A Primer on Payment Cards*

Jean-Charles Rochet

Jean Tirole

IDEI, Toulouse

– Report Prepared for the Portuguese Competition Authority –

Final Version

July 22, 2005

*Acknowledgement: We benefited from the comments of Luis Cabral and Pedro Pereira on a first version of this report.

1 Introduction

Payment cards are gradually becoming the most popular non-cash payment instrument around the world [see Evans-Schmalensee 2003]. Different types of payment cards co-exist, offering distinct functionalities and operating according to distinct business models: credit cards, charge cards, debit cards, retailer cards. The networks providing these payment card services also vary a lot in their organization. They can be:

- proprietary networks (Amex),
- open-access international networks that can be for profit (MasterCard since 2003) or not-for-profit (MasterCard until 2003, Visa),
- domestic networks run by banks, either as associations (GIE Carte Bancaire in France) or for-profit corporations (Rede Multibanco or Redunicre in Portugal).

The intensity of competition in downstream markets (issuers/cardholder banks and acquirers/merchant banks) also varies from country to country. In some countries like the USA, the acquiring market is competitive while in others, like Portugal for credit cards and Israel, it is monopolistic and integrated with the network operator.

In order to operate these networks, downstream banks have entered into different forms of cooperative agreements (multilateral interchange fees, honor-all-cards rules, no surcharge rules,...). Some of these cooperative agreements have repeatedly been attacked by retailer associations and competition authorities in several regions of the world (USA, UK, Israel, EU, France, Australia,...).

The purpose of this report is to survey the recent academic literature on the economics of payment cards, stressing the indirect externalities inherent to this industry (and other "two-sided markets"), and providing guidelines for potential regulatory and antitrust interventions. It is organized as follows:

- Section 2 explains the indirect externalities involved in the payment card industry, distinguishing between usage externalities, related to the choice (typically made by the buyer) of a payment instrument among those jointly available to the buyer and the seller, and membership externalities, related to the prior decisions of the buyer to purchase a particular payment card and of the seller to join one or several payment card networks and invest in the associated POS terminal.
- Section 3 provides a model of the payment card industry that is flexible enough to encompass the different situations described above: credit cards vs. payment cards,

proprietary networks vs. open networks, competitive acquiring vs. monopolistic acquiring.

- Section 4 uses this model to analyze the determinants of interchange fees and merchant fees, and develops criteria for deciding whether these fees can be considered "excessive".
- Section 5 discusses the robustness of the conclusions of the benchmark model to changes in the assumptions.
- Section 6 discusses regulatory and competition issues, such as the potential rationale for regulating interchange fees, and that for treating payment card networks as "essential facilities".
- Section 7 looks at the implications of duality.
- Section 8 examines foreclosure and tie-ins, with a particular attention to the case of tie-ins of debit and credit cards.
- Section 9 provides preliminary answers to specific questions posed by the Portuguese Competition Authority.

2 The Fundamental Externalities Inherent to the Payment Card Industry

The choice of a payment instrument to settle a transaction affects the costs and benefits of both parties to the transaction and therefore involves a fundamental externality. It is standard to distinguish between membership externalities (do I have a phone?) and usage externalities (do I call or turn it on?). In the case of payment systems, the distinction is on the merchant side muddled by (one feature of) the Honor-All-Cards (HAC) rule that forces a merchant to accept all payments by card within a given system or none. To illustrate usage externalities, it is therefore useful to envision an hypothetical world in which a merchant is allowed to discriminate in acceptance of a *given* card (that is, the merchant accepts the card as he wishes, depending on the circumstances).

If the buyer[1] insists on paying by cash (which is legal tender), the seller incurs the cost of handling and holding the cash until it can be deposited in a bank. On the other hand, a cash payment allows the seller to save the merchant fee charged by his bank for managing card payments. Symmetrically, a seller who refuses a payment by card typically forces the

[1]For simplicity, we will call these parties the buyer and the seller, instead of using the technical terms: payor and payee.

buyer to incur the cost of finding an ATM and withdrawing cash. He may also prevent the buyer from receiving the benefits that are often associated with a card payment (such as deferred debit or free interest period, frequent flyers miles or cash back bonuses). He may even make the transaction infeasible altogether if the card is a credit card and the buyer needs the credit facility to buy the good or service sold by the seller. This externality can be characterized as a "usage" externality (see Rochet and Tirole 2004 for more detail).

