## WORKING PAPER NO. 03-16 COST SAVINGS FROM ELECTRONIC PAYMENTS AND ATMS IN EUROPE

David Humphrey
Florida State University and Visiting Fellow, Payment Cards Center, Federal Reserve Bank of Philadelphia

Magnus Willesson
Göran Bergendahl
Ted Lindblom
University of Gothenburg
August 2003

## Federal Reserve Bank of Philadelphia

# Cost Savings from Electronic Payments and ATMs in Europe* 

David Humphrey<br>Florida State University, Tallahassee, FL, U.S.A.<br>and Visiting Fellow, Payment Cards Center, Federal Reserve Bank of Philadelphia<br>Magnus Willesson<br>University of Gothenburg, Gothenburg, Sweden<br>Göran Bergendahl<br>University of Gothenburg, Gothenburg, Sweden<br>Ted Lindblom<br>University of Gothenburg, Gothenburg, Sweden

August 2003


#### Abstract

Electronic payments are considerably cheaper than their paperbased alternatives. Similarly, ATMs are a more cost-efficient way to deliver certain depositor services than are branch offices. As the share of electronic payments in 12 European countries rose from 0.43 in 1987 to 0.79 in 1999 and ATMs expanded while the number of branch offices was constant, bank operating costs are estimated to be $\$ 32$ billion lower than they otherwise might have been, saving $0.38 \%$ of the 12 nations' GDP. Our results are robust to the form of cost function estimated-composite, Fourier, or translog. (93 words)


Key Words: payments, ATM, bank costs, Europe
JEL Classification Code: E41 C53

* We have benefited from comments of seminar participants at the Universities of Valencia and Granada as well as the Federal Reserve Banks of Chicago and Philadelphia. Financial support for the first author from the Foundation BBVA in Spain is gratefully acknowledged. The views expressed are those of the authors alone and not those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.


## 1 Introduction.

Most people pay little attention to their nation's payment system, the banking infrastructure which transfers funds among individuals and organizations in a country. Their only concern is the bank fees they may pay to "use their own money." When direct costs for users of payment services are related to the underlying differential expense of electronic versus paper-based transactions, a nation's payment costs can fall in real terms. If a country is able to shift from an all paper-based to an all electronic-based payment system, annual savings of perhaps $1 \%$ of GDP can be realized. This is because electronic payments, depending on the application (point-of-sale, bill payment, or employee disbursement), are from one-half to two-thirds lower than their alternative paper-based non-cash instrument.

Little information exists regarding the cost of a nation's payment system despite the fact that such expenses may absorb upwards of 2 to $3 \%$ of GDP. No time-series of bank-specific or national aggregate data on payment costs are available (Norway excepted) to determine how a country may have benefited by a shift to lower cost electronic payments and expanded use of ATMs. Our purpose is to provide such an estimate for most of Europe. We use cross-country panel data in an "output characteristics" cost function to do this. Specifically, we search for the production function that relates six output characteristics to the annual operating cost of each of 12 European countries' banking sectors over 1987-1999. Interest expenses are excluded and a cost function is estimated that relates total operating cost to the total annual number of check, paper giro, electronic giro, and card transactions in each country along with the number of ATMs and (standardized, size-adjusted) branch offices while controlling for differences in input prices across countries. As the vast majority of bank operating expenses derive from processing and accounting for payments, delivering cash through ATMs, and taking deposits and disbursing loans at branch offices, the above specification allows us to illustrate how payment costs and ATM and branch service delivery expenses have varied over time.

The usual approach for identifying cost relationships in banking relates total operating plus interest cost to the value of the stock of loans, deposits, securities, etc., under the assumption that changes in these stocks reflect changes in the underlying flow of banking services. A time dummy variable then identifies (disembodied) technical change or, less often, (embodied) technical change is assumed to be reflected in changes in the cost share or price of certain inputs. Our approach is quite different. Pairing newly available payment transaction data with existing information on ATM and
branch operations, we can measure more precisely the set of banking services that are the most costly for banks to provide. Changes in the composition of these services directly reflects the adoption of newer technology. Scale effects associated with these services are determined from cross-country variations in predicted unit costs as payment and service delivery production functions are basically the same across countries. The effect of technical change from altering payment and service delivery modes is captured in our specified variables rather than combined into a single set of linear, quadratic, or time-specific dummy variables.

In what follows, we illustrate in Section 2 how changes in bank operating costs have been affected by changes in the use of different payment instruments and service delivery methods over 1987-1999 for 12 European countries. ${ }^{1}$ Our "output characteristics" cost models are specified in Section 3. While the composite cost function underlies our analysis (Pulley and Braunstein, 1992; Pulley and Humphrey, 1993), the robustness of our results is demonstrated by contrasting our main estimated operating cost changes with those from more commonly used translog and Fourier cost models.

Estimates of how banking costs have changed from expanded electronic payments and greater reliance on ATMs are reported in Section 4. Overall, the 12 European countries may have saved $\$ 32$ billion by shifting from paper-based to electronic payments and relying on ATMs rather than costly branch offices to deliver certain depositor-related services. This savings is $0.38 \%$ of their aggregate GDP in 1999. Countries that have progressed further in shifting to electronic from paper-based payments and to ATMs from branch offices have benefited the most. Two checks on the robustness of our approach are noted in Section 5 while our main results are summarized in Section 6, which concludes the paper. It is likely that the same payment use and service delivery trends shown to have benefited Europe will also apply to other nations, including the U.S., who are at an earlier stage of this technology substitution process.

## 2 Changes in Payment and Service Delivery Mix.

In the banking industry, the ratio of operating expenses to the value of total assets (OC/TA) is an accepted indicator of unit operating costs. ${ }^{2}$ As

[^0]seen in row 1 of Table 1, this indicator of bank unit costs has fallen by $24 \%$ over 1987-1999 for Europe. ${ }^{3}$ Smaller reductions were experienced by Denmark, Sweden, the Netherlands, and Belgium (from - $13 \%$ to $-17 \%$ ) while larger reductions occurred for Spain, Norway, and the U.K. ( $-34 \%$ to $-43 \%$ ). Reductions close to the overall average were experienced by Germany, Italy, and France ( -22 to $-24 \%$ ). ${ }^{4}$ Although the mix of payment instruments and service delivery channels often differ markedly among our 12 countries, all of them expanded their share of electronic payments and their supply of ATMs relative to branch offices.

