Cost Savings from Electronic Payments in Europe

March, 2003

Abstract

Electronic payments are considerably cheaper than their paper-based alternatives. As the share of electronic payments in 12 European countries rose from .43 in 1987 to .79 in 1999, bank payment costs may be $18 billion lower than they otherwise might have been. While scale economies serve to reduce costs of an expanding volume of electronic payments, they work in reverse for a contracting volume of paper-based transactions. Another $8 billion in savings is attributed to expanded use of ATMs (rather than branch offices) to deliver banking services. Our results are reasonably robust to the form of cost function estimated–composite, Fourier, or translog.

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

Corresponding Author:
Dave Humphrey, Department of Finance, Florida State University, Tallahassee, FL 32306-1042; (850) 644-7899; dhumphr@garnet.acns.fsu.edu

1 Introduction.

Most people pay little attention to their nation's payment system. Their only concern is the bank fees they may pay to "use their own money". When payment fees are related to the underlying differential cost of electronic versus paper-based transactions, a nation's payment costs can fall in real terms. As a country shifts from an all paper-based to an all electronic-based payment system, annual savings of 1-2% 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 to 3% of GDP. No time-series of bank level or 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. 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 relate the annual operating (not total) cost of each of 12 European countries' banking sectors over 1987-1999 to the total annual number of check, paper giro, electronic giro, and card transactions in each country along with the number of ATMs and (standardized, size-adjusted) branch offices while controlling for differences in input prices across countries. As the vast majority of banking 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 separate payment costs from ATM and branch service delivery expenses over time.

The usual approach for identifying cost savings specifies (disembodied) technical change as a time-specific indicator variable added to a standard total cost function or, less often, determines (embodied) technical change from changes in the cost share or price of certain inputs. Our approach is quite different. We directly measure six main indicators of the direct effect of technical change rather than combine them all into a single set of linear, quadratic, or time-specific dummy variables. In this manner we are able to distinguish between payment-related and service delivery-related aspects of technical change as well as scale effects of their operation.

In what follows, we illustrate in Section 2 how changes in bank operating costs have been affected by the use of different payment instruments and service delivery methods over 1987-1999 for 12 European countries.[1] Our "output characteristic" 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 estimated operating cost curves with those from more commonly used translog and Fourier cost models.

Estimates of the effect of expanded electronic payments on banking costs for Europe are reported in Section 4. The 12 European countries may have saved some $18 billion by shifting from paper-based to electronic payments and perhaps an additional $8 billion in service delivery costs by adopting ATMs. Together, these two developments are estimated to have saved the European banking system about $26 billion (or 25 billion Euros). Countries that have progressed further in shifting to electronic from paper-based payments, and to ATMs from branch offices for dispensing cash and other services, have benefited the most from the associated reduction in banking industry operating expenses. Our main results are summarized in Section 5 which concludes the paper. It is likely that the same payment use and service delivery trends shown to have benefited Europe may also apply to other nations 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 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

[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, which affect total costs but not operating expenses, are not the same across countries nor constant over time.

[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 OC/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).

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%

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

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.

No country has 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 electronic (electronic giro and debit and credit card transactions) and two that are paper-based (check and paper-based giro payments).[5] Overall, the number of non-cash transactions in Europe rose from 24.6 billion in 1987 to 46.5 billion in 1999, a 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 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 .43 to .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

[4] While slight increases were experienced by Switzerland and Finland (4% to 8%) between 1987 and 1999, this is misleading. The ratio OC/TA initially rose after 1987 in these two countries, reached a peak in the early to mid-1990s, and then fell by -21% and -42%, respectively, by 1999.

[5] Giro transactions include direct debits and credit transfers.

[6] While the trend is upward, the levels of non-cash use 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.

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 reduction in unit operating expense shown in Table 1 (-24%) 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. Other banking functions, such as loan origination and monitoring, liquidity management, and off-balance-sheet activities generate little labor or capital cost but, of course, bring in the majority of 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 by 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.[8] 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%), it is likely that the reduction in bank operating expenses outlined above may be related in part to the rise in ATMs (which increased by 318%).

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] Re-

[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.

[8] During this period, the number of branch offices per 10,000 inhabitants fell somewhat in 10 countries but rose in 2 (Italy and Spain). Thus the primary 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

gardless 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 effects on bank cost from the shift to electronic payments and ATM use in Europe. This requires a statistical analysis which 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.

