Payment Scale Economies, Competition, and Pricing

ECB-DNB Conference on Retail Payments

May, 2009

David Humphrey
Florida State University, USA

Main Points

1. Traditional banking scale economies are small (cost elasticity $\approx$ .95) when total cost (a flow) is regressed on assets/loans (a stock). Regressing operating cost (a flow) on payment transactions (a flow) shows large scale effects (at .27).

   Doubling payment volume raises payment operating cost by only 27%

2. Measures of banking competition, which should be positively correlated with one another, are often negatively related (giving inconsistent results).

   Separating loan-deposit rate spread competition from non-interest income (and payment) competition using a revenue frontier yields relatively small differences across European countries (in contrast to EC Report, 2007).

3. Transaction-based pricing of payment services would improve consumer choice. It also "automatically" raises revenues to cover costs as payment volume expands even as unit costs fall, benefiting banks. Finally, it would permit a clearer assessment of price competition in payment services.

Percent Changes: 11 European Countries 1987-2004

	Operating Cost (2004, Mil PPP)	OC/TA	Point of Sale	Bill Payment	ATMs	Branches
France	$82,850	.02%	78%	185%	280%	1.4%
Germany	77,247	-40	501	115	601	14
U.K.	63,972	-52	117	214	160	-25
Italy	50,204	-29	121	117	809	133
Netherlands	34,157	-33	330	128	1,593	-50
Spain	32,120	-50	714	390	858	22
Belgium	12,070	-23	136	98	802	-48
Sweden	5,637	-38	685	8	70	-33
Denmark	4,112	-39	206	333	522	-38
Finland	2,783	-59	1,057	136	11	-46
Norway	2,160	-60	757	67	70	-38
All Countries Together:		-34%	140%	151%	434%	9.8%

If delete Italy, number of branches falls by 2% over time period.

Usual Approach to Scale Measurement: Flow to Stock

Typical banking cost function relates <u>flow</u> of total cost (TC) to <u>stock</u> of assets:

TC = f (value of <u>stock</u> of loan, security, balance sheet assets; input prices)

Total cost = operating costs (labor, physical capital) + interest expenses:

Average operating costs are main source of scale economies
(spreading labor and capital expenses over a larger banking "output")

Average interest expenses vary by deposit/liability composition
(little variation by bank size or "output" level)

Operating costs reported separately from interest expenses so don't need
to separate using covariation with different "outputs" in a regression.

Our Approach to Scale Measurement: Flow to Flow

Our approach captures better measurable physical aspects of banking "output".

Relate operating cost <u>flow</u> (OC) to <u>flow</u> of payments and <u>stock</u> of ATMs, BR:

$OC = f$ *(POS volume; bill payment volume; ATMs; BR; input prices)*

POS volume = number of debit card + check transactions (declining)
Bill payment volume = number of electronic & paper giros (declining)

More accurate local identification of scale economies than traditional bank cost function specification since cost drivers are directly specified.
(Bolt & Humphrey, 2007)

Translog/Fourier Cost Function

$$\ln OC = \alpha_0 + \sum_{i=1}^{4} \alpha_i \ln Q_i + 1/2 \sum_{i=1}^{4} \sum_{j=1}^{4} \alpha_{ij} \ln Q_i \ln Q_j + \sum_{i=1}^{4} \sum_{k=1}^{2} \delta_{ik}$$

$$\ln Q_i \ln P_k + \sum_{k=1}^{2} \beta_k \ln P_k + 1/2 \sum_{k=1}^{2} \sum_{m=1}^{2} \beta_{km} \ln P_k \ln P_m$$

$$+ \sum_{n=1}^{4} [\tau_n \cos(\ln Q_n^*) + \omega_n \sin(\ln Q_n^*)]$$

$$+ \sum_{n=1}^{4} \sum_{q=n}^{4} [\tau_{nq} \cos(\ln Q_n^* + \ln Q_q^*) + \omega_{nq} \sin(\ln Q_n^* + \ln Q_q^*)]$$

$$+ \sum_{n=1}^{4} [\tau_{nnn} \cos(\ln Q_n^* + \ln Q_n^* + \ln Q_n^*)$$

$$+ \omega_{nnn} \sin(\ln Q_n^* + \ln Q_n^* + \ln Q_n^*)]$$

$$S_k = \beta_k + \sum_{m=1}^{2} \beta_{km} \ln P_m + \sum_{i=1}^{4} \delta_{ik} \ln Q_i$$

