ANO 2002/6

Oslo
August 30, 2002

Working Paper

Research Department

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

by

Kjersti-Gro Lindquist

NORGES BANK

Working papers fra Norges Bank kan bestilles over e-post:
posten@norges-bank.no
eller ved henvendelse til:
Norges Bank, Abonnementsservice
Postboks 1179 Sentrum
0107 Oslo
Telefon 22 31 63 83, Telefaks 22 41 31 05

Fra 1999 og senere er publikasjonene tilgjengelige som pdf-filer
på www.norges-bank.no, under "Publisert".

Working papers inneholder forskningsarbeider
og utredninger som vanligvis
ikke har fått sin endelige form.
Hensikten er blant annet at forfatteren
kan motta kommentarer fra kolleger
og andre interesserte.

Synspunkter og konklusjoner i arbeidene
står for forfatternes regning.

Working papers from Norges Bank can be ordered by e-mail:
posten@norges-bank.no
or from Norges Bank, Subscription service,
P.O.Box. 1179 Sentrum
N-0107 Oslo, Norway.
Tel. +47 22 31 63 83, Fax. +47 22 41 31 05

Working papers from 1999 onwards are available as pdf-files on the bank's
web site: www.norges-bank.no, under "Published".

Norges Bank's working papers present
research projects and reports
(not usually in their final form)
and are intended inter alia to enable
the author to benefit from the comments
of colleagues and other interested parties.

Views and conclusions expressed in working papers are
the responsibility of the authors alone.

ISSN 0801-2504
ISBN 82-7553-196-9

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

Kjersti-Gro Lindquist

Norges Bank (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 countries, payment services in banking have shifted from paper-based giro and cheque payments to electronic giro and debit card payments. This paper analyses the effect of this change in payment technology 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 production technology across banks. The results show that the move towards electronic payment services has decreased average costs and increased the economies of scale in bank intermediation. An output composition effect can explain that decreasing economies of scale is often found in analyses of banks.

Keywords: Banking industry; Electronic payments; Technical change; Panel data
JEL classification: C33, G21, O33

Acknowledgement: The author wishes to thank Erik Biørn, Martin Hellwig, Almas Heshmati, Tor Arnt Johnsen, Aris Kaloudis, David Llewellyn, Terje Skjerpen, Bent Vale, Jukka Vesela, Grete Øwre, and participants at the 23rd SUERF Colloquium in Brussels 2001, the International Conference on Panel Data in Berlin 2002 and ESEM in Venice 2002 for valuable comments and discussions.

1. Introduction

In many OECD-countries, particularly in the Nordic area, payment services in banking have shifted significantly from paper-based giro and cheque 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 2001 this had risen to 83 per cent. This change was - at least partly - spurred by banks' pricing policy, see Humphrey, Kim and Vale (2001). Banks' motivation to offer electronic payment services may be due to both cost-saving efforts and competition. Once the necessary investment is made, electronic payments cost banks much less to produce than paper-based payments, see Humphrey and Berger (1990), Flatraaker and Robinson (1995) and Wells (1996). If customers find electronic payment services more convenient than paper-based services, and hence prefer the former, electronic payment services may be viewed as strategic variables in the competition for customers. This paper focuses on the impact of new electronic payment services, which clearly represent new technology, on banks' average costs, scale properties and input demand in the production of deposits and loans.

In analyses of bank behaviour, the definition of output is in general a major challenge, and one may argue that payment services, both electronic, paper based and cash payment services, should be treated as outputs. However, although increasingly important, 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 in order 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 clearinghouse and settlement bank to execute the transactions. The latter is particularly important with respect to giro payments, card payments and automatic teller machine (ATM) transactions. The banks are charged for their share of the costs in these systems. Within the chosen approach, one may argue that we ideally should treat payment services as an input similarly to other variable inputs. We are, however, unable to do 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. Hence, treating payment services as an additional input would involve double counting of costs.

Due to its nature, electronic payment systems are characterised by scale economies, see Humphrey (1981) for an analysis of automated clearinghouses. When a system is established, increasing the number of transactions costs very little. In contrast, according to Humphrey (1982), cheque-processing operations face diseconomies of scale. Therefore, our a priori belief is that the observed change in payment technology has reduced costs and affected scale economies in bank intermediation. 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, due, for instance, to the increase in automatic instalment on 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 desirable for banks to introduce this new technology, if the competition for customer aspect is important. Hence, the effect on operating costs of the intermediation process from the shift to electronic payment services is an empirical question, which, to our knowledge, has not been addressed in the literature before.

In order to approach this issue econometrically, we formulate a five-equation system consisting of a cost function in four (primary) variable inputs, i.e. labour, physical capital, materials and fund (deposits and money market borrowing), and the corresponding cost-share equations. A multiple-output approach is applied, where the NOK volumes of deposits and loans are treated as two separate outputs. Our specification is consistent with the intermediation approach, since focus is on the intermediation function of banks and the model includes fund as input. See Humphrey (1985) for a discussion of the intermediation versus the production approach.

Technical change in banking is captured by two variables; First, to represent the implementation of new technology in payment services, we include, as a separate argument in the cost function, electronic payments as a share of 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 analyse the effect on banks' average costs and scale properties, we also test if this change in technology has affected input demand asymmetrically, i.e. non-neutrally. 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. The latter 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 within the dual approach. Using a deterministic trend as a proxy variable for technical change makes us unable to reveal the sources of technical change, however. A second disadvantage is that a time trend in addition captures the effects of "non-technical change" variables that are trending and not explicitly included in the cost function, such as deregulation. Therefore, estimated time trend effects must in general be interpreted with care. An important

[1] We focus on the change from paper based to electronic based payments and not, in addition, on the change from cash payments to paper based payments for two reasons: First, we lack data on cash payments until very recently. Second, the major change from cash payments to paper based payments, i.e. cheque and paper-based giro payments, occurred prior to our observation period.

contribution of this analysis is the identification of the effect of a major technological innovation in banking.

