

08.02.02. Preliminary. All comments are welcome.

The Effect of New Technology in Payment Services on Banks' Intermediation

Kjersti-Gro Lindquist

Central Bank of Norway, P.O.Box 1179 Sentrum, N-0107 Oslo, Norway.

Tel: +47 22 31 68 73. Fax: +47 22 42 60 62. E-mail:kjersti-gro.lindquist@norges-bank.no

Abstract

In many OECD-countries, payment services in banking have shifted from paper-based giro and check payments to electronic giro and debit card payments. This paper analyses the effect of this change in payment technology on banks' scale properties and input demand within a multiple-output framework using Norwegian bank-level panel data. The dual approach with four variable inputs is applied, and the general model includes random coefficients to capture heterogeneity in the production technology across banks. The results show that electronic payments have increased the economies of scale in the production of deposits more than in the production of loans. Furthermore, the move towards electronic payment services has affected input demand asymmetric, i.e. non-neutral, and the input ratio between labour and both physical capital and materials have decreased.

Keywords: Banking industry; Electronic payments; Technical change; Panel data

JEL classification: G21, O33

Acknowledgement: The author wishes to thank Erik Biørn, Martin Hellwig, Tor Arnt Johnsen, Aris Kaloudis, Terje Skjerpen, Bent Vale, Grete Øwre, and participants at the 23rd SUERF Colloquium in Brussels October 2001 for valuable comments and discussions.

1. Introduction

Over the last couple of decades, banks in Norway, as in most OECD-countries, have gone through important changes. Market deregulation, increased competition, technological innovations and financial crisis have spurred mergers and acquisitions – both domestic and across borders, investments in new technologies and probably also changed banks' strategic behaviour. While several of these issues have been analysed and discussed in the literature, little has been done on the importance of new electronic payment systems, which clearly represent new technology.

In many OECD-countries, particularly in the Nordic area, payment services have shifted dramatically from paper-based giro and check payments to electronic giro and debit card payments, see Koskenkylä (2001). According to Norwegian data, only 14 per cent of non-cash payments were in electronic form in 1987, but by 2000 this had risen to 78 per cent. This change was - at least partly - spurred by banks' pricing policy, see Humphrey et al. (2001). Banks' motivation to offer electronic payment services may be due to both cost saving and competition; Once the necessary investment is made, for the same type of transaction, electronic payments cost banks much less to produce than paper-based payments, see Flatraaker and Robinson (1995), Wells (1996) and Brooz et al. (1998). If customers find electronic payment services more convenient than paper-based services, and hence prefer the first, electronic payment services may be viewed as strategic variables in the competition for customers.

This paper focuses on the importance of new technologies, as represented by new electronic payment services, for the development in banks' scale properties and input demand in the production of loans and deposits. As in general in analyses of bank behaviour, the definition of bank output is a major challenge, and one may argue that payment services, both electronic, paper based and cash payment services, should be treated as outputs. However, payment services contribute marginally to banks' income, and our view is that these services have the character of being inputs in the production of loans and deposits, because banks are obliged to offer these services to attract depositors in particular.

Payment services are either "intermediate inputs" produced within banks by the primary inputs labour, physical capital and materials, or banks use a clearing house and settlement bank to execute the transactions. The latter is particularly important with respect to giro payments, card payments and ATM transactions. The banks are charged for their share of the costs in these systems. One may argue that we ideally should treat electronic payment services as an input similarly to other variable inputs. We are, however, prevented from this, because it

is not possible to disentangle inputs used to produce payment services within banks from inputs used directly in the production of loans and deposits.

Due to its nature, electronic payment systems are probably characterised by positive economies of scale, see Humphrey (1981). When a system is established, increasing the number of transactions costs very little. Our a priori belief is that the observed change in payment technology has affected costs in Norwegian banks negatively – i.e. reduced average variable costs in banking. Since, in general, payment services are more strongly connected to deposits than to loans, we expect changes in payment technology to affect the production of deposits in particular. An effect on loans may be present, however, due, for instance, to the increase in automatic payments on loans. It is clearly of interest to analyse if positive economies of scale within electronic payment systems have affected the scale properties of producing deposits and also loans. The expected cost reducing effect will not be present, however, if banks do not take the opportunity to rationalise. Even if a rationalisation effect is not present, it may still be optimal for banks to introduce this new technology, if the competition for customer aspect is important.

We estimate a system consisting of a cost function in four (primary) variable inputs, i.e. labour, physical capital, materials and funds, and the corresponding cost-share equations. A multiple-output approach is applied, where loans and deposits are treated as two separate outputs. Technical change in banking is captured by two variables; First, to represent the implementation of new technology in payment services, we include the development in electronic payments relative to total non-cash payments in volume terms, i.e. measured in number of transactions.¹ We specify a rather general model in this variable, and in addition to testing for effects on banks' scale properties, we also test if this change in technology has affected input demand asymmetric, i.e. non-neutral. Second, to capture technical change that is not part of the shift from paper-based to electronic payment services, we include a deterministic time trend. This variable is assumed to capture effects on costs from experience, increased knowledge, other innovations and improved production techniques. The inclusion of a time trend is rather common in studies of technical change, see e.g. Hunter and Timme (1986, 1991) for analyses using bank data, but one major weakness of this approach is that we are unable to reveal the sources of technical change. A second disadvantage is that a time trend in general also captures the effects of “non-technical change” variables that are trending and not explicitly included in the cost function, such as deregulation. Therefore, estimated

¹ We focus on the change from paper based to electronic based payments and not on the change from cash payments to paper based payments for two reasons: First, we do not have data on cash payments until very recently. Second, the major change from cash payments to paper based payments, i.e. giro payments, came previous to our observation period.

time trend effects must in general be interpreted with care. An important contribution of this analysis is the identification of the effect of a major technological innovation in banking.