Translog cost function = rows 1 and 2, plus cost share equation
Fourier cost function   = all rows, plus cost share equation
(very little difference in results, so focus on simpler translog form)

Translog: Predicted Unit Payment Cost
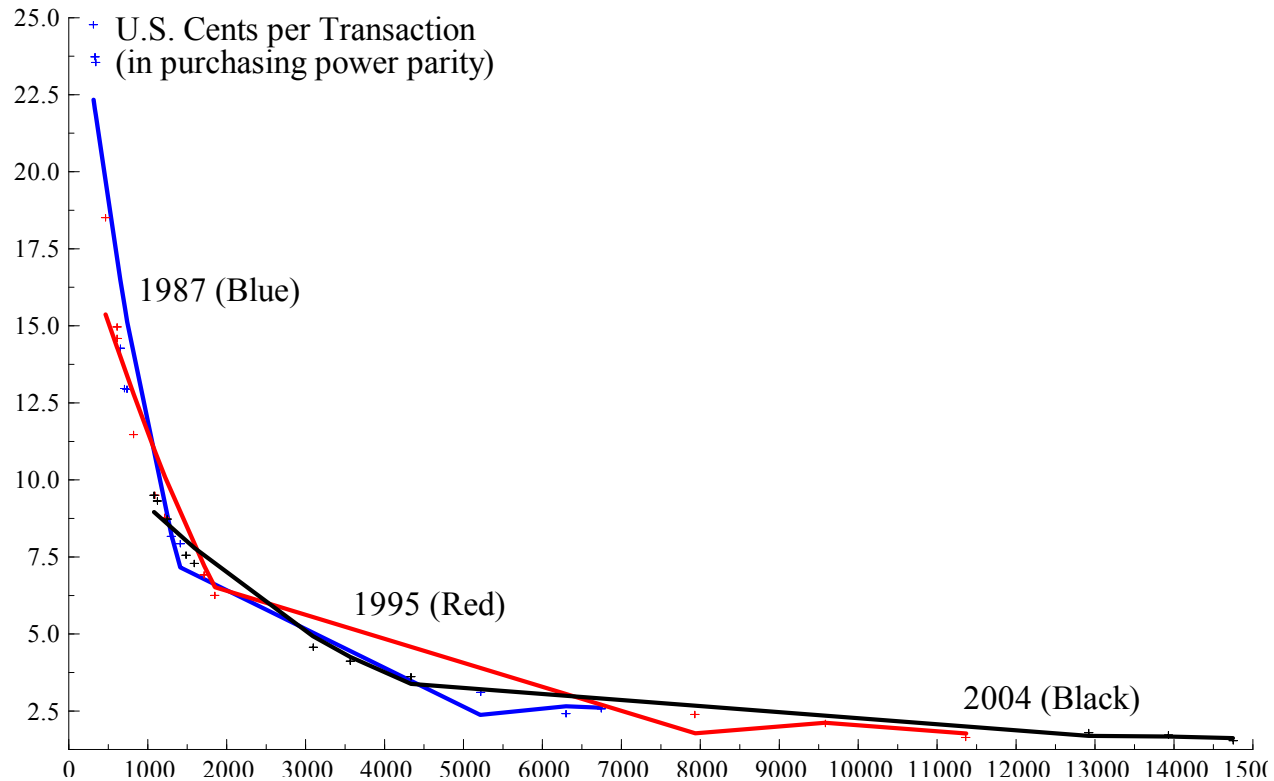
Slopes indicate scale effect. <u>Not</u> an average cost curve since level of curve necessarily includes mean average cost of ATMs and branch operations.

