Bank

Fidjestøl, A., D. Flatraaker, and C.J. Vogt (1989b), "Intekter, kostnader og prissettning i betalninsformidlingen" (Income, costs and pricing in the payment system), *Penger og Kreditt* no. 4/89, Norges Bank

Finansinspektionen (2002a), "Från konto till konto", Rapport 2002:2

Finansinspektionen (2002b), "Från konto till konto IV", Rapport 2002:9

Gresvik, O. and G. Øwre (2003), "Costs and Income in the Norwegian Payment System 2001. An Application of the Activity Based Costing Framework", Working Paper 2003/8, Norges Bank

Humphrey, D. and A. Berger (1990), "Market Failure and Resource Use: Economic Incentives to Use Different Payment Instruments", *The U.S. Payment System: Efficiency, Risk and the Role of the Federal Reserve. Proceedings from a Symposium on the U.S. Payment System Sponsored by the Federal Reserve Bank of Richmond.* Ed. D. Humphrey. 1990, Kluwer Academic Publishers, pp.45-86.

Humphrey, D. and L. Pulley (1997), "The check is in the mail: Why the US lags in the adoption of cost-saving electronick payments", Florida State University.

Humphrey, D., M. Kim and B. Vale (1998), "Realizing the Gains from Electronic Payments: Costs, Pricing and Payment Choice", Arbeidsnotat 1998/1, Norges Bank.

Nyberg, L. and G. Guibourg (2003), "Card payments in Sweden", *Economic Review* 2003:2, Sveriges riksbank, pp. 29-39.

Robinson, P. E. and D-I. Flatraaker (1995a), "Costs in the payment system", *Economic Bulletin*, Vol. 2, Norges Bank, 1995, pp. 207-216.

Robinson, P. E. and D-I. Flatraaker (1995b), "Income, costs and pricing in the payment system", *Economic Bulletin*, Vol. 66, Norges Bank, 1995, pp.321-332.

Rochet, J. and J. Tirole (2003), "Platform Competition in Two-Sided Markets, Journal of the European Economic Association, 1(4), pp. 990-1029.

Wells, K. E. (1996) "Are Checks Overused?", *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 20(4), pp. 2-12.

Wright, J. (2004), "The Determinants of Optimal Interchange Fees in Payment Systems", *Journal of Industrial Economics*, 52(1), pp. 1-26.

Earlier Working Papers:

For a complete list of Working Papers published by Sveriges Riksbank, see www.riksbank.se