As in general in empirical cost or production function analyses applying micro data, the representation of heterogeneity is important. It is rather common to assume that variation in size, age, management, employees' education, technology etc., can be represented by a firm specific fixed or random intercept term.² Most likely, however, such differences will manifest themselves not only as a permanent, i.e. constant variation in efficiency across firms, but will also result in heterogeneity in slope coefficients. In this case, the standard modelling approach may lead to inefficient estimation and invalid inference, see Biørn et al. (2001). In this paper, a random coefficient approach is applied, which specifies heterogeneity in both intercept term and some slope coefficients. Hence, rather than assuming a priori that all coefficients are constant, i.e. equal, across all banks, we make assumptions about the distribution from which the bank specific coefficients are drawn. The expectation vector in this distribution represents the coefficient of an average bank, while the covariance matrix measures the degree of heterogeneity that is due to the random coefficient variation.

We have access to annual bank-level data for most of the variables in our system over the period 1987-1999.³ With respect to electronic payments, we only have the number of transactions at the industry level, however, while we ideally should have had the number of transactions at the bank level. On the other hand, we expect the share of electronic payments in total non-cash payments measured in number of transactions to be rather common across banks. Because banks participate in a jointly owned system for clearing, and the necessary infrastructure in general is available to all banks, banks can rather easily offer their customers electronic payment services. For instance, the first Norwegian bank to offer e-banking was a small rural saving bank.

The empirical results show that the change in technology in payment services has increased the economies of scale in the production of deposits more than in the production of loans, as was expected. Furthermore, the move towards electronic payments has affected input demand asymmetric, i.e. non-neutral, and the input ratio between labour and both physical capital and materials have decreased.

² In their analysis of large U.S. commercial banks, Hunter and Timme (1991) split their sample according to bank size to capture size effects on scale or technical change properties in particular.

³ The data are measured at the overall bank (firm) level rather than at the bank office (plant) level, which is common when applying accounting data. See Berger et al. (1986) for a discussion of this distinction.

Section 2 presents the econometric model, and the empirical results are presented in Section 3. The main conclusions are summarised in Section 4.

2. The cost-share equation system

Banks' technology is represented by a translog cost function with variable returns to scale as put forward by Christensen et al. (1971, 1973). With this functional form, the impact of technical change, as represented by the introduction of new electronic payment services, can be specified in a general way. The translog cost function is flexible and can be interpreted as a quadratic approximation to a general continuous twice-differentiable cost function in logs that satisfies linear homogeneity in prices.

While Hunter et al. (1990) and Lawrence (1989) find that the standard translog specification provides an adequate fit of bank cost data, Shaffer (1998) shows that the translog cost function disfavors large banks, since economies of scale is exhausted already for relatively small banks. See also the references in Shaffer (1998). In principle, however, the general translog cost function captures even increasing economies of scale with respect to bank size. Scale properties are determined by the sign and size of estimated coefficients in combination with data. One may therefore argue that this is an empirical question, and if increasing or non-decreasing economies of scale with respect to bank size is an important feature of the data, we should expect to find this.

Our most general translog cost function with two outputs and four inputs is given in equation (1).⁴ Deposits (D) and loans (L) are treated as two separate outputs, and labour (W), physical capital (K), materials incl. energy (M) and funds (F) are inputs. Subscript f denotes bank.

$$\begin{aligned} \ln C_f = & \alpha_{0f} + \sum_i \alpha_{if} \ln P_{if} + 1/2 \sum_i \sum_j \beta_{ij} \ln P_{if} \ln P_{jf} + \sum_m \gamma_m \ln X_{mf} + 1/2 \sum_m \sum_n \gamma_{mn} \ln X_{mf} \ln X_{nf} \\ & + \sum_i \sum_m \gamma_{im} \ln P_{if} \ln X_{mf} + \gamma_E \ln EP + 1/2 \gamma_{EE} (\ln EP)^2 + \sum_i \gamma_{iE} \ln P_{if} \ln EP \\ & + \sum_m \gamma_{mE} \ln X_{mf} \ln EP + \gamma_\tau \ln \tau + u_{Cf}, \end{aligned} \quad (1)$$

$i, j = W, K, M, F, \quad m, n = D, L,$

$$C_f = \sum_i P_{if} \cdot V_{if}, \quad i = W, K, M, F, \quad (2)$$

where C_f is total variable costs of bank f; P_{if} is the bank specific price of input i; X_{Df} is total deposits in bank f; X_{Lf} is total loans in bank f; EP is the share of electronic payments in total non-cash payments in the banking industry; τ is a deterministic trend variable included to capture technical change that is not attributed to the shift in payment technology. Due to

⁴ For a discussion of the multi-product representation within the translog cost function, see Caves et al. (1980).

multicollinearity, a very simple specification in τ was chosen; u_{cf} is an added disturbance; V_{if} is the quantity of input i used by bank f .