The representation of heterogeneity is of general importance in empirical cost or production function analyses applying micro data. It is rather common to either pool the data and hence fit the same model to all banks, or to assume that variation in size, age, management, employees' education, technology etc., can be represented by a bank specific fixed or random intercept term. Most likely, however, such differences will manifest themselves not only as a permanent, i.e. constant variation in efficiency across banks, 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, Lindquist and Skjerpen (2002). In this paper, a random coefficient approach, which specifies heterogeneity in both intercept term and slope coefficients, is applied. Hence, rather than assuming a priori that all coefficients are constant, i.e. equal, across 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. This relatively rich specification of heterogeneity is a second important contribution to the existing empirical literature on bank cost analyses.

We have access to annual bank-level data for most of the variables in our system over the period 1987-1999.[2] 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, it is not unreasonable to expect the share of electronic payments in total non-cash payments measured in number of transactions to be fairly stable 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. This is true for both small and large banks, and, in general, both small and large banks provide electronic payment services to their customers.

[2] The data are measured at the overall bank (firm) level rather than at the bank branch (plant) level, which is common when applying accounting data. See Berger, Hanweck and Humphrey (1986) for a discussion of this distinction. We follow Hunter and Timme (1991) and do not include the number of branches as an additional variable in our system. They argue that the number of branches is a function of total output, which is included in our model, and hence the chosen approach can be interpreted as a reduced form representation of total costs. Furthermore, including the number of branches in our general model would increase the number of coefficients to be estimated significantly.

The empirical results show that the move towards electronic payment services has reduced average costs and increased economies of scale in bank intermediation. As expected, the positive economies of scale effect is more due to an effect on the production of deposits than on the production of loans. The results also reveal rather large economies of scale in the production of deposits, while the scale elasticity in the production of loans is closer to one. This can explain why analyses of banks often find decreasing economies of scale.[3] In general, the deposits to loans ratio is smaller in large banks than in small banks, which most likely gives an output composition effect on estimated scale elasticities. We also find that the move towards electronic payments has affected input demand asymmetrically, i.e. non-neutrally, but this does not explain the observed decrease in the input ratio between labour and both physical capital and materials input. These changes in input ratios are rather due to substitution effects and non-homotheticity in bank intermediation.

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

[3] For recent reviews of the scale economies literature, see Berger, Hunter and Timme (1993) and Hughes (1999).

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, Jorgenson and Lau (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 flexible way. The translog cost function can be interpreted as a second order approximation to a general continuous twice-differentiable cost function in logs that satisfies linear homogeneity in prices.

While Hunter, Timme and Yang (1990) and Lawrence (1989) find that the standard translog specification provides an adequate fit to bank cost data, Shaffer (1998) shows that the translog form tend to impose a spurious U-shaped average cost structure in the case of monotonically declining true average cost data. The problem is sensitive to the range of firm sizes and to the distribution of firms within the range. See also McAllister and McManus (1993) and Mitchell and Onvural (1996). To alleviate this shortcoming of the translog cost-function, one can either add more flexibility to the functional form or introduce heterogeneity in coefficients across banks or groups of banks.[4] In this paper we introduce flexibility by the heterogeneity-in-coefficients approach, but rather than applying the more common approach and split the sample in subgroups, we assume random coefficients, which implies coefficient heterogeneity at the bank level. In particular, our general model includes a random coefficient in the scale elasticity, which means that this elasticity varies across banks due to variation in both the data and the coefficient vector. We should also add that even if the described short-coming of the translog function is true within the single-output case, in a multiple-output case with scope economies, which probably is present in bank intermediation, the conclusion may change.

Our most general translog cost function with two outputs and four inputs is given in Eq. (1).[5] Deposits and loans are outputs, and labour, physical capital, materials and fund (deposits and money market borrowing) are inputs. Subscript f denotes bank.

$$\begin{aligned} \ln C_f = {} & \alpha_{0f} + \Sigma_i \alpha_{if} \ln P_{if} + 1/2\, \Sigma_i \Sigma_j \beta_{ij} \ln P_{if} \ln P_{jf} + \Sigma_m \gamma_{mf} \ln X_{mf} + 1/2\, \Sigma_m \Sigma_n \gamma_{mn} \ln X_{mf} \ln X_{nf} \\ & + \Sigma_i \Sigma_m \gamma_{im} \ln P_{if} \ln X_{mf} + \gamma_{Ef} \ln EP + 1/2\, \gamma_{EE} (\ln EP)^2 + \Sigma_i \gamma_{iE} \ln P_{if} \ln EP \\ & + \Sigma_m \gamma_{mE} \ln X_{mf} \ln EP + \gamma_\tau \ln\tau + u_{Cf}, \quad i,j=W,K,M,F; \; m,n=D,L, \end{aligned} \tag{1}$$

[4] See Berger, A.N., and L.J. Mester (1997) for a recent survey of alternative functional forms and specification of heterogeneity when estimating cost and profit functions using US bank-level data, and Humphrey and Vale (2002) for an analysis using Norwegian bank-level data.

[5] For a discussion of the multiple-output representation within the translog cost function, see Caves, Christensen and Tretheway (1980).

where C_f is total operating costs plus interest expenses of bank f; P_{if} is the bank specific price/user cost of input i, i=W (labour), K (physical capital), M (materials incl. energy), F (fund); X_{Df} is the volume of deposits in bank f; X_{Lf} is the volume of 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 multicollinearity, a very simple specification in τ was chosen; u_{Cf} is an added disturbance. The Appendix gives a more detailed discussion of the data.

Within the chosen framework, deposits are treated as both output and input – or more precisely as part of an input. Hughes, Mester and Moon (2002) argue that whether to treat deposits as an output, an input or as both is a technological question, and the choice should depend on whether the data are consistent with the technological role of outputs and/or inputs. We will return to this issue in the empirical part of the paper.

Application of Shephard's lemma gives the cost-share (S_{if}) equations in (2).

$$S_{if} = \partial \ln C_f / \partial \ln P_{if} = (P_{if} \cdot V_{if})/C_f = \alpha_{if} + \Sigma_j \beta_{ij} \ln P_{jf} + \Sigma_m \gamma_{im} \ln X_{mf} + \gamma_{iE} \ln EP + u_{if}, \quad i,j=W,K,M,F; \; m,n=D,L, \tag{2}$$

where V_{if} is the quantity of input i used by bank f and u_{if}, i=W,K,M,F, are disturbances in the cost-share equations. Because the cost shares always sum to unity, that is $\Sigma_i S_{if} = 1$, any cost-share equation can be derived from 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 when estimating the model.