Few countries have accurate information on the value, much less the number, of cash transactions. Consequently, the four payment instruments we focus on are non-cash transactions, two of which are paper-based (check and paper giro payments) and two that are electronic (electronic giro and debit and credit card transactions). ${ }^{5}$ Overall, the number of non-cash transactions in Europe rose from 24.6 billion in 1987 to 46.5 billion in 1999, an $89 \%$ increase. This rise far exceeded the rate of population growth so the number of non-cash transactions made per person per year shown in row
which affect total costs but not operating expenses, are not the same across countries nor constant over time. The OC/TA ratio is an indicator of the labor, materials, outsourcing, and capital consumption costs of producing the flow of banking services which underlies the value of total assets.
${ }^{3}$ In order to reflect properly the cost experience for Europe as a whole, our OC/TA measure is computed as the sum of all 12 countries' bank operating expenses (OC) divided by the sum of the value of their banking assets (TA). A simple average of each country's $\mathrm{OC} / \mathrm{TA}$ ratio would weight equally each country even though their level of operating expense and value of assets are quite different. All ratios in Table 1 therefore treat the 12 countries as if they were a single entity (i.e., they are the sum of the numerator divided by the sum of the denominator). While it is possible that the $-24 \%$ change in the $\mathrm{OC} / \mathrm{TA}$ ratio was due to banks' expanded reliance on purchased funds from other banks or other liabilities used for funding rather than a reduction in payment and service delivery costs, this effect was small. The mean ratio of deposits to total assets across the 12 countries was very stable. It was 0.908 in $1987,0.887$ in 1988 , ended up at 0.881 in 1999 , so only perhaps around two percentage points of the $-24 \%$ change could be attributed to this cause.
${ }^{4}$ While slight increases were experienced by Switzerland and Finland ( $4 \%$ to $8 \%$ ) between 1987 and 1999, this is misleading. The OC/TA ratio initially rose after 1987 in these two countries (probably due to initially high investment outlays and capital consumption expenses), reached a peak in the early to mid-1990s, and then fell by $-21 \%$ and $-42 \%$, respectively, by 1999 (reflecting a lower level of needed investment).
${ }^{5}$ Giro transactions include direct debits and credit transfers. The number of internet payments is so small relative to the four payment instruments we use that it was not separately reported for any of our 12 countries in the ECB or BIS payment system statistical documents during our time period.

Table 1: Bank Operating Cost, Payments, and Service Delivery in Europe

|  | 1987 | 1989 | 1991 | 1993 | 1995 | 1997 | 1999 | Change |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| OC/TA | .021 | .020 | .020 | .019 | .018 | .017 | .016 | $-24 \%$ |
| Non-Cash/POP | 72 | 78 | 89 | 99 | 108 | 118 | 131 | $82 \%$ |
| Ele/Non-Cash | .43 | .48 | .53 | .60 | .65 | .72 | .79 | $84 \%$ |
| ATM/BR | .28 | .41 | .54 | .67 | .85 | .99 | 1.19 | $325 \%$ |
| Sour |  |  |  |  |  |  |  |  |

Source: OECD, ECB, BIS, and own calculations (data are rounded).

2 of Table 1 rose from 72 to 131 over this period. ${ }^{6}$
The processing of payment transactions, debiting and crediting deposit accounts, and safekeeping of funds generates the vast majority of bank "back-office" labor and capital (including branch and computer) expenses. As seen in row 3 of Table 1, the share of electronic payments in all non-cash transactions in Europe expanded from 0.43 to 0.79 over the period (an $84 \%$ rise). Electronic debit card or giro payments for point-of-sale transactions, consumer bill payments, and employee disbursements are typically much cheaper than their paper-based alternatives (a check or paper giro transaction). For these types of transactions survey information and cost estimates suggest that electronic payments are often from one-half to two-thirds lower than their paper-based alternatives (Flatraaker and Robinson, 1995; Wells, 1996; Humphrey, Willesson, Lindblom, and Bergendahl, 2003). ${ }^{7}$ It is expected that the $24 \%$ reduction in unit operating expense shown in Table 1 is associated with the rise in electronic giro $(192 \%)$ and card-based payments ( $671 \%$ ) along with the reduction in check $(-10 \%)$ and paper giro transactions (-79\%).

In terms of delivering banking services to customers, ATMs and branch offices generate most of the labor and capital costs associated with bank "front office" expenses. Services such as loan origination and monitoring, liquidity management, and off-balance-sheet activities generate little labor or capital cost but, being risk-taking activities, bring in the majority of

[^1]revenues (with fee income representing the remainder). The rapid expansion of ATMs in Europe during the last half of the 1980s indicates that, for the range of services provided (cash withdrawal, account transfer, balance inquiry), ATMs have replaced the traditional banking office for a large and growing segment of depositors.

Evidence of this shift is seen in row 4 of Table 1, which shows that the number of ATMs per branch office increased $325 \%$ over 1987-1999 in Europe or from about 1 ATM for each of 3.5 offices in 1987 to 1.2 ATMs per office in 1999. Had ATMs not been invented, branch offices would have expanded in rough proportion to population growth. As the growth in the number of branch offices in Europe over 1987-1999 was minuscule (at only $0.3 \%$ or 500 new offices), it is likely that the reduction in bank operating expenses outlined above may be related in part to the rise in ATMs (which increased $318 \%$ or by 156,000 terminals) and represent a cheaper way to deliver a major depositor service-cash acquisition-than building new branches. ${ }^{8}$

Although the number of branch offices per unit of population used to deliver banking services can differ considerably across countries, this primarily reflects differences in the average size of banking offices in a country. ${ }^{9}$ Regardless of the number of branches or their average size, the number of bank employees per 10,000 inhabitants fell in all but Germany, the Netherlands, and the U.K. over our 13 -year period, suggesting that ATM use has likely conserved on bank labor costs. ${ }^{10}$

In what follows, we attempt to determine the change in bank cost from the shift to electronic payments and ATM use in Europe. This requires a statistical analysis that relates cross-country national banking system operating costs to national information on the transaction volume of four different types of payment instruments, numbers of ATMs and banking offices, as well as labor and capital input prices in a panel data set.

[^2]
## 3 Using Output Characteristics to Determine Cost Effects of Scale and Technical Change.

Instead of measuring the flow of banking payment, deposit account maintenance, cash accessibility, liquidity, and loan initiation and monitoring services directly, it has been common in academic studies (due to a lack of data) to assume that these service flows are proportional to the value of the stock of bank deposits, loans, and securities in the balance sheet. Inferences on how costs may vary by size of bank are obtained by relating total operating plus interest expenses across banks and over time to the value of their deposits, loans, and security holdings (or some other combination of on- or off-balance-sheet positions). As information does not normally exist regarding the adoption of specific technical and other cost-saving innovations in banking, the default has been to assume that unknown technical change occurs linearly (or quadratically) with the passage of time and/or is somehow associated with (embodied in) the cost share or price of particular inputs.

An alternative approach, and the one used here, is to relate banking costs to newly available measurable physical characteristics of banking output associated with payment processing and service delivery levels and mix. This achieves two goals. First, the number-but not necessarily the mix-of transactions being processed on behalf of bank customers, along with the number of bank branches and ATMs-but not necessarily their mix, are directly associated with the size of a bank and its labor, capital, and materials operating cost from which scale economies can be determined. ${ }^{11}$ Second, changes over time in the mix of electronic to paper-based transactions or in the mix of ATMs to branches, along with improvements in their associated technology, represent an alternative and more specific way to identify the cost effect of technical change in banking. ${ }^{12}$

Paper-based and electronic payment transactions are jointly processed in the deposit accounting function while aspects of service delivery are jointly

[^3]produced via branches and ATMs. Thus payment and service delivery functions can be considered functionally separable. About the only interaction would be consumers and businesses depositing (a declining number of) checks at a branch office and perhaps, on a one-time basis, filling out documents to pay recurring bills by electronic giro or applying for a debit/credit card. After establishing a giro account, bill payments occur automatically, as do all card payments, without branch or ATM intervention.