Application of Shephard's lemma gives the cost-share (S_{if}) equations in (3).

$$S_{if} = \partial \ln C_f / \partial \ln P_{if} = (P_{if} \cdot V_{if}) / C_f = \alpha_{if} + \sum_j \beta_{ij} \ln P_{jf} + \sum_m \gamma_{im} \ln X_{mf} + \gamma_{iE} \ln EP + u_{if}, \quad (3)$$

$i, j = W, K, M, F, m, n = D, L,$

where u_{if} , $i=W, K, M, F$, are disturbances in the cost-share equations. Because the cost shares always sum to unity, that is $\sum_i S_{if} = 1$, any cost-share equation can be expressed in terms of the other equations by using the adding up restrictions (listed below). This also implies a singular error-covariance matrix, but estimation may proceed with the arbitrarily deletion of one cost-share equation. We exclude the cost-share equation of materials.

Theory requires the cost function to be homogeneous of degree one in input prices and cross-price effects to be symmetric. These theoretical restrictions, in addition to the adding up restrictions, are imposed on the general model that we estimate:

Price homogeneity and symmetry: $\sum_j \beta_{ij} = 0$ for all i ; $\beta_{ij} = \beta_{ji}$ for all i and j where $i \neq j$.

Adding up conditions: $\sum_i \alpha_{if} = 1$; $\sum_i \beta_{ij} = 0$ for all j ; $\sum_i \gamma_{ik} = 0$ for $k=D, L, E$.

The intercept term in the cost function, α_{0f} , the slope coefficients α_{if} $i=W, K, M, F$ and γ_{mEf} $m=D, L$, are all specified as random coefficients and capture heterogeneity in the production technology across banks. With this specification, with the exception of the time trend, all the cost function elasticities (first order derivatives with respect to the logarithm of each explanatory variable) include a random coefficient.⁵ Let θ_f denote the column vector containing all the random coefficients of bank f in the model, and let θ denote the common expectation vector of θ_f for all banks, with $\alpha_0 = E(\alpha_{0f})$, $\alpha_W = E(\alpha_{Wf})$, etc.,

$$\theta_f = [\alpha_{0f}, \alpha_{Wf}, \alpha_{Kf}, \alpha_{Mf}, \alpha_{Ff}, \gamma_{DEF}, \gamma_{LEf}]', \quad (4)$$

$$\theta = E(\theta_f) = [\alpha_0, \alpha_W, \alpha_K, \alpha_M, \alpha_F, \gamma_{DE}, \gamma_{LE}]'. \quad (5)$$

The random coefficients are specified as

$$\theta_f = \theta + \delta_f, \quad (6)$$

⁵ More general random coefficient specifications were estimated, but we quickly run into a problem of non-convergence. Our general model is on the limit of what is manageable with the chosen software.

where δ_f is a zero mean vector specific to bank f . We assume that the explanatory variables, the genuine error terms (u_{Cf} , u_{if}) and the bank specific vector δ_f , are mutually independent. We make the following assumptions about the reduced system that we estimate,

$$[u_{Cf}, u_{Wf}, u_{Kf}, u_{Ff}]' \sim \text{IIN} [0, \Omega^u], \quad \delta_f \sim \text{IIN} [0, \Omega^\delta], \quad (7)$$

where IIN signifies independently, identically, normally distributed. Beyond symmetry, there are no restrictions imposed on the covariance matrix, Ω^u . The genuine error terms are assumed to be homoskedastic across banks and not autocorrelated within banks.

Own-price and cross-price elasticities of substitution in input demand are given in (8) and (9). These elasticities are defined as the Slutsky analogues, i.e. as output-constrained price elasticities of input quantities. The price elasticities depend on data as well as estimated coefficients, and will in general vary both over time and across banks.

$$\varepsilon_{iif} = \beta_{ii} / S_{if} + S_{if} - 1 \quad \text{for all } i, \quad (8)$$

$$\varepsilon_{ijf} = \beta_{ij} / S_{if} + S_{jf} \quad \text{for } i \neq j. \quad (9)$$

Grant (1993) shows that the elasticities of substitution in the translog function case may be evaluated at any expansion point, including points of sample means, as long as the theory restrictions of price homogeneity and Slutsky-symmetry hold. The cross-price elasticities are in general not symmetric. The translog cost function has been criticised because the area where the regularity conditions are met can be narrow. Particularly own-price elasticities have attained focus, because positive values - even within sample - have been revealed. We check the within sample properties.

The elasticity of costs with respect to electronic payments and the output specific scale elasticity, which equals the inverse elasticity of costs with respect to an output, are defined in (10) and (11) respectively. As with the price elasticities, within the general model, these elasticities depend on both data and estimated coefficients.

$$\varepsilon_E = \partial \ln C_f / \partial \ln EP = \gamma_E + \gamma_{EE} \ln EP + \sum_i \gamma_{iE} \ln P_{if} + \sum_m \gamma_{mE} \ln X_{mf}, \quad (10)$$

$$\varepsilon_m = (\partial \ln C_f / \partial \ln X_{mf})^{-1} = (\gamma_m + \gamma_{mm} \ln X_{mf} + \gamma_{mn} \ln X_{nf} + \sum_i \gamma_{im} \ln P_{if} + \gamma_{mE} \ln EP)^{-1}, \quad (11)$$

$i, j = W, K, M, F, \quad m, n = D, L.$

We expect ε_E to be negative, i.e. that the increase in the share of electronic payments in total non-cash payments has decreased variable costs. But as already discussed, a cost reduction effect need not be present, if the banks offer electronic payment services primarily to attract customers.