Payment card networks are also characterized by a more classical network externality. Indeed, a seller who decides to join a payment card network, and thus implicitly commits to accepting the cards issued by the members of this network, increases the potential utility of buyers who hold such cards by offering them a new opportunity for using their cards. This is similar to the positive externality generated when a new user joins a telecom network. This externality is called a "membership" externality (see Rochet and Tirole 2004). It perhaps becomes less and less important as the network matures, i.e., when most potential users have joined[2]. By contrast, even in a mature network (where most buyers hold cards and most sellers accept them), the usage externality identified above remains important: the choice of the payment instrument is ultimately a decision of the buyer, that impacts the net costs of the seller. Our objective in this section is to study the consequences of this usage externality.

To simplify the analysis, we assume that transactions can only be settled either by using a card or cash and we start with the case where the number of transactions is fixed. Thus, the efficiency of card usage is determined by answering the following question: which transactions are (and should be) settled using a card rather than cash?

The answer to this question depends on the difference in the net utilities accruing to buyers and sellers for a card payment and for a cash payment. Let b^B denote this difference in gross utility for the buyer and b^S that for the seller, for a typical transaction. Let c denote the total cost of a card payment for the two banks who provide the payment service:[3] the bank of the buyer, called the issuer, and the bank of the seller, called the acquirer. Assuming that the issuer and the acquirer (or anyone else for that matter) incur no direct cost when the transaction is settled in cash,[4] c can also be interpreted as the

[2]Consider indeed a mature network, where almost all merchants accept payment cards and almost all consumers own a card. The last merchants to refuse cards are probably the ones for which consumers are the least interested in paying by card. Symmetrically, the last consumers without cards are probably the ones with the weakest purchasing power. For a dynamic analysis of network externalities, see Katz and Shapiro (1985, 86).

[3]c may also include the marginal cost of the card network, but we neglect it in this analysis.

[4]Consumers and merchants handle cash though their banks. To be certain, the banks' cost of handling cash may at the margin not be equal to what they charge for the service. Similarly, the cost of ATM services may or may not at the margin be aligned with their pricing to ATM cardholders. If cash services are priced below cost, then card payments should be further encouraged in comparison to the no-distortion-of-pricing-of-the-cash-service benchmark studied in this paper; The analysis must be amended as explained in the next footnote (which is set in the context of check services).

incremental cost of card versus cash. Thus social welfare is maximized (i.e. the use of cards is socially efficient) whenever card payments occur if and only if:

$$b^B + b^S \geq c. \tag{1}$$

Baxter (1983) was the first to emphasize that perfect competition among banks does not, on its own, lead to this condition (i.e. to an efficient choice of payment instruments) in the absence of an interchange fee. Indeed, perfect competition without transfers implies that p^B (the unit price of card services for buyers) equals c^B (the issuers' marginal cost) and similarly that p^S (the unit price of card services for sellers) equals c^S (the acquirers' marginal cost), where $c^S + c^B = c$. In the absence of side-payments and strategic considerations by sellers (both aspects are studied later) a card payment takes place if and only if both parties agree, i.e.:

$$b^B \geq p^B = c^B \quad \text{and} \quad b^S \geq p^S = c^S. \tag{2}$$

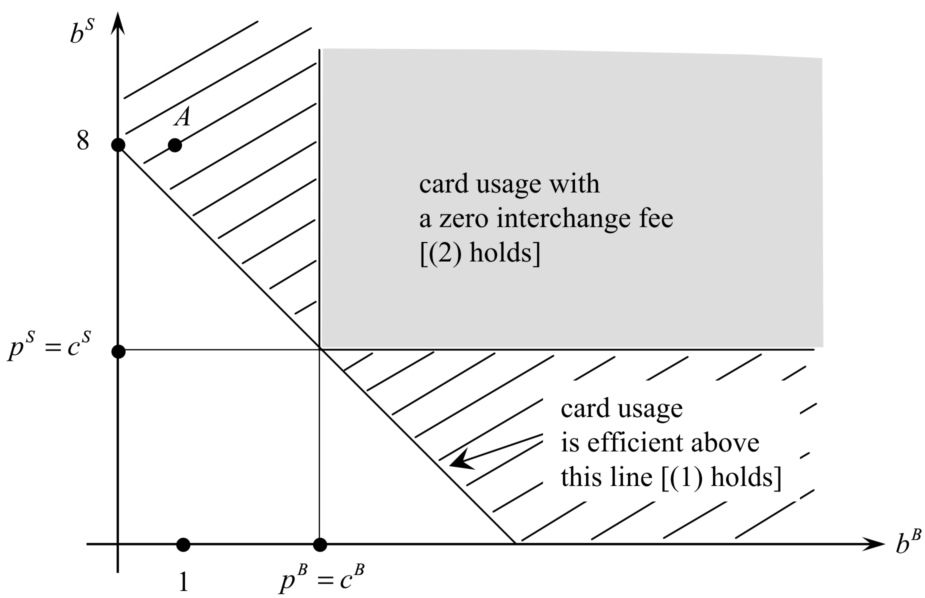


Figure 1: Under-efficient usage of card services when the interchange fee is zero. The hatched area represents foregone card transactions that would have been efficient.