Average incremental cost (Baumol, Panzar, and Willig, 1982) has same problem.

Fourier: Predicted Unit Payment Cost

U.S. Cents per Transaction
(in purchasing power parity)

1987 (Blue)

1995 (Red)

2004 (Black)
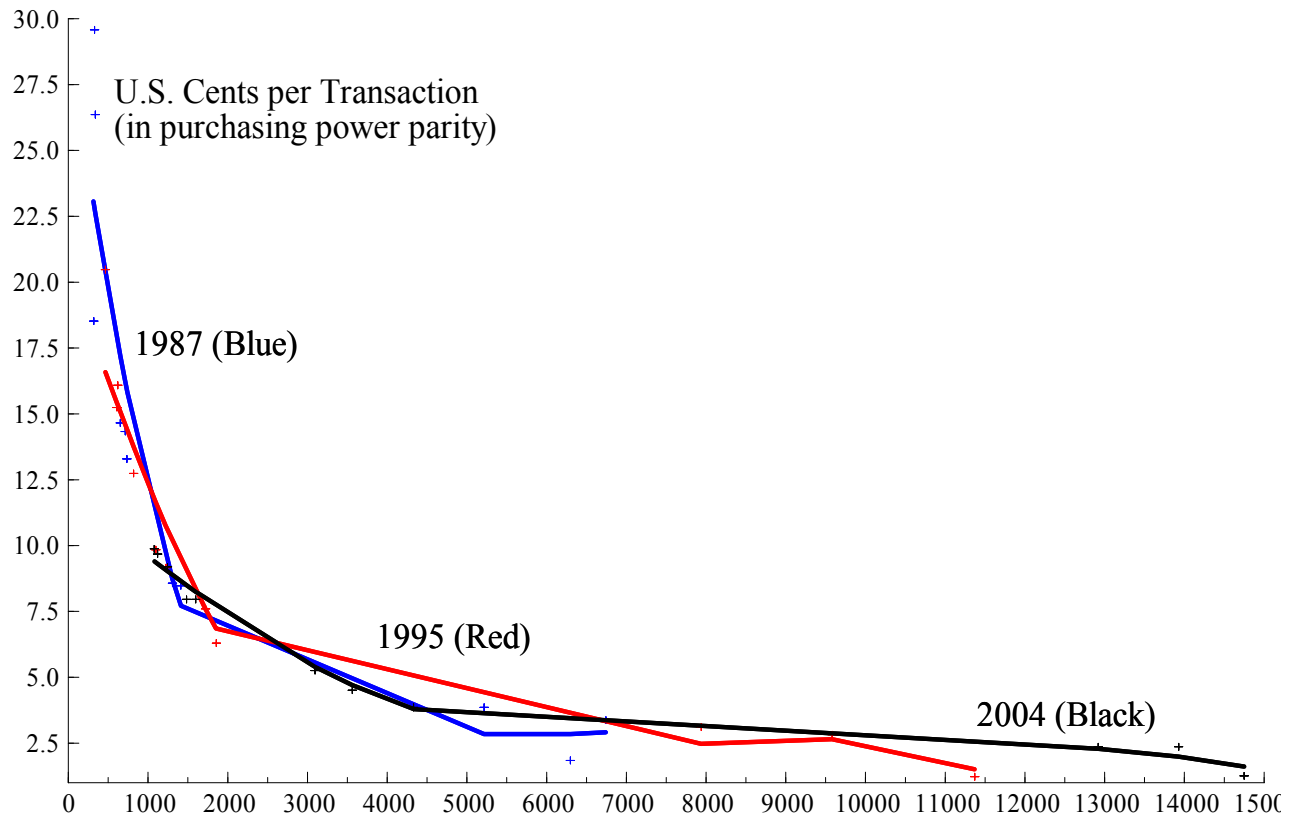
Little difference between translog and Fourier functional forms.