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 to assume that these service flows are proportional to the value of the stock of bank deposits, securities, and loans 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 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

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 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%.

with the size of a bank and its labor, capital, and materials operating cost. When mix is constant and technology is unchanged, levels of these activities reflect bank size from which scale economies can be determined. 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.[11]

Paper-based and electronic payment transactions are jointly processed while service delivery is jointly 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.

Panel data on 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 are used in a non-linear, functionally separable, 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.[12]

The composite model 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. This is because the level of banking output in a composite function is 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 small–than if the log of output is related to the log of operating cost.[13] As well, by specifying the log of input prices,

[11] To circumvent the impossibility of separating scale effects from technical change with only 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.

[12] 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.

[13] As illustrated in Pulley and Braunstein (1992), this can occur when one or more

it is possible to impose the theoretical condition of linear homogeneity in input prices in estimation.[14]

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

$$\begin{aligned} OC^{(\phi)} &= f^{(\phi)}(\underline{Q}, \underline{\ln P}) \\ &= \{[\alpha_0 + \sum_{i=1}^{6} \alpha_i Q'_i + 1/2 \sum_{i=1}^{6} \sum_{i=1}^{6} \alpha_{ij} Q'_i Q'_j] \bullet \exp[\beta_0 \\ &\quad + \sum_{k=1}^{2} \beta_k \ln P_k + 1/2 \sum_{k=1}^{2} \sum_{m=1}^{2} \beta_{k,m} \ln P_k \ln P_m]\}^{(\phi)} \qquad (1) \\ S_k &= \beta_k + \sum_{m=1}^{2} \beta_{k,m} \ln P_k \end{aligned}$$

where:

OC = total operating expenses, composed of all labor, capital, and materials costs (but no interest expenses);

$Q'_{i,j}$ ij = six output characteristics composed of four payment processing alternatives–the number of checks ($CHECK$), paper or electronic giro payments ($PGIRO, EGIRO$), and debit and credit card transactions ($CARD$)– along with two service delivery alternatives–the number of automated teller machines (ATM) and the number of standardized, size-adjusted, bank branches (BR). In (1), $Q' = Q - 1$;

$P_{k,m}$ k, m = two input prices referring to 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 = the cost share for the labor input (the capital/materials input share is deleted to avoid singularity).

---

outputs is less than 10% of total output. This occurs for two countries in our sample for ATMs (as a percent of ATMs plus branches) and for all 12 countries for at least one payment instrument (as a percent of all four instruments) for some years.

[14] 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.

It is expected that operating costs not directly associated with the type of payment or mode 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 $GM^{\phi-1}$be the geometric mean of operating cost OC, then the separable quadratic form of the composite model is estimated from the "pseudo model" (2):[15]

$$
\begin{aligned}
D &= [-(OC^{(\phi)}/GM^{\phi-1}) + f^{(\phi)}(\underline{Q}, \underline{\ln P})/GM^{\phi-1} \\
&= [-\{(OC^{\phi} - 1)/\phi GM^{\phi-1}\} + (\{[\alpha_0 + \sum_{i=1}^{6} \alpha_i Q'_i \\
&\quad +1/2 \sum_{i=1}^{6} \sum_{i=1}^{6} \alpha_{ij} Q'_i Q'_j] \bullet \exp[\beta_0 + \sum_{k=1}^{2} \beta_k \ln P_k \qquad (2) \\
&\quad +1/2 \sum_{k=1}^{2} \sum_{m=1}^{2} \beta_{k,m} \ln P_k \ln P_m]\}^{\phi} - 1)/\phi GM^{\phi-1} \\
S_k &= \beta_k + \sum_{m=1}^{2} \beta_{k,m} \ln P_k
\end{aligned}
$$

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 comparable. Otherwise, differences in the 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 cross-country estimation since we specify the number of branch offices as an output characteristic. France,

[15] It is generally not feasible to estimate both α_0 and β_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 α_0 in estimation.

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.[16]

3.2 Alternative Translog and Fourier Cost Functions.

To illustrate the robustness of our results, we also estimate translog and Fourier cost functions. A translog function may generate biased results, compared to the composite form, when levels of some outputs are small and outputs are specified in logs.[17] Even so, as these two additional functions are often used in cost analyses, it is useful to compare their results with those from our composite form.

