BANCA D'ITALIA

Temi di discussione

del Servizio Studi

Cost efficiency in the retail payment networks: first evidence from the Italian credit card system

by Guerino Ardizzi

Number 480 - June 2003

COST EFFICIENCY IN THE RETAIL PAYMENT NETWORKS: FIRST EVIDENCE FROM THE ITALIAN CREDIT CARD SYSTEM

by Guerino Ardizzi*

Abstract

In this paper, a parametric cost frontier for the credit card market is specified ("stochastic frontier approach" - SFA) and robustness checks of the main results are performed. The aim is to provide some clues to: the *x*-inefficiency problem and determinants in a retail payment circuit; the main technical characteristics of the industry (i.e. scale economies, cost structure, factor substitution); policy implications. The Italian case study indicates that: the credit card industry could benefit from significant increasing returns to scale, but bigger the network (in terms of transactions handled) the more intermediaries tend to veer away from their efficient cost frontier. Moreover, the cost structure borne by firms is strongly dependent on intra-network agreements. Possible solutions to the *x*-inefficiency problem might come from the "theory of incentives", which provides pricing (or cost recovery) mechanisms computed on the basis of costs expected under efficiency conditions; however, more studies are needed to investigate the topic within the context of "self-regulated" payment networks.

JEL classification: L11, G2, L8.

Keywords: credit card industry, interchange fees, technology, efficiency, cost frontier.

.

Contents

* Bank of Italy, Payment System Oversight Office and University of Rome "La Sapienza", Department of Economics.

1. Introduction[1]

A payment system is a set of rules, technical mechanisms and institutions for the transfer of money between payers and payees. It can deal with large-value operations, usually for monetary settlement where big companies or financial institutions are involved, or handle large numbers of low-value transactions for different end-users, such as consumers and small retailers, in a widespread network.

Cash is the most common retail payment instrument. Other instruments are cheques, credit and debit transfers, and payment cards. Advances in information technology and the increasing adoption of electronic fund transfer procedures have allowed operators to reduce their production costs and the time associated with non-cash payment services. Nevertheless, we usually experience a sort of “monetary illusion” about the “social production cost” of payment services (Humphrey-Berger 1990), especially due to a lack of information, cross-subsidy phenomena and implied pricing mechanisms.

Some economic characteristics mark a “payment system” as a “network industry”. First of all, the payment service (Rochet-Tirole 1999) is usually jointly offered to two parties (the debtor and the creditor) by two other parties (the payer's bank and the payee's bank) interacting through a common infrastructure. The total cost of the service is the sum of the “two-sided” production costs, to be covered by the “two-sided” market demand of end-users, but there is no reason why both banks should break even on the same transaction. The “winning” bank on one “side” could therefore compensate the “losing” bank on the other “side” for faciliting the provision of the joint service: for instance, in the payment card

[1] I am particularly grateful to David B. Humphrey for his helpful suggestions and comments on a preliminary version of the paper. I would also like to thank Paolo Angelini, Mark Armstrong, David Balto, Gerardo Coppola, Riccardo De Bonis, Jean-Charles Rochet, Stefano Siviero, Jean Tirole, Luigi Ventura and two anonymous referees for their useful comments. I have also benefited from comments by participants at the conference on “The Economics of Payment Networks”, IDEI Université de Toulouse I, June 10th and 11th 2002 and from suggestions on a previous version of the paper made by participants at the seminar presentation held at the Research Department of the Bank of Italy on January 10th 2001. I also thank Nives Coppari for editorial assistance. All the remaining errors are mine. The views herein are personal and do not involve the responsibility of the Bank of Italy.

scheme bank used by the merchant usually pays an “interchange fee” to the cardholder's bank to share the issuing costs on the card transaction[2].

Secondly, a payment instrument is of value only if it is accepted by other parties. The wider its acceptance, the greater its value to consumers. The value of a credit card accepted worldwide or of an ATM card (automated teller machine for cash withdrawals) that can be used whenever or wherever a consumer needs cash are clear examples of network benefits. A payment card issued by one bank may be used at another bank's terminals only if the card and the terminal share common technology and security standards. Hence, inter-operability and compatibility are important issues for network goods and services. These issues take on even greater importance because the formation of a network usually depends on attaining “critical mass” and scale economies. However, a typical “chicken and egg problem" is inherent in most network goods and services and often makes it difficult for new networks to achieve a critical mass of consumers (or transactions). For example, a consumer's incentive to buy and utilize a payment card depends directly on the number of merchants who accept the card. At the same time, a merchant's willingness to accept a payment card depends directly on the number of cardholders who use the card.

The consequence of this is that co-operation and standardization are needed between two or more companies, especially in the intermediate markets, in order to form a payment network. However, that could also determine cross-subsidy mechanisms and a “low-powered incentive environment” (Leibenstein 1966). Hence, a typical *x*-inefficiency problem may emerge: one or more firms (or production units) would not produce at the feasible minimum cost.

Until now efficiency problems in self-regulated retail payment systems have been of primary interest for the ex post (eventual) action of the antitrust authorities[3]. However,

[2] Conceptually, there is a similar problem in the sector of public utilities, when the commercial revenues of a company are generally not high enough to cover operating costs and a transfer is then required to balance the budget (see Gagnepain-Ivaldi 1998). Historically, the business model for payment cards has typically favoured cardholders over merchants who accept the cards despite high merchant fees. Hence, some card-issuing costs are covered by an “interbank exchange fee” from the merchant's acquiring bank instead of directly charging the consumer (see Rochet-Tirole 1999 and 2002).

[3] For instance, the European Commission has recently exempted multilateral interchange fees for cross-border payments made with Visa cards from EU antitrust rules - as such a fee is useful for the network

recent studies, such as those conducted in the United Kingdom, Australia and Italy, have emphasized the need to establish ex ante principles for network pricing (or cost recovery) schemes in payment systems in order to guarantee greater transparency and efficiency over time; it is foreseen that central banks will play a for this[4] .

In this paper we leave out some considerations of welfare economics: for example, accepting a given level of cost inefficiency because it is necessary for the start or the expansion of a payment circuit. Furthermore, in such cases, *x*-inefficiency might be considered as "seigniorage" income for the creation of a new means of payment (see Frankel 1998); this is a typical case in monetary and banking history. On the other hand, specialized literature shows that criticism is being directed at the traditional approach, which tends to justify the maintenance of some "protection mechanism", if necessary to foster the growth of the network considered as a "collective good".[5] Moreover, we know how the Leibenstein *x*-inefficiency concept has been criticized by economists such as Stigler (1976), who stress that "measured inefficiency may be a reflection of a failure to incorporate the right variables and the right constraints and to specify the right economic objective of the production unit" (Fried-Lovell-Schmidt 1993), but that criticism does not provide a solution in the empirical analysis.

In this study, after a general description of the "product" (section 2), a parametric cost frontier for the Italian credit card market (sections 3, 4 and 5) is specified. To this end, the Battese-Coelli (1995) econometric model for efficiency analysis ("stochastic frontier

---

provision of the service - but at the same time the interchange fee will be capped by Visa EU to reflect the level of costs for the main services provided by issuing banks to retailers as determined by a cost study to be carried out by Visa and audited by an independent accountant (see EC comp/D1/29373 Visa International).

[4] See the following surveys: "Competition in UK Banking: A Report of the Chancellor of the Exchequer" (2000), by Don Cruickshank; "Debit and Credit Card Schemes in Australia: A Study of Interchange Fees and Access" (2000) by the Reserve Bank of Australia and the Australian Competition and Consumer Commission. In particular, the central bank has recently ordered sweeping changes in the way the credit card business operates in Australia - cutting interchange fees by about 40 per cent - in order to lower costs and promote competition (see "Reform of Credit Card Schemes in Australia, 2002", RBA). For Italy, see "The Italian Case Study: Interchange Fee, Market Structure and Cost Efficiency in the Retail Payment System" (2002) by Ardizzi-Coppola (Bank of Italy, paper presented at the conference on "The Economics of Payment Networks", IDEI Université de Toulouse I, June 2002). In general, for the role of central banks in retail payment systems see: "Policy Issues for Central Banks in Retail Payments" - BIS (2002).

[5] Katz (2001), for example, points out that above a certain level of saturation the marginal benefit deriving from accession to the circuit by new user would tend to disappear. As a consequence, it would be no longer

approach" - SFA) allows us to explain both the firm and time specific effects in terms of cost inefficiency. Furthermore, the robustness of the results on the inefficiency scores is tested through deterministic methods.

This paper is only an introduction to the problem of *x*-inefficiency in retail payment circuits. It aims to provide a better understanding of: a) the study of the main technical characteristics of the "payments industry" (i.e. scale economies, cost structure, factor substitution); and b) the cost inefficiency determinants in a "self-regulated" retail payment circuit.

2. Payment by cards

The use of payment cards[6] in Europe, especially debit cards, is growing faster than any other form of payment. Transaction volumes are increasing at around 20 per cent per year, with about 35 card payments per inhabitant registered in 1999. In the USA, the average annual growth is 12 per cent, with about 95 transactions per capita per year. The volume of payment by credit card is significantly higher than debit cards in the US. Credit cards are the most widely used instrument for paying over the Internet in business-to-consumer transactions, both in Europe and in the USA.

The payment card circuit (credit or debit) works as follows (Evans-Schmalensee 1999): 1) a bank, or other financial company named Issuer, issues a card linked to a "network brand" (i.e. Visa Card, MasterCard, Diners, American Express, etc.) to a consumer (cardholder) who may use the card to charge purchases; 2) a merchant can accept the card as payment if another intermediary (or the same issuing institution), named Acquirer, associated with the same "network brand", provides a "point of sale" terminal (P.O.S.) to

---

necessary for network externalities to be internalized through side payments set on a co-operative basis (i.e. so-called interchange fees).

[6] We define "credit card" as a card whose holder has been granted a credit line, enabling the holder to make purchases and/or draw cash up to a pre-arranged limit; the credit can be settled in full by the end of a specified period (charge card/travel and entertainment card), usually the month, or in part, with the balance taken as extended credit (revolving card) in which case an interest is charged to the cardholder in addition to the usual annual fee. The "debit card" is a card enabling the holder to have purchases directly charged to funds in an account at a deposit-taking institution (usually combined with another function: i.e., cash card for ATM cash withdrawals or cheque guarantee card). See BIS (1999). Both credit and debit cards exploit a network system (issuing, acquiring and technical infrastructures) to provide payment services to consumers and merchants.

authorize and process the transaction; 3) the merchant will be credited for the purchase amount, net a merchant discount fee, and the cardholder will be debited for the same amount. Whenever the Issuer and the Acquirer are different parties (involving so-called 'off-us' transactions), the circuit is a “four-party” system (issuer, cardholder, acquirer, merchant). The acquirer generally pays the issuer an explicit (flat or percentage) interchange fee[7] , set collectively by the institutions belonging to the system or by the self-regulatory body managing the inter-network relationships (i.e., Visa International). If the issuer and the acquirer are the same party, the network is a “three party” system.

Some authors further differentiate between two types of credit card schemes (Gans-King 2001):

a) open loop systems (i.e., VISA International and EuroCard/MasterCard) - the most widely adopted by the financial operators - are self-regulated organizations in the form of membership associations, managed by an international board and established to promote and develop payment cards (or other instruments). They provide operating and access regulations and manage a world-wide electronic network handling authorization and transmission of clearing and settlement data. The member institutions (i.e., banks or other financial institutions) carry out issuing and acquiring activities as licensees of a network brand. At present, this scheme is the most widely adopted in the world in terms of transactions handled.

b) Closed loop systems (i.e., American Express, Diners Club International): financial companies which manage a proprietary circuit and perform issuing and acquiring services directly, as well as some brand promotion activities.

The banking system plays an important role in network governance, in advertising the payment instrument with consumers, and in settling monetary transactions.