As explained above, the translog cost function is criticised because it disfavours large banks with respect to scale economies. However, the formula for the scale elasticities shows that, when keeping the level of output n constant and using the expected values of the random coefficients, the scale elasticity with respect to output m declines as output grows only if $(\gamma_{mm} + \gamma_{mn}) > 0$. If $(\gamma_{mm} + \gamma_{mn}) < 0$ or $(\gamma_{mm} + \gamma_{mn}) = 0$, the scale elasticity increases or is independent of bank size respectively. If X_D and X_L grow at the same pace, the overall scale elasticity will decline as output grows if $(\gamma_{DD} + \gamma_{LL} + 2 \cdot \gamma_{DL}) > 0$. If $(\gamma_{DD} + \gamma_{LL} + 2 \cdot \gamma_{DL}) < 0$ or $(\gamma_{DD} + \gamma_{LL} + 2 \cdot \gamma_{DL}) = 0$, the overall scale elasticity will increase with or be independent of bank size respectively.

There are a number of questions concerning the properties of the cost function – i.e. the production function - that can be analysed on the basis of our general model. We are, however, particularly interested in (i) the effect on variable costs from the increase in electronic payments, (ii) the effect on scale elasticities, and (iii) if new technology, as represented by new electronic payment services, has affected input demand asymmetric, i.e. non-neutral.

With respect to the first issue, if $\epsilon_E < 0$, average variable costs declines as the share of electronic payments increases. With respect to the scale elasticities, if the coefficients γ_{DE} and γ_{LE} are zero, then no interaction terms between output and electronic payments enter the cost function, and we conclude that the increase in electronic payments has not affected the scale elasticity in intermediation. It also would imply that the effect of electronic payments on variable costs is independent of the level of output. If, on the other hand, γ_{DE} and/or γ_{LE} are found to be negative (positive), this implies an increase (decrease) in the scale elasticity with respect to the production of deposits and/or loans as the share of electronic payments increases for an average bank.

The question of input demand neutrality can be analysed by testing restrictions on the coefficients γ_{iE} , $i=W, K, M, F$. If, for all i , $\gamma_{iE} = 0$, we conclude that the increased use of electronic payments has affected input demand neutrally, i.e. input volumes have changed proportionally and input ratios are not affected by this shift in technology. If, on the other hand, $\gamma_{iE} \neq 0$, then the effect is biased, since relative cost-shares, and hence relative input volumes change. Within the chosen approach, we can calculate the effect of increased use of electronic payments on input ratios by using equation (12). ΔIR_{ij} denotes the change in the input ratio between inputs i and j .

$$\Delta IR_{ij} = \partial(S_{if}/S_{jf})/\partial \ln EP = (\gamma_{iE} S_{jf} - \gamma_{jE} S_{if})/(S_{jf})^2 \quad (12)$$

If ΔIR_{ij} is positive (negative) we conclude that the input ratio between input i and j increases (decreases) as the share of electronic payments increases. Since electronic payment services are assumed to replace more expensive labour intensive payment services, we are particularly interested in checking the impact of electronic payments on the input ratios between labour and

physical capital and between labour and materials. Electronic payments are - to some degree - contingent on investments in new capital. Furthermore, the costs faced by banks when using jointly owned clearing and settlement systems to execute electronic payment transactions are included in materials input. Hence, we expect the isolated effect on the input ratios between labour and both physical capital and materials from an increase in electronic payments to be negative.

If all cost shares are independent of the output level, i.e. $\gamma_{iD}=\gamma_{iL}=0$, we conclude that the production technology is homothetic, and factor ratios remain constant when the level of output changes. Homotheticity in addition to the absence of price effects in the cost-share equations, i.e. $\beta_{ij} = 0$ for all i,j , imply a Cobb-Douglas production technology. As in Jorgenson (1986), we define a positive (negative) effect of output growth on a cost share as a positive (negative) scale bias.

3. Empirical results

We now present the results from estimating the cost function in (1) and the cost-share equations in (3). We exclude the cost-share equation of materials for reasons explained earlier. We use bank-level panel data from Norwegian banks combined with industry-level information from National accounts. A more detailed presentation of the data and empirical variables is given in the Appendix. The panel includes 2102 annual observations of 226 banks over the period 1978-1999. The panel is unbalanced, and 133 banks are observed in the maximum 13 years, which is 82 per cent of the observations. Likelihood estimates of the coefficients in the simultaneous system with cross-equation restrictions are obtained by using the PROC MIXED-procedure in the SAS/STAT software (SAS, 1992).^{6,7}

As in general in cost function analyses, we may face an endogeneity problem with respect to the explanatory variables. Wages vary across banks, which may suggest that banks are not price takers in the labour market. However, this variation may reflect differences in seniority and level of education, and one can argue that due to a centralised wage formation system and relatively high degree of unionisation in Norway, it is plausible to assume that wages are weakly exogenous. Because we only have access to the industry-level price on materials input (incl. energy), any possible endogeneity problem with respect to this price is reduced. The price of funds is a weighted average of interest rates on deposit accounts and a money market interest rate, the latter is clearly exogenous to the banks. Although it can be argued that the implementation of new technology in banks in general should be treated as an endogenous process, two arguments can be put forward that defend the weak exogeneity assumption on the share of electronic payments in total non-cash payments. First, we only have access to the industry-level share, which is exogenous to each bank, and second, the change in this share is determined not only by the banks but also largely by the bank customers.⁸

Table 1 gives the results from estimating the most general model (M_G) and a reduced model (M_R) with two zero-restrictions on the coefficient vector.

⁶ Various applications of this procedure are discussed in Verbeke and Molenberghs (1997).

⁷ For simplicity, the unbalancedness of our data is suppressed when presenting the model in section 2. A precise representation of a similar model with random coefficients on an unbalanced panel is given in Biørn et al. (1998).