As illustrated by Figure 1, condition (2) is more restrictive than condition (1) and so the usage of card services is suboptimal even if banks are perfectly competitive.[5] This is due to the externality described above: consider for example the case where $b^B = 1$,

[5] This is a fortiori true if the alternative payment instrument is a check, of total cost c_0. The

$b^S = 8$, $c^B = 4$, $c^S = 4$ (this corresponds to point A in Figure 1). By not using the card (this is rational for the buyer since $b^B < p^B = c^B$) the buyer inflicts a negative externality ($c^S - b^S = -6$) on the seller, and prevents a socially efficient card payment (since $b^B + b^S = 9 > c^B + c^S = 8$).

When sellers are homogenous (i.e. b^S is the same for all sellers) and banks are perfectly competitive, Baxter (1983) shows that this inefficiency can be eliminated by having the acquirer pay an interchange fee of $a_0 = b^S - c^S$ to the issuer. For this interchange fee, the merchants are indifferent between a cash and a card transaction as the merchant discount, $a_0 + c^S$, is equal to their gross surplus, b^S, from the transaction. The net cost of the issuer becomes $c^B + c^S - b^S = c - b^S$. And so issuers internalize the total cost of a card transaction net of the seller's benefit. Under perfect competition between issuing banks, this change in the issuer's cost is fully passed through to buyers (p^B becomes $c - b^S$) and the card transaction takes place if and only if:

$$b^B \geq c - b^S, \tag{3}$$

which is equivalent to (1), the social efficiency condition. In a perfectly competitive context, an interchange fee set at the optimal level a_0 allows the internalization of the fundamental externality described above and restores efficiency of card usage.[6,7]

These conclusions are illustrated in Figures 2 and 3:

---

incremental cost of cards becomes negative, whereas checks are often free of charge for users. In this case, the condition for efficient usage of cards becomes $b^B + b^S \geq c - c_0$, whereas in the competitive equilibrium without interchange fees, the condition for card usage remains $\{b^B \geq c^B$ and $b^S \geq c^S\}$ when check payments are not charged by banks. If checks were charged at their unit costs, say c_0^B and c_0^S, the reasoning of the text would apply, provided that the costs of card payments are replaced by the incremental costs of card versus checks. Notice that in this case an interchange fee on checks could suffice to restore efficiency. In practice, more than two payment instruments are available, which complicates the analysis. In particular, an efficient use of payment modes requires that interchange fees be set appropriately for all payment instruments.

[6] If b^S varies across sellers but is observable, Baxter's result remains valid if the network sets differentiated interchange fees.

[7] Notice that no network effect is involved in this example. Network effects only become important when fixed fees and fixed costs are introduced, or when one of the parties (typically the seller) has to commit ex-ante to accept cards. In this case, his acceptance decision is influenced by the number of cardholders.