### 3.1 A Composite Cost Function.

Our panel data consists of total operating cost, the number of check, paper and electronic giro, and card transactions, the number of ATMs and branch offices, plus data on banking industry labor and capital input prices for 12 countries annually over 1987-1999. This is used in a non-linear composite cost function. The purpose is to estimate the effect that increasing electronic payments and expanded ATM terminal availability may have had on the cost of banking services in Europe. ${ }^{13}$

The composite model can accommodate zero or low values of banking outputs (which occurs in some countries for checks and paper giro transactions in the late 1990s). As a result, it approximates better the scope-type joint cost effects that are associated with altering the mix of how banking services are delivered and the types of payments processed. The levels of banking output in a composite function are not in logs, although input prices are. By keeping output in absolutes, we specify a direct relationship between output and operating costs that is likely more accurate-for prediction purposes when one or more outputs are very small-than if the log of output is related to the log of operating cost. ${ }^{14}$ As well, by specifying the log of input prices, it is possible to impose the theoretical condition of linear homogeneity in input prices in estimation. ${ }^{15}$

[^4]The composite cost function (1), in its non-separable quadratic form, is estimated jointly with $\mathrm{n}-1$ cost share equations. The Box-Cox (1964) transformation is represented by a superscripted parameter in parenthesis $(\phi)$ where $O C^{(\phi)}=\left(O C^{\phi}-1\right) / \phi$ for $\phi \neq 0$ and $O C^{(\phi)}=\ln O C$ for $\phi=0$ in:

$$
\begin{align*}
O C^{(\phi)}= & f^{(\phi)}(\underline{Q}, \underline{\ln P}) \\
= & \left\{[ \alpha _ { 0 } + \sum _ { i = 1 } ^ { 6 } \alpha _ { i } Q _ { i } ^ { \prime } + 1 / 2 \sum _ { i = 1 } ^ { 6 } \sum _ { j = 1 } ^ { 6 } \alpha _ { i , j } Q _ { i } ^ { \prime } Q _ { j } ^ { \prime } + \sum _ { i = 1 } ^ { 6 } \sum _ { k = 1 } ^ { 2 } \delta _ { i , k } Q _ { i } ^ { \prime } \operatorname { l n } P _ { k } ] \bullet \operatorname { e x p } \left[\beta_{0}\right.\right. \\
& \left.\left.+\sum_{k=1}^{2} \beta_{k} \ln P_{k}+1 / 2 \sum_{k=1}^{2} \sum_{m=1}^{2} \beta_{k, m} \ln P_{k} \ln P_{m}+\sum_{i=1}^{6} \sum_{k=1}^{2} \mu_{i, k} Q_{i}^{\prime} \ln P_{k}\right]\right\}^{(\phi)}  \tag{1}\\
S_{k}= & \left(\sum_{i=1}^{6} \delta_{i, k} Q_{i}^{\prime}\right) \bullet\left[\alpha_{0}+\sum_{i=1}^{6} \alpha_{i} Q_{i}^{\prime}+1 / 2 \sum_{i=1}^{6} \sum_{j=1}^{6} \alpha_{i, j} Q_{i}^{\prime} Q_{j}^{\prime}+\sum_{i=1}^{6} \sum_{k=1}^{2} \delta_{i, k} Q_{i}^{\prime} \ln P_{k}\right]^{-1} \\
& +\beta_{k}+\sum_{m=1}^{2} \beta_{k, m} \ln P_{m}+\sum_{i=1}^{6} \mu_{i, k} Q_{i}^{\prime}
\end{align*}
$$

where:
$O C=$ total operating cost, composed of all labor, materials, outsourcing, and capital consumption costs (but no interest expenses); ${ }^{16}$
$Q_{i, j}^{\prime} \quad=$ six output characteristics $(i=j=1, \ldots, 6)$ composed of four payment processing alternatives and two service delivery modes. The four payment processing alternatives are the number of checks ( CHECK ), paper or electronic giro payments (PGIRO, EGIRO), and debit and credit card transactions ( $C A R D$ ). The two service delivery alternatives are represented by the number of automated teller machines (ATM) and the number of standardized, size-adjusted, bank branches $\left(B R^{S T D}\right)$. In (1), $Q^{\prime}=Q-1$;
$P_{k, m}=$ prices of two inputs-the average labor cost per bank employee and an opportunity cost approximation to the price of bank physical capital and materials inputs represented by each country's market interest rate; and
$S_{k} \quad=$ the cost share for the labor input (the capital/materials input cost share is deleted to avoid singularity).

[^5]It is expected that operating costs not directly associated with the type of payment or mode and level of service delivery will be represented in the intercept term.

The composite function is non-linear and is estimated iteratively. Following Pulley and Braunstein (1992), let $\underline{D}=\underline{0}$ and $G M^{\phi-1}$ be the geometric mean of operating cost $O C$, then the composite model is estimated from the "pseudo model" (2): ${ }^{17}$

$$
\begin{align*}
D= & {\left[-\left(O C^{(\phi)} / G M^{\phi-1}\right)+f^{(\phi)}(\underline{Q}, \underline{\ln P}) / G M^{\phi-1}\right] } \\
= & {\left[-\left\{\left(O C^{\phi}-1\right) / \phi G M^{\phi-1}\right\}+\left(\left\{\left[\alpha_{0}+\sum_{i=1}^{6} \alpha_{i} Q_{i}^{\prime}+1 / 2 \sum_{i=1}^{6}\right.\right.\right.\right.} \\
& \left.\sum_{j=1}^{6} \alpha_{i, j} Q_{i}^{\prime} Q_{j}^{\prime}+\sum_{i=1}^{6} \sum_{k=1}^{2} \delta_{i, k} Q_{i}^{\prime} \ln P_{k}\right] \bullet \exp \left[\beta_{0}+\sum_{k=1}^{2} \beta_{k} \ln P_{k}\right.  \tag{2}\\
& \left.\left.\left.\left.+1 / 2 \sum_{k=1}^{2} \sum_{m=1}^{2} \beta_{k, m} \ln P_{k} \ln P_{m}+\sum_{i=1}^{6} \sum_{k=1}^{2} \mu_{i, k} Q_{i}^{\prime} \ln P_{k}\right]\right\}^{\phi}-1\right) / \phi G M^{\phi-1}\right] \\
S_{k}= & \left(\sum_{i=1}^{6} \delta_{i, k} Q_{i}^{\prime}\right) \bullet\left[\alpha_{0}+\sum_{i=1}^{6} \alpha_{i} Q_{i}^{\prime}+1 / 2 \sum_{i=1}^{6} \sum_{j=1}^{6} \alpha_{i, j} Q_{i}^{\prime} Q_{j}^{\prime}+\sum_{i=1}^{6} \sum_{k=1}^{2} \delta_{i, k} Q_{i}^{\prime} \ln P_{k}\right]^{-1} \\
& +\beta_{k}+\sum_{m=1}^{2} \beta_{k, m} \ln P_{m}+\sum_{i=1}^{6} \mu_{i, k} Q_{i}^{\prime}
\end{align*}
$$