[7] Generally, “interchange fees” are used to balance (or transfer) costs between banks or other intermediaries for the joint provision of payment services (see Baxter 1983, Rochet-Tirole 1999 and Balto 2000). A bank pays an interchange fee to another bank for services rendered in a payment scheme involving more than one party, as in the payment card networks. Usually, the interchange fee covers the issuing costs (mark-up included) for clearing and authorisation, debiting cardholder accounts, fraud and default, interest free period. In this paper, we consider both the typical interchange fees and any other intra-network payments to guarantee co-operation and the provision of network transaction services (i.e., outsourcing, partnership agreements, etc.).

2.1 Credit card network industry

In Italy there were 590 million annual payment card operations in 2000, representing 19 per cent of total non-cash payments that year (compared with 3 per cent in 1990). In the last decade, P.O.S. card payments have expanded rapidly, with an average annual growth of over 30 per cent, but the “gap” with other industrialized countries is still large. In Italy there are only 10 transactions per capita annually. Credit card transactions have increased by about 15-20 per cent annually on average in recent years, although paying with debit cards is growing faster in Italy than other cards. Credit card payments[8] amount to 46 per cent of total card payments and experts predict the use of credit cards will increase in the near future.

The credit card industry is concentrated in Italy (see Table 1), although in recent years there has been growing competition between suppliers of payment services (Blue Book 1999). The CartaSì card dominates the domestic issuing and acquiring market (Bank of Italy, Antitrust Decision 135/A/2001). It is issued by Servizi Interbancari (a financial company owned by the leading Italian banks), associated with the international “open loop” circuits (Visa and MasterCard network) and distributed by the member banks of the scheme (about 800 Italian credit institutions). Deutsche Bank is the second largest operator in Italy with the BankAmericard card. The “closed loop” circuits American Express and Diners Club (travel & entertainment cards) have smaller market shares. Moreover, a number of individual banks - representing small shares of the market - have launched proprietary cards directly linked to the international networks (such as IntesaBCI-Setefi with Moneta). Finally, schemes typically present important domestic outsourcers (owned by banks) for the technical management of transaction data.[9]

[8] In Italy the large majority of credit cards are charge cards (one month delay before debiting the cardholder), without charging interest to the cardholder. Most credit cards are issued with pre-authorized direct debit and though the cardholder enjoys an interest-free grace period, the amount charged is automatically debited against the cardholder’s current account at the end of the billing period.

[9] In Europe, card processing arrangements remain largely nationally based and co-operative in nature (see Banking Automation Bulletin for Europe n. 187/2002).

TABLE 1: CREDIT CARDS IN ITALY
(market shares)

Card name (Issuer)	Issuing: % value of purchases by card	Acquiring: % value of transactions at p.o.s.
CartaSì (Servizi Interbancari)	56.9	46.3
BankAmericard (Deutsche Bank)	17.5	14.8
Amex (American Express)	8.6	13.0
TopCard (Bnl)	4.6	2.1
Moneta (IntesaBCI/Setefi)	6.2	8.7
Diners (Diners Club)	3.3	4.6
Other cards	2.9	10.5
	100.0	100.0

Source: Bank of Italy – antitrust decision 2001 n. 135/A and Antitrust Authority opinion 7/2002.

The revenues (R) of a financial company or bank that operates as both issuer and acquirer may be calculated by the following formula:

(1) $$R = q_I if + q_A mdf + c - Qif$$

(1-bis) $$R = (q_I - Q) if + q_A mdf + c$$

where R is total revenues; q_I is the value of transactions processed by the company as issuer but acquired by others; q_A is the total value of transactions (including those executed with the bank's own cards) processed by the bank as acquirer; and Q is the value of transactions processed by other issuers for payments acquired by the bank in question; if is the interchange fee (multilaterally set as percentage of the transacted value) and mdf is the (gross) merchant discount fee, which is determined by the acquirer according to a "mark-up" rule to cover the if costs, other acquiring costs and a profit margin; c is the annual fee charged to the cardholders.

The empirical evidence seems to show that the company would have a strong incentive to operate in both the issuing and acquiring markets[10] and to maximize the value or the

[10] For instance, in UK the four largest banks handle over the 80 per cent of both issuing and acquiring transactions. Generally, according to the rules set by the self-regulatory bodies (such as Visa and MasterCard),

number of transactions processed so as to reap the interchange fee flows (see the first variable of the equation, with $q_I > Q$), as well as to retain the "implicit" interchange fee for transactions executed with its cards at merchants "acquired" by the issuing bank (included in the second term $q_A\, mdf$). Given that the scheme is structured so that the entire interchange fees plus other acquiring costs are passed on to the final fee (merchant discount[11]) and If is constant (multilaterally set), greater revenues would be guaranteed if q_I and q_A increased over time, both in absolute terms and in relative terms, compared with other intermediaries (the negative effect of Q is also reduced).

In Italy, too, the issuing and acquiring markets are controlled by the same operators (see Table 1) and, at the domestic level, multilateral interchange fees on the intermediate market (see note 7) serve less as a redistribution or balancing tool than as a reference fee for final prices (i.e., merchant discount fee). This is true also for the "closed loop" systems (such as American Express and Diners Club) where the interchange fee is only implicit, the "company is vertically integrated" and performs the roles of issuer, system and acquirer.[12] These fees are mainly set to balance (or transfer) costs and profits between intermediaries and to guarantee inter-firm co-operation or the provision of network services (i.e., outsourcing, partnership agreements, commercial distribution, etc.). Nevertheless, the methods used by the self regulatory bodies to calculate uniform interchange fees for payment services are not clear and do not usually take into account technical progress and gains in productivity over time (Cruickshank Report 2000 and European Commission-Visa 2001)[13]

---

acquiring financial companies must also be card issuers (so-called "no acquiring without issuing" rule). See EC (2001), "Decision, Comp/D1/29373-Visa International" and the following surveys: "Competition in UK Banking: A Report to the Chancellor of the Exchequer" (2000), by Don Cruickshank; "Debit and Credit Card Schemes in Australia: A Study of Interchange Fees and Access" (2000) by the Reserve Bank of Australia and the Australian Competition and Consumer Commission. For a description of the Italian case study, see Ardizzi-Coppola (2002).

[11] The merchant discount fee (including the interchange fee) could be difficult to reduce over time as the interchange fee would represent a cost constraint (or access price) for the firms operating on the acquiring side of the market (see Cruickshank Report 2000).

[12] See Rochet-Tirole 2001, p. 30. We could also have explicit "cross-border" interchange fees for proprietary circuits or three-party schemes, when the institutional arrangement provides different national financial companies issuing domestic cards with the same international brand: i.e. Diners Club Italy, France, US etc.. In this case, cross-border transactions are treated as 'off-us transactions' within the circuit.

[13] For instance, multilateral interbank fees simply based on the average costs of different issuing banks could lead to a series of problems (see Cruickshank Report 2000), such as: a) less efficient operators are subsidized by more efficient operators and are not subjected to internal competition within the network; b)

(i.e., telecommunications, electricity provision, etc.). Moreover, in this paper, all other costs associated with intra-network relationships and co-operation are considered.

Tables 2 and 3 give a synthetic description of the services provided (issuing and acquiring) and the running costs in the credit card industry, respectively. In particular, the leading Italian non-bank credit card companies are considered (Servizi Interbancari, American Express, Diners Club d'Italia and Setefi-IntesaBCI) where specific cost information is available.

TABLE 2: CREDIT CARD SERVICES

ISSUING	ACQUIRING
Manages card applications and evaluates customer risk	Manages merchant database
Produces and mails cards and PINs	POS terminal support
Authorization	Authorization
Accounting and settlement services	Transaction management services
Produces and mails cardholders' billing statements	Clearing services
Payments by cardholders	Merchant account statements
Customer assistance	Risk management, charge-backs and retrievals, credit collection and customer service

Source: www.ssb.it

arbitrariness in the allocation of costs or, rather, informational asymmetries in the level of effective costs may exacerbate distortions created by inefficient behaviour. Thus, the need to ensure greater efficiency in sectors like network industries or public utilities has led to a gradual shift from pricing schemes based on the traditional full costing coverage mechanism (so called “cost plus” schemes or low-powered incentive schemes), to “fixed price” schemes (e.g., the “price cap regulation” in the UK) computed on the basis of costs expected under efficiency conditions taking into account technical-organisational progress and the related average gains in productivity across time. See Laffont-Tirole (1993) and Kwoka (1993). For an empirical analysis of the impact of fixed-price regulation schemes versus the traditional cost-plus mechanism on *x*-efficiency levels, see Piacenza (2001), a recent study concerning local public transportation in Italy.

TABLE 3: CREDIT CARD INDUSTRY (FINANCIAL COMPANIES)

COSTS	% ON TOTAL RUNNING COSTS
Personnel	10.00
Administrative	16.00
Marketing	4.50
Other running costs	7.50
Commission expenses	49.00
Interest expenses	10.00
Amortization costs	3.00
Total	100.00

Source: financial companies - annual balance sheets (average data).

Commission expenses represent a large share of total costs, as shown in Table 3. Such expenses include: a) the costs of the "domestic" and "cross-border" interchange fees paid as acquirer to other issuers (see above); b) the charges due to partner banks (i.e. for commercial or sales aids, co-branding etc.); c) the access fees and other expenses to connect with international circuits. In short, such fees belong to the credit card network agreements. Furthermore, other non-personnel costs include outsourcing expenses related to the technical processing arrangements.

Finally, table 3b (above) summarises the operating cost trend for the Italian credit card market in the last decade: we can observe a significant growth of the production costs and a quite stable level of the same costs with respect of the value of processed transactions during the time.

TABLE 3B: COST DYNAMICS OF THE CREDIT CARD INDUSTRY IN ITALY
(financial companies)

Year	Running costs per euro 100 of transactions	Deflated running costs in euros (index number)
1992	5.0	100
1993	5.0	106
1994	5.0	117
1995	5.4	150
1996	5.1	169
1997	5.2	205
1998	5.1	235
1999	4.5	281
2000	4.0	304
2001	4.0	322

Source: financial companies, annual balance sheets

3. Empirical studies of payment services and Stochastic Frontier Approach (SFA)

In the economic literature the main empirical analyses of cost efficiency in non-cash payment circuits, aimed at measuring the ‘distance’ between observed cost and minimum potential cost, were mostly published in the early 1990s. These studies examined the payment services supplied through the Federal Reserve Payment System to identify inefficient behaviour in “public management”, as well as to observe the effects of scale economies and technological change on production costs. Furthermore, some economists applied cost function models to retail payment services, generally those provided by commercial banks to their customers (i.e., ATM cash withdrawals; see Walker 1978 and Humphrey 1994), but without dealing with the *x*-inefficiency problem (see Sivakumar and Shaffer 2002). In the case of credit institutions, the main problems arise from the paucity of analytical information on the cost of bank payment services and from the multi-product nature of the banking sector: it is difficult to carry out an efficiency analysis of different services provided (payments, loans and finance). Moreover, at present there is no empirical evidence on cost (or production) efficiency for non-bank institutions specialized in the provision of payment services, such as credit card companies.

The parametric specification of a traditional cost function [$\min_x$ {w'x} so that y=f(x;β) ⇒ C(y,w;β)] allows us to study the technology underlying the production process (β-coefficients which measure scale economies, technical change, factor demand, etc.). If only the white noise error term ν (independent of the firm's behaviour and with zero mean) is considered, a typical (log)-stochastic cost function [C(y,p;β)exp{ν}] can be estimated. The cost function approach assumes that each observed production unit or firm is efficient and produces the output vector at the minimum feasible (predicted) cost, given the input prices. By contrast, the estimation of a best cost practice frontier allows us to consider the deviations from the minimum cost, better known as *x*-inefficiency deviations (technical and allocative inefficiency).[14]

A general frontier cost specification can be as follows[15]:

$$C_{it} = C(y_{it}, p_{it}, t_{it}; b)\exp\{e_{it}\}, \qquad [2]$$

[14] Usually, the inefficiency of a firm consists of two components (see Fried-Lovell-Schmidt 1993): technical efficiency, which reflects the ability of a firm to obtain maximum output from a given set of inputs; and allocative efficiency, which reflects the ability of a firm to use the inputs in optimal proportions, given their respective prices. These two measures are combined to provide a measure of total economic efficiency or, also, *x*-efficiency according to Leibenstein. The cost function (or frontier) stems from an optimisation problem which provides both the price and technical constraints - while the production frontier considers only the technical constraint and technical inefficiency; hence the solution for cost minimisation implies equality of the marginal rate of input substitution with their respective prices and the economic efficiency (technical and allocative). However, to separate estimations of allocative and technical inefficiency in the stochastic frontier approach, we must consider the input share equations in the system as well, including the one-sided allocative inefficiency term (deviation between the observed and optimal input shares). In the context of stochastic frontiers, too many restrictions must be imposed in the model and therefore it becomes difficult to obtain consistent parameter estimations (Greene 1993).