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 and symmetry in the cross output effect, are imposed on the general model that we estimate:

Price homogeneity and symmetry:	$\Sigma_j \beta_{ij} = 0$ for all i; $\beta_{ij} = \beta_{ji}$ for all i and j where $i \neq j$.
Adding up conditions:	$\Sigma_i \alpha_{if} = 1$; $\Sigma_i \beta_{ij} = 0$ for all j; $\Sigma_i \gamma_{ik} = 0$ for k=D,L,E.
Output symmetry:	$\gamma_{mn} = \gamma_{nm}$ for $m \neq n$.

The intercept term in the cost function, α_{0f}, the slope coefficients α_{if} i=W,K,M,F, γ_{mf} m=D,L and γ_{Ef} 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.[6] 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 α_0=E(α_{0f}), α_W=E(α_{Wf}), etc.,

$$\theta_f = [\alpha_{0f}, \alpha_{Wf}, \alpha_{Kf}, \alpha_{Mf}, \alpha_{Ff}, \gamma_{Df}, \gamma_{Lf}, \gamma_{Ef}]', \tag{3}$$

$$\theta = E(\theta_f) = [\alpha_0, \alpha_W, \alpha_K, \alpha_M, \alpha_F, \gamma_D, \gamma_L, \gamma_E]'. \tag{4}$$

The random coefficients are specified as

$$\theta_f = \theta + \delta_f, \tag{5}$$

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,

$$[u_{Cf}, u_{Wf}, u_{Kf}, u_{Ff}]' \sim IIN\,[0, \Omega^u], \quad \delta_f \sim IIN\,[0, \Omega^\delta], \tag{6}$$

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 homoscedastic across banks and not autocorrelated within banks.

Own-price and cross-price elasticities of substitution are defined as the Slutsky analogues, i.e. as output-constrained price elasticities of input quantities:

$$\varepsilon_{iif} = \beta_{ii} / S_{if} + S_{if} - 1 \qquad \text{for all } i, \tag{7}$$

$$\varepsilon_{ijf} = \beta_{ij} / S_{if} + S_{jf} \qquad \text{for } i \neq j. \tag{8}$$

These price elasticities depend on data as well as estimated coefficients and vary, in general, both over time and across banks. The cross-price elasticities are in general not symmetric. 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 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

[6] 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.

sample - have been revealed. We therefore check the within sample properties of the price elasticities.

The elasticity of costs with respect to electronic payments share and the output-specific scale elasticity, which equals the inverse elasticity of costs with respect to each output, are defined in Eq. (9) and Eq. (10) respectively. The global scale elasticity is defined in Eq. (11). As with the price elasticities, within the general model, these elasticities depend on both data and estimated coefficients, and in particular random coefficients.

$$\varepsilon_E = \partial \ln C_f / \partial \ln EP = \gamma_{Ef} + \gamma_{EE} \ln EP + \Sigma_i \gamma_{iE} \ln P_{if} + \Sigma_m \gamma_{mE} \ln X_{mf}, \quad (9)$$

$$\varepsilon_m = (\partial \ln C_f / \partial \ln X_{mf})^{-1} = (\gamma_{mf} + \gamma_{mm} \ln X_{mf} + \gamma_{mn} \ln X_{nf} + \Sigma_i \gamma_{im} \ln P_{if} + \gamma_{mE} \ln EP)^{-1}, \quad (10)$$

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

$$\varepsilon_X = (\partial \ln C_f / \partial \ln X_{Df} + \partial \ln C_f / \partial \ln X_{Lf})^{-1}. \quad (11)$$

We expect ε_E to be negative, i.e. that the increase in the share of electronic payments in total non-cash payments has reduced (average) costs. But, as already discussed, if the banks offer electronic payment services primarily to attract customers, a cost reduction effect need not be present. As explained above, the translog cost function has been 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 $\gamma_{mm}>0$, m=D,L. If $\gamma_{mm}<0$ or $\gamma_{mm}=0$, the scale elasticity increases or is independent of bank size respectively. If X_D and X_L grow proportionally, the global 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 global scale elasticity will increase with or be independent of bank size respectively. We expect scope economies to be present in bank intermediation, which is true if $\gamma_{DL}<0$. Hence, scope economies may reduce the tendency of decreasing economies of scale with bank size.

There are a number of questions concerning the properties of the cost function that can be analysed on the basis of our general model. We are particularly interested in (i) the effect on average 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 asymmetrically, i.e. non-neutrally.

With respect to the first issue, if $\varepsilon_E<0$, (average) costs decline 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 economies of scale in bank intermediation. It also implies that the effect of electronic payments on costs is independent of the output level. 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 the formula

$$\Delta_E IR_{ij} = \partial(V_{if}/V_{jf})/\partial EP = (P_{jf}/P_{if})\,[(\gamma_{iE}\,S_{jf} - \gamma_{jE}\,S_{if})/(EP \cdot S_{if}^2)], \qquad (12)$$

where $\Delta_E IR_{ij}$ denotes the change in the ratio between inputs i and j as EP increases. If $\Delta_E IR_{ij}$ is positive (negative) we conclude that the ratio between inputs 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, see Vesela (2000), 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 would 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 levels, i.e. $\gamma_{iD}=\gamma_{iL}=0$, i=W,K,M,F, we conclude that bank intermediation is homothetic, and input ratios remain constant as these activities expand. 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 technology in intermediation. As in Jorgenson (1986), we define a positive (negative) effect of output growth on a cost share as a positive (negative) scale bias. If the technology is not homothetic, we can calculate the effect of changing output levels on the input ratios by using Eq. (13). The interpretation of $\Delta_m IR_{ij}$ is similar to that of $\Delta_E IR_{ij}$.

$$\Delta_m IR_{ij} = \partial(V_{if}/V_{jf})/\partial X_m = (P_{jf}/P_{if})\,[(\gamma_{im}\,S_{jf} - \gamma_{jm}\,S_{if})/(X_m \cdot S_{if}^2)]\,, \quad m=D,L. \qquad (13)$$

3. Empirical results

We now present the results from estimating the cost function and the cost-share equations in (1) and (2). 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 1987-1999. The panel is unbalanced, and 133 banks are observed in the maximum 13 years, which is 82 per cent of the observations. Maximum-likelihood estimates of the random and fixed coefficients in the simultaneous system with cross-equation restrictions are obtained by using the PROC MIXED-procedure in the SAS/STAT software (SAS, 1992).[7,8] Although this procedure allows us to estimate a rather complicated model using unbalanced panel-data, one shortcoming is that we are not able to test for dynamic mis-specification or heteroscedastic error terms.