The translog cost function (3) is estimated jointly with n-1 cost share equations:

$$\begin{aligned} \ln OC &= \alpha_0 + \sum_{i=1}^{6} \alpha_i \ln Q_i + 1/2 \sum_{i=1}^{6} \sum_{i=1}^{6} \alpha_{ij} \ln Q_i \ln Q_j + \sum_{i=1}^{6} \sum_{k=1}^{2} \delta_{i,k} \\ & \quad \ln Q_i \ln P_k + \sum_{k=1}^{2} \beta_k \ln P_k + 1/2 \sum_{k=1}^{2} \sum_{m=1}^{2} \beta_{k,m} \ln P_k \ln P_m \qquad (3) \\ S_k &= \beta_k + \sum_{m=1}^{2} \beta_{k,m} \ln P_k + \sum_{i=1}^{6} \delta_{i,k} \ln Q_i \end{aligned}$$

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

[16] Specifically, the number each of country's banking offices (BR) was adjusted as follows: $(BR^{STD}) = BR[(L/BR)/16.04]$, where L/BR 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 (BR^{STD}) which is used for each country in the estimations, not BR. For example, in one year the U.K. 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 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 = .43 which reduces the number of "standard" Spanish offices by close to 60%.

[17] This problem exists in our data set. See Footnote 13.

orthogonal over the $[0, 2\pi]$ interval. The Fourier function (4) is estimated jointly with the cost shares:[18]

$$\begin{aligned}
\ln TC &= \text{Translog Cost Function} \\
&+ \sum_{n=1}^{6} [\tau_n \cos(\ln Q_n^*) + \omega_n \cos(\ln Q_n^*)] \\
&+ \sum_{n=1}^{6} \sum_{q=1}^{6} [\tau_{nq} \cos(\ln Q_n^* + \ln Q_q^*) + \omega_{nq} \sin(\ln Q_n^* + \ln Q_q^*)] \\
&+ \sum_{n=1}^{6} [\tau_{nnn} \cos(\ln Q_n^* + \ln Q_n^* + \ln Q_n^*) \qquad (4) \\
&+ \omega_{nnn} \sin(\ln Q_n^* + \ln Q_n^* + \ln Q_n^*)] \\
S_k &= \beta_k + \sum_{m=1}^{2} \beta_{k,m} \ln P_k + \sum_{i=1}^{6} \delta_{i,k} \ln Q_i
\end{aligned}$$

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.[19] 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 ϕ in the composite form is .16, 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

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

[19]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 and so was not used as the unit of account.

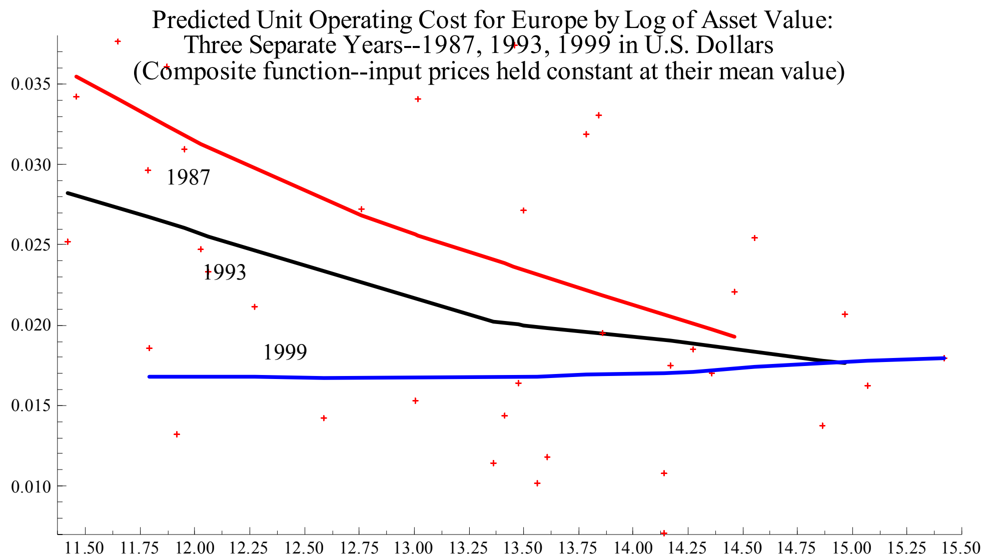

Figure 1:

in absolutes and prices in logs (when $\phi = 1.0$).[20] Even so, the estimated model is significantly different from either of these alternatives since ϕ is significantly different from zero or one.