[15] Frontiers have been estimated using many different methods over the past 40 years. The two principal methods are stochastic frontier and D.E.A. or Data Envelopment Analysis, which involve econometric methods and mathematical programming (without stochastic error terms or random noises), respectively. For a comparative survey, see Fried-Lovell-Schmidt 1993. In this paper, we partially leave out the D.E.A. approach (but see the robustness analysis in section 6.2) for the following reasons: first of all, this method does not allow us to calculate jointly the inefficiency levels and determinants and less consistent "two step" procedures would be used (see below note 19); second, the cross-section sample adopted here (number of firms for each period; see section 4) is too small to obtain more accurate results through a DEA method with technical change analysis over time; moreover, DEA is more sensitive to "outliers" and "noises" in the data. Finally, many studies have shown a strong rank-order correlation between inefficiency scores estimated or computed with different methods of cost frontier analysis (for instance see: Hjalmarsson-Kumbhakar-Heshmati 1996, Bauer-Hancock 1993 and Resti 1997). For this purpose, it is possible to test the robustness of our results on the inefficiency analysis (see below section 6).

with [3] $$e_{it} = v_{it} + u_{it},$$

where C denotes total cost, the scalar y represents output, p is an m × 1 vector of prices of variable factors, τ indicates the year of the observation, and β is a k × 1 vector of technology parameters to be estimated. For all variables the subscript i indexes firm (i = 1, …, I), and the subscript t indexes observation (t = 1, …, Ti). The subscript i on T is used to indicate the unbalanced nature of the panel. For all i, $1 \le T_i \le T$, with T denoting the maximum number of observations available for a firm.16 The error term e_{it} is separated into two components: (*i*) the *white noise component* ~ *i.i.d.* $N(0, s_v^2)$, v_{it}, which can take both positive and negative values and captures exogenous shocks to the production process; (*ii*) the non negative inefficiency term, u_{it}, representing *firm* and *time-specific* cost inefficiency, which indicates the amount by which the logarithm of cost of the i^{th} firm in the t^{th} period exceeds the logarithm of the stochastic frontier, $\ln C(.) + v_{it}$ due to x-inefficiency. Therefore, the cost inefficiency for the i^{th} firm in the t^{th} period can be measured as a ratio of observed cost to potential (stochastic) cost:

[4] $$inefficiency = \exp\{u_{it}\} = \frac{C_{it}}{C\ (y_{it}, p_{it}, t_{it}; b)\exp\{v_{it}\}} \geq 1$$

which assumes values from one (u_{it} = 0 or 'zero inefficiency') to + ∞ (for large enough u_{it})[17].

[16] Briefly speaking, not each firm or production unit is observed in each period, a so-called “unbalanced” panel.

[17] A cost frontier is a function $C(y,p,b) = \min_x \{\mathbf{p}^T\mathbf{x}: \mathbf{x}\ \hat{I}\ L(\mathbf{y})\} = \min_x \{\boldsymbol{p}^T\mathbf{x}: : D_I(\mathbf{y}, \mathbf{x'}) \geq 1\}$, where L(**y**) is the feasible production set and D_I (**y**, **x’**) represents the distance function which measures the maximum radial contraction from the observed input-output combination (**y**, **x’**) to the efficient production set. To be consistent with the optimal solution and the duality characteristic of cost and production function (two faces of the same coin, although the latter considers only the technical inefficiency), the following mathematical property for a cost function $C(y,p,b)$ must be satisfied (cfr. Kumbhakar-Lovell, 2000):

- *C* is non-negative;
- *C* is homogenous of degree 1 in the input price vector *p*;
- *C* is non-decreasing in *y*;
- *C* is non-decreasing in *p;*
- *C* is concave in *p*;
- *C* is continuous in *p;*
- *C* is lower hemicontinuous in vector *y* (or continuous in *y* scalar).

In the past, empirical applications have been made of cost frontier analyses for payment services in the U.S. Federal Reserve System (see Table 4). To this end, typical parametric models[18] for panel data have been proposed: a) the *Fixed Effect (*or *LSDV*, *Least Squares Dummy Variable) model*; b) the *GLS or FGLS* (*Feasible Generalized Least Squares) model*, *ML* (*Maximum Likelihood) model* with specific hypothesis on the one-sided term distribution (i.e., the half-normal random variable); c) the "*thick frontier model*" (Berger-Humphrey 1990) which compares the mean inefficiency between different groups of firms or production units (i.e., the more efficient firms in the first quartile of the distribution can be treated as a "thick frontier" to make comparisons with other firms). These kinds of stochastic frontier, which are useful for measuring both the technological characteristics of the production process and the level of *x*-inefficiency, do not include the inefficiency determinants in the model. In the nineties, the main exponents of the *SFA* (see Battese-Coelli, 1993, 1995; Kumbhakar, 1991; Kumbhakar-Lovell, 2000; Huang-Liu,1994) proposed estimating a single stage ML model where[19], beside the frontier cost specification, it would be possible to insert explanatory variables (*firm-specific* or *time effects)* for the one-sided error term u (x-inefficiency). In this paper a simplified version of the Battese-Coelli model (1995) for panel data is applied, according to its general x-inefficiency term specification:

$$[5] \qquad u_{it} = d'z_{it} + w_{it} = \sum_q d_q z_{qit} + w_{it}$$

where the q subscript on d and z_{it} indexes explanatory variables, including a constant term ($q = 0, \ldots, Q$), while w_{it} is defined by the truncation of the normal distribution with zero mean and variance s_u^2, so that the (variable) truncation point is $-d'z_{it}$. That is consistent with u_{it} being $\sim N^+(d'z_{it}, s_u^2) \Rightarrow u_{it} \geq 0 \Rightarrow w_{it} \geq -d'z_{it}$. The u_{it}s are non-negative random variables, which are assumed to be independently but not identically distributed, with one parameter,

[18] In Bauer e Hancock (1993) we also find a non parametric application to measure efficiency: the so-called *Free Disposal Hull Filter*, a linear programming technique to derive a deterministic "thick frontier" from the observed data.

[19] Initially, in parametric analysis the determinants of inefficiency were studied with a two-stage approach: in the first stage, the inefficiency deviation is assumed to be independently and identically distributed before decomposing the global deviation into the inefficiency and random noise components; in the second stage, the predicted inefficiency is assumed to be a function of a number of firm-specific factors, contradicting the previous hypothesis of the inefficiency term distribution (see Battese-Coelli 1993).

the mode ($d'z_{it}$ in the case of the u_{it} distribution), characterising the placement[20] and the other (here s_u^{2}) representing the spread of the density function.

The simultaneous estimation of the parameters of the stochastic frontier [2]-[3] and the cost inefficiency term [5] is made through the ML method[21], given the hypothesis regarding the distribution of v_{it} and u_{it}.

[20] Battese and Coelli note that, unlike the Reifschneider and Stevenson (1991) model, here the mode $d'z_{ft}$ of the u_{ft} distribution is not required here to be non-negative for each observation, so that $w_{ft} \leq 0$ is possible (i.e., if $d'z_{ft} > 0$).

[21] Using the above distributional assumptions on v_{it} and u_{it} in equations [2]-[5], the *log-likelihood function* for the sample observations, $\ln C = (\ln C_{11}, \ldots, \ln C_{1T1}; \ln C_{21}, \ldots, \ln C_{2T2}; \ldots, \ln C_{ITI})$, can be written as (see Battese-Coelli 1993):

$$L(b, d, s^{2}, g; \ln C) = -\frac{1}{2}\left(\sum_{i=1}^{I} T_i\right)\left\{\ln 2p + \ln s^{2}\right\}$$
$$-\frac{1}{2}\sum_{i=1}^{I}\sum_{t=1}^{T_i}\left\{[\ln C_{it} - \ln C(y_{it}, p_{ft}, z_{ft}, t_{ft}; b) - d'z_{it}]^{2} / s^{2}\right\}$$
$$-\sum_{i=1}^{I}\sum_{t=1}^{T_i}\left\{\ln \Phi[d_{it}] - \ln \Phi[d_{it}^{*}]\right\},$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, $d_{it} = d'z_{it}/(gs^{2})^{1/2}$, and $d_{it}^{*} = [(1-g)d'z_{it} + g(\ln C_{it} - \ln C(y_{it}, p_{it}, z_{it}, t_{it}; b))]/[g(1-g)s^{2}]^{1/2}$, with $s^{2} \equiv (s_v^{2} + s_u^{2})$ and $g \equiv s_u^{2}/(s_v^{2} + s_u^{2})$. These ML estimators are consistent for T ® ¥ (see Kumbhakar-Lovell 2000, p. 169) and "an unbalanced panel causes no real problems for the Battese-Coelli method" (Horrace-Schmidt 1996, p. 262). However, little work has been done to formally establish the asymptotic properties of these ML estimators (see Battese-Broca 1996).

TABLE 4: MAIN STUDIES ON THE STOCHASTIC COST FRONTIER FOR THE PAYMENT SERVICES

<table>
<tr><th>AUTHORS (DATA)</th><th>DATA-SET</th><th>OUTPUT DEFINITION</th><th>ESTIMA-TION METHOD</th><th>FUNCTIONAL FORM</th><th>MEAN INEFFICIENCY (1)</th><th>RETURNS TO SCALE (2)</th><th>TECHNICAL CHANGE (3)</th><th>LABOUR SHARE</th></tr>
<tr><td>Bauer (1993)</td><td>panel 1983-1990: 47 Fed processing site</td><td>Number of processed cheques "return" and "forward"</td><td>GLS random effects</td><td>translog multiproduct</td><td>36%</td><td>1.14</td><td>1.0%</td><td>47%</td></tr>
<tr><td rowspan="4">Bauer-Hancock (1993)</td><td rowspan="4">panel 1979-1990: n. 47 Fed processing site, n. 2256 obs</td><td rowspan="4">Number of processed cheques</td><td>LSDV fixed effects</td><td rowspan="4">translog monoproduct</td><td>85%</td><td rowspan="4">1.17</td><td>-8.0%</td><td rowspan="4">45%</td></tr>
<tr><td>GLS random effects</td><td>41%</td><td>-2.7%</td></tr>
<tr><td>MLE (u half normal)</td><td>11%</td><td>-1.0%</td></tr>
<tr><td>Thick frontier</td><td>49%</td><td>-4.9%</td></tr>
<tr><td rowspan="3">Bauer-Ferrier (1996)</td><td>panel 90-94: n. 240 obs</td><td>Number of Fedwire operations</td><td rowspan="3">LSDV fixed effects</td><td rowspan="3">Translog (and Fourier) monoproduct</td><td>60%</td><td>0.88 -1.26</td><td>-6.0%</td><td>14%</td></tr>
<tr><td>panel 90-94: n. 931 obs</td><td>Number of processed cheques</td><td>37%</td><td>1.03</td><td>1.7%</td><td>49%</td></tr>
<tr><td>panel 90-94: n. 240 obs</td><td>Number of Automated Clearing House operations</td><td>69%</td><td>2.08</td><td>-11.0%</td><td>21%</td></tr>
</table>

Note:

(1): computed as exp(u) -1 ≥ 0, where 1 means full efficiency.

(2): $s = (\partial y/\partial c)(c/y)$: where $s > 1$ for increasing returns to scale.