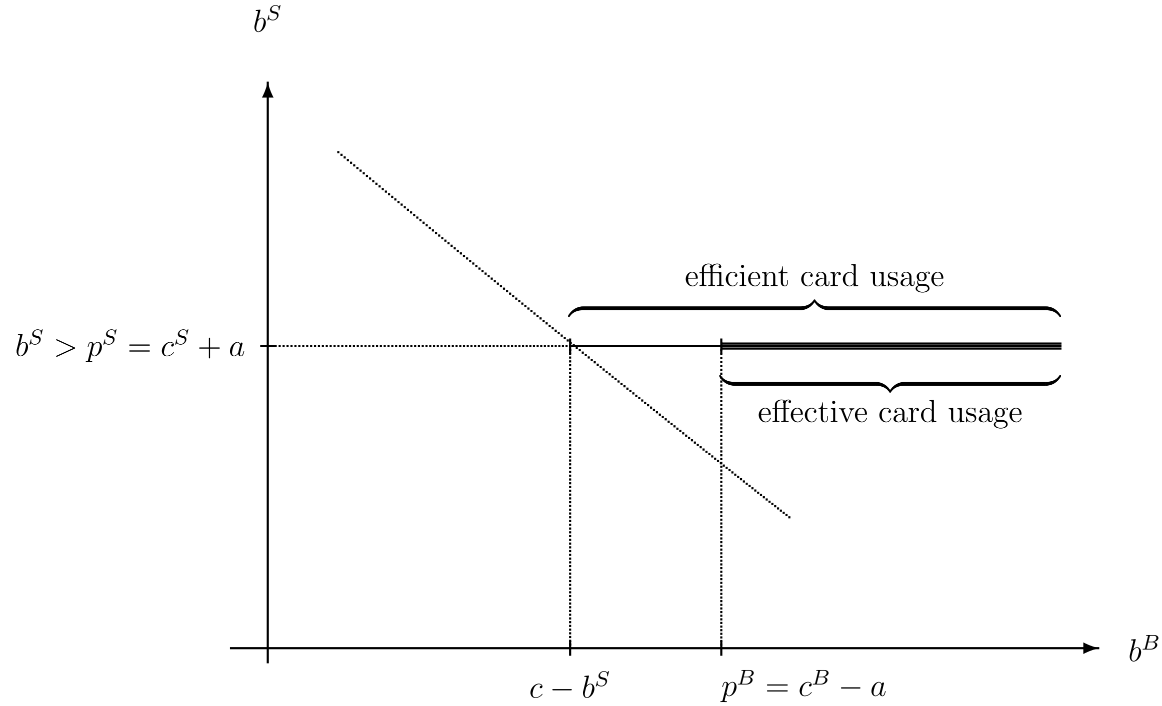


Figure 2: Card usage with homogenous sellers
and an inefficiently low interchange fee $a < a_0 = b^S - c^S$.

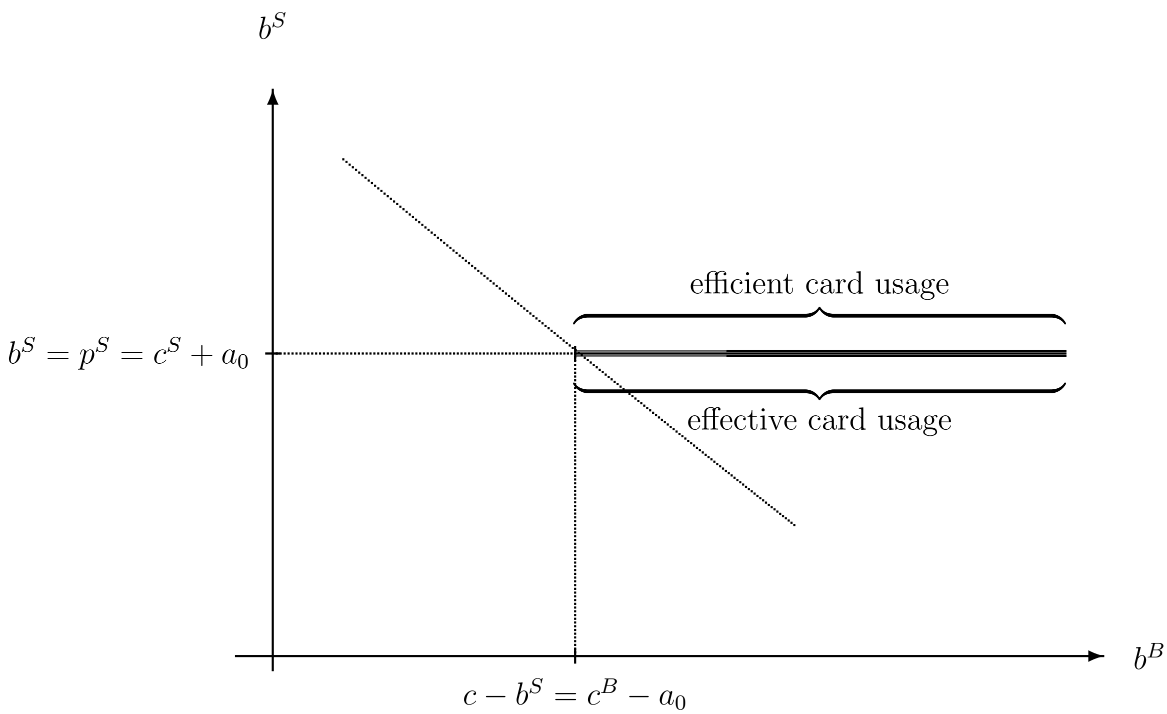


Figure 3: Card usage is efficient when $a = a_0 = b^S - c^S$
(Baxter's interchange fee).