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 three years shown.[21] Thus Figure 1 illustrates both the scale effect (cross-country slope) as well as technical change (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 .023, .019, and .018, respectively, and are close to the OC/TA ratios computed using observed data in row 1 of Table 1 above. The reduction in this ratio between 1987 and 1999 was -22%.[22] The overall operating cost

[20] The estimated parameters of the composite function underlying this figure are presented in a short Appendix.

[21] Bank size on the X-axis is indicated by the natural log of asset value of each of the 12 countries' banking systems.

[22] 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. Thus the set of 12 countries is treated as if it were a single entity. If we had instead used

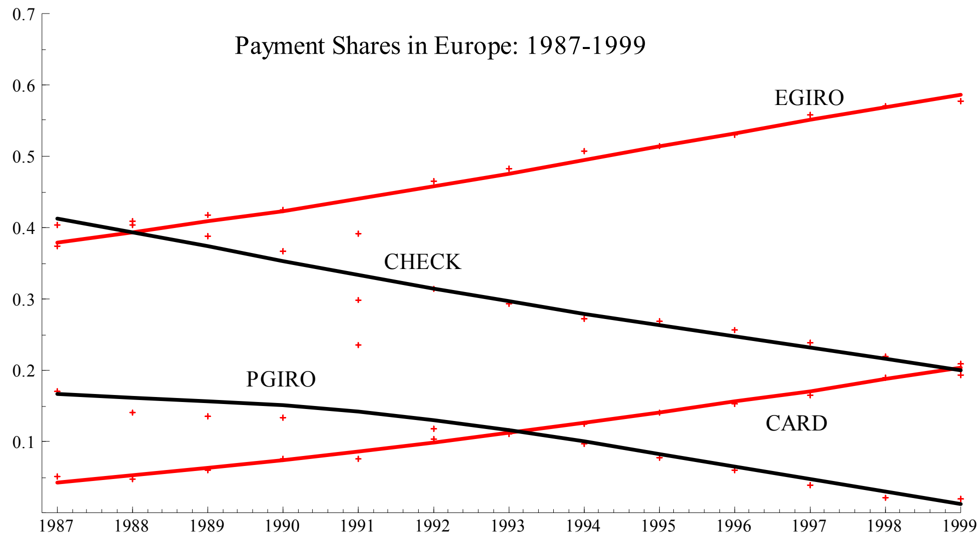

Figure 2:

scale economy value during this period is .89 so that a 10% expansion in all six output characteristics is associated with only a 8.9% rise in operating cost (so average operating cost would fall). 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.

Since total bank operating cost for our 12 countries in 1987 was $162.9 billion (or around 155.1 billion Euros), this suggests that operating expenses could have been $36 billion (.22 times $162.9 billion) higher in 1999 than they were if there were no scale or technical change effects to reduce operating costs from their ratio to assets in 1987. Put differently, unit operating cost associated with changes in payment volume and service delivery levels and mix across our 12 countries appears to have fallen by some $2.8 billion a year in Europe.

---

a simple average of each of 12 separate country ratios, the result would have been .027, .021, and .017 for these three years and the change between 1987 and 1999 would have been -36%. In either case, inflation affects the numerator and denominator of the ratio but, since the rates may not be the same, probably is not fully canceled out.