(3): downward (< 0) or upward (> 0) cost frontier shifts

4. A stochastic cost frontier model for the credit card industry

At present, no specific econometric application for the payment card circuits using the efficient frontier approach have been published in Italy. In this explanatory investigation, we use an unbalanced panel data for the four leading credit card issuers and acquirers (non-banks) operating in Italy (see Table 1: Servizi Interbancari SPA, American Express, Diners Club, Setefi) over the period 1990 to 2001, for a total of 40 observations[22]. The database contains information on productive and cost structures extracted from annual reports or contained in the Supervisory Returns Database of the Bank of Italy. The decision was made to use a limited sample of non-bank intermediaries mainly for the following reasons: first, this sample is quite homogeneous in a credit (charge) card business context - while other small non-bank operators carry on other core business (i.e., "consumer credit" representing the 80-90 per cent of their commercial activity); second, the four firms represent over 70 per cent of the national credit card market; third, whenever important bank operators (such as Deutsche Bank) are included, no specific production cost data are available to catch the specific x-inefficiency effects and the related accounting standards might be less clear, true and fair by comparison with a public balance sheet.

In our specification, we use a *translog single output*[23] total cost frontier, without fixed (or quasi-fixed) inputs. A cost frontier approach is adopted as it considers both technical and allocative efficiency (see note 14) and the cost minimization objective is more appropriate than output maximization (underlying the production frontier approach) when the input prices (rather than input quantities) are exogenous and the output is demand driven and not storable (see Kumbhakar-Lovell 2000). The credit card system - like other network industries (telecommunications, electricity generation etc.) - seems to satisfy these

[22] The panel of firms adopted (40 observations): 1990 and 1991 (2 production units); 1992-1995 (3 units); 1996-2001 (4 units). For parametric cost frontiers with limited sample data see: Thiry-Lawaree 1987, Thiry-Tulkens 1998, Kumbhakar 1991, Levaggi 1994, Hjarmasson, Kumbhakar, Heshmati 1996.

[23] The translog single output cost frontier (or function) may be viewed as a second-order Taylor series approximation (usually around the mean value) to any arbitrary twice-differentiable cost function (relation between costs, output and input prices) satisfying the appropriate regularity conditions (Humphrey, 1981a). The translog cost model is the most widely adopted in payment services analysis (see Table 4).

exogeneity criteria both for the output and the input prices[24]. Moreover, the cost structure borne by credit card companies would be strongly dependent on network agreements (see above Table 3).

In particular (see eq. 6-7), for each firm in each period[25] the proposed stochastic frontier model provides: 1) a composite output (y), given by the sum of the number of transactions processed as issuer and as acquirer[26]; 2) the input prices, labour (p_L) and capital (p_k), respectively given by the cost of personnel divided by the number of employers and by the other running/operating costs (EDP costs, administrative costs, commission expenses, interest expenses, amortization costs) standardized with the composite output y; 3) a temporal (annual) dummy variable t, to catch technical Hicks (downward or upward) neutral shifts of the cost frontier. Finally, v and u represent the white noise and inefficient term respectively.

[6]
$$\ln\left(\frac{C_{it}}{p_{Lit}}\right) = b_0 + b_y \ln y_{it} + b_k \ln\left(\frac{p_{Kit}}{p_{Lit}}\right) + b_{ky} \ln\left(\frac{p_{Kit}}{p_{Lit}}\right) \ln y_{it} + \frac{1}{2} b_{yy} (\ln y_{it})^2 + \frac{1}{2} b_{kk} \left(\ln\left(\frac{p_{Kit}}{p_{Lit}}\right)\right)^2 + b_t t + v_{it} + u_{it}$$

[24] In addition, an appropriate production function requires the “capital” quantities to be independent variables, and measuring “fixed” capital in a credit card company is a difficult task; indeed, the technical infrastructures and distribution channels are not usually directly owned by the firms. Useful information on capital input quantities might come from the network access point statistics (i.e., credit cards in circulation, point of sales, automated teller machines); but reliable and comparable data for the last decade are not always available for each firm (further studies could better investigate such a topic; for telecommunications cost analysis see Shin-Ying 1992).

[25] All the monetary variables are deflated by the GNP implicit price deflator in Italy.

[26] We do not consider a multi-output cost specification since the issuing and acquiring number of transactions tend to correspond or to be strongly correlated in a vertically integrated market such as the credit card system (see above par. 2.1): so we gain more degrees of freedom in the translog specification and reduce the "collinearity" problems in the estimation procedure.

The standard symmetry and linear homogeneity conditions are imposed.[27] The one-sided error term specification is the following:

[7] $u_{it} = d_0 + d_1 \ln(Issuing_{it})$

The first parameter on the right-hand side of equation [7] is the constant, the second represents the *x*-inefficiency effect expected to be correlated to the (log) value of transactions handled by the firm as issuer (see par. 3). This variable has been chosen for the following reasons: a) the issuer is paid a multilateral interchange fee (explicit or implicit) as percentage of the transacted value which represents a stable source of revenues not necessarily reflecting underlying costs (see par. 2.2, eq. 1; see also “The Cruickshank Report” 2000), as well as a fixed annual fee from the cardholder; b) dealing with integrated sectors and given that the same operators control both the issuing and the acquiring sides of the credit card market, the issuing ‘turnover’ may be a good proxy of market power and the revenue share of each firm within the credit card industry.[28]

The thesis is that the credit card pricing scheme - considering both the market structure and the multilateral *network* agreements in the intermediate market - does not provide a connection between revenues and potential ('efficient') expenses. That determines a “low powered incentive environment” for the firms and, hence, an *x*-inefficiency problem for the payment circuit.

[27] See note 17 as well. In this specific context, the symmetry condition for twice-differentiable cost function is: $b_{kl} = b_{lk}$. The normalization of the monetary variables, C and P_k with respect to the price of labour, P_l, is made to ensure the linear homogeneity of the cost function in input prices, involving the following restrictions on the coefficients:

$b_k + b_l = 1$; $b_{kk} + b_{kl} = 0$; $b_{ll} + b_{kl} = 0$; with b_l e $b_{kl} = 0$, as the explanatory variable 'price of labour' is zero after its normalization and log-transformation. The other duality conditions for a cost/frontier function are to be verified *ex post*. In particular, for the ‘average firm’ (see note 15) the non-decreasing of C in y e p (monotonicity) requires the non-negative sign of estimated value of b_y e b_k (in the case of other production units, we must also consider the second order coefficients and compute the function value for each production level). The second order duality condition (cost concavity) of the frontier is more complicated to test: it requires that the Hessian matrix [$\partial^2 C / \partial p_K \partial p_L$] be negative semi-definite. A necessary condition for cost concavity is that the own price input demand elasticity be negative (downward sloping input demand); see Humphrey 1981.

[28] Reliable information on the value of transactions processed as acquirer are not always available for the last decade. However, at the domestic level there is a large share of “on-us” transactions since the Servizi Interbancari scheme (Visa/MasterCard) represents about 60 per cent of both the issuing and acquiring side of the credit card market and the other financial companies - such as Amex and Diners Club - are “closed loop circuits” (the issuer and the acquirer are the same firm). Hence, the “issuing” variable could be considered a good proxy to catch the output revenue effects (both for issuing and acquiring) of the firms (interchange fee flows included). Moreover, in the present model, the issuing variable (both in

5. Results

The parameter estimates and their t-statistics for the translog frontier cost function appear in Table 5 (ML estimation of the model [6]-[7]).[29] Almost all of the estimated coefficients are strongly significant and the related signs are consistent with expectations. The main empirical results - computed for the average firm[30] - are discussed below; (also see Table 7 and Graph 2 for a short description of the technological characteristics of the estimated cost frontier).

TABLE 5: ML ESTIMATION: STOCHASTIC COST FRONTIER PARAMETERS

Regressor (1)	Parameter	Coefficient	Std error	T-ratio	(2)
const	beta 0	8.194	0.044	188.281	***
Y	beta y	0.706	0.020	36.087	***
Pk	beta k	0.877	0.040	21.704	***
YY	beta yy	-0.047	0.006	-7.605	***
Pkk	beta kk	-0.046	0.022	-2.088	**
YPk	beta ky	-0.051	0.022	-2.348	**
T	beta t	0.009	0.003	2.928	***
const	delta 0	0.275	0.026	10.744	***
issuing	delta 1	0.222	0.018	12.363	***
	sigma-squared	0.004	0.001	4.670	***
	gamma	0.932	0.032	28.700	***
log likelihood function		76.121			

(1) All independent variables have been divided by their sample means before the transformation in logarithmic

(2) Statistically different from zero at, respectively: *** 99%, ** 95%, * 90%.

the *x*-efficiency term and in the 'product' specification), includes minor credit card services provided to the cardholder, such as ATM cash withdrawals (about 5 per cent of the total operations handled by the issuer institutions).

[29] The estimation has been conducted with the *software* Frontier 4.1 (Coelli 1996). In Table 6 some descriptive statistics of the sample are shown.

[30] The "average firm" is a hypothetical firm which exhibits sample average values for each variable of the cost model. As we have normalized all independent variables on their respective sample mean before the transformation in logarithms, parameters related to the first-order output terms provide a direct estimate of the average firm cost elasticity and those related to the first-order input prices provide the average firm input cost-shares.

TABLE 6: SAMPLE STATISTICS
(monetary variable deflated)

Variables	Mean	Std deviation
Total costs C (million euros)	133.5	174.1
Output Y (volume in millions)	65.9	92.6
“Capital” price (euro)	2.0	0.9
Labour price (euro)	36,899.6	12,694.0
Value of transactions - issuing (million euros)	2,944.2	4,143.7
“Capital” share	90%	8.2%
Labour share	10%	8.2%

5.1 Output cost elasticity

The estimated parameter b_y provides the output cost elasticity (0.706) and shows the existence of increasing returns to scale for the industry (s = 1.42).[31] In Graph 1, the average and marginal estimated cost curves related to the efficient frontier are illustrated, given the input prices of the ‘average firm’.[32] Likewise, in other network industries such as the *public utilities* regulated sectors (e.g., telecommunications, transportation, energy, gas etc.), the rule "price=marginal cost" would determine a loss for the firm.

[31] In the case of single-output (y is a scalar) the inverse of the output cost elasticity represents the return to scale s (Humphrey 1981b; Levaggi 1994):

$$s = \frac{AC}{MC} = \frac{C}{y.(\partial C / \partial y)} = \left[\frac{\partial \ln C}{\partial \ln y}\right]^{-1}$$

If $s > 1$ (increasing return to scale), then marginal costs (MC = ¶C/¶y) are less than average costs ($AC = C/y$). The other two cases, $s < 1$ and s = 1, indicate, respectively, decreasing returns to scale (diseconomies) and constant returns to scale. Likewise, in other network industries, it might be possible to consider: the “economies of density” where we compute the unit cost reduction due to output expansion (i.e. number of transactions) given the network dimension for each firm, and the “economies of size” where the degree of return to scale is related both to the output and network expansion (i.e., number of p.o.s. or credit cards in circulation). In this paper we compute only the global economies of size, since we do not include network variables in the model. For more details, see Braeutigam (1999).

[32] The curve points have been computed fixing the average firm input prices and varying only the output level of each firm or production unit. We have derived the "individual" returns to scale degree, given the average firm input prices. We have observed the presence of increasing returns to scale for each production level of the sample (with a minimum and maximum value of 1.16 and 1.85 respectively), as well as a positive correlation between economies of scale and firm "size" (in terms of number of processed operations). On this point, for the Italian banking system, see Resti (1996).

5.2. Input demand and factor substitution

The parameter b_k returns the estimated "capital" *factor share*[33] (non-personnel costs on total costs), confirming the "capital intensive" nature of the sector (about 90 per cent, see also Table 6[34]). Table 7 below reports the *Morishima*[35] and the own-price input elasticity.[36] In particular, the *Morishima* elasticity is greater than one (high level of substitution between personnel and non-personnel costs, probably due to *outsourcing*, partnership commercial agreements, etc.), while the own-price input elasticities are both negative (as expected) and show the inelastic demand of "capital" with respect to labour. That is coherent with the cost structure of the intermediaries (Table 3) – as pointed out in the previous paragraph - which is strongly constrained by network relationships.