⁸ In Lindquist (2001), a comparable single-output version of the model with only fixed effects in the cost function, i.e. no random coefficients, and without the trend variable is estimated. A three-stage least squares (3SLS) procedure, which takes into account that explanatory variables may not be weakly exogenous, is used. A comparison of our results with that in Lindquist (2001) shows that the results are relatively robust. We therefore conclude that the weak exogeneity assumptions seem not to be important for the results.

Table 1. *The estimated cost function and cost-share equations^a*

The general model (M_G)				The reduced model (M_R)			
Coefficient	Estimate	Std.error	Pr > t	Coefficient	Estimate	Std.error	Pr > t
α_0	3.721	0.580	< .0001	α_0	3.260	0.579	< .0001
α_W	0.656	0.027	< .0001	α_W	0.648	0.027	< .0001
α_K	0.169	0.011	< .0001	α_K	0.168	0.009	< .0001
α_M^b	-0.394	0.026	< .0001	α_M^b	-0.423	0.026	< .0001
α_F	-0.431	0.041	< .0001	α_F	-0.393	0.041	< .0001
β_{WW}	0.060	0.003	< .0001	β_{WW}	0.058	0.003	< .0001
β_{WK}	-0.000	0.001	.9881	β_{WK}^c	0		
β_{WM}^b	0.013	0.003	< .0001	β_{WM}^b	0.013	0.003	< .0001
β_{WF}	-0.073	0.002	< .0001	β_{WF}	-0.072	0.002	< .0001
β_{KK}	0.012	0.001	< .0001	β_{KK}	0.013	0.001	< .0001
β_{KM}^b	-0.009	0.001	< .0001	β_{KM}^b	-0.009	0.001	< .0001
β_{KF}	-0.003	0.001	.0005	β_{KF}	-0.004	0.001	< .0001
β_{MM}^b	0.072	0.004	< .0001	β_{MM}^b	0.062	0.004	< .0001
β_{MF}^b	-0.075	0.002	< .0001	β_{MF}^b	-0.066	0.002	< .0001
β_{FF}	0.151	0.003	< .0001	β_{FF}	0.141	0.004	< .0001
γ_D	0.924	0.136	< .0001	γ_D	0.767	0.084	< .0001
γ_L	-0.837	0.132	< .0001	γ_L	-0.666	0.080	< .0001
γ_{DD}	0.125	0.018	< .0001	γ_{DD}	0.134	0.008	< .0001
γ_{LL}	0.273	0.029	< .0001	γ_{LL}	0.269	0.013	< .0001
γ_{WD}	-0.018	0.003	< .0001	γ_{WD}	-0.017	0.003	< .0001
γ_{WL}	-0.031	0.003	< .0001	γ_{WL}	-0.031	0.003	< .0001
γ_{KD}	-0.003	0.001	.0149	γ_{KD}	-0.003	0.001	.0176
γ_{KL}	-0.005	0.001	< .0001	γ_{KL}	-0.005	0.001	< .0001
γ_{MD}^b	0.003	0.003	.2344	γ_{MD}^b	0.005	0.003	.0879
γ_{ML}^b	-0.032	0.003	< .0001	γ_{ML}^b	-0.033	0.003	< .0001
γ_{FD}	0.017	0.005	.0009	γ_{FD}	0.015	0.005	.0030
γ_{FL}	0.068	0.005	< .0001	γ_{FL}	0.068	0.005	< .0001
γ_E	1.687	0.152	< .0001	γ_E	0.972	0.143	< .0001
γ_{EE}	-0.210	0.032	< .0001	γ_{EE}	-0.324	0.032	< .0001
γ_{WE}	0.011	0.003	< .0001	γ_{WE}	0.012	0.003	< .0001
γ_{KE}	-0.002	0.001	.0433	γ_{KE}	-0.002	0.001	.0219
γ_{ME}^b	0.006	0.002	.0252	γ_{ME}^b	0.013	0.003	< .0001
γ_{FE}	-0.014	0.004	.0006	γ_{FE}	-0.022	0.004	< .0001
γ_{DL}	-0.174	0.022	< .0001	γ_{DL}	-0.176	0.010	< .0001
γ_{DE}	-0.088	0.051	.0866	γ_{DE}	-0.098	0.010	< .0001
γ_{LE}	-0.055	0.049	.2626	γ_{LE}^c	0		
γ_τ	-0.020	0.002	< .0001	γ_τ	-0.023	0.002	< .0001
lnL	14884.3			lnL	14800.3		

^a The cost-share equation of materials (incl. energy) is excluded. The estimated expected values of the random coefficients are reported. lnL is the value of the log-likelihood function.

^b Coefficient calculated using the adding up restrictions.

^c Restricted a priori.

Testing the restriction $\beta_{WK}=0$ on the general model using the likelihood-ratio test gives $\chi^2(1)=0.0$, and the restriction is clearly not rejected.⁹ Testing restrictions on random coefficients, as γ_{LE} , are more complicated, since these test statistics are not asymptotically χ^2 -distributed under the null hypothesis of coefficient homogeneity. In this case, parameters of the covariance matrix Ω^δ are on the border of the admissible parameter space. Table 1 shows that, according to the simple t-test, we must accept a 28 per cent significance level to reject the zero-restriction on $E(\gamma_{LE})$. However, we are not able to implement a zero restriction on this expectation without at the same time to implement zero restrictions on Ω^δ . Hence, the zero-restriction on γ_{LE} reduces the degree of heterogeneity captured by the model, and this explains why the value of the log-likelihood function is much lower for model M_R than model M_G . Despite the drop in the log-likelihood, a reduction of the number of random coefficients may seem appealing, since, as already explained, the random coefficient specification of model M_G is on the limit of what is admissible due to problems with convergence of the model.