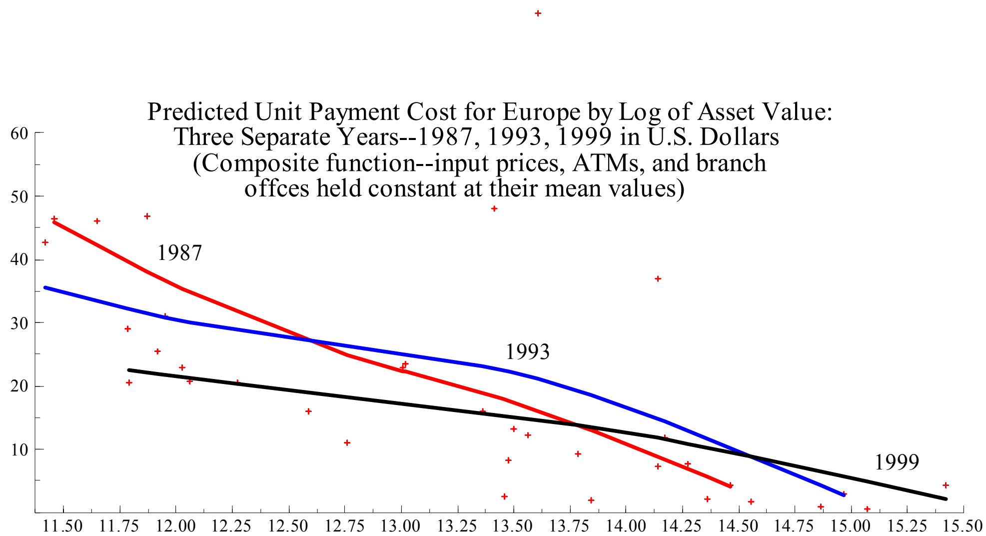

Figure 3:

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 .40 in 1987 to .19 in 1999, while the reduction in paper giro payments was almost ninety percent, falling from .17 to .02 over the period. The share of electronic giro payments rose by fifty-four percent (from .37 in 1987 to .58 in 1999) while the card share rose by over three hundred percent (.05 to .21). Overall, the share of electronic payments rose from .43 to .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.[23]

[23] Predicted payment operating costs were obtained from an estimated composite function where the levels and mix of the four payment instruments were allowed to vary each year across the 12 countries but input prices and the number of ATMs and branch offices were held constant at their mean values in the panel data set. The resulting payment operating cost estimates were then divided by the observed level of the sum of all four instruments, giving an indicator of unit payment costs.

It is emphasized that these unit cost indicators are not average cost curves. Strictly speaking, it is not possible to obtain accurate average cost estimates from a multiple output cost function.[24] Our predicted payment operating costs are determined by evaluating an estimated (composite) cost function with observed payment transaction volumes for each of the four payment instruments while holding constant the number of ATMs, the number of (adjusted) branch offices, and input prices at their sample mean values over 1987-1999. The predicted cost estimates are then divided by the sum of the observed transaction volumes for the four payment instruments. 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 volumes affects operating cost over time and across countries. Consequently, the dollar costs shown on the Y-axis necessarily includes more than just payment costs and so is not a measure of the level of average payment cost alone. However, 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.[25] As seen, unit payment costs are lower for countries with larger banking systems and appear to fall over time. Although we do not have accurate information on the level of average payment costs (as just explained), it is still possible to determine approximately the value of the cost savings from the change in payment levels, mix, and technical change. From 1987 to 1999, the savings is estimated to be $17.6 billion. If spread evenly over 12 countries, this suggests savings of some $1.5 billion each over a 13 year period.[26]

[24] This was pointed out in Baumol, Panzer, and Willig (1982) who propose a measure of average incremental cost instead.

[25] 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) a country's banking system asset size for each of the three years shown.

[26] The average predicted payment cost in 1987, which includes the mean cost of ATMs, branches, and input prices, is $16.6 billion. In determining earlier the predicted change in operating cost over 1987-1999, the 1987 operating cost/total asset ratio was compared to its value in 1999. If we keep the ratio of payment cost to assets constant at its 1987 level, the 124% rise in asset value over 13 years implies that payment costs could have been $37.2 billion (=2.24 x $16.6 billion) in 1999. The average predicted payment cost in 1999 is $19.6 billion and includes the same mean cost effects from ATMs, branches, and input prices as used for the 1987 estimate. Subtracting this value from the projected cost of $37.2 billion yields an apparent cost savings of $17.6 billion. This suggests that predicted payment costs for 1999 could have been 90% higher than they were. Expanding payment costs using the 89% rise in payment volume, rather than the 124% rise in asset value, gives the savings as $11.8 billion. This latter figure assumes no inflation of payment costs while the former, which includes the inflation of asset values, presumes that payment