One data measurement problem required correction. It is clear that a single payment transaction in one country, whether by check, giro, or card, will be measured as a single payment transaction in another country. The same basically holds for an ATM even though newer models may be somewhat more efficient. However, this is not the case for banking offices across countries. The size of a branch-measured by the number of workers per office-were often quite different across countries making it necessary to standardize them according to some benchmark to make them more comparable. Otherwise, differences in the (statistically allocated) operating cost of a single branch office in one country, compared to the operating cost of a single branch in another, could differ due both to their possible difference in efficiency (which is acceptable) as well as to their different sizes (comparability problem). Using non-comparable branch office data would bias our

[^6]cross-country estimation since we specify the number of branch offices as an output characteristic. France, with an average of 16.04 workers per office over our 13 -year period, was selected as the benchmark and other countries were adjusted accordingly. ${ }^{18}$

### 3.2 Alternative Translog and Fourier Cost Functions.

To illustrate the robustness of our analysis, we also estimate translog and Fourier cost functions. However, either function may generate biased results compared to the composite form when levels of some outputs are very small since outputs are specified in logs. ${ }^{19}$ As well, neither model can be estimated when some outputs are zero. Since paper giro transactions in some years were zero for some countries, paper giro payments and checks were aggregated into a single category of paper-based payments. For symmetry, electronic giro and card transactions were aggregated into a single category of electronic payments. This gives four (rather than six) output characteristics when these two models are estimated. As this aggregation had little effect on results derived from the six characteristic composite model (2), the composite results reported below are from this more general (unaggregated) model.

The translog cost function (3) with four output characteristics is estimated jointly with $\mathrm{n}-1$ cost share equations:

[^7]\[

$$
\begin{align*}
\ln O C= & \alpha_{0}+\sum_{i=1}^{4} \alpha_{i} \ln Q_{i}+1 / 2 \sum_{i=1}^{4} \sum_{j=1}^{4} \alpha_{i j} \ln Q_{i} \ln Q_{j}+\sum_{i=1}^{4} \sum_{k=1}^{2} \delta_{i, k} \\
& \ln Q_{i} \ln P_{k}+\sum_{k=1}^{2} \beta_{k} \ln P_{k}+1 / 2 \sum_{k=1}^{2} \sum_{m=1}^{2} \beta_{k, m} \ln P_{k} \ln P_{m}  \tag{3}\\
S_{k}= & \beta_{k}+\sum_{m=1}^{2} \beta_{k, m} \ln P_{m}+\sum_{i=1}^{4} \delta_{i, k} \ln Q_{i}
\end{align*}
$$
\]

where the variables have been defined above.
The Fourier form we use adds sin and cos terms to the translog cost function. As our main concern is to allow for greater flexibility in the local identification of output effects on operating costs, the sin and cos terms are applied to the output $(Q)$ measure. The Fourier form is a globally flexible approximation since the respective sin and cos terms are mutually orthogonal over the $[0,2 \pi]$ interval. The Fourier function (4) is estimated jointly with the cost shares: ${ }^{20}$

$$
\begin{align*}
\ln T C= & \text { Translog Cost Function } \\
& +\sum_{n=1}^{4}\left[\tau_{n} \cos \left(\ln Q_{n}^{*}\right)+\omega_{n} \sin \left(\ln Q_{n}^{*}\right)\right] \\
& +\sum_{n=1}^{4} \sum_{q=1}^{4}\left[\tau_{n q} \cos \left(\ln Q_{n}^{*}+\ln Q_{q}^{*}\right)+\omega_{n q} \sin \left(\ln Q_{n}^{*}+\ln Q_{q}^{*}\right)\right] \\
& +\sum_{n=1}^{4}\left[\tau_{n n n} \cos \left(\ln Q_{n}^{*}+\ln Q_{n}^{*}+\ln Q_{n}^{*}\right)\right.  \tag{4}\\
& \left.+\omega_{n n n} \sin \left(\ln Q_{n}^{*}+\ln Q_{n}^{*}+\ln Q_{n}^{*}\right)\right] \\
S_{k}= & \beta_{k}+\sum_{m=1}^{2} \beta_{k, m} \ln P_{m}+\sum_{i=1}^{4} \delta_{i, k} \ln Q_{i}
\end{align*}
$$

[^8]
## 4 Cost Effects from Changes in Payment and Service Delivery Levels and Mix.

### 4.1 Composite Function Results.

Predicted unit operating cost for 1987, 1993, and 1999 (in U.S. dollars) are shown in Figure 1 for our panel of 156 observations on 12 countries over 13 years using a composite function. ${ }^{21}$ The levels and mix of check, giro, and card payment volumes as well as the number of ATMs and branch offices are specific to the year indicated but vary across the 12 countries (giving the slope) while input prices are held constant at their overall mean value in the panel data set. As $\phi$ in the composite form is 0.26 , the estimated model is closer to a specification which includes the log of output as well as input prices (when $\phi=0.0$ ) than it is to a specification with output in absolutes and prices in logs (when $\phi=1.0$ ). Even so, the estimated model is significantly different from either of these alternatives since $\phi$ is significantly different from zero or one. ${ }^{22}$

The curves shown in Figure 1 are cubic splines of the predicted values and illustrate how unit operating cost varies by (the log of) banking sector asset size for each of the years noted. ${ }^{23}$ Thus Figure 1 illustrates both the scale effect and mix (cross-country slope) as well as technical change and change in mix (time-series shift) associated with back office payment processing and front office service delivery cost changes. In 1987, 1993, and 1999, the predicted operating cost per dollar of observed assets, as a weighted average across 12 countries, was $0.024,0.019$, and 0.017 , respectively, and are close to the OC/TA ratios computed using observed data in row 1 of Table 1 above. The change in this ratio between 1987 and 1999 was - $30 \% .{ }^{24}$ As operating

[^9]

Figure 1:
cost averaged $26 \%$ of total cost (TC) over the whole period, a $30 \%$ reduction in operating expenses translates into a $7.8 \%$ reduction in total cost $(=0.30$ x 0.26). With constant interest rates, this suggests that an indicator of average cost (measured as TC/TA) and hence real bank prices could have fallen by $7.8 \%$, benefitting users of banking services.

Total bank operating cost for our 12 countries was $\$ 162.9$ billion in 1987. If operating cost had expanded at the same rate as the growth of all payment transactions over 1987-1999 (89\%), it would have grown to $\$ 307.7$ billion by 1999. Subtracting the observed level of operating cost in 1999 ( $\$ 275.6$ billion) from this projected value implies that the overall cost savings may be $\$ 32.1$ billion. This savings estimate is $0.38 \%$ of the 12 nations' aggregate GDP in 1999. Put differently, operating cost associated with changes in payment volume and service delivery levels and mix, as well as possibly other influences across our 12 countries, may have fallen by some $\$ 2.5$ billion a

Thus the set of 12 countries is treated as if it were a single entity. Using a simple average of each of 12 separate country ratios the result would have been $0.026,0.022$, and 0.018 for these three years and the change between 1987 and 1999 would have been $-34 \%$. All results shown have been rounded off. Inflation affects the numerator and denominator so cross-country differences in inflation should not affect the ratio.