According to Table 1, in the general model M_G , the coefficients γ_{mE} , $m=D,L$, are both negative, which implies that the elasticity of scale in the production of both deposits and loans increases as the share of electronic payments increases. The effect on the production of deposits is larger than on the production of loans, as was expected, since payment services are more closely connected to deposits than to loans. The effect on loans is 60 per cent less than the effect on deposits. The restriction $\gamma_{LE}=0$ in model M_R cultivates this a priori expectation, and in this case new technology in payment services affects the scale elasticity in the production of deposits only. In addition to the simplification argument for the zero-restriction on γ_{LE} put forward above, a second argument in favour of this restriction is that correlation between loans and deposits may hamper the identification of the “true” effect of new electronic payment services on each of the two products. For this reason, one may argue that we should exploit our a priori knowledge about banks and restrict the general model in accordance with this knowledge. On the other hand, the zero-restriction on γ_{LE} is probably too strict, and testing this restriction within a model where γ_{LE} is treated as a fixed, i.e. non-random, coefficient, rejects the restriction. We find $\chi^2(1)=4.8$, and we must choose a significance level close to 2.5 to not reject this restriction in that case.

Concerning the question of input-demand neutrality, we study the coefficients γ_{iE} , $i=W,K,M,F$. According to Table 1, in both models, the simple t-test rejects that any of these coefficients are zero. We test the joint restriction that $\gamma_{iE}=0$ for all $i=W,K,F$ on model M_G . The likelihood ratio test gives $\chi^2(3)=195.5$, and the joint restriction is clearly rejected. We conclude that the move towards electronic payment services has caused the cost-shares of both labour and materials to increase, while the cost-shares of physical capital and funding have decreased. We find that

⁹ If not stated otherwise, a five per cent significance level is used throughout the paper.

$\Delta IR_{LK} < 0$ and $\Delta IR_{LM} < 0$, which imply that the input ratio between labour and both physical capital and materials has decreased as the share of electronic payments has increased. This is consistent with our a priori beliefs; new electronic payment services have particularly substituted out paper-based and labour-intensive methods, cf. Vesela (2000). We also find that $\Delta IR_{KM} < 0$, which implies that the input ratio between physical capital and materials has declined due to this change in technology.

The results do not support homotheticity, and we find a negative scale bias with respect to both deposits and loans for labour and physical capital. I.e., the cost shares of these two inputs decrease with the level of both outputs. For funds the scale bias is positive with respect to both products, while the results for materials imply a positive scale bias with respect to deposits while it is negative with respect to loans.

In Table 2 we present the own-price and cross-price elasticities predicted by the two alternative models in Table 1. The elasticities are calculated at the overall empirical means of the cost shares.

Table 2. *Own- and cross-price elasticities calculated at the overall empirical means of the cost shares*

Own-price elasticities	Estimate		Cross-price elasticities	Estimate	
	M _G	M _R		M _G	M _R
ϵ_{WW}	-0.436	-0.448	ϵ_{WK}	0.043	0.043
ϵ_{KK}	-0.666	-0.656	ϵ_{WM}	0.202	0.207
ϵ_{MM}	-0.250	-0.339	ϵ_{WF}	0.191	0.198
ϵ_{FF}	-0.082	-0.096	ϵ_{KW}	0.142	0.142
			ϵ_{KM}	-0.101	-0.104
			ϵ_{KF}	0.625	0.618
			ϵ_{MW}	0.256	0.261
			ϵ_{MK}	-0.038	-0.039
			ϵ_{MF}	0.032	0.117
			ϵ_{FW}	0.039	0.040
			ϵ_{FK}	0.038	0.038
			ϵ_{FM}	0.005	0.019

Table 3 shows that the price elasticities are very robust to the choice of model specification. Evaluated at the overall empirical means of the cost shares, all own-price elasticities have the correct sign, this is true also at the annual means but not at all sample points. Hence, we have some problems with the “concavity in prices” condition. It should be remembered, though, that

no cleaning of the data has been executed. We have only excluded five observations due to missing observations on explanatory variables. On average, all four inputs are inelastic in demand, but the own-price elasticity of funds is small. Furthermore, for the average bank, most inputs are substitutes in demand. The exception is physical capital and materials, which are complements.

Table 3 gives the elasticity of costs with respect to the share of electronic payments in total non-cash payments, cf. Eq. (10), and the elasticities of scale, cf. Eq. (11). The results are very close in the two alternative models. According to Table 3, for an average bank, the elasticity of costs with respect to electronic payments, ϵ_E , is negative. Hence, we conclude that variable costs decrease over time due to the change in technology in payment services. An increase in the share of electronic payments by one per cent decreases variable costs by around 0.3 per cent for the average bank. In addition, Table 1 shows that the estimated trend coefficient is negative. Hence, additional cost reducing processes are present in the banking industry. This negative trend effect probably captures effects on costs from experience, increased knowledge, other innovations than the shift towards electronic payments, as well as improved production techniques.

Table 3. *The elasticity of costs with respect to electronic payments (ϵ_E) and the elasticities of scale (ϵ_D, ϵ^a)*

	M_G	M_R
ϵ_E	-0.318	-0.268
ϵ_D	3.185	3.653
ϵ_L	1.050	1.013

^a Calculated at overall empirical means of the variables. The estimated expectations of the random coefficients are used.