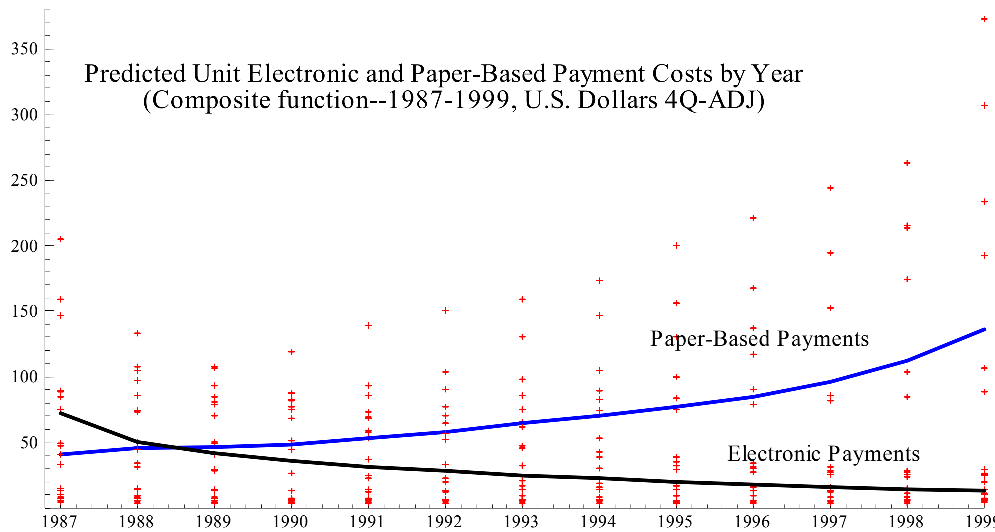

Figure 4:

Not all payment costs are falling. Indeed, the reduction in unit payment expense seen 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 the predicted costs include the mean values of the cost of ATMs, branch offices, and input prices. In addition, the predicted costs shown 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).

transactions are directly related to the level of banking assets. Neither assumption is fully correct. Evaluating a cost function with mean values of input prices overstates actual 1987 payment costs and understates their actual 1999 value while card transactions have increasingly replaced cash payments at the point-of-sale and so are not in a fixed relationship to bank asset value.

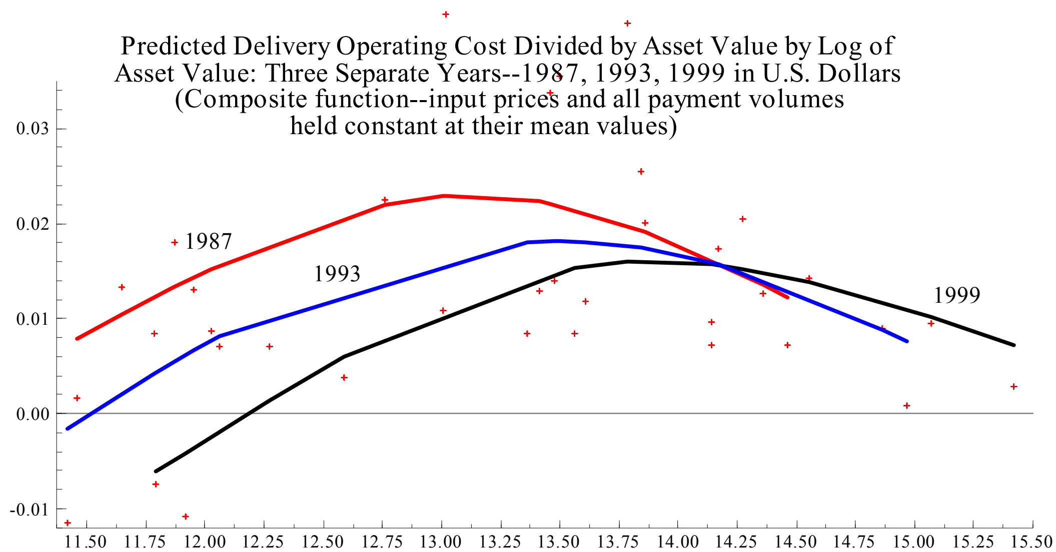

Figure 5:

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 since 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. Across the 12 countries, service delivery expenses first rise and then fall as a ratio to asset value. The predictive accuracy here is not very good since, for banking systems in the smallest countries, predicted delivery costs as a ratio to asset value are seen to be negative in two out of the three years shown. 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 contrast, the number of (actual, not standardized) branch offices rose by only 0.3%, from 172,400 in 1987