Figure 2:
year in Europe. ${ }^{25}$
The mean operating cost scale economy over 1987-1999 is estimated to be 0.90 . A $10 \%$ expansion in all six output characteristics is thus associated with only a $9.0 \%$ rise in operating expenses (so average operating cost would fall). ${ }^{26}$ Consequently, countries with larger banking systems as measured on the X -axis in Figure 1 experience lower unit costs for the payment and service delivery products they produce due to scale and mix effects. ${ }^{27}$

[^10]

Figure 3:

### 4.2 Processing Costs: Check, Giro, and Card Transactions.

The change in payment shares among four types of non-cash payment instruments in Europe is shown in Figure 2. The share of check transactions dropped by half, from 0.40 in 1987 to 0.19 in 1999, while the reduction in paper giro payments was almost ninety percent, falling from 0.17 to 0.02 over the period. The share of electronic giro payments rose by 54 percent (from 0.37 in 1987 to 0.58 in 1999) while the card share rose by over 300 percent (from 0.05 to 0.21 ). Overall, the share of electronic payments rose from 0.43 to 0.79 in Europe over our 13-year period. Since electronic payments have greater scale economies than paper-based payments and are markedly cheaper to process, the rising share of electronic payments in Europe should reduce banking system costs.

Predicted unit payment costs are shown for three years in Figure 3. The level of predicted payment operating costs divided by the total volume of

[^11]payment transactions for four payment instruments on the Y-axis is arrayed against the log of the value of total banking assets for each of 12 countries for 1987, 1993, and 1999 on the X-axis. Predicted payment operating costs are determined by evaluating an estimated (composite) cost function with observed payment transaction volumes for all of the four payment instruments while holding constant the number of ATMs, the number of (standardized) branch offices, and input prices at their sample mean values over 1987-1999. It is emphasized that the unit payment cost indicators shown are not average cost curves. Strictly speaking, it is not feasible to obtain accurate average cost estimates for any subset of outputs from a multiple output cost function. ${ }^{28}$ The curves in Figure 3 reflect the values of the sum of the mean cost of ATMs, branch offices, and input prices along with how the variation in payment transaction volume affects operating cost over time and across countries. While the level shown on the Y-axis necessarily includes more than just payment costs, the slope of the curves indicate how payment costs on a per transaction basis varies across our 12 countries while shifts in the curves indicate how these costs have changed over time. ${ }^{29}$ Since we do not have accurate information on the predicted level of total payment operating cost (or the level of average cost per payment as just explained), it is not possible to determine separately the cost savings from the shift to electronic payments. ${ }^{30}$

As seen in Figure 3, unit payment costs are lower for countries with larger banking systems and appear to fall over time. Strong scale effects exist since the scale economy estimate is 0.09 from the composite form. This implies

[^12]

Figure 4:
that a $10 \%$ rise in payment volume is associated on average across countries with only a $0.9 \%$ rise in allocated operating cost, so unit cost falls. ${ }^{31}$

Not all payment costs are falling. Indeed, the reduction in overall unit payment expense illustrated in Figure 3 is composed of rising predicted check and paper giro unit costs and falling electronic giro and card predicted unit costs over time. The changes in predicted paper-based and electronic payments shown in Figure 4 illustrate more clearly the direction and degree of the cost changes experienced. As in Figure 3, however, the unit cost levels indicated on the Y-axis are not average costs since these predicted costs include the mean values of the cost of ATMs, branch offices, and input prices. In addition, the predicted costs shown in Figure 4 for paper-based (electronic) payments include costs associated with the mean values of electronic (paper-based) transactions. That said, it is clear that the cost of paper-based payments is rising as their use declines (the reverse of scale economies) while the cost of electronic payments continues to fall (due to scale effects and technical change).

[^13]Predicted Delivery Operating Cost Divided by Asset Value by Log of Asset Value: Three Separate Years--1987, 1993, 1999 in U.S. Dollars
(Composite function--input prices and all payment volumes
held constant at their mean values)


Figure 5:

### 4.3 Service Delivery Costs: ATMs and Branch offices.

Service delivery costs represent operating expenses associated with ATMs and branch offices, holding input prices and four types of payment volumes constant at their mean values over the period. Predicted delivery operating expenses from a composite function are divided by the observed value of total assets as it would make little sense to deflate them by the sum of ATMs and the number of branches. The resulting predicted values are shown for three years in Figure 5. ${ }^{32}$ Across the 12 countries, service delivery expenses are flat or rise somewhat and then fall as a ratio to asset value. More detailed analysis (not shown) indicates that almost all the reduction in predicted costs associated with the downward shift in these curves is the result of reductions in predicted ATM unit costs (since predicted branch expenses were relatively stable over 1987-1999). Over our 13-year period, ATMs more than quadrupled, rising from 49,000 in 1987 to 205,000 in 1999. In

[^14]

Figure 6:
contrast, the number of (actual, not standardized) branch offices rose by only $0.3 \%$, from 172,400 in 1987 to 172,900 in 1999. As the number of branch offices was effectively constant over the period, the estimated delivery scale economy value of 0.81 reflects only the growth of ATMs and indicates that a $10 \%$ expansion is associated with a $8.1 \%$ increase in allocated operating cost. ${ }^{33}$ The reduction in predicted delivery costs as a ratio to asset value is seen more clearly by year in Figure 6.

## 5 Robustness Check and Country-Specific Effects.

### 5.1 Translog and Fourier Function Results.

As paper giro transactions were zero for some countries at the end of the 1990s, neither the translog nor the Fourier functions could be estimated until the set of six output characteristics used so far were combined into four. ${ }^{34}$

[^15]Predicted Unit Operating Cost for Europe by Log of Asset Value: 4 Payments Aggregated into 2--All Electronic \& All Paper-Based (Fourier function--input prices held constant at their mean value)


Figure 7:

Re-estimating the composite function with four output characteristics gave scale and cost results quite similar to those noted above and illustrated in Figure 1 with six characteristics. In addition, the mean scale and cost change results for both the translog and the Fourier functions are similar to those just reported for the composite form. As seen for the Fourier results in Figure 7, the predicted unit operating cost curves all have pretty much the same slope as those for the composite form in Figure 1. ${ }^{35}$ The main difference is that cost reductions over time appear to be somewhat larger using the composite form than that identified with the Fourier or translog.