With respect to the production of deposits, we find clear evidence of economies of scale for the average bank. We also find that this scale elasticity is an increasing function of produced deposits, i.e. $(\gamma_{DD} + \gamma_{DL}) < 0$. Holding produced loans constant, an increase in all inputs by one per cent increases produced deposits by more than 3 per cent, or, alternatively, a separate increase in produced deposits by one per cent increases total variable costs by only around 0.3 per cent on average within our sample. According to Table 3, the scale elasticity with respect to loans is close to one, and holding produced deposits constant, an increase in the production of deposits by one per cent increases total variable costs by a little less than one per cent for the average bank. This scale elasticity is a declining function of loans, i.e. $(\gamma_{LL} + \gamma_{DL}) > 0$. The economic interpretation of these results is plausible. It costs a bank relatively little to increase its production of deposits, and the increase in costs is smaller the larger the bank already is. Setting up a deposit account costs little in general, sending out information With respect to loans, the bank must

evaluate the risk involved with each customer, and this process probably keeps the economies of scale down. A declining economies of scale with respect to size may reflect that large banks in general grant big loans to some customers which also involve relatively high variable costs for those banks.

The elasticity of scale in the production of deposits increases over time in both M_G and M_R , partly due an increase in produced deposits and partly due to the increase in the share of electronic payments. The elasticity of scale in the production of loans is relatively stable over time in both models, however, despite the isolated negative effect from the increase in loans.

IV. Conclusions

This paper analyses the importance of new technologies, i.e. new electronic payment services, for the development in banks' scale properties and input demand in the production of loans and deposits. A four-factor translog cost function and the corresponding cost-share equations are estimated simultaneously using an unbalanced panel data with data for Norwegian banks over the years 1987-1999. The inputs are labour, physical capital, materials and funds. We apply a random coefficient approach, which allows for heterogeneity across banks in both intercept term and slope coefficients.

To represent the implementation of new technology in the payment systems in the Norwegian banking industry, we include the share of electronic payments in total non-cash payments measured in number of transactions. A rather general model is specified in this variable. We also include a deterministic time trend to capture technical change that is not captured by this electronic payments variable.

The results show that variable costs have decreased and the elasticity of scale has increased as the share of electronic payments has increased, and as was expected, the effect is strongest in the production of deposits. The move towards electronic payment services has affected input demand asymmetric, i.e. non-neutral, causing the cost-shares of both labour and materials to increase, while the cost-shares of physical capital and funds decrease. Calculations on the estimated models show that the input ratios between labour and both physical capital and materials decrease as the share of electronic payments increases. This is consistent with our a priori beliefs; new electronic payment systems have particularly substituted out paper-based and labour-intensive methods.

Appendix

The data and definition of variables

Primarily we use annual bank-level data which banks are obliged to report to the Central Bank. This is combined with bank-level information on employment from Statistics Norway. In addition, we apply some price indices at the industry level from Statistics Norway, i.e. we use data for industry 63: Bank and insurance. The panel includes 2102 annual observations of 226 banks over the period 1978-1999. The panel is unbalanced, but 141 banks are observed in the maximum 13 years, which is 82 per cent of the observations.

Table A1. *Definition of the empirical variables*

C_f	Total variable costs in bank f, 1000 NOK
P_{Wf}	Wage in bank f, NOK per man-hour
P_K	User cost of physical capital, NOK
P_M	Price of materials incl. energy in bank f, NOK 1997=1
P_{Ff}	Interest rate on funds in bank f
X_{Df}	Volume of produced deposits in bank f, 1000 1997-NOK
X_{Lf}	Volume of total loans in bank f, 1000 1997-NOK
EP	Share of electronic payments in total non-cash payments in the Norwegian banking industry. Measured in number of transactions
τ	A time trend, 1978=1

Total variable costs are calculated as the sum of reported costs on labour, physical capital costs, costs on materials input and interest paid on funds.

We calculate data on two physical capital stock components, i.e. Machinery, fixtures and transport equipment and Buildings. The two components are aggregated to total physical capital stock. Data in value terms are the sum of deflated book values multiplied by a calculated user cost. We include rented capital in our measure. The user cost of physical capital is a weighted average of the user cost of each component. Each user cost is a function of the corresponding industry level investment price from the Norwegian national account, the one period ahead rate of change in this price, the 10 years Norwegian government bond rate, and a depreciation rate. The latter is based on information on the service life from Norwegian national accounts and geometric depreciation. The service life is 9 years for Machinery etc. and 60 years for Buildings. The user cost is common across banks.

The price index of materials incl. energy is a weighted average of the industry level price of materials, electricity and fuels from Norwegian national accounts. This price is common across

banks. Materials input includes the use of clearing house and settlement bank to execute electronic payment transactions.

Fund costs equal banks' interest payments on deposits and money market funds. The price of funds is - for each bank - calculated as the weighted average of the bank's interest rate on deposits and a money market rate, using ordinary deposits and loans in the money market as weights.

The volumes of loans and deposits are calculated by deflating nominal stock data, as registered per 31.12 each year, by the consumer price index (CPI). With respect to business loans, one may argue that the GDP-price index is a more appropriate deflator, but the GDP-price index and CPI follow a common trend. We use nominal stock data per 31.12 each year as a measure of output the same year for two reasons: First, using the average of the previous and this years nominal stocks per 31.12 reduces the number of observations. Second, one can argue that banks' annual variable costs reflect the production of loans and deposits that build up to the stocks registered by the end of the year.