---

[33] Applying the Shephard Lemma to the (log)cost function , the *input share* S_i are derived:

$$\frac{\partial \ln C}{\partial \ln P_i} = \frac{P_i x_i}{C} = S_i \qquad i \in \{L, K\} \qquad \text{where } x_i = \text{optimal demand for } input\ i$$

[34] If we do not include the commission charges due to other intermediaries, the average *labour factor share* rises to 20 per cent, confirming the "*capital intensive*" nature of the 'payments industry', as shown in other similar studies on the US *Fed Payment System* (see Table 4), with some exceptions in the cheque processing circuit.

[35] *Morishima* elasticity (Blackorby and Russell, 1989 and Seldom, Jewell and O'Brien 2000) measures the curvature of the isoquant when adjustments are made in inputs *i* and *j* in response to a change in the price ratio (P_i/P_j) due to an increase in P_j . This indicator has more flexible properties than the *Allen* input elasticity (see below) such as asymmetry (s_{ij} ¹ s_{ji}) and is computed as follows:

$$s_{ij}^{M} = S_j(s_{ij}^{A} - s_{jj}^{A}) \qquad i, j \in \{L, K\}$$

where

$$s_{ij}^{A} = \frac{b_{ij} + S_i S_j}{S_i S_j} \qquad i, j \in \{L, K\};\ i \neq j \qquad \text{and}$$

$$s_{ii}^{A} = \frac{b_{ii} + S_i^2 - S_i}{S_i^2} \qquad i \in \{L, K\}$$

where s_{ij}^{A} and s_{jj}^{A} are respectively the partial own-Allen and cross-Allen elasticity substitution, S_i and S_j are the cost *input* shares; b_{ij} , b_{ii} represent the second-order estimated parameter for the input price variables.

[36] They are computed: $h_{iPi} = S_i\, s_{ii}^{A}$; $i \in \{L, K\}$ (see Humphrey 1981)

5.3. Technological change

The estimated coefficient $b_t = \partial \ln C / \partial t$ is positive[37] (+0.009, see Table 5). This indicates a slightly upward neutral shift of the frontier over the period 1990-2001. As pointed out in other studies (Bauer and Ferrier 1996)[38], this kind of shift might be due to an excessive growth of "capital" expenses during the period. Moreover, the same authors underline that "because the output measures are not quality adjusted, the 'regress' may be reflecting costs associated with the capital intensive quality improvement". In this context, such an improvement may reflect: automation level, safety and fraud prevention standards, speediness of payments, customer assistance. In Italy, enhanced electronic P.O.S. transactions have risen significantly in the last decade with respect to 'signature' or 'off-line' transactions.

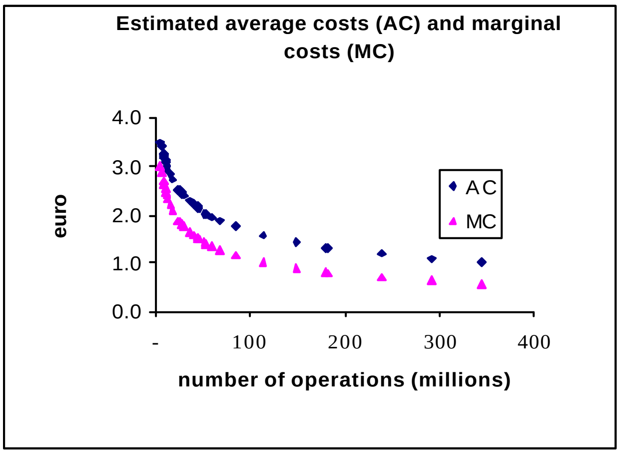
TABLE 7: ESTIMATED TECHNOLOGICAL PARAMETERS FOR THE TRANSLOG COST FRONTIER
(average firm)

		coeff.	Std error
Returns to scale		1.42	0.0277
Factor shares : K = 0.88; L= 0.12			
Morishima elasticity of factor substitution			
		j = K	j = L
	i = K	-	1.9
	i = L	2.1	-
Own-price factor elasticity			
	i = K	-0.4	
	i = L	-1.6	

[37] The LR statistic rejects the null hypothesis of $b_t = 0$ (see Table 8).

[38] See also Bauer-Hancock 1993 for the US cheque processing system (Table 4 in the text). No similar studies of credit card circuits are available.

GRAPH 1



5.4. Cost efficiency

The coefficient d_1 associated with the explanatory variable in the inefficiency model is strongly significant and positive, indicating that the larger the amount of handled transactions the more x-inefficiency increases, according to the proposed model. That is coherent with the hypothesis of the existence of "subsidy effects" due to credit card pricing (and cost recovery) schemes, where revenues are mainly linked to the value of transactions and not necessarily to the "best practice" costs, given the market structure. The estimation procedure replaces the parameters s_v^2 and s_u^2 for the *log likelihood* function with the parameter $g \equiv s_u^2/(s_v^2 + s_u^2)$, which indicates the relative contributions of u_{it} and v_{it} to e_{it} , given the error term hypothesis of distributions (see Battese and Corra 1977). In short, it explains the deviation "share" from the frontier due to x-inefficiency. As $g \rightarrow 0$ the symmetric noise component, v_{it}, dominates the one-sided cost inefficiency term, u_{it}; when $g \rightarrow 1$, instead, we return to a deterministic frontier model, with no random noise included.[39]

[39] Aigner-Chu (1968), Afriat (1972) and Schmidt (1976) derive deterministic production frontiers assuming that the discrepancies between the estimated function and the production situations observed capture exclusively inefficiencies. The ML method "with a particular distribution assumption for the u one-sided residual" (see Thiry-Tulken 1989) may also be adopted for deterministic frontiers with no random noise term. Indeed, adjustment by means of maximum likelihood using a truncated normal law matches the quadratic programming adjustment and the distinction between deterministic statistical and non statistical frontier estimates seems to fade. Of course, the main interest in using the ML method is revealed in the statistical properties of the obtained estimates, not present in a mathematical programming model. However, in our

In our model, the estimated g is equal to 0.93, strongly significant and quite consistent with the stochastic frontier approach. Table 8 presents some generalized likelihood-ratio tests of the hypothesis on the cost efficiency model, calculated as follows:

$$[8] \qquad LR^{40} = -2\{\ln[l(H_0)/l(H_1)]\} = -2[L(H_0) - L(H_1)]$$

In particular, the hypothesis that inefficient effects are absent (g, d_0, d_1 = 0) or present a non-stochastic distribution (g = d_0 =0) - reducing the model to a traditional mean response function in which the inefficiency explanatory variables are included in the cost function - are rejected[41].

Finally, the predicted inefficiency scores - for each period - provided by the software FRONTIER Version 4.1. (Coelli 1996a) - are shown in Table 9 (on average equal to 18.2 per cent and coherent with the inefficiency levels estimated in other studies of payment systems; see above Table 4), together with the value of transactions handled by the networks. They are reported as: $[(C-C^*)/C)]$, where C^* and C represent the minimum cost and the observed cost for each firm, respectively. Moreover, yearly, the weighted means of the inefficiency scores take into account the different firm "size" within the sample in terms of total running costs. Hence, according to our results, the larger the network of a credit card company or the more it increases over time, the *x*-inefficiency problems emerge, making it difficult to benefit from increasing returns to scale (also see above section 5.1).

estimation (see above in the text) the random noise component is strongly dominated by the one-sided cost inefficiency term u, it would be reasonable to suppose our results may be approximately confirmed within the context of deterministic goal programming techniques. See below for a robustness analysis of our results through a "standard costs" method and a simple "Data Envelopment Analysis" (DEA) application.

[40] $L(H_0)$ and $L(H_1)$ represent respectively the *log(likelihood)* function values computed in the null and alternative hypothesis. The test is distributed as a chi-square (c^2) random variable with degree of freedom equal to the number of parameter restrictions imposed. For restrictions on the parameter g , which is to be non-negative, the generalized LR statistic is a mixed c^2-distribution, whose critical values are obtained from Table 1 in Kodde and Palm (1986), for inequality restriction joint hypothesis tests.

[41] The "time specific" dummy variable (d_t) is not present in the inefficiency model, since the test t and LR accept the null hypothesis $d_t = 0$. Nevertheless, the explanatory variable "issuing" already contains a strong temporal trend component which affects the inefficiency level over time (see Table 9).

TABLE 8: LIKELIHOOD-RATIO TESTS

Null Hypothesis	Log-likelihood	χ2-statistic	Decision
$H_0 : \gamma = \delta_0 = \delta_I = 0$	57.020	38.200(*)	Reject** H_0
$H_0 : \gamma = \delta_0 = 0$	67.943	16.356(*)	Reject** H_0
$H_0 : \delta_I = 0$	57.210	38.202	Reject** H_0
$H_0 : \delta_0 = 0$	59.669	32.904	Reject** H_0
$H_0 : \delta_I = \beta_t = 0$	59.783	32.676	Reject** H_0
$H_0 : \beta_t = 0$	62.647	26.948	Reject** H_0

Note: * In this case the LR test statistic is asymptotically distributed as a mixture of chi-square distributions with degrees of freedom equal to the number of parameters assumed to be equal to zero in the null hypothesis H_0, provided H_0 is true. The critical values for this mixed χ^2-distribution are obtained from Table 1 in Kodde and Palm (1986).

** Test significant at 99% ($\chi 2_{0.99}$).

TABLE 9: ESTIMATED COST INEFFICIENCY ON THE ACTUAL COSTS
(% yearly average)

Year	Simple	Weighted
1990	9.0	12.9
1991	11.8	17.0
1992	8.5	16.6
1993	9.7	19.2
1994	9.5	18.7
1995	14.1	27.8
1996	18.1	29.6
1997	20.0	33.0
1998	22.0	36.3
1999	22.6	38.5
2000	24.9	39.7
2001	27.0	40.9
Overall	18.2	27.5

6. Robustness

Since little work has been done in the literature to establish formally the asymptotic properties of the MLEs in the stochastic frontier analysis with finite samples[42] (Coelli 1995, Battese-Broca 1996), tests of robustness of *x*-efficiency results by deterministic techniques are made. In particular, we take into consideration the complete (balanced) side of the panel (years: 1996-2001, where all credit card companies are observed in all time periods) and make comparisons between the stochastic frontier estimation and other techniques of calculation of the *x*-inefficiency scores: a descriptive “standard cost" method and a simple mathematical linear programming application (Data Envelopment Analysis).

6.1. "Standard cost" approach

It is possible to evaluate the potential efficiency gain of a firm just by examining the comparisons with a “standard”, such as economic performance of other competing firms with homogeneous environmental conditions, according to the “yardstick competition” theory (Tirole 1990). In Table 10, for each year a “standard running cost” for the credit card market (financial companies) is calculated multiplying the capital and labour input prices of the most efficient firms by the effective input quantities index of each production unit. The theoretical efficiency gain ratio is computed as a percentage deviation of the standard costs from the actual expenses in the balance sheets.

[42] For instance, Coelli (1995) uses a “Monte Carlo” experimentation to investigate such properties for the half-normal (truncation at zero) stochastic frontier production function and the results indicate substantial “bias” in the MLE only “when the percentage contribution of inefficiency in the composed error (denoted γ) is small” and that such a method should be used in preference to “corrected” OLS estimations when γ (see above section 5.4) is greater than 50 per cent. See also above note 39. In econometric analysis, under regularity conditions, ML estimators are most attractive because of their large-sample or asymptotic properties. However, “the occasional statement that the properties of the MLE are only optimal in large samples is not true” (Greene 2000, p. 127). Indeed, when sampling provides sufficient statistics on the “population” distribution, “MLE will be a function of them, which means that when minimum variance unbiased estimators exist, they will be MLEs”. Moreover, the “limited” longitudinal sample adopted in this paper actually represents almost the entire credit card market in Italy. An interesting topic for further empirical studies with limited samples could be to use “Monte Carlo” experimentation to investigate the MLE properties for the more general case of truncated normal distributions for the *x*-efficiency term.