As in general in cost function analyses, we may face a problem of endogenous explanatory variables. Wages vary across banks, which may suggest that banks are not price takers in the labour market. However, this variation probably reflects 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 endogeneity problem with respect to this price is highly reduced. The price on fund is a challenge, since banks use both deposits and the money market to fund their loans, and the interest rates on deposits cannot be assumed to be exogenous to the banks. We therefore use a money market interest rate, which is clearly exogenous, as the price of fund rather than the weighted average of the money market interest rate and the interest rates on deposit accounts. We re-estimate the model using a weighted average of the money market interest rate and the interest rates on deposit accounts, however, to see if choosing different information sets is of importance.[9] Although it can be argued that the implementation of new technology in banks in general should be treated as an endogenous process, arguments can be put forward that defend the weak exogeneity assumption on the share of electronic payments in total non-cash payments. First, although the choice to implement new technology or not in principle is made by the bank, this choice is not the major driving force behind the development in the share-variable used in the analysis. Rather, the major driving force is the customers' choice to shift payment

[7] Various applications of this procedure are discussed in Verbeke and Molenberghs (1997).
[8] For simplicity, the unbalance of our data was suppressed when presenting the model in section 2. A precise representation of a similar model with random coefficients applied on an unbalanced panel is given in Biørn, Lindquist and Skjerpen (2003).
[9] The price of fund is in both cases calculated as the interest rate multiplied by the consumer price index. This gives consistency between this variable and the output variables, see the Appendix.

technology. Second, we only have access to the industry-level share, which is exogenous to each bank.[10]

Table 1 gives the results from estimating the most general model (M_G) and a reduced model (M_R), the latter includes four restrictions on the coefficient vector, i.e. $\gamma_{KD}=\gamma_{FD}=E(\gamma_{Ef})=\gamma_{LE}=0$. Not all the estimated coefficients in model M_G are significantly different from zero according to the approximate t-test, see the p-values, and this motivates a reduction of the model. In the reduction process we focus on the estimated and not on the calculated coefficients. To put restrictions on coefficients calculated from the adding up conditions, such as coefficient γ_{ME}, we would need to transform the data set. Testing restrictions on the fixed, i.e. non-random, coefficients, can be done by using the χ^2-form of the likelihood-ratio test. Testing homogeneity restrictions on random coefficients, i.e. that coefficients are equal across banks, implies a zero-restriction on the variance. This is 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 boundary of the admissible parameter space. Restrictions on the expectation vector that do not involve restrictions on the covariance matrix can be tested by using a standard likelihood-ratio test, however.

All the restrictions imposed on model M_R are supported by the likelihood ratio test when tested both sequentially and against model M_G.[11] Testing the joint restriction, i.e. model M_R, against model M_G gives $\chi^2(4)=3.4$, and the joint restriction is clearly not rejected. As explained above, testing the zero restriction on the random coefficient γ_{Ef} can be more complicated, depending on whether the test includes a zero restriction on the variance, i.e. $var(\gamma_{Ef})=0$, or not. Testing the restriction that $E(\gamma_{Ef})=0$ while keeping $var(\gamma_{Ef})\neq0$ gives $\chi^2(1)=0.1$, and, as already explained above, the restriction is not rejected. If we in addition include the restriction that $var(\gamma_{Ef})=0$, the value of the log-likelihood function declines from 16515.6 (model M_G in Table 1) to 16383.4. While the restriction $E(\gamma_{Ef})=0$ implies that this coefficient is zero for the average bank, with the additional zero-restriction on the variance, this coefficient is zero for all the banks in our sample. Hence, this latter restriction reduces the degree of heterogeneity of the model, and the large decline in the value of the log-likelihood function shows that this affects the explanatory power of the model very much.

[10] In Lindquist (2002), 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 (2002) shows that the results are relatively robust. Hence, the weak exogeneity assumptions seem not to be very important for the results.