Overall, the composite form yielded a $-30 \%$ change in the OC/TA ratio over 1987-1999 while the translog and Fourier forms gave the change as $33 \%$ and $-28 \%$, respectively. In addition, the operating cost scale economy values were 0.90 (composite), 0.92 (translog), and 0.84 (Fourier) and the cost concavity condition was satisfied. From this perspective, our main conclusions above drawn from the composite form are seemingly robust to

[^16]the form of the cost function specified. ${ }^{36}$

## 5.2 "Standard Approach" with Time Dummy for Technical Change.

As an additional check on our results, we used the same data in a more standard banking model specification where operating cost is a function of: (1) balance sheet values of two "outputs" loans and securities (or loans and deposits); (2) the same prices of labor and capital and materials inputs used here; and (3), a linear time dummy variable that is typically assumed to reflect the influence of exogenous technical change over time. ${ }^{37}$ The resulting scale economy estimates from the composite (0.91), translog (0.91), and Fourier ( 0.80 ) standard models are very similar to those obtained when output characteristics are directly measured as four types of payment transactions plus ATMs and branch offices.

Changes in the ratio of predicted operating cost to total assets (OC/TA) for these three standard cost models between 1987 and 1999 were $-12 \%$ (composite), $-26 \%$ (translog), and $-21 \%$ (Fourier). This is a combination of scale effects, unspecified endogenous productivity advances, and unspecified exogenous technical change. ${ }^{38}$ In our analysis above, the unspecified productivity and technical changes in the standard approach are identified as being the shift to electronic payments and the expanded use of ATMs. The cost changes associated our more detailed specification are larger, at $-30 \%$. $-33 \%$, and $-28 \%$ for the same three cost functions, than those obtained with the standard cost model that uses a time dummy variable. In our view this

[^17]is a richer specification that focuses on the likely sources of cost reduction rather than leaving it unidentified.

Lastly, we note that many things may have affected the shift to electronic payments and ATMs. This ranges from banking mergers, deregulation, or even changes in competition on one side to price incentives for payment users and cost considerations by management for service delivery on the other. Our analysis looks only at the end effect of these influences-namely the actual shift to electronic payments and the adoption of ATMs versus a continued expansion of branches-and estimates the resulting cost benefit. Our personal opinion is that the usual suspects so often invoked in banking analyses-mergers, deregulation, and competition-have played a very minor role here. ${ }^{39}$

### 5.3 Country-Specific Cost Effects.

Country-specific effects on bank unit operating cost can be approximated from Figure 1 while Figures 3 and 5, respectively, indicate the approximate effect on unit cost from the substitution of electronic for paper-based payments and the expansion of ATMs relative to bank branches. As the X-axis in these figures reflects the log of total assets of each country's banking system, the order of our 12 countries on this axis (from smallest to largest) is Finland, Norway, Denmark, Sweden, Belgium, Spain, Netherlands, Switzerland, Italy, U.K., France, and Germany.

Although operating cost scale economies seem to exist for almost all countries in Figure 1, this benefit-which combines payment and service delivery effects-seems to have been eliminated for the very largest countries by 1999. And while reductions in unit costs over time (downward shifts in the curves in Figure 1) appear to have been fairly equally experienced during the 1990s, the exception again is for the largest countries late in the decade. These results from Figure 1 are due to the fact that the largest countries shown in Figure 3 (e.g., France, Germany) experienced little or no cost reduction in unit payment cost during the 1990s (other than that due to scale effects) while smaller countries (particularly in Scandinavia) consistently reduced these expenses over time. As shown in Figure 5, smaller countries also consistently experienced lower delivery expenses relative to

[^18]assets as ATMs expanded-but experienced almost no scale benefits-while just the opposite occurred in the larger countries. Service delivery scale effects in larger countries seem to have been very strong even though unit cost appears to have shifted up over time rather than down. Such differential cost effects are expected. As not all countries have switched to electronic payments or adopted ATMs at the same rate during the 1990s, the associated cost effects will also differ. ${ }^{40}$

## 6 Summary and Conclusions.

The operating cost of providing bank payment and other services has fallen by $24 \%$ relative to banks' total assets in 12 European countries over 19871999. This reduction is associated with the ongoing replacement of paperbased payment instruments (checks and paper giro payments) with lower cost electronic alternatives (debit cards and electronic giro payments) as well as the expanded use of lower cost and more convenient ATMs relative to branch offices to deliver cash and certain other services to depositors and borrowers. While the banking literature properly emphasizes the role of intermediation activities due to their importance for investment and economic growth, the vast majority of operating cost incurred by banks-the approximate two-fifths of total operating plus interest expenses going to labor, materials, outsourcing, and capital consumption-are associated with back office payment activities and front office cash acquisition, deposit safekeeping, trust, and loan initiation and monitoring activities. Back office payment activities are reflected in newly available data on transaction volumes for four types of payment instruments while front office services are associated with the number of ATMs and branch offices. ${ }^{41}$

Survey information from multiple sources indicate that an electronic payment only costs from one-third to one-half as much as its substitute paper-

[^19]based alternative (c.f., Humphrey, Willesson, Lindblom, and Bergendahl, 2003). Thus bank operating costs would be expected to fall as the share of electronic payments in total non-cash payment transactions in 12 European countries rose from 0.43 to 0.79 over 1987-1999. Similarly, ATMs-a cost-efficient way to deliver heavily used cash acquisition and certain other depositor services-have expanded from 49,000 to 205,000 over the same period (a $325 \%$ increase) while the number of branch offices rose from 172,400 to only 172,900 (a rise of only $0.3 \%$ ).

The effect on bank operating cost from these six output characteristics incorporate both scale influences and technical change. A statistical model based on these output characteristics relates operating cost to service delivery and payment levels and mix to determine how changes in these characteristics have affected operating costs across the banking sectors of 12 European countries. It does not matter much whether a composite, translog, or Fourier cost model is used in this analysis.

Overall, bank operating costs may be $\$ 32$ billion lower in 1999 due to the shift to electronic payments and the expanded use of ATMs which permitted banks to conserve on building new branch offices. The estimated $\$ 32$ billion in cost savings is equal to $0.38 \%$ of the 1999 GDP of the 12 European nations in our sample. As a result, banks may have experienced a $7.8 \%$ reduction in an indicator of their real average total (operating plus interest) cost, and consumers will have benefited to an equal degree if all the cost reduction was passed on in the form of price adjustments.

## 7 Bibliography

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

Berger, A., and L. Mester (1997). "Inside the black box: What explains differences in the efficiencies of financial institutions?" Journal of Banking and Finance, 21: 895-947.

Box, G., and D. Cox (1964). "An analysis of transformations," Journal of the Royal Statistical Society, Series B, 26: 211-246.

Carbó Valverde, S., D. Humphrey, and R. Lopez del Paso (2003). "Effects of ATMs and electronic payments on banking costs," Working Paper, Fundacion de las Cajas de Ahorros Confederadas, Madrid, May.

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

Humphrey, D., M. Willesson, T. Lindblom, and G. Bergendahl (2003). "What does it cost to make a payment?" Review of Network Economics, 2: 159-174.

Mitchell, K., and N. 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.

Norges Bank (2002). Annual report on payment systems: 2001, Oslo, July.

Pulley, L., and Y. Braunstein (1992). "A composite cost function for multiproduct firms with an application to economies of scope in banking," Review of Economics and Statistics, 74: 221-230.

Pulley, L., and D. Humphrey (1993). "The role of fixed costs and cost complementarities in determining scope economies and the cost of narrow banking proposals," Journal of Business, 66: 437-462.