TABLE 10: STANDARD COST

Year	Num.of firms	Credit card payments in mln euros (issuing)	Average input prices (a)		Actual tunning costs (mln euros)	Average capital factor share on total costs (%)		Standard input prices (a) (b)		Standard running costs (c) mln euros	Theoretical efficiency gain average(%) (d)	
			labour	capital		simple	weighted	labour	capital		simple	weighted
1996	4.0	9,463	39.8	1.8	**450**	78.2	87.0	*26.9*	*1.3*	***308***	*23.0*	*31.5*
1997	4.0	11,544	38.7	1.5	**533**	78.4	88.9	*27.5*	*1.0*	***275***	*31.7*	*48.5*
1998	4.0	14,584	36.8	1.5	**612**	81.4	90.4	*26.0*	*0.8*	***301***	*33.7*	*50.8*
1999	4.0	17,047	34.6	1.5	**709**	83.6	91.5	*24.0*	*0.9*	***382***	*29.9*	*46.2*
2000	4.0	20,453	33.8	1.5	**768**	85.0	92.5	*22.6*	*0.7*	***392***	*33.6*	*48.9*
2001	4.0	23,521	34.8	1.5	**848**	85.2	93.0	*25.4*	*0.8*	***448***	*34.6*	*51.4*
overall	4.0	16,102	36.4	1.5	**653**	82.0	90.6	*25.4*	*0.9*	***351***	*31.1*	*46.2*

(a) - Labour price = personnel costs per employer in thousands of euros; Capital price = other operating costs in euros per transaction .

(b) - Computed as the minimum input price of the firms in the sample.
(c) - Computed as the sum of the products of the input quantities by each firm for the "standard" input price.

(d) - Computed as the % deviation (on a yearly basis) of the standard costs from the real total costs.

The results obtained with this descriptive method seem to be consistent with the ML estimation (see Table 9): the x-inefficiency levels remain high and tend to rise with the “network” size (see the weighted means too). Moreover, it is worth noting the growth in the capital input share of the credit card companies, which is similar on average to the value estimated on the stochastic frontier (Table 7 in section 5.4).

6.2. DEA application

The DEA (Data Envelopment Analysis) is a mathematical programming approach to estimate frontier functions and to calculate efficiency measures. In short, the DEA approach provides a solution to the following linear programming cost minimization problem (Coelli, Rao, Battese 1998):

$$\mathrm{Min}_{\lambda, x_i} \; p_{it}' x_{it}^*$$

$$\text{subject to} \quad -y_{it} + Y\lambda \geq 0,$$

$$x_{it}^* + X\lambda \geq 0,$$

$$N1'\lambda = 1, \; \lambda \geq 0,$$

where p_{it} is a vector of input prices for the i-th firm at time t and $x_{it}*$ is the cost-minimising vector of input quantities for the i-th firm, given the input prices p_{it} and the output levels y_{it}, at time t; λ is a $N\times 1$ vector of constants and $X\lambda$ and $Y\lambda$ are linear combinations of input and output quantities, respectively.

In this case, too, the total cost efficiency or economic efficiency score of the i-th firm in the sample at time t is calculated as: $(p_{it}'x_{it} - p_{it}'x_{it}*) / (p_{it}'x_{it})$; that is, the ratio of the difference between actual cost and minimum (best practice) cost to observed (actual) cost, which is equal to zero for the efficient “unit”.

Generally in the DEA applications, “it is not obvious how to handle panel data in order to get models comparable with the parametric models” (Hjalmarsson et al. 1996). Two option models are considered here in the hypothesis of variable returns to scale: "the first is to compute a sequential frontier, i.e., efficiency computed each year on the basis of all observations generated up to that year", i.e., the first year consists of the 1996 cross-section (4 firms), while the last year the data set covers all observations 1996-2001. The second alternative is an intertemporal frontier (again see Hjalmarsson et al.): we merge the data for all years into one set and calculate efficiency scores for the entire data set. For the cost minimization calculations[43], the DEAP computer program Version 2.1. is adopted (Coelli 1996b).

Finally, Table 11 reports the Pearson ratio and Spearman and Kendall tau-b[44] rank-order correlation coefficients to determine the association between the efficiency indexes for each observation, computed with the sequential deterministic frontier (dea4), the intertemporal deterministic frontier (dea24) and the previous ML stochastic frontier method (sfa40). The latter still considers the estimation procedure over the whole sample (incomplete panel) so as

[43] As in the stochastic frontier function specification, we consider 1) a composite output (y), given by the sum of the number of transactions processed as issuer and as acquirer ; 2) the input prices, labour (p_l) and capital (p_k), respectively given by the cost of personnel divided by the number of employers (x_l) and by the other running/operating costs standardized with an index measuring the "capital" input quantities (x_k). In particular, for the period considered (1996-2001), more reliable network access point statistics are available (e.g., credit cards in circulation, points of sales, automated teller machines). Hence, the "capital" input quantity index chosen for each firm in the DEA application is given by: $(card^2+pos^2)^{1/2}$ where "card" and "pos" are the number of cards in circulation and p.o.s. machines handled by the firms, respectively.

[44] The Kendall index is concerned with differences in relative rankings, while the Spearman coefficient captures the differences in absolute rankings.

to stress the robustness of the ranking analysis with different sample structures as well. Moreover, Table 12 shows the annual simple mean efficiency scores with the different computing methodologies.

As pointed out in the economic literature (Fried, Lovell and Schmidt, 1993), deterministic cost frontiers assume that the discrepancies between the estimated function and the production situations observed capture exclusively inefficiencies - without considering stochastic error terms or random noises - and they are more sensitive to "outliers" in the data. However, although the absolute level of efficiency can vary considerably depending on the technique chosen, “there is broad agreement among the methods about the relative efficiency rankings of production units” (Bauer-Hancock, 1993). To this purpose, the high level of rank correlation coefficient reported in Table 11 between different approaches with respect to each observation in the sample, confirms once again the robustness of the main results of the *x*-inefficiency problem in the credit card industry obtained through the stochastic frontier specification (note also the high rank order correlation between *x*-inefficiency scores and environmental variable "issuing").

TABLE 11: PEARSON, SPEARMAN AND KENDALL TAU-B CORRELATION COEFFICIENTS ON THE INEFFICIENCY SCORES PER FIRM

	dea4+	dea24	sfa40	Issuing
1) Pearson				
dea4+	1			
dea24	0.933	1		
sfa40	0.965	0.905	1	
Issuing	0.794	0.767	0.907	1
2) Spearman				
dea4+	1			
dea24	0.910	1		
sfa40	0.869	0.823	1	
Issuing	0.845	0.857	0.882	1
3) Kendall tau-b				
dea4+	1			
dea24	0.799	1		
sfa40	0.644	0.591	1	
Issuing	0.620	0.664	0.732	1

Legend:
dea4+ = DEA inefficiency scores per firm computed yearly on the balanced panel (years 1996-2001).

dea24 = DEA inefficiency scores per firm computed cross-sectionally on the balanced panel (24 obs).
sfa40 = SFA inefficiency scores per firm estimated on the unbalanced panel.

Issuing = transacted value by card per firm.

TABLE 12: ESTIMATED INEFFICIENCY

(simple means)

Year	dea4+	dea24	sfa40
1996	37.8	43.8	18.1
1997	39.4	47.8	20.0
1998	40.0	50.9	22.0
1999	37.8	49.4	22.6
2000	36.8	52.7	24.9
2001	40.7	40.7	27.9
Overall	38.8	47.6	22.3

Legend:

dea4+ = DEA inefficiency scores per firm computed yearly on the balanced panel (years 1996-2001).

dea24 = DEA inefficiency scores per firm computed cross-sectionally on the balanced panel (24 obs).

sfa40 = SFA inefficiency scores per firm estimated on the unbalanced panel.

7. Conclusions

This paper provides one of the first empirical analyses of the cost inefficiency problem in private retail payment circuits. Regarding the Italian credit card market in the last decade, the following aspects emerge: the credit card industry would have been benefited from significant increasing returns to scale but the more extensive the "network" (in terms of value of transactions handled) the more intermediaries tend to veer from their best practice cost frontier. Moreover, the cost structure borne by companies is strongly dependent on network agreements, and “interchange” commission flows represent significant intermediate input costs for the firms, playing an important role as a component of the final price in a concentrated and vertically integrated market. These factors tend to be related to the “network size” managed by each company and would determine a “low powered incentive environment”: the link between revenues and efficient production costs is no longer ensured over time.

Until now, efficiency problems in the retail payments systems have been of primary interest for the ex post and potential intervention of the antitrust authorities. Nevertheless, recent studies have also emphasized the need to establish ex ante systematic rules and principles for intra-network pricing schemes in the payment systems in order to guarantee greater transparency and efficiency over time.

In this context, the “theory of incentives” (see Laffont-Tirole 1993) might offer some solutions to the problem of fostering the low-cost provision of services considered to be essential to citizens and consistent with the objectives of welfare economics in the presence of information asymmetries, scale and/or network economies. For instance, pricing or cost recovery methods on the intermediate market might be based on costs expected under efficiency conditions according to a “best practice” frontier of the system, taking into account average gains in productivity across the time. However, more theoretical and empirical studies are required to investigate the topic more fully within the context of self-regulated payment networks.

References

Afriat, S. N. (1972), "Efficiency Estimation of Production Functions", *International Economic Review*, Vol. 13, No. 3, 568-598.

Aigner, D. J., and S. F. Chu (1968), "On Estimating the Industry Production Function", *American Economic Review*, Vol. 58, No. 4, 826-839.

Ardizzi, G. (2002) "Efficienza nei sistemi di pagamento 'retail': indagine esplorativa sul circuito delle carte di credito in Italia", Università degli Studi di Roma "La Sapienza", Dipartimento di Scienze Economiche, Discussion Paper No. 39.

Ardizzi, G. and G. Coppola (2002), "The Italian Case Study: Interchange Fee, Market Structure and Cost Efficiency in the Retail Payment System", paper presented at the Conference on *The Economics of Payment Networks*, IDEI Université de Toulouse I, June 10th and 11th 2002.

Balto D. A. (2000), "The Problem of Interchange Fees: Costs Without Benefits", *European Competition Law Review*, Vol. 21, No. 4, 215-224.

Bank for International Settlements (2002), "*Policy Issues for Central Banks in Retail Payments*", Basel.

Bank for International Settlements (2000), *Red Book*, Basel.

Banking Automation Bulletin for Europe (2002), "Half a Billion Payment Cards in Europe", March, No. 187.

Battese, G. E. and T. Coelli (1995), "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data", *Empirical Economics*, Vol. 20, 325-332.

Battese, G. E. and T. Coelli (1993), "A Stochastic Frontier Production Function Incorporating a Model for Technical Inefficiency Effects", Department of Econometrics, University of New England, Armidale, Working Paper No. 69.

Battese, G. E. and S. S. Broca (1996), "Functional Forms of Stochastic Frontier Production Functions and Model for Technical Inefficiency Effects", Centre for Efficiency and Productivity Analysis (CEPA), Department of Econometrics, University of New England, Armidale, Working Paper No. 4.

Bauer, W. and G. D. Ferrier (1996), "Scale Economies, Cost Efficiencies, and Technological Change in Federal Reserve Payments Processing", *Journal of Money and Banking*, Vol. 4, No. 4, 1004-1039.

Bauer, P.W. and D. Hancock (1993), "The Efficiency of the Federal Reserve Check Processing Facilities", *Journal of Banking and Finance*, Vol. 17, 287-311.

Bauer, P.W. and D. Hancock (1995), "Scale Economies and Technological Change in Federal Reserve ACH Payment Processing", *Federal Reserve Bank of Cleveland Economic Review*, No. 3, 14-29.

Baxter W. F. (1983), “Bank Interchange of Transactional Paper: Legal and Economic Perspectives”, *Journal of Law and Economics*, Vol. 26, 541-588.

Berechman, P. (1983), “Analysis of Costs, Economies of Scale and Factor Demand in Bus Transit”, *Journal of Transport Economic and Policy,* Vol. 17, 1-16.