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

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	p-value	Coefficient	Estimate	Std.error	p-value
α_0	3.752	0.871	< .0001	α_0	3.835	0.748	< .0001
α_W	0.682	0.026	< .0001	α_W	0.679	0.025	< .0001
α_K	0.148	0.011	< .0001	α_K	0.143	0.010	< .0001
α_F	-0.354	0.040	< .0001	α_F	-0.342	0.036	< .0001
α_M[b]	-0.476	0.026	< .0001	α_M[b]	-0.480	0.024	< .0001
β_{WW}	0.055	0.003	< .0001	β_{WW}	0.055	0.003	< .0001
β_{KK}	0.014	0.001	< .0001	β_{KK}	0.014	0.001	< .0001
β_{FF}	0.192	0.004	< .0001	β_{FF}	0.192	0.004	< .0001
β_{MM}[b]	0.076	0.004	< .0001	β_{MM}[b]	0.076	0.004	< .0001
β_{WK}	0.004	0.001	.0006	β_{WK}	0.004	0.001	0.0004
β_{WF}	-0.087	0.002	< .0001	β_{WF}	-0.087	0.002	< .0001
β_{WM}[b]	0.028	0.003	< .0001	β_{WM}[b]	0.028	0.003	< .0001
β_{KF}	-0.009	0.001	< .0001	β_{KF}	-0.009	0.001	< .0001
β_{KM}[b]	-0.008	0.002	< .0001	β_{KM}[b]	-0.008	0.002	< .0001
β_{MF}[b]	-0.095	0.003	< .0001	β_{MF}[b]	-0.095	0.003	< .0001
γ_D	0.910	0.125	< .0001	γ_D	0.960	0.120	< .0001
γ_L	-1.059	0.120	< .0001	γ_L	-1.120	0.109	< .0001
γ_{DD}	0.127	0.016	< .0001	γ_{DD}	0.129	0.015	< .0001
γ_{LL}	0.293	0.025	< .0001	γ_{LL}	0.296	0.025	< .0001
γ_{WD}	-0.012	0.003	< .0001	γ_{WD}	-0.011	0.002	< .0001
γ_{WL}	-0.034	0.003	< .0001	γ_{WL}	-0.034	0.002	< .0001
γ_{KD}	-0.002	0.001	.1971	γ_{KD}	0[c]		
γ_{KL}	-0.006	0.001	< .0001	γ_{KL}	-0.007	0.001	< .0001
γ_{FD}	0.002	0.005	.6924	γ_{FD}	0[c]		
γ_{FL}	0.076	0.005	< .0001	γ_{FL}	0.077	0.002	< .0001
γ_{MD}[b]	0.012	0.003	< .0001	γ_{MD}[b]	0.011	0.002	< .0001
γ_{ML}[b]	-0.037	0.003	< .0001	γ_{ML}[b]	-0.037	0.002	< .0001
γ_E	-0.049	0.103	.6377	γ_E	0[c]		
γ_{EE}	-0.299	0.018	< .0001	γ_{EE}	-0.290	0.015	< .0001
γ_{WE}	0.016	0.002	< .0001	γ_{WE}	0.016	0.002	< .0001
γ_{KE}	-0.006	0.001	< .0001	γ_{KE}	-0.006	0.001	< .0001
γ_{FE}	-0.012	0.004	.0008	γ_{FE}	-0.012	0.004	.0009
γ_{ME}[b]	0.002	0.002	.4463	γ_{ME}[b]	0.002	0.002	.4806
γ_{DL}	-0.170	0.022	< .0001	γ_{DL}	-0.172	0.021	< .0001
γ_{DE}	-0.067	0.027	.0132	γ_{DE}	-0.035	0.002	< .0001
γ_{LE}	0.035	0.026	.1751	γ_{LE}	0[c]		
γ_τ	-0.015	0.001	< .0001	γ_τ	-0.015	0.001	< .0001
LnL		16515.6		lnL		16513.9	

[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. Restrictions on random coefficients do not involve restrictions on the covariance matrix.

We will now discuss how electronic payment services affect the economies of scale in bank intermediation. According to Table 1, the coefficient γ_{DE} is significantly negative, while the coefficient γ_{LE} is not significantly different from zero. Hence, the scale elasticity in the production of deposits increases as the share of electronic payments increases, while the scale elasticity in the production of loans is unaffected by this change in payment technology. This is consistent with our a priori belief, i.e. that this shift in technology affects the production of deposits more than the production of loans, since payment services are more closely connected to deposits than to loans.

The results also imply that economies of scope are present in bank intermediation, since $\gamma_{DL}<0$. Hence, banks benefit from their production of deposits in their production of loans and vice versa. Information about bank customers as depositors may help banks in their evaluation of risk in connection with applications for loan. And furthermore, the production of deposits and loans probably involve similar work tasks, computer programs, advertising strategies and customer handling in general. Learning from one side of the intermediation process therefore affects the productivity of the other side of bank intermediation. Scope economies, in our two-output case, imply that the output-specific scale elasticities increase with the level of the other output, and the isolated effect on the global scale elasticity is also positive.

Concerning the question of input-demand neutrality, we study the coefficients γ_{iE}, i=W,K,M,F. According to Table 1, with the exception of the coefficient calculated from the adding up condition, i.e. γ_{ME}, the simple t-test rejects that these coefficients are zero, as shown by the p-values. We test the joint restriction that $\gamma_{iE}=0$ for all i=W,K,F on model M_G, which implies that also $\gamma_{ME}=0$. The likelihood ratio test gives $\chi^2(3)=88.2$, and the joint restriction is clearly rejected. We conclude that the move towards electronic payment services has caused the cost-shares of labour to increase, while the cost-shares of physical capital and funds have decreased. The cost-share of materials is basically unaffected.

According to the data, labour input has decreased relatively to the input of physical capital and materials, i.e. both the labour-physical capital ratio and the labour-materials input ratio have declined over time. We do not find evidence for the hypothesis that this has been caused by the transition to electronic payments services, however. I.e. we do not find that $\Delta_E IR_{LK}<0$ and $\Delta_E IR_{LM}<0$. This may partly reflect that we measure labour input in number of man-hours rather than in number of employees. The decline in number of employees is stronger than the decline in number of man-hours, and the number of hours per employee has clearly increased. This is interesting, since, at the same time, the number of hours per full-time position as stated by tariff agreements has decreased. There is also a clear positive trend in employees' average level of education in the banking industry, which suggests that the employees, to some degree, have been replaced over time. Banks may well have reduced the number of employees as a result of the shift in technology in payment services, this is probably particularly true for unskilled employees with routine work. At the same time, however, banks have taken the opportunity to increase the number of hours worked by skilled employees and focus on more sophisticated customer services and operations. An expansion of the econometric model to include heterogeneous labour will most likely increase the problem of model convergence and is beyond the scope of this analysis.[12]

According to the estimated model, substitution effects due to changes in relative prices explain much of the observed decline in the labour-physical capital input ratio. While labour costs per man-hour has increased rather smoothly over time, the user cost of physical capital has declined in periods with declining interest rates. Price effects have also caused the input ratio between labour and materials to decline. In addition to the price effects, the observed decline in the input ratios between labour and both physical capital and materials is due to non-homotheticity in bank intermediation. I.e., these changes in input ratios are partly due to the increase in the volumes of deposits and loans. More specifically, we find that $\Delta_m IR_{WK}<0$ for m=D,L and $\Delta_D IR_{WM}<0$, while $\Delta_L IR_{WM}>0$. Hence, higher volumes of outputs, in particular deposits, can be produced with relatively less labour input.