Röller, L.-H. (1990). "Proper quadratic cost functions with an application to the Bell System," Review of Economics and Statistics, 72: 202-210.

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

## 8 Appendix A: Parameter Estimates for the Composite Cost Function.

Number of observations $=156 . \log$ likelihood $=478.893$. Standard errors were computed from a heteroscedastic-consistent matrix (Robust-White). Durbin-Watson $=0.378767$.



```
\beta}\mp@subsup{\beta}{11}{}\mathrm{ B11 .171812 30.0887 [.000]
\mu11 M11 -.252346E-04 -7.40451 [.000]
\mu31
\mu61 M51 3.26612 8.08715 [.000]
```

${ }^{*}$ Note: $\delta_{21}, \delta_{41}, \delta_{51}$, and $\mu_{21}, \mu_{41}, \mu_{51}$ (representing the input price/output level interaction terms for paper giros, cards, and ATMs) were all set to zero and not estimated in the final model.


[^0]:    ${ }^{1}$ From smallest to largest in terms of total banking assets in 1999, the 12 countries are: Finland, Norway, Denmark, Sweden, Belgium, Spain, Netherlands, Switzerland, Italy, U.K., France, and Germany.
    ${ }^{2}$ The alternative of using the ratio of operating cost to total cost will not accurately portray how operating expenses have changed in Europe. This is because interest rates,

[^1]:    ${ }^{6}$ While the trend is upward, the number of non-cash transactions per person across countries can be quite different. In 1999, the total number of non-cash transactions per person ranges at the lower end from 42 to 55 payments a year for Italy, Spain, and Switzerland while at the higher end it is 165 to 178 annually for the Netherlands, U.K., Germany, and France. The remaining countries were in the middle with between 116 and 147 payments per person.
    ${ }^{7}$ This is largely due to the fact that electronic payments incur lower labor costs and experience greater scale economies than paper-based transactions. In addition, advances in computer and telecommunications technology over time have lowered the absolute cost of processing electronic payments at all scales of operation.

[^2]:    ${ }^{8}$ During this period, the number of branch offices per 10,000 inhabitants fell somewhat in 10 countries, rose in 2 (Italy and Spain), but on balance was flat overall. Thus the sole reason for the rise in the ATM/branch ratio was the rapid rise in ATMs.
    ${ }^{9}$ In 1999 Spain and Belgium provided 9.8 and 14.2 offices per 10,000 inhabitants while the other 10 countries provided 5.0 or less (the U.K. provided only 1.9). The offset to providing many offices is that there were only 6 workers per banking office in Spain and 5 in Belgium. The other 10 countries all had more than twice as many workers (12) per office while the U.K. had 36.
    ${ }^{10}$ The reduction in bank workers per 10,000 of population over 1987-1999 varied from $-2 \%$ to $-5 \%$ for Spain, Belgium, Italy, and Switzerland at the low end to around $-50 \%$ for Norway and Finland at the other extreme. Of the remaining countries, three experienced $-15 \%$ to $-29 \%$ reductions while three had increases of $6 \%$ to $29 \%$.

[^3]:    ${ }^{11}$ The provision of deposit and loan services not directly associated with the number of payments or ATMs should be associated with the number of banking branches in a country. This is similar to the common assumption made in the literature where the stock of the value of deposits, loans, or assets is assumed to reflect the underlying flow of associated banking services. Indeed, the $\mathrm{R}^{2}$ between the number of (standardized) branch offices and the value of total banking assets or the value of total deposits across our 12 countries and over time is .76 in both cases.
    ${ }^{12}$ To circumvent the impossibility of separating scale effects from technical change with time-series data, it has been common practice to use panel data so that the cross-section component identifies scale while the time-series component identifies technical change.

[^4]:    ${ }^{13}$ EFT-POS terminal availability is associated with the volume of electronic card payments-a variable we already use-and thus is not separately specified in the model. Our 13 year period should mitigate the effect of having initially high investments in electronic payment arrangements that fall to lower levels over time.
    ${ }^{14}$ As illustrated in Pulley and Braunstein (1992), this can occur when one or more outputs is less than $5 \%$ of total output. This occurs for two countries in our sample for ATMs (as a percent of ATMs plus branches in the late 1980s) and for all 12 countries for at least one payment instrument (as a percent of all four instruments in the late 1990s).
    ${ }^{15} \mathrm{~A}$ similar function (CES-quadratic) was used by Röller (1990) to determine scope effects of local and long-distance telephone costs for the Bell System while Pulley and Humphrey (1993) used a composite form to assess the cost effects of separating risky loan assets from deposit liabilities into two separate "banks", funding the former with uninsured CDs and investing the latter in safe assets.

[^5]:    ${ }^{16}$ As is usual, fixed costs are reported as capital consumption expenses rather than the current value of all fixed assets. Depending on ATM ownership and access agreements, in some countries OC will include the capital expense of a bank's own ATMs plus costs incurred by their customers who use other banks' ATMs. This generates some unknown amount of double counting since the revenue received does not reduce the reported cost of the other banks' ATMs.

[^6]:    ${ }^{17}$ It is generally not feasible to estimate both $\alpha_{0}$ and $\beta_{0}$ intercepts. As we are more interested in output quantities than input prices, and on the basis of fit, we set $\beta_{0}=0$ and retain $\alpha_{0}$ in estimation.

[^7]:    ${ }^{18}$ Specifically, the number of each country's banking offices $(B R)$ was adjusted as follows: $B R^{S T D}=B R[(L / B R) / 16.04]$, where $L / B R$ is the observed labor/branch ratio in each country for each year and 16.04 workers per office is the standardized size of each office. This gives the number of standardized, size-adjusted branches $\left(B R^{S T D}\right)$ which is used for each country in the estimations, not $B R$. For example, in one year the U.K. may have had 32.9 workers per branch office (actually, this is the average over 1987-1999). Dividing this value by the French benchmark gives $32.9 / 16.04=2.05$ which effectively doubles the number of "standard" U.K. branches used in the analysis. In contrast, in one year Spain may have had 6.9 workers per branch office (this too is the average over 1987-1999). Dividing this value by the French benchmark gives $6.9 / 16.04=0.43$ which reduces the number of "standard" Spanish offices by close to $60 \%$.
    ${ }^{19}$ This problem exists in our data set. See Footnote 14.

[^8]:    ${ }^{20}$ The new terms are $\ln Q^{*}=\ln Q \cdot Y Q+Z Q, Y Q=(.8 \cdot 2 \pi) /(\max \ln Q-\min \ln Q)$, $Z Q=.2 \pi-\min \ln Q \cdot Y Q$, and $\pi=3.141593 \ldots$, so that $\ln Q^{*}$ is essentially expressed in radians (Mitchell and Onvural, 1996; Berger and Mester, 1997). Our Fourier specification follows Berger and Mester.