Berger A.N., and D.B. Humphrey (1990), “The Dominance of Inefficiencies over Scale and Product Mix Economics in Banking”, Federal Reserve System Board of Governors, Washington, Discussion Paper No. 107.

Berger A. N., R. S. Demsetz and P. E. Strahan (1999), “The Consolidation of the Financial Services Industry: Causes, Consequences, and Implications for the Future”, *Journal of Banking and Finance*, Vol. 23, No.2-4, 135-700.

Blackorby, C. and R.R. Russell (1989), “Will the Real Elasticity of Substitution Please Stand Up? (A Comparison Of The Allen/Uzawa and Morishima Elasticities)”, *The American Economic Review*, Vol. 79, 882-888.

Braeutigam R.R., “Learning About Transport Cost”, *Techniques of Transportation Analysis*, 1999, 57-97.

Caves, D. W., L. R. Christensen, M. W. Tretheway and R. Windle (1985), “Network Effects and the Measurement of Returns to Scale and Density for U.S. Railroads”, in Daughety, A. (ed.), *“Analytical Studies in Transport Economics”*, Cambridge, Cambridge University Press.

Coelli, T. (1995), “Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis”, *Journal of Productivity Analysis*, Vol. 6, 247-268.

Coelli, T. (1996a), “A Guide to Frontier Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation”, Centre for Efficiency and Productivity Analysis (CEPA), University of New England, Armidale, Working Paper No. 7.

Coelli, T. (1996b), “A Guide to DEAP Version 2.1: A Data Envelopment Analysis Computer Program”, Centre for Efficiency and Productivity Analysis (CEPA), University of New England, Armidale, Working Paper No. 8.

Coelli, T., D. S. Rao and E. Battese (1998), *“An Introduction to Efficiency and Productivity Analysis”*, Norwell, Kluwer Academic Publishers.

Crew M. A., (ed.) (1994), *“Incentive Regulation for Public Utilities”*, Boston, Kluwer Academic Publishers.

Cruickshank D. (2000), *“Competition in UK Banking: A Report to the Chancellor of the Exchequer*”, London, Banking Review.

European Central Bank (2000), *“Blue Book*”, Frankfurt.

European Commission (2001), “Decision, Comp/D1/29373-Visa International”, Bruxelles.

Evans D. S. and R. L. Schmalensee (1999), *“Paying with Plastic”*, Cambridge, MIT Press.

Ferrier, G. D. and C. A. Lovell (1990), "Measuring Cost Efficiency in Banking: Econometrics and Linear Programming Evidence", *Journal of Econometrics,* Vol. 46, No. 1-2, 229-245.

Frankel A. S. (1998), "Monopoly and Competition in the Supply and Exchange of Money", *Antitrust Law Journal*, Vol. 66, No. 2, 313-361.

Gans J. S. and S. P. King (2001), "The Role of Interchange Fees in Credit Card Associations: Competitive Analysis and Regulatory Issues", *American Business Law Review*, Vol. 20, 94-101.

Gagnepain, P. and M. Ivaldi (1998), "Stochastic Frontiers and Asymmetric Information Models", Mimeo, Gremaq - IDEI, Université des Sciences Sociales de Toulouse, France.

Greene, W.G. (2000) "Econometric Analysis", IV edition, New Jersey, Prentice Hall.

Greene, W.G. (1993) "An Econometric Approach to Efficiency Analysis", in Fried H. and Lovell, C. and S. Schmidt (Eds) (1993).

Hjalmarsson, L., S. C. Kumbhakar and A. Heshmati (1996), "DEA, DFA and SFA: A Comparison", *Journal of Productivity Analysis,* Vol. 7, No. 2-3, 303-327.

Horrace, W.C. and P. Schmidt (1996), "Confident Statements for Efficiency Estimates from Stochastic Frontier Models", *Journal of Productivity Analysis*, Vol.7, No. 2-3, 257-282.

Huang, C. J. and J. T. Liu (1994), "Estimation of a Non-Neutral Stochastic Frontier Production Function", *Journal of Productivity Analysis*, Vol. 5, 171-180.

Humphrey, D.B. (1994), "Delivering Deposit Services: ATMs Versus Branches. Federal Reserve", *Bank of Richmond Economic Quarterly*, 59-81.

Humphrey, D.B. (1981a), "Economies to Scale in Federal Reserve Check Processing Operations", *Journal of Econometrics*, Vol. 15, 155-173.

Humphrey, D.B. (1981b), "Scale Economies at Automated Clearing Houses", *Journal of Bank Research*, Vol. 12, 71-81.

Humphrey D. B. and A. Berger (1990), "Market Failure and Resource Use: Economic Incentives to Use Different Payment Instruments, in "*The U.S. Payment System: Efficiency, Risk and the Role of the Federal Reserve*", RBR, Richmond.

Katz, M. L. (2001), "*Network Effect and No-Surcharge Rule in the Australian Credit and Charge Card Market*", Reserve Bank of Australia, Sidney.

Kumbhakar, S. C. (1990), "Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency", *Journal of Econometrics*, Vol. 46, No. 1-2, October/November, 201-212.

Kumbhakar, S. C. and C.A. Lovell (2000), "*Stochastic Frontier Analysis*", Cambridge University Press, Cambridge.

Kwoka J. (1993), "Implementing Price Caps in Telecommunications", *Journal of Policy Analysis and Management*, Vol. 12, No. 4, 726-752.

Laffont, J. J. and J. Tirole (1993*) "A Theory of Incentives in Procurement and Regulation"*, MIT Press, Cambridge.

Leibenstein, H. (1966), "Allocative Efficiency Versus «*X*-Efficiency»*", American Economic Review*, Vol. 56, 392-415.

Levaggi, R. (1994), "Parametric and Non-Parametric Approach to Efficiency: the Case of Urban Transport in Italy", *Studi Economici*, Vol. 53, 67-88.

Fried, H. O., C. A. Lovell and S. S. Schmidt (1993), "*The Measurement of Productive Efficiency: Techniques and Applications*", Oxford Press, Oxford.

Piacenza, M. (2001), "Regulatory Constraints and Cost Efficiency of the Italian Public Transit Systems. An Exploratory Stochastic Frontier Model", Relazione al *I Convegno Hermes: Regolamentazione e Incentivi nel Trasporto Pubblico Locale: Esperienze a Confronto e Proposte per il Caso Italiano, 3/12/01*", Moncalieri, Turin.

Reserve Bank of Australia (2002), "*Reform of Credit Card Schemes in Australia*", RBA, Sidney.

Reserve Bank of Australia and Australian Competition and Consumer Commission (2000), "*Debit and Credit Card Schemes in Australia, a Study of Interchange Fees and Access*", RBA-ACCC, Sidney.

Resti, A. (1997), "Linear Programming and Econometric Methods for Bank Efficiency Evaluation: an Empirical Comparison Based on a Panel of Italian Banks", Institute of European Finance, Research Paper in Banking and Finance, Working Paper No. 1.

Rochet J. and J. Tirole (1999), "Cooperation Among Competitors: the Economics of Payment Card Associations", CEPR, Discussion Paper No. 2101.

Rochet J. and J. Tirole (2002), "Platform Competition in Two-Sided Markets", LSE, Discussion Paper No. 409.

Schmidt, P. (1976) "On the Statistical Estimation of Parametric Frontier Production Functions", *Review of Economics and Statistics*, Vol. 58, No. 2, 238-239.

Shin, R. and J. S. Ying (1992), "Unnatural Monopolies in Local Telephone", *Rand Journal of Economics*, Vol. 23, No. 2, 183-193.

Seldon, B. J. and R. T. Jewell and D. M. O'Brien (2000), "Media Substitution and Economies of Scale", *International Journal of Industrial Organization*, Vol. 18, No. 8, 1153-1180.

Sivakumar, K. and S. Shaffer (2002), "Cost efficiency among credit card banks", *Journal of Economics and Business*, Vol. 54, 595-614.

Stigler, G. J. (1976), "The Xistence of *X*-Efficiency", *American Economic Review*, March.

Thiry, B. and J. Lawarree (1987), "Productivité, Cout et Caracteristiques Technologiques des Sociétés Belges de Transport Urbain" in *Annales de l'Economie Publique, Sociale et Cooperative*, Vol. 4, 368-396.

Thiry, B. and H. Tulkens (1988), "Allowing for Technical Inefficiency in Parametric Estimates of Production Functions, with an Application to Urban Transit Firms", Center Operation Research and Econometrics (CORE), Catholic University of Louvain, Discussion Paper No. 8841.

Thiry, B. and H. Tulkens (1989), "Productivity, Efficiency and Technical Progress", University of Lovain, Center Operation Research and Econometrics (CORE) Reprint, University Catholic of Lovain, Discussion Paper No. 835.

Tirole, J. (1990), "*The Theory of Industrial Organization*", MIT Press, Cambridge.

Tresoldi, C. (2000), "European Retail Payments – A Task for Central Banks?", *Payment Systems Worldwide*, Vol. 11, No. 3, 9-11.

Varian, H.R. (1984), "*Microeconomic Analysis*", II Edition, Norton and Company, New York.

Walker, D.A. (1978), "Economics of Scale in Electronic Funds Transfer Systems", *Journal of Banking and Finance*, Vol. 2, 65-78.

Visa International (2001), "*Credit Card Schemes in Australia, a Response to the Reserve Bank of Australia and Australian Competition and Consumer Commission Joint Study*", Visa.

(*) Requests for copies should be sent to:
Banca d’Italia – Servizio Studi – Divisione Biblioteca e pubblicazioni – Via Nazionale, 91 – 00184 Rome (fax 0039 06 47922059). They are available on the Internet at www.bancaditalia.it

"TEMI" LATER PUBLISHED ELSEWHERE

1999

L. GUISO and G. PARIGI, *Investment and demand uncertainty*, Quarterly Journal of Economics, Vol. 114 (1), pp. 185-228, **TD No. 289 (November 1996)**.

A. F. POZZOLO, *Gli effetti della liberalizzazione valutaria sulle transazioni finanziarie dell'Italia con l'estero*, Rivista di Politica Economica, Vol. 89 (3), pp. 45-76, **TD No. 296 (February 1997)**.

A. CUKIERMAN and F. LIPPI, *Central bank independence, centralization of wage bargaining, inflation and unemployment: theory and evidence*, European Economic Review, Vol. 43 (7), pp. 1395-1434, **TD No. 332 (April 1998)**.

P. CASELLI and R. RINALDI, *La politica fiscale nei paesi dell'Unione europea negli anni novanta,* Studi e note di economia, (1), pp. 71-109, **TD No. 334 (July 1998)**.

A. BRANDOLINI, *The distribution of personal income in post-war Italy: Source description, data quality, and the time pattern of income inequality*, Giornale degli economisti e Annali di economia, Vol. 58 (2), pp. 183-239, **TD No. 350 (April 1999)**.

L. GUISO, A. K. KASHYAP, F. PANETTA and D. TERLIZZESE, *Will a common European monetary policy have asymmetric effects?*, Economic Perspectives, Federal Reserve Bank of Chicago, Vol. 23 (4), pp. 56-75, **TD No. 384 (October 2000)**.

2000

P. ANGELINI, *Are Banks Risk-Averse? Timing of the Operations in the Interbank Market*, Journal of Money, Credit and Banking, Vol. 32 (1), pp. 54-73, **TD No. 266 (April 1996)**

F. DRUDI and R: GIORDANO, *Default Risk and optimal debt management,* Journal of Banking and Finance, Vol. 24 (6), pp. 861-892, **TD No. 278 (September 1996)**.

F. DRUDI and R. GIORDANO, *Wage indexation, employment and inflation,* Scandinavian Journal of Economics, Vol. 102 (4), pp. 645-668, **TD No. 292 (December 1996)**.

F. DRUDI and A. PRATI, *Signaling fiscal regime sustainability,* European Economic Review, Vol. 44 (10), pp. 1897-1930, **TD No. 335 (September 1998)**.

F. FORNARI and R. VIOLI, *The probability density function of interest rates implied in the price of options,* in: R. Violi, (ed.) , Mercati dei derivati, controllo monetario e stabilità finanziaria, Il Mulino, Bologna, **TD No. 339 (October 1998)**.