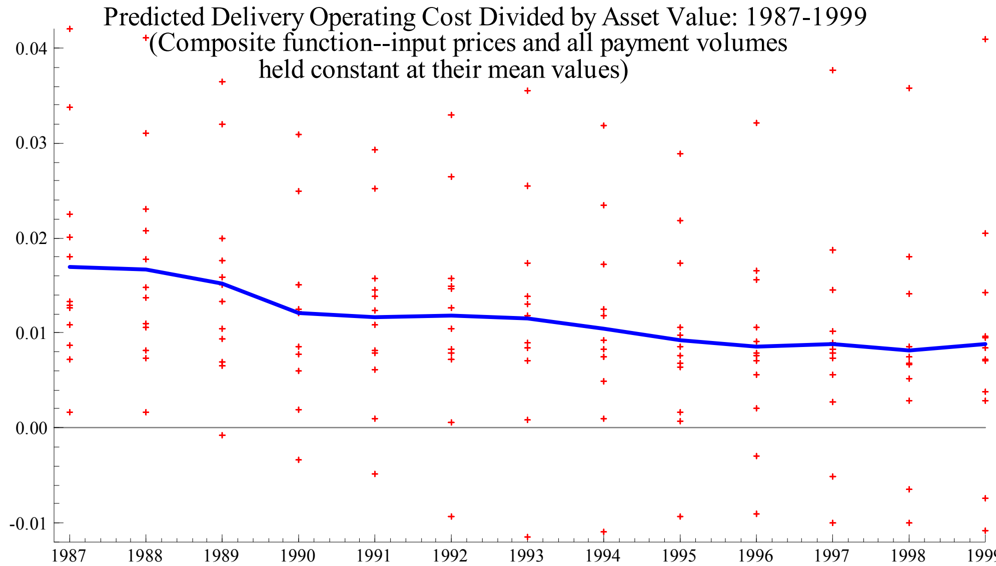

Figure 6:

to 172,900 in 1999. The reduction in predicted delivery costs as a ratio to asset value is seen more clearly by year in Figure 6. The estimated savings in delivery cost from substituting ATMs for traditional branch offices is $8.5 billion over 1987-1999.[27]

4.4 Translog and Fourier Function Results.

The predicted unit operating cost curves using either the translog (3) or Fourier (4) cost function equations look almost identical so only the Fourier

[27] The average predicted delivery cost in 1987, which includes the mean cost of all payment instruments and input prices, is $10.5 billion while the increase in total banking asset value over 1987-1999 was 124%. If the mix of ATMs and branch offices had stayed the same as it was in 1987, and if the number of ATMs and branches grew in the same proportion as did banking assets over 1987-1999, and if we continue to hold the volume of payment transactions and input prices constant at their mean values, then the projected delivery cost in 1999 would be $23.5 billion (=2.24 x $10.5 billion). The average predicted delivery cost in 1999 is $15.0 billion and includes the same mean cost effects from payment transactions and input prices as is used for the 1987 estimate. Subtracting this value from the projected cost of $23.5 billion yields an apparent cost savings of $8.5 billion. This suggests that our predicted delivery costs for 1999 could have been 57% higher if the mix of ATMs and branch offices–but not their number–had remained as it was in 1987.