With respect to electronic payments, we only have the number of transactions at the industry level. Table A2 gives summary statistics for the variables used in the regressions. According to the empirical means, the input share of funds is well above the other variable inputs.

Table A2. Summary statistics for the endogenous and explanatory variables in the analysis

	Mean	St. dev.	Min.	Max.
Total costs, Mill. NOK	395	1521	859	19100
Cost share of labour	0.157	0.065	0.005	0.776
Cost share of physical capital	0.037	0.020	0.0002	0.217
Cost share of materials incl. energy	0.139	0.059	0.009	0.745
Cost share of funds	0.667	0.119	0.002	0.964
Wage, NOK per man-hour	225	112	6	3588
User costs of physical capital ¹	0.164	0.078	0.042	0.336
Price index of materials incl. energy ²	0.911	0.069	0.798	1.036
Interest rate on funding	8.453	15.670	0.978	561.662
Deposits, Mill. 1997-NOK	3635	13700	0.1	211000
Loans, Mill. 1997-NOK	3978	15200	1.6	233000
Share of electronic payments ³	0.382	0.189	0.137	0.734
Trend variable	6.824	3.822	1	13

¹ Equal for all banks.

² Annual National account data, 1997=1. Equal for all banks.

³ The share of electronic payments in total non-cash payments in the banking industry. Equal for all banks.

References

- Berger, A.N., G.A. Hanweck and D.B. Humphrey (1986): Competitive viability in banking: Scale, scope, and product mix economies, Washington, D.C.: Federal Reserve Board.
- Biørn, E., K.-G. Lindquist and T. Skjerpen (1998): Random Coefficients and Unbalanced Panels: An Application on Data from Norwegian Chemical Plants, Discussion Papers no. 235, Statistics Norway.
- Biørn, E., K.-G. Lindquist and T. Skjerpen (2001): Heterogeneity in Returns to Scale: A Random Coefficient Analysis with Unbalanced Panel Data, *Journal of Productivity Analysis*, forthcoming.
- Brooz, Allen and Hamilton (1998): Developments in Banking.
- Caves, D.W., L.W. Christensen and M.W. Tretheway (1980): Flexible Cost Functions for Multiproduct Firms, *Review of Economics and Statistics*, 62, 477-481.
- Christensen, L.R., D.W. Jorgenson and L. Lau (1971), Conjugate duality and the transcendental logarithmic production function, *Econometrica*, 39, 255-256.
- Christensen, L.R., D.W. Jorgenson and L. Lau (1973), Transcendental logarithmic production frontiers, *Review of Economics and Statistics*, 55, 28-45.
- Flatraaker, D.-I. and P. Robinson (1995): Income, Costs, and Pricing in the Payment System, *Norwegian Central Bank Economic Bulletin*, 66, 321-332.
- Grant, J.H. (1993), The translog approximate function. Substitution among inputs in manufacturing evaluated at sample means, *Economics Letters*, 41, 235-240.
- Hancock, D. (1985): The Financial Firm: Production with Monetary and Nonmonetary Goods, *Journal of Political Economy*, 93, 859-880.
- Humphrey, D.B. (1981): Scale Economies at Automated Clearinghouses, *Journal of Bank Research*, (Summer 1981), 71-81.
- Humphrey, D.B., M. Kim and B. Vale (2001): Realizing the Gains from Electronic Payments: Costs, Pricing, and Payment Choice, *Journal of Money, Credit and Banking*, 33, 216-234.
- Hunter, W.C. and S.G. Timme (1986): Technical Change, Organizational Form, and the Structure of Bank Production, *Journal of Money, Credit and Banking*, 18, 152-166.

- Hunter, W.C. and S.G. Timme (1991): Technological Change in Large U.S. Commercial Banks, *Journal of Business*, 64, 339-362.
- Hunter, W.C., S.G. Timme and W. K. Yang (1990): An Examination of Cost Subadditivity and Multiproduct Production in Large U.S. Commercial Banks”, *Journal of Money, Credit and Banking*, 22, 504-525.
- Jorgenson, D.W. (1986), Econometric methods for modelling producer behaviour. Chapter 31 in: Griliches, Z. and M.D. Intriligator (eds.), *Handbook of Econometrics Volume III*. Amsterdam: North-Holland.
- Koskenkylä, H. (2001): Opportunities and Threats of European Banks, Bank of Finland, Working paper no. 6/2001.
- Lawrence, Colin (1989): Banking Costs, Generalized Functional Forms, and the Estimation of Economies of Scale and Scope, *Journal of Money, Credit and Banking*, 21, 368-379.
- Lindquist, K.-G. (2001): Electronic Payments and Scale Properties in the Banking Industry, SUERF Colloquium Volume.
- SAS (1992): SAS Technical Report P-229, SAS/STAT Software: Changes and Enhancements, Release 6.07, Cary, NC: SAS Institute, Inc.
- Shaffer, S. (1998): Functional forms and declining average costs, *Journal of Financial Services Research*, 14, 91-115.
- Verbeke, G. and G. Molenberghs (eds.) (1997): *Linear Mixed Models in Practice. A SAS Oriented Approach*. New York: Springer.
- Vesela, J. (2000): *Technological Transformation and Retail Banking Competition: Implications and Measurement*, Bank of Finland Studies E:20.
- Wells. K. (1996): Are Checks Overused?, *Federal Reserve Bank of Minneapolis Quarterly Journal Review*, 20, 2-12.