D. J. MARCHETTI and G. PARIGI, *Energy consumption, survey data and the prediction of industrial production in Italy,* Journal of Forecasting, Vol. 19 (5), pp. 419-440, **TD No. 342 (December 1998)**.

A. BAFFIGI, M. PAGNINI and F. QUINTILIANI, *Localismo bancario e distretti industriali: assetto dei mercati del credito e finanziamento degli investimenti,* in: L.F. Signorini (ed.), Lo sviluppo locale: un'indagine della Banca d'Italia sui distretti industriali, Donzelli, **TD No. 347** (**March 1999)**.

A. SCALIA and V. VACCA, *Does market transparency matter? A case study*, in: Market Liquidity: Research Findings and Selected Policy Implications, Basel, Bank for International Settlements, **TD No. 359 (October 1999)**.

F. SCHIVARDI, *Rigidità nel mercato del lavoro, disoccupazione e crescita*, Giornale degli economisti e Annali di economia, Vol. 59 (1), pp. 117-143, **TD No. 364 (December 1999)**.

G. BODO, R. GOLINELLI and G. PARIGI, *Forecasting industrial production in the euro area,* Empirical Economics, Vol. 25 (4), pp. 541-561, **TD No. 370 (March 2000)**.

F. ALTISSIMO, D. J. MARCHETTI and G. P. ONETO, *The Italian business cycle: Coincident and leading indicators and some stylized facts,* Giornale degli economisti e Annali di economia, Vol. 60 (2), pp. 147-220, **TD No. 377 (October 2000)**.

C. MICHELACCI and P. ZAFFARONI, *(Fractional) Beta convergence*, Journal of Monetary Economics, Vol. 45, pp. 129-153, **TD No. 383 (October 2000)**.

R. DE BONIS and A. FERRANDO, *The Italian banking structure in the nineties: testing the multimarket contact hypothesis*, Economic Notes, Vol. 29 (2), pp. 215-241, **TD No. 387 (October 2000)**.

2001

M. CARUSO, *Stock prices and money velocity: A multi-country analysis,* Empirical Economics, Vol. 26 (4), pp. 651-72, **TD No. 264 (February 1996)**.

P. CIPOLLONE and D. J. MARCHETTI, *Bottlenecks and limits to growth: A multisectoral analysis of Italian industry*, Journal of Policy Modeling, Vol. 23 (6), pp. 601-620, **TD No. 314 (August 1997)**.

P. CASELLI, *Fiscal consolidations under fixed exchange rates,* European Economic Review, Vol. 45 (3), pp. 425-450, **TD No. 336 (October 1998)**.

F. ALTISSIMO and G. L. VIOLANTE, *Nonlinear VAR: Some theory and an application to US GNP and unemployment,* Journal of Applied Econometrics, Vol. 16 (4), pp. 461-486, **TD No. 338 (October 1998)**.

F. NUCCI and A. F. POZZOLO, *Investment and the exchange rate*, European Economic Review, Vol. 45 (2), pp. 259-283, **TD No. 344 (December 1998)**.

L. GAMBACORTA, *On the institutional design of the European monetary union: Conservatism, stability pact and economic shocks,* Economic Notes, Vol. 30 (1), pp. 109-143, **TD No. 356 (June 1999)**.

P. FINALDI RUSSO and P. ROSSI, *Credit costraints in italian industrial districts,* Applied Economics, Vol. 33 (11), pp. 1469-1477, **TD No. 360 (December 1999)**.

A. CUKIERMAN and F. LIPPI, *Labor markets and monetary union: A strategic analysis,* Economic Journal, Vol. 111 (473), pp. 541-565, **TD No. 365 (February 2000)**.

G. PARIGI and S. SIVIERO, *An investment-function-based measure of capacity utilisation, potential output and utilised capacity in the Bank of Italy's quarterly model*, Economic Modelling, Vol. 18 (4), pp. 525-550, **TD No. 367 (February 2000)**.

F. BALASSONE and D. MONACELLI, *Emu fiscal rules: Is there a gap?,* in: M. Bordignon and D. Da Empoli (eds.), Politica fiscale, flessibilità dei mercati e crescita, Milano, Franco Angeli, **TD No. 375 (July 2000)**.

A. B. ATKINSON and A. BRANDOLINI, *Promise and pitfalls in the use of "secondary" data-sets: Income inequality in OECD countries*, Journal of Economic Literature, Vol. 39 (3), pp. 771-799, **TD No. 379 (October 2000)**.

D. FOCARELLI and A. F. POZZOLO, *The determinants of cross-border bank shareholdings: An analysis with bank-level data from OECD countries*, Journal of Banking and Finance, Vol. 25 (12), pp. 2305-2337, **TD No. 382 (October 2000)**.

M. SBRACIA and A. ZAGHINI, *Expectations and information in second generation currency crises models*, Economic Modelling, Vol. 18 (2), pp. 203-222, **TD No. 391 (December 2000)**.

F. FORNARI and A. MELE, *Recovering the probability density function of asset prices using GARCH as diffusion approximations,* Journal of Empirical Finance, Vol. 8 (1), pp. 83-110**, TD No. 396 (February 2001)**.

P. CIPOLLONE, *La convergenza dei salari manifatturieri in Europa,* Politica economica, Vol. 17 (1), pp. 97-125, **TD No. 398 (February 2001)**.

E. BONACCORSI DI PATTI and G. GOBBI, *The changing structure of local credit markets: Are small businesses special?*, Journal of Banking and Finance, Vol. 25 (12), pp. 2209-2237, **TD No. 404 (June 2001)**.

G. MESSINA, *Decentramento fiscale e perequazione regionale. Efficienza e redistribuzione nel nuovo sistema di finanziamento delle regioni a statuto ordinario*, Studi economici, Vol. 56 (73), pp. 131-148, **TD No. 416 (August 2001)**.

2002

R. CESARI and F. PANETTA, *Style, fees and performance of Italian equity funds*, Journal of Banking and Finance, Vol. 26 (1), **TD No. 325 (January 1998)**.

C. GIANNINI, *"Enemy of none but a common friend of all"? An international perspective on the lender-of-last-resort function*, Essay in International Finance, Vol. 214, Princeton, N. J., Princeton University Press, **TD No. 341 (December 1998)**.

A. ZAGHINI, *Fiscal adjustments and economic performing: A comparative study*, Applied Economics, Vol. 33 (5), pp. 613-624, **TD No. 355 (June 1999)**.

F. ALTISSIMO, S. SIVIERO and D. TERLIZZESE, *How deep are the deep parameters?,* Annales d'Economie et de Statistique,.(67/68), pp. 207-226, **TD No. 354 (June 1999)**.

F. FORNARI, C. MONTICELLI, M. PERICOLI and M. TIVEGNA, *The impact of news on the exchange rate of the lira and long-term interest rates,* Economic Modelling, Vol. 19 (4), pp. 611-639, **TD No. 358 (October 1999**).

D. FOCARELLI, F. PANETTA and C. SALLEO, *Why do banks merge?,* Journal of Money, Credit and Banking, Vol. 34 (4), pp. 1047-1066, **TD No. 361 (December 1999)**.

D. J. MARCHETTI, *Markup and the business cycle: Evidence from Italian manufacturing branches,* Open Economies Review, Vol. 13 (1), pp. 87-103, **TD No. 362 (December 1999**).

F. BUSETTI, *Testing for stochastic trends in series with structural breaks*, Journal of Forecasting, Vol. 21 (2), pp. 81-105, **TD No. 385 (October 2000)**.

F. LIPPI, *Revisiting the Case for a Populist Central Banker,* European Economic Review*,* Vol. 46 (3), pp. 601-612, **TD No. 386 (October 2000)**.

F. PANETTA, *The stability of the relation between the stock market and macroeconomic forces*, Economic Notes, Vol. 31 (3), **TD No. 393 (February 2001)**.

G. GRANDE and L. VENTURA, *Labor income and risky assets under market incompleteness: Evidence from Italian data,* Journal of Banking and Finance, Vol. 26 (2-3), pp. 597-620, **TD No. 399 (March 2001)**.

A. BRANDOLINI, P. CIPOLLONE and P. SESTITO, *Earnings dispersion, low pay and household poverty in Italy, 1977-1998,* in D. Cohen, T. Piketty and G. Saint-Paul (eds.), The Economics of Rising Inequalities, pp. 225-264, Oxford, Oxford University Press, **TD No. 427 (November 2001)**.

2003

F. SCHIVARDI, *Reallocation and learning over the business cycle,* European Economic Review, , Vol. 47 (1), pp. 95-111, **TD No. 345 (December 1998)**.

P. CASELLI, P. PAGANO and F. SCHIVARDI, *Uncertainty and slowdown of capital accumulation in Europe*, Applied Economics, Vol. 35 (1), pp. 79-89, **TD No. 372 (March 2000).**

E. GAIOTTI and A. GENERALE, *Does monetary policy have asymmetric effects? A look at the investment decisions of Italian firms,* Giornale degli Economisti e Annali di Economia, Vol. 61 (1), pp. 29-59, **TD No**. **429 (December 2001)**.

FORTHCOMING

A. F. POZZOLO, *Research and development regional spillovers, and the localisation of economic activities*, The Manchester School, **TD No. 331 (March 1998)**.

L. GAMBACORTA, *Asymmetric bank lending channels and ECB monetary policy*, Economic Modelling, **TD No. 340 (October 1998)**.

F. LIPPI, *Strategic monetary policy with non-atomistic wage-setters,* Review of Economic Studies, **TD No. 374 (June 2000)**.

P. ANGELINI and N. CETORELLI, *Bank competition and regulatory reform: The case of the Italian banking industry,* Journal of Money, Credit and Banking, **TD No. 380 (October 2000)**.

P. CHIADES and L. GAMBACORTA, *The Bernanke and Blinder model in an open economy: The Italian case,* German Economic Review, **TD No. 388 (December 2000)**.

P. PAGANO and F. SCHIVARDI, *Firm size distribution and growth,* Scandinavian Journal of Economics, **TD. No. 394 (February 2001)**.

M. PERICOLI and M. SBRACIA, *A Primer on Financial Contagion,* Journal of Economic Surveys, **TD No. 407 (June 2001)**.

M. SBRACIA and A. ZAGHINI, *The role of the banking system in the international transmission of shocks,* World Economy, **TD No. 409 (June 2001)**.

L. GAMBACORTA, *The Italian banking system and monetary policy transmission: Evidence from bank level data*, in: I. Angeloni, A. Kashyap and B. Mojon (eds.), Monetary Policy Transmission in the Euro Area, Cambridge, Cambridge University Press, **TD No. 430 (December 2001).**

M. EHRMANN, L. GAMBACORTA, J. MARTÍNEZ PAGÉS, P. SEVESTRE and A. WORMS, *Financial systems and the role of banks in monetary policy transmission in the euro area,* in: I. Angeloni, A. Kashyap and B. Mojon (eds.), Monetary Policy Transmission in the Euro Area, Cambridge, Cambridge University Press. **TD No. 432 (December 2001)**.

F. SPADAFORA, *Financial crises, moral hazard and the speciality of the international market: further evidence from the pricing of syndicated bank loans to emerging markets*, Emerging Markets Review, Vol. 4 (2), pp. 167-198, **TD No. 438 (March 2002)**.

D. FOCARELLI, *Bootstrap bias-correction procedure in estimating long-run relationships from dynamic panels, with an application to money demand in the euro area,* Economic Modelling, **TD No. 440 (March 2002)**.

D. FOCARELLI and F. PANETTA, *Are mergers beneficial to consumers? Evidence from the market for bank deposits*, American Economic Review, **TD No. 448 (July 2002)**.

A. BAFFIGI, R. GOLINELLI and G. PARIGI, *Bridge models to forecast the euro area GDP*, International Journal of Forecasting, **TD No. 456 (December 2002)**.

F. BUSETTI and A. M. ROBERT TAYLOR, *Testing against stochastic trend and seasonality in the presence of unattended breaks and unit roots*, Journal of Econometrics, **TD No. 470 (February 2003)**.