Translog Payment and Service Delivery Scale Economies

|	Payment Volume 2004	Average Payment SCE	Point of Sale SCE	Bill Payment SCE	ATM SCE	Branch SCE
Germany	14,748	.23	.06	.17	.22	.59
France	13,926	.30	.08	.22	.31	.36
U.K.	12,919	.35	.11	.24	.36	.27
Spain	4,335	.30	.10	.20	.23	.48
Netherlands	3,563	.17	.09	.09	.24	.65
Italy	3,094	.21	.05	.16	.17	.62
Belgium	1,594	.20	.10	.10	.26	.59
Sweden	1,488	.33	.18	.15	.39	.37
Finland	1,244	.35	.19	.16	.40	.34
Norway	1,117	.34	.19	.15	.40	.34
Denmark	1,081	.24	.12	.12	.28	.52
Average	5,374	.27	.11	.16	.30	.47

Comparison: Non-cash transaction volume in 2004 totals 59.1 billion for Europe versus 84.5 billion in the U.S. Europe uses more cash than U.S.

Other Payment Scale Economy Estimates (Different Methods)

	Cash	Card
Norway 1994-2001		
d(payment costs)/d(volume)		0.43
Netherlands 2002		
Marginal Cost/Average Cost	0.37	0.39
Belgium 2003		
Marginal Cost/Average Cost	0.25	0.39
U.S. 2005 (approximate)		
d(payment costs)/d(volume)		0.31 – 0.39
Netherlands 1997-2005		
Bank Data, Econometric Model		0.27 – 0.31

Sources: Gresvik and Øwre (2002); Brits and Winder (2005), Table 4.3; Quaden (2005), Table 3; First Annapolis Consulting (2006); Bolt and Humphrey (2008b).

Translog: Predicted Unit Delivery Cost
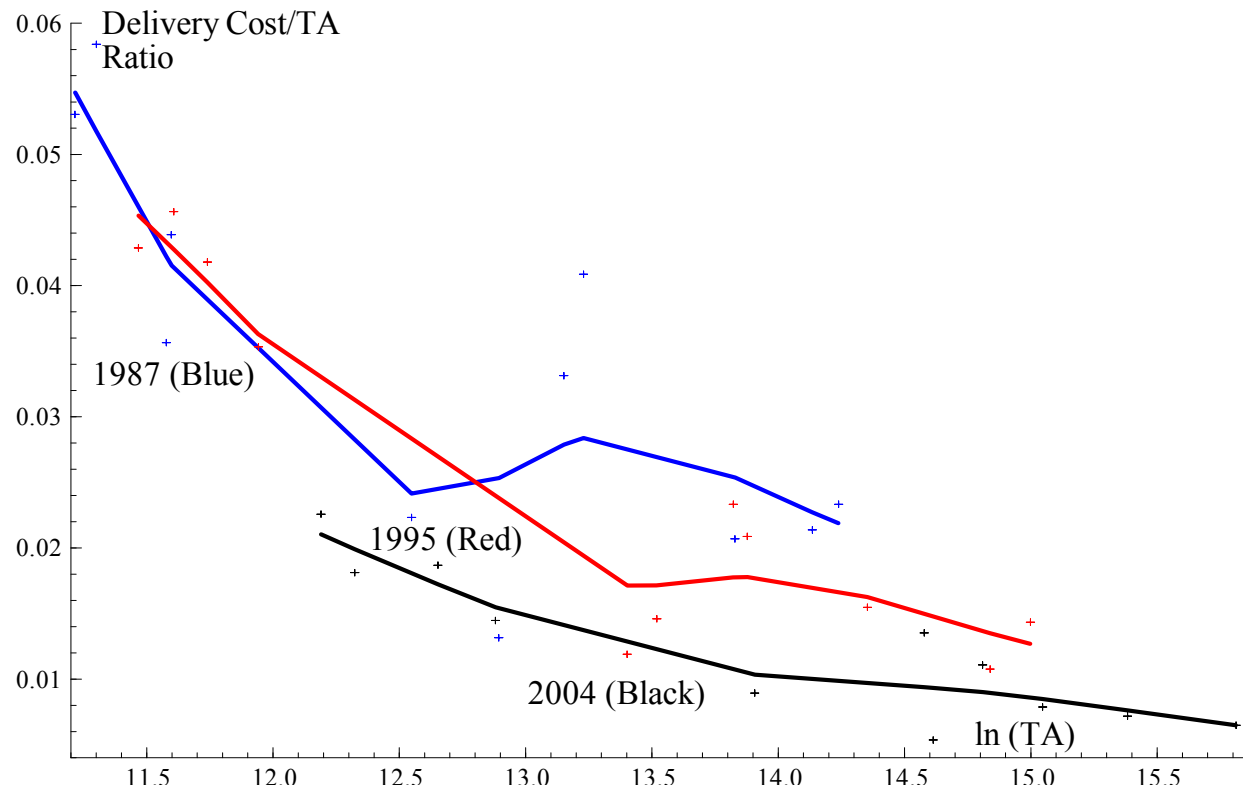
Conclude:

Operating cost <u>shifts</u> over time due to replacement of branches with ATMs;
Scale economies exist for both payment and service delivery operations.

Total "Simple Sum" Versus Realized Scale Economies

	Total SCE	Realized SCE	
Germany	1.05	.31	Standard procedure has been to add up
France	.97	.47	individual SCEs to obtain a total SCE.
U.K.	.99	.54	
Spain	1.01	.45	This is only appropriate when all the
Netherlands	1.06	.24	banking "outputs" experience similar
Italy	.99	.30	rates of growth.
Belgium	1.05	.26	
Sweden	1.09	.21	Since ATMs expanded by 434% while
Finland	1.09	.20	payment transactions expanded by 140% to
Norway	1.08	.23	151%, but branches only rose by 9.8%, a
Denmark	1.04	.37	weighted relationship is needed.
Average	1.04	.40	Our weighted relationship gives us a
			realized SCE.

A "simple sum" of individual SCEs suggests constant returns/slight diseconomies to scale. Adjusting for unequal changes in banking "output" gives SCEs that were actually realized, showing strong operating cost scale economies for Europe.

Competition

EC report (2007) outlined how specific bank fees for payment and other services differ markedly across Europe (but countries rarely identified; confidential).

Among others, the report compared:
- account maintenance fees
- consumer transaction fees
- market rate/deposit interest rate spread
- merchant card interchange fees
- account switching fees
- tying loans to first having a deposit
- cross-country banking profitability

Found price differences "too large" to be due solely to cost differences, suggesting markedly different competition levels across Europe.

Competition

Comparing individual service prices can be misleading. A high price for one bank service compared to other countries can be offset by a low price for a different service or low interest paid on deposits at the same bank.

Price data are limited. Use revenue information instead.

Recent theoretical IO literature suggests that revenue focus can be superior to price information in assessing competition (Boone, 2008a, 2008b).

Importance of retail banking in Europe:
  - accounts for ≈ 50% of total banking activity or € 275 billion
  - generates around 2% of European GDP and 3 million jobs.

A Competition Frontier

Bank fee/service pricing can differ due to:
    - lack of competition (the thrust of the EC report) or
    - service cost/productivity differences across countries.

Model framework (Bolt & Humphrey, 2008a):

    retail banking revenues = f (cost, productivity, competition)

Maintained hypothesis:
    - if can "explain" cost/productivity effects on banking revenues, then
      unexplained revenues "left over" reflects competition + normal error.
    - use Distribution Free frontier model (Berger, 1993) to average normal
      error close to zero, leaving average effect of competition on revenue.

Standard Measures of Competition

The *usual* approach "explaining" cross-country service price/profit differences uses:

banking prices/profits = g (competition measure, controls)

Herfindahl-Hirschman Index: banking market concentration.

Lerner Index: mark-up of price over marginal cost.

H-Statistic: correlation of output price with input prices.