[^9]:    ${ }^{21}$ Unit operating cost is the ratio of predicted operating cost to observed asset value and is an indicator of average operating cost. Value data in each country's domestic currency was translated into U.S. dollars at market exchange rates for each year. As our time period starts in 1987, the Euro did not exist until much later and so was not used as the unit of account.
    ${ }^{22}$ The composite cost function was concave and all but two marginal costs (for checks and paper giros) were positive at their mean values. In the final form estimated, the interaction terms between input prices and paper giro, card, and ATMs were set to zero. While the cost concavity condition was not met when these parameters were in the model, the output relationships reported below were unaffected. Estimated parameters are presented in a short Appendix.
    ${ }^{23}$ Bank size on the X -axis is indicated by the natural $\log$ of asset value of each of the 12 countries' banking systems.
    ${ }^{24}$ These ratios are computed as the sum of the predicted value of bank operating expense for 12 countries divided by the sum of the value of total bank assets for the 12 countries.

[^10]:    ${ }^{25}$ Other ways to estimate the cost savings exist but are higher. For example, if operating cost had instead grown at the same rate as total assets (124\%)-which we believe is unlikely, it would have grown to $\$ 364.3$ billion by 1999. The implied cost savings here are $\$ 88.7$ billion or $1.06 \%$ of aggregate GDP. Alternatively, if there were no growth in the 1987 level of operating cost but we apply our -0.30 change in predicted unit cost to this value, the cost savings would be $\$ 48.9$ billion or $0.59 \%$ of aggregate GDP. We feel that the estimate in the text is the most reasonable of the three.
    ${ }^{26}$ Simplifying the composite specification so there is only one electronic payment category (composed of electronic giro and card transactions) and one paper payment category (check and paper giro transactions), yields a very similar scale estimate (0.88).
    ${ }^{27}$ A fixed effects model is not appropriate here. This is because payment transaction volumes, the number of ATMs and branch offices, and input prices are the major reasons why bank operating costs vary across our 12 countries-and these are already specified in the model. Adding 12 country dummy variables to (2)-and dropping the intercept-was highly collinear with our included variables so that the resulting scale and technical change

[^11]:    effects previously noted were dramatically altered (e.g., a negative scale elasticity) with little improvement in explanatory power. In addition, there is little reason for our estimated parameters to vary across countries: payment processing production functions are the same across countries, an ATM is an ATM in all countries, and we have standardized the size of branch offices (see Footnote 18).

[^12]:    ${ }^{28}$ This was pointed out in Baumol, Panzer, and Willig (1982) who propose a measure of average incremental cost instead. However, average cost curves can be determined when all outputs are allowed to vary (as in Figure 1).
    ${ }^{29}$ As in Figure 1, the curves in Figure 3 are cubic splines of the predicted values and illustrate how unit payment operating cost varies by (the log of) the asset size of a country's banking system for each of the three years shown.
    ${ }^{30}$ For example, let P87 (P99) = the true level of payment operating cost in 1987 (1999) and $\mathrm{D}=$ the mean cost level of ATMs, branches, and input prices from the estimated cost function. Our estimated payment operating costs in these two years will be P87+D and P99+D, respectively. To derive an estimate of payment cost savings we could assume that $\mathrm{P} 87+\mathrm{D}$ would have increased at the same rate as did total payment volume $(89 \%)$ if there had been no shift to electronic payments or ATMs over the period. This suggests that payment costs in 1999 could have been (P87+D) x 1.89, rather than its estimated value from the model of $\mathrm{P} 99+\mathrm{D}$. Were it not for the unknown value D , a cost savings estimate could be obtained from (P87+D) x $1.89-(\mathrm{P} 99+\mathrm{D})$. Rearranging, we have (P87 x 1.89) - P99 + (D x 1.89) - D. While the first term (P87 x 1.89) - P99 could represent an estimate of the savings in payment operating costs, the unknown second term ( $\mathrm{D} \times 1.89$ ) - D will overstate this estimate by 0.89D or 0.89 times the mean cost of ATMs, branches, and input prices.

[^13]:    ${ }^{31}$ Specifying one electronic and one paper-based payment category in the composite model yields a payment scale economy of . 18 .

[^14]:    ${ }^{32}$ As explained above, the predicted cost curves in Figure 5 (or 6) are not average cost curves. While the slopes indicate the extent of changes in costs across countries and/or over time, their level includes the mean cost of payments and input prices. Correspondingly, it is not possible to determine the cost savings due to changes in service delivery modes separate from the shift to electronic payments (see Footnote 30).

[^15]:    ${ }^{33}$ With one electronic and one paper-based payment category, the delivery scale economy is 0.71 .
    ${ }^{34}$ As outlined earlier, the four characteristics are: paper-based payments (composed of checks plus paper giro transactions); electronic payments (electronic giro plus card

[^16]:    transactions); and the number of ATMs and (standardized) branch offices.
    ${ }^{35}$ The cost curves for the translog are almost identical to those for the Fourier in Figure 7 and thus are not shown.

[^17]:    ${ }^{36}$ The statistical fit of the three models was similar as all had $\mathrm{R}^{2}$ values for the cost function above 0.95 . Differences in results that exist between the composite and the translog or the Fourier estimates are likely due to the fact that the log of operating cost is being predicted in the latter two models rather than the unlogged value as occurs with the composite form.
    ${ }^{37}$ Adding a linear time dummy to the 4 output characteristic translog (3) and Fourier (4) equations suggested that this variable significantly raises operating cost by about $1 \%$ a year. This anomalous result occurs because the time dummy is highly correlated with the time-series component of our payment and ATM variables.
    ${ }^{38}$ Exogenous technical change is identified through the time dummy variable. As this variable does not interact with any of the other specified "output" or input price variables, its estimated parameter is set to zero and the OC/TA ratios are recalculated. Changes in this second set of ratios between 1987 and 1999 should approximate the cost change from scale and endogenous effects alone. For the same three models, these cost reductions were $-10 \%,-15 \%$, and $-13 \%$. Thus, using a standard cost model, scale effects and endogenous productivity advances appear to dominate the exogenous changes. Had the time dummy been allowed to interact with the other variables-a likely better specification-its approximate separate influence on costs could not be determined.

[^18]:    ${ }^{39}$ While banking mergers in Europe probably led to some reduction in operating cost at these institutions due to scale economies in payment operations, the removal of restrictions on branching within various countries-the most common form of deregulation over this period-likely added to operating costs through short-run overbranching or excess ATM expansion to garner a larger deposit market share for the long run.

[^19]:    ${ }^{40}$ More country-specific information is available when our cost models are applied to panel banking data in particular countries, as has been done for Spain (Carbó Valverde, Humphrey, and Lopez del Paso, 2003). Although more difficult and expensive, payment cost data can be collected directly from banks as has been done in Norway (Norges Bank, 2002). Both of these detailed approaches reinforce the conclusions offered here for Europe as a whole regarding the large savings that can be obtained by greater use of electronic payments.
    ${ }^{41}$ The $R^{2}$ between the number of bank offices and the value of total assets or total deposits are both 0.76 . Thus the number of bank offices closely reflects the stock of assets, deposits, or loans (which are highly correlated with assets or deposits) that are often used in the literature to represent the underlying flow of banking deposit and loan services.