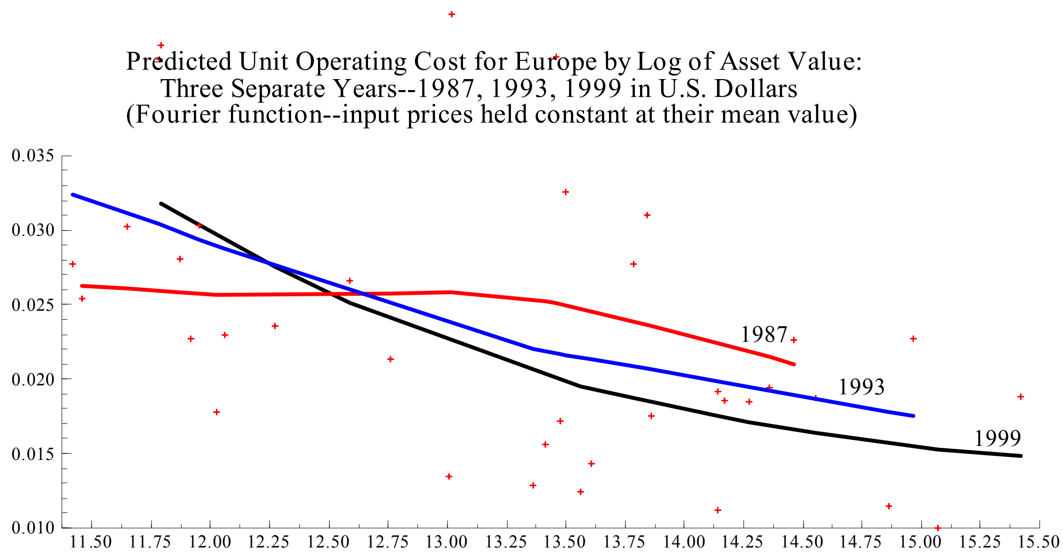

Figure 7:

results are shown in Figure 7. Since the number of paper giro transactions are zero for some countries toward the end of our time period, neither the translog nor the Fourier functions could be estimated until the number of paper and electronic giro payments were aggregated (so positive values exist for all five (not six) output characteristics). Doing the same aggregation and reestimating the composite form does not importantly alter the predicted unit operating cost curves shown in Figure 1. Thus, we can contrast the predicted values for the composite form in Figure 1 with those shown here for the Fourier form in Figure 7.

Overall, all three cost functions (composite, translog, and Fourier) give very similar mean values for predicted unit operating costs for 1987, 1993, and 1999. As a result the estimated cost savings over time from shifting to electronic payments, altering the mix of service delivery methods, and the scale effects from expanded payment volumes would be similar. From this perspective, our conclusions above drawn from the composite form are seemingly robust to the cost function form specified. However, there are some differences between Figures 1 and 7 in terms of the slope of the estimated curves for 1987 and 1999. While the composite form suggests that there is little difference in unit operating costs between countries with small or large banking systems, the Fourier (and translog) forms suggest that larger countries still have an apparent cost advantage.

5 Summary and Conclusions.

6 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.

Carbo, S., D. Humphrey, and R. Lopez del Paso (2001). "Future use of cash and card payments in Spain", Working Paper, Florida State University, Tallahassee, FL.

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*, (forthcoming).

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.

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.

7 Appendix A: Parameter Estimates for the Composite Cost Function.

Number of observations = 156. Log likelihood = 437.148. Standard Errors computed from heteroscedastic-consistent matrix (Robust-White). Durbin-Watson = 1.95306. Cost function concavity condition met and all six estimated marginal costs were positive at their mean values.

Parameter	Estimate:	t-statistic:
ϕPHI	.161600	6.82262
α_0A0	-189.556	-1.20370
α_1A1	-.614657	-.675965
α_{2P}A2P	-.660369	-1.22050
α_{2E}A2E	-.474660E-02	-.020513
α_3A3	.472959	1.01389
α_4A4	12270.4	.199856
α_5A5	.151519E+07	3.30331
α_{11}A11	-.948424E-03	-1.03751
α_{11P}A22P	-.668608E-03	-.190355
α_{22E}A22E	.552645E-03	1.45809
α_{33}A33	.286311E-02	1.80642
α_{44}A44	.141775E+08	2.72795
α_{55}A55	-.882883E+08	-2.88707
α_{12P}A12P	-.014199	-2.58463
α_{12E}A12E	-.263694E-02	-1.60629
α_{13}A13	-.172982E-02	-.860038
α_{14}A14	531.862	2.37510
α_{15}A15	393.232	1.71106
α_{2P2E}A2P2E	-.372769E-03	-.437776
α_{23P}A23P	-.234757E-02	-.578621
α_{23E}A23E	-.208086E-02	-1.46344
α_{24P}A24P	191.991	1.10306
α_{24E}A24E	67.5139	.925551
α_{25P}A25P	1091.23	2.54514
α_{25E}A25E	121.988	1.25112
α_{34}A34	-250.363	-.856234
α_{35}A35	577.234	1.60791
α_{45}A45	-.101762E+08	-.827220
α_{51}D51	17.5369	7.3914
β_1B1	-.960841	-17.5797

β_{11}B11	.168117	32.6467