Main problem

Expect these 3 "standard" indicators to be positively correlated as they purport to measure the same thing—namely, competition.

Analysis shows that correlations are low (R^2s ≈ .05 to .09) and often negative (H-statistic adjusted so higher value implies less competition).

Occurs across 14 European countries (1,912 banks). Similar results within countries over time (Carbo, Humphrey, Maudos, Molyneux, 2009).

Thus choice of competition measure can affect the conclusions obtained

Our Approach

Use frontier efficiency analysis to indicate competition:

$$\ln revenue = f(\ln cost, \ln productivity) + \ln e + \ln u$$

<u>Assumption</u>: random error ($\ln e$) ≈ 0 when averaged across separate cross-section estimations in panel data while the average of $\ln u$ that remains will reflect the average effect of (unspecified) competition.

Bank/country with lowest average residual ($\ln u$) is the bank/country where variation in underlying cost/productivity explains the greatest amount of revenue variation.

This minimum value determines the frontier. Competitive efficiency (CE) of a given bank/country is measured relative to this frontier:

$$CE_i = \bar{u}_i / \bar{u}_{\min} - 1$$

Industrial Organization Literature

Indicator of competition (plus fixed costs) using relative profit concept
(Boone, 2008a, 2008b)

Calculation from reported data, not estimation:
Profit = Total reported revenues – Total variable costs

Here profit combines actual profits with unknown return to fixed inputs.
Higher values reflect less competition for all firm operations.

We use statistical cost analysis to assign revenues to fixed and variable inputs,
leaving an indicator of "excess" revenue to measure effect of competition.

$$CE_i = \bar{u}_i / \bar{u}_{min} - 1$$

11 countries over 20 years: 1987 to 2006

Belgium	Germany	Spain
Denmark	Italy	Sweden
Finland	Netherlands	U.K.
France	Norway	

Can determine the degree of banking market competition in one country relative to another. Can not determine the absolute level of competition, even for the most competitive country.

Bank Revenues and Operating Costs

Banking activities/revenues included in competition model:
    Non-interest income revenues (NII) includes fee income & payments
    Loan-deposit rate spread (SPREAD)

Securities revenues already determined in a competitive market (excluded)

Over 1987-2006:
    Non-interest income (NII) + SPREAD revenues rose by 160%
    Bank total operating costs rose by 128%.

Share of NNI revenues was 20% now is 44%. SPREAD revenue share fell.

Two Equation SUR Model

Translog equations regress revenues on cost and productivity.

Residuals averaged across countries over 3 time periods to compute CE_i

$$\ln(NII/OC) = \alpha_0 + \sum_{i=1}^{5} \alpha_i \ln X_i + 1/2 \sum_{i=1}^{5} \sum_{j=1}^{5} \alpha_{ij} \ln X_i \ln X_j + \sum_{i=1}^{5} \sum_{k=1}^{2} \delta_{ik}$$

$$\ln X_i \ln P_k + \sum_{k=1}^{2} \beta_k \ln P_k + 1/2 \sum_{k=1}^{2} \sum_{m=1a}^{2} \beta_{km} \ln P_k \ln P_m \tag{1}$$

$$\ln(SPREAD/OC) = \theta_0 + \sum_{i=1}^{5} \theta_i \ln X_i + 1/2 \sum_{i=1}^{5} \sum_{j=1}^{5} \theta_{ij} \ln X_i \ln X_j + \sum_{i=1}^{5} \sum_{k=1}^{2} \psi_{ik}$$

$$\ln X_i \ln P_k + \sum_{k=1}^{2} \phi_k \ln P_k + 1/2 \sum_{k=1}^{2} \sum_{m=1}^{2} \phi_{km} \ln P_k \ln P_m \tag{2}$$

where:

$X_i = L/DEP, ATM/DEP, GAP, PC, ATMC,$

$P_k = PL, PK.$

Cost/Productivity Influences on Revenue Variation

Cost Influences:
   PL = average price of all labor inputs
   PK = average opportunity cost of capital (market interest rate)
   PC = index of unit payment costs (from earlier scale economy estimate)
   ATMC = index of unit ATM costs (from earlier scale economy estimate)

Productivity Influences:
   L/DEP = labor/output ratio (labor productivity in supporting deposits)
   ATM/DEP = capital/output ratio (capital productivity)

Business cycle influence reflects exogeneous loan demand, default risk
   GAP = GDP output gap
   CAPITAL = (equity + reserves)/loans

ATMs have replaced expensive branch offices.