In Table 2 we present the own-price and cross-price elasticities predicted by alternative specifications of the model. A comparison of the elasticities from models M_G and M_R in Table 1 shows that these are almost identical, and we therefore present the elasticities from model M_R only. To evaluate the importance of including random coefficients to capture heterogeneity in the coefficient vector, we estimate the M_R model with two reductions on the number of random coefficients. The calculated elasticities are included in Table 2. The reported t-values of the elasticities are calculated using a first-order Taylor expansion, starting the expansion at the global sample mean of the exogenous variables [cf. Kmenta (1986, p. 486)].[13] In model M_{RI} the intercept term is the only random coefficient, while model M_{RN} includes no random coefficients, i.e. all coefficients are specified as equal across all banks. As explained earlier, in our basic specification, the price of funds (P_F) is based on a money market interest rate. We re-estimate model M_R, however, and replace P_F with a price based on the weighted average of the money market interest rate and the interest rates on deposits. The latter cannot be assumed to be exogenous to the banks. This alternative specification is denoted M_{RF}, and the elasticities predicted by this specification is included in Table 2.

[12] For an analysis and discussion of labour adjustment within Swedish banks, see Kumbhakar, Heshmati and Hjalmarsson (2002). Labour heterogeneity is not included in the study, however.
[13] These calculations, for which we use the TSP 4.3 software (see Hall (1996)), are only done for model M_R.

Table 2. *Own- and cross-price elasticities predicted by alternative specifications of the reduced model. Calculated at the overall empirical means of the cost shares*

Own-price	Estimate				Cross-price	Estimate			
	M_R	M_{RI}	M_{RN}	M_{RF}		M_R	M_{RI}	M_{RN}	M_{RF}
ε_{WW}	-0.468 (-92.75)	-0.480	-0.453	-0.431	ε_{WK}	0.071 (1.83)	0.083	0.070	0.067
ε_{KK}	-0.635 (-1.57)	-0.872	-0.867	-0.660	ε_{WM}	0.307 (5.53)	0.178	0.175	0.297
ε_{MM}	-0.208 (-10.19)	-0.150	-0.193	-0.356	ε_{WF}	0.090 (67.72)	0.220	0.208	0.067
ε_{FF}	-0.024 (-29.94)	-0.074	-0.081	-0.054	ε_{KW}	0.237 (2.14)	0.274	0.231	0.223
					ε_{KM}	-0.086 (-0.98)	-0.039	-0.036	-0.137
					ε_{KF}	0.485 (2.25)	0.637	0.672	0.574
					ε_{MW}	0.387 (12.78)	0.224	0.221	0.376
					ε_{MK}	-0.033 (-3.38)	-0.015	-0.014	-0.052
					ε_{MF}	-0.146 (-11.49)	-0.059	-0.015	0.033
					ε_{FW}	0.018 (19.00)	0.045	0.042	0.013
					ε_{FK}	0.029 (1.88)	0.039	0.041	0.035
					ε_{FM}	-0.023 (-8.20)	-0.009	-0.002	0.005

t-values are given in brackets. M_R is the reduced specification from Table 1. In M_{RI} and M_{RN} the vector of random coefficients are reduced. Subscript RI denotes that only the intercept term is defined as a random coefficient, while subscript RN denotes that no random coefficients are included. Subscript RF refers to an alternative information set where the weighted average of the money market interest rate and the interest rate on deposits multiplied by CPI is used as the price of funds.

A comparison of the price elasticities in Table 2 shows that the predictions from the alternative model specifications are surprisingly close. We conclude that, in general, the two alternative definitions of the price of fund do not affect the price elasticities for the average bank, see models M_R and M_{RF}. Hence, although the weak exogeneity assumption with respect to P_F most likely is violated when estimating model M_{RF}, this does not affect our main conclusions with respect to the price effects. Reducing the rather complicated structure of random coefficients in model M_R, as in models M_{RI} and M_{RN}, influence the price elasticities somewhat more, but not dramatically.

Reducing the number of random coefficients or choosing the alternative information set do affect the value of the log-likelihood function (lnL), however. With respect to the specifications included in Table 2 we have lnL=16513.9 in M_R, lnL=13824.8 in M_{RI}, lnL=13214.9 in M_{RN}, and lnL=15230.3 in M_{RF}. Reducing the number of random coefficients reduces the value of the log-likelihood function, as was expected. The decrease in lnL when choosing the alternative information set is more thought-provoking, since the weighted average of the money market interest rate and the interest rate on deposits better reflects the true average interest rate on fund. This may suggest that, on the margin, if banks want to increase their funding, they pay a price close to the money market interest rate rather than the weighted average of this money market rate and the interest rate on bank deposits. In addition to the weak exogeneity argument, this is an argument for choosing the money market interest rate alone as the price of fund.

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 only exclude five observations due to incomplete data. For the average bank, labour is substitutable with all other inputs. Also physical capital and fund are, on average, substitutes in demand. Materials and both physical capital and fund are complements in demand for the average bank according to these results. We will not pay more attention to the price elasticities, however, since these are not our main interest.

Table 3 gives the elasticity of costs with respect to the share of electronic payments in total non-cash payments, cf. Eq. (9), and the elasticities of scale, cf. Eq. (10) and (11). Again we do not present the elasticities predicted by model M_G, because they are very close to those predicted by model M_R. To calculate the t-values we use the same methodology as in Table 2. While the price elasticities are relatively robust across the alternative specifications, the conclusion is a bit more complicated with respect to the elasticities in Table 3. However, for an average bank, the elasticity of costs with respect to electronic payments, ε_E, is negative across all four specifications. An increase in the share of electronic payments by one per cent decreases costs by around 0.3 per cent for the average bank. We conclude that (average) operating cost decreases over time due to the change in technology in payment services. 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 scale elasticities (ε_D, ε_L, ε_X) predicted by alternative specifications of the reduced model*[a]

	M_R		M_{RI}	M_{RN}	M_{RF}
ε_E	-0.293	(11.72)	-0.293	-0.296	-0.325
ε_D	3.011	(9.23)	4.503	2.498	4.794
ε_L	0.986	(28.29)	1.284	1.401	0.890
ε_X	0.743	(57.23)	0.999	0.898	0.751

[a] Calculated at overall empirical means of the variables. The estimated expectations of the random coefficients are used. t-values are given in brackets. M_R is the reduced specification from Table 1; In M_{RI} and M_{RN} the vector of random coefficients are reduced. Subscript RI denotes that only the intercept term is defined as a random coefficient, while subscript RN denotes that no random coefficients are included. Subscript RF refers to an alternative information set where the weighted average of the money market interest rate and the interest rate on deposits multiplied by CPI is used as the price of funds.