Cross-Country Competition Efficiency

	Non-Interest Income	Rate Spread Revenues
U.K.	0.0% (frontier)	4.9%
Spain	3.1	0.1
France	3.2	0.0 (frontier)
Netherlands	4.6	2.1
Norway	4.7	2.9
Denmark	5.3	1.9
Finland	5.9	2.2
Italy	6.9	2.5
Belgium	7.1	0.7
Germany	7.2	3.1
Sweden	13.5	1.3
Average	5.6%	2.0%

Lower spread CE consistent with greater cross-country competition.
Higher non-interest income CE for priced services (including payments).

Most and Least Competitive Countries

	CE$_{NII}$	CE$_{SPREAD}$	H-Statistic	Profit/ Revenue	CR-3
Most Competitive	U.K. Spain France	France Spain Belgium	Netherlands U.K. Germany	Belgium Netherlands Germany	Spain Italy U.K.
Least Competitive	Belgium Germany Sweden	Norway Germany U.K.	Finland Denmark Italy	Sweden Finland Spain	Belgium Netherlands Finland

U.K. and Spain are most competitive; Finland is least competitive (3 out of 5).
Belgium and Germany ranked as most <u>and</u> least competitive (2 out of 5).

Pricing of Payment Services

In the statistical model:
   lower unit payment cost is associated with
   a *lower* (non-interest income)/(operating cost) ratio.

Payment operating costs rise more than non-interest income even though payment scale economies are large and unit payment cost is falling.

As transaction-based pricing is relatively rare, revenues do not rise "automatically" with payment volume to cover higher total operating costs.

Higher payment-related operating costs likely leads to higher non-transaction bank service prices and may look like the exercise of market power.

Pricing of Payment Services

Banks do not rely on per transaction pricing of payment services. Instead, higher payment operating costs covered by raising prices/fees unrelated to payment volume and their falling unit costs.

Transaction pricing would tie payment prices closer to costs so that:
- management has a clearer picture of where its profits are generated
- consumers can balance better costs with payment service benefits
- authorities could monitor better market price competition.

Implementation could follow Norway's example:
- banks agree on date when transaction pricing could be implemented
- but no agreement on the price to charge (which could be zero).
  (Enge & Øwre, 2006)

Conclusions 1

Bank payment and service delivery scale economies are large and help explain the average 34% reduction in the operating cost/asset ratio over 1987 to 2006.

Payment and ATM scale economies and productivity measures are used in a revenue-based measure of bank competition. Countries are ranked by their dispersion from a "competition efficiency" frontier.

Differences in competition play a relatively small role in explaining variation in cross-country bank revenues. Cost/productivity influences explain 78% to 95% of the revenue variation (R^2).

Rather than look at the entire bank, we look at 2 main revenue sources:
> non-interest income (includes payments)
> loan-deposit rate spread
> (securities returns are determined in "competitive" markets).

Conclusions 2

Non-interest income activities rose from 20% to 44% of revenues.
These fee-based activities (including payments) associated with less competition.

Pricing payments on a transaction basis would:
- generate revenues as payment volumes expand
- transform payments into a "profit center" rather than being a "cost center"
  where it can be difficult to cover costs as volume expands

Antitrust policy:
- relying on a single direct competition indicator may lead to inconsistencies
- when prices/fees are not tied to expanding volume, they can rise to cover
  higher total operating cost even though unit costs are falling due to scale
  economies. This can look like the exercise of market power but need not be.