With respect to the production of deposits, we find clear evidence of scale economies for the average bank.[14] Concentrating on model M_R and holding produced loans constant, an increase in all inputs by one per cent increases produced deposits by 3 per cent for the average bank within our sample. This scale elasticity varies quite a lot from about 2½ to 5 per cent depending on the specification, however. The scale elasticity with respect to loans ranges from 0.9 to 1.4 per cent for the average bank, and according to model M_R it is close to one. Both these product specific scale elasticities are declining functions of the product in question, but this negative effect is significantly stronger for loans than for deposits. The scope effect is relatively important, however, and the product specific scale elasticities are increasing functions of the other product. For the average bank, the global scale elasticity is clearly below one in model M_R and M_{RF}, i.e. in both models that include a rich specification of random coefficients. According to the models with a reduced random-coefficient specification, the global scale elasticity is close to one. The results implies a declining global scale elasticity, since $(\gamma_{DD}+\gamma_{LL}+2\cdot\gamma_{DL})=0.080>0$.

To understand the development in scale elasticities over time, we have calculated these elasticities at the annual means of the variables over 1987-1999. These calculations show that both the deposit specific and global scale elasticity increase over time, while the loan specific scale elasticity is relatively stable. The increase in the two former elasticities are partly due to the shift in payment technology, since the increase in electronic payment services has a significant positive effect on the deposit specific and hence also the global scale elasticity. This shift in

[14] Since $\partial \ln C/\partial \ln X_D=(\varepsilon_D)^{-1}$ is positive, the results are consistent with our treatment of deposits as an output, see Hughes et al. (2001). And furthermore, Table 1 and the significant own price elasticity of fund in Table 2 support our treatment of fund (deposits and money market borrowing) as an input in the cost function.

technology has not affected the scale elasticity in the production of loans. In addition, a general decline in the deposits to loans ratio has increased the relative importance of the positive scope effect in the deposit specific scale elasticity, while the opposite is true in the loan specific scale elasticity. This effect on the loan specific scale elasticity has more or less been cancelled out by price effects, however, cf. Eq. (10).

We find the economic interpretation of the results with respect to the scale elasticities plausible. In general, increasing the volume of deposits, by either increasing the volume of an existing deposit account or opening a new account, costs little. The bank does not need to evaluate ordinary depositors, and deposit account conditions are largely pre-defined and common to most customers. In addition, depositor services are largely automated. With respect to loans, the bank must evaluate the risk involved with each application, and this is true for both new and old customers. This rather labour intensive process probably keeps scale economies down.

The conclusion that bank intermediation is characterised by rather large economies of scale in the production of deposits and a scale elasticity closer to one in the production of loans can explain why analyses of banks often find decreasing economies of scale. In general, the deposits to loans ratio is smaller in large banks than in small banks. Since the data applied in general reflect total operating costs, one should expect an effect from output composition on estimated scale elasticities, if the true data generating process is characterised by larger economies of scale in the production of deposits than in the production of loans.

IV. Conclusions

This paper analyses the importance of new technologies, i.e. primarily new electronic payment services, for the development in banks' average costs, 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 for the years 1987-1999. The inputs are labour, physical capital, materials and fund (deposit and money market borrowing). We apply a random coefficient approach, which allows for heterogeneity across banks in both intercept terms 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.

According to the results, average costs have decreased and both the deposit specific and global scale elasticity have increased as the share of electronic payments has increased. The scale elasticity in the production of loans is unaffected by this change in technology. A stronger effect on deposits than on loans was expected, since payment services are more closely connected to deposits than to loans. According to the results, bank intermediation is characterised by economies of scope, rather large economies of scale in the production of deposits and a scale elasticity closer to one in the production of loans. The results with respect to the product specific scale elasticities may help explain why analyses of banks often find decreasing economies of scale. In general, the deposits to loans ratio is smaller in large banks than in small banks, and one should expect an effect from output composition on estimated scale elasticities.

The move towards electronic payment services has affected input demand non-neutrally, causing the cost-share of labour to increase, while the cost-shares of physical capital and fund decrease. The cost share of materials is basically unaffected. Calculations on the estimated model show that the input ratios between labour and both physical capital and materials decrease due to substitution effects and non-homotheticity of the production function. We do not find support for the hypothesis that the shift in payment technology has caused the decline in these input ratios. This may, however, be due to our measurement of labour input. We use the number of man-hours rather than the number of employees, and the former has declined much less than the number of employees. Our data do not reflect the probably most important effect of new technology in the payment systems on employment, namely the reduction in the number of unskilled employees with routine work. At the same time, banks have employed skilled workers, who in general work more hours per worker than unskilled.

Appendix

The data and definition of variables

Primarily we use annual bank-level data which banks are obliged to report to Norges 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 1987-1999. The panel is unbalanced, but 141 banks are observed in the maximum 13 years, which is 82 per cent of the observations. No cleaning of the data has been executed, we only exclude five observations due to incomplete data.

Table A1. *Definition of the empirical variables*

C_f	Total operating costs plus interest expenses in bank f, 1000 NOK
P_{Wf}	Total labour costs per man hour in bank f, NOK
P_{Kf}	User cost of physical capital in bank f, NOK
P_M	Price of materials incl. energy, 1997=1
P_{Ff}	Price of fund 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	Deterministic time trend, 1987=1

Total operating costs plus interest expenses are calculated as the sum of reported costs on labour, user costs on physical capital, costs on materials and interest payments on deposits and loans in the money market.

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 of each component is common across banks, but the aggregate user cost varies due to variation in the share of the components.

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.

Interest payments on deposits and loans in the money market are defined as funding costs. Consistent with the definition of loans and deposits, we calculate the price of fund as the interest rate paid on fund multiplied by the consumer price index (CPI). Two alternative prices of fund are calculated: In alternative one we use the weighted average of banks' interest rate on deposits and a money market interest rate, using ordinary deposits and loans in the money market as weights. In the second alternative we only take into account the money market interest rate.

The volumes of loans and deposits are calculated by deflating nominal stock data, as registered per 31.12 each year, by CPI. I.e., we do not follow the rather common approach and replace output with its value, since the validity of this approach hinge on rather strong assumption within the chosen neo-classical approach. To avoid the measurement error problem completely, however, we need price information that unfortunately is not available. For a more thorough discussion of this, see Sarkis (1999). With respect to business loans, one may argue that the gross domestic product (GDP) price index is a more appropriate deflator, but the GDP-price index and CPI follow a common trend, so this is not of great importance. 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 degrees of freedom. Second, one can argue that banks' annual operating costs may just as well reflect the production of loans and deposits that build up to the stocks registered by the end of the year as a crude average measure.

With respect to electronic payments, we only have the number of transactions at the industry level. We calculate the share of electronic payments in total non-cash payments.

Table A2 gives summary statistics for the variables. According to the empirical means, the input share of fund is well above the other variable inputs.

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

	Mean	St. dev.	Min.	Max.
Total operating costs, Mill. NOK[1]	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 fund	0.667	0.119	0.002	0.964
Labour costs, NOK per man-hour	225	112	6	3588
User costs of physical capital	0.164	0.078	0.042	0.336
Price index of materials incl. energy[2]	0.911	0.069	0.798	1.036
Price of fund, Alt. I[3]	7.396	13.647	0.978	480.839
Price of fund, Alt. II[3]	8.174	2.693	3.850	11.627
Deposits, Mill. 1997-NOK[4]	3635	13700	0.1	211000
Loans, Mill. 1997-NOK[4]	3978	15200	1.6	233000
Share of electronic payments[5]	0.382	0.189	0.137	0.734
Trend variable	6.824	3.822	1	13

[1] Including interest expenses.

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

[3] Alt. I is based on the weighted average of the money market interest rate and the interest rate on deposits, while Alt. II is based on the money market interest rate. We multiply both interest rates by CPI.

[4] Measured per 31. December.

[5] The share of electronic payments in total non-cash payments measured in number of transactions 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.

Berger, A.N. and L.J. Mester (1997): Inside the Black Box: What Explains Differences in the Efficiencies of Financial Institutions?, *Journal of Banking and Finance*, 21, 895-947.

Berger, A.N., W.C. Hunter and S.G. Timme (1993): The Efficiency of Financial Institutions: A Review and Preview of Research Past, Present, and Future, *Journal of Banking and Finance*, 17, 221-249.

Biørn, E., K.-G. Lindquist and T. Skjerpen (2003): Random Coefficients and Unbalanced Panels: An Application on Data from Norwegian Chemical Plants, *Annales d'Economie et de Statistique*, forthcoming.

Biørn, E., K.-G. Lindquist and T. Skjerpen (2002): Heterogeneity in Returns to Scale: A Random Coefficient Analysis with Unbalanced Panel Data, *Journal of Productivity Analysis*, 18, 39-57.

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.E. Robinson (1995): Income, Costs, and Pricing in the Payment System, *Norges 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.

Hall, B.H. (1996): *TSP Reference Manual. Version 4.3,* TSP International.

Hughes, J.P. (1999): Incorporating Risk into the Analysis of Production, *Atlantic Economic Journal*, 27, 1-23.

Hughes, J.P., L.J. Mester and C.-G. Moon (2001): Are Scale Economies in Banking Elusive or Illusive? Evidence Obtained by Incorporating Capital Structure and Risk-Taking into Models of Bank Production, *Journal of Banking and Finance*, 25, 2169-2208.

Humphrey, D.B. (1981): Scale Economies at Automated Clearinghouses, *Journal of Bank Research*, (Summer 1981), 71-81.

Humphrey, D.B. (1982): Costs, Scale Economies, Competition, and Product Mix in the U.S. Payments Mechanism, *Federal Reserve Bulletin*, 68, 215.

Humphrey, D.B. (1985): Cost and scale economies in bank intermediation. In R. Aspinwall and R. Eisenbeis (eds.), *Handbook for Bank Strategy*. New York: Wiley.

Humphrey, D.B and A.N. Berger (1990): Market failure and resource use: Economic incentives to use different payment instruments. In D.B. Humphrey (ed.), *The U.S. payment system: Efficiency, risk and the role of the Federal Reserve: Proceedings of a symposium on the U.S. payment system sponsored by the Federal Reserve Bank of Richmond*, pp. 45-86. Boston: Kluwer Academic.

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.

Humphrey, D.B. and B. Vale (2002): Scale economies, bank mergers, and electronic payments: A spline function approach, Norges Bank, Unpublished paper.

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.

Kmenta, J. (1986): *Elements of Econometrics*, second edition. New York: Macmillan.

Koskenkylä, H. (2001): Opportunities and Threats of European Banks, Bank of Finland, Working paper no. 6/2001.

Kumbhakar, S.C., A. Heshmati and L. Hjalmarsson (2002): How Fast Do Banks Adjust? A Dynamic Model of Labor-Use with an Application to Swedish Banks, *Journal of Productivity Analysis*, 18, 79-102.

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. (2002): Electronic Payments and Scale Properties in the Banking Industry, SUERF Colloquium Volume.

McAllister, P.H. and D. McManus (1993): Resolving the Scale Efficiency Puzzle in Banking, *Journal of Banking and Finance*, 17, 389-405.

Mitchell, K. and N.M. Onvural (1996): Economies of Scale and Scope at Large Commercial Banks: Evidence from the Fourier Flexible Functional Form, *Journal of Money, Credit and Banking*, 28, 178-199.

Sarkis, Z. (1999): *General Cost Structure Analysis. Theory and Application to the Banking Industry*, Boston: Kluwer Academic.

SAS (1992): SAS Technical Report P-229, SAS/STAT Software: Changes and Enhancements, Release 6.07, Cary, NC: SAS Istitute, 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.

Vesala, 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.

Kjersti-Gro Lindquist: The Effect of New Technology in Payment Services on Banks' Intermediation

KEYWORDS:

Banking industry
Electronic payments
Technical change
Panel data

Working Paper 2002/6