In their criticism of Baxter's analysis, Carlton and Frankel (1995) argue that this externality can alternatively be internalized, in the absence of interchange fee, by allowing sellers to charge differentiated prices for cash and card payments. Indeed, if sellers themselves are perfectly competitive and if they and their customers incur no transaction costs from a dual (cash/card) price system, they fully pass through their net cost $c^S - b^S$ of card payments to buyers by surcharging (or discounting) card payments relative to cash payments. Buyers then use their card if and only if

$$b^B - p^B \geq c^S - b^S, \tag{4}$$

In the case of competitive issuers ($p^B = c^B$), this condition is equivalent to the efficiency condition :

$$b^B + b^S \geq c^B + c^S = c.$$

In practice, however,[8] in countries where it is allowed, differentiated pricing for cash and card payments is seldom used by sellers (this phenomenon is called "price coherence" by Frankel (1998)), probably because of transaction costs. Casual evidence from the UK or some US States, where there is no No Discrimination Rule, shows that in practice such transaction costs prevent a large proportion of merchants from charging differentiated prices for cash and cards. More systematic survey evidence from the Netherlands and Sweden also points at a small amount of surcharging.[9]

But if the majority of sellers do not surcharge buyers, the volume of payments (which depends both on the proportion of sellers who accept cards and on the proportion of buyers who are willing to pay by cards) is a function of both prices p^B and p^S and not only on their sum. In this case, the level of the interchange fee matters.

A second criticism can be leveled at the Carlton-Frankel analysis: Suppose that transaction costs were low enough so that the merchants would differentiate prices and pass through to cardholders any increase in the interchange fee and therefore in the merchant discount. On the cardholder side, the cost charged by the issuer for using the card decreases by the increase in the interchange fee, but the price paid to the merchant also increases by the same amount: The interchange fee is therefore "irrelevant" or "neutral", in that its level has no impact on card usage.[10] Consequently, in the world depicted by Carlton and Frankel, the interchange fee should not be a topic of interest to competition authorities, retailers, or any other player in the industry.

[8]This is acknowledged by Carlton and Frankel.

[9]See ITM (2000). In the US, surcharges for card payments are forbidden (at least in some states: see Chakravorti and Shah (2001) for a detailed discussion of surcharges and discounts in the US) but cash discounts are allowed. In the UK any form of restriction is illegal, yet differentiated pricing is not frequently observed.

[10]This intuition was verified in a restricted setting by Rochet and Tirole (2002). Gans and King (2003) prove it in an extremely general setup.

Returning to the absence of price differentiation, we next ask whether credit card networks choose an appropriate level of the interchange fee or there is a systematic bias toward under- or over-provision of card services? Unfortunately, Baxter's framework cannot provide an answer to this question, because in a perfectly competitive world, banks make no profit regardless of the level of the interchange fee. Because he further focuses on not-for-profit networks, it is therefore impossible to predict how the payment card network will select its interchange fees. In order to predict the interchange fee, one must allow the network owners' profit be affected by the interchange fee, either because they are imperfectly competitive users of the network, or because the network is for-profit, or both. Indeed, banking industries are not perfectly competitive in many countries. Carlton and Frankel (1995) have expressed the concern that interchange fees could be used as collusive devices when banks have market power. The model presented in the next section allows us to examine this question.

3 A Model of the Payment Card Industry

Rochet and Tirole (2002) provided the first fully-fledged model of an imperfectly competitive payment card industry, allowing a comparison between privately optimal and socially optimal interchange fees. With respect to Baxter's analysis, two important features of the payment card industry were added: imperfect competition between issuers, and strategic behavior of sellers. For tractability, Rochet and Tirole made some simplifying assumptions: competition is perfect among acquirers, the total number of transactions (cash plus card) is fixed and normalized to one, and sellers are homogenous. A possible motivation for the highly-competitive-acquirers assumption is the merchants' low search cost and limited brand loyalty; this is however a strong assumption, and certainly one that is grossly violated in countries with monopolized acquiring (Portugal, Israel). The second assumption amounts to focusing on the choice of payment instruments for a given volume of transactions and therefore neglecting the impact of price variations in payment services on the demand for final goods and services: it also seems reasonable in a first step.[11]

We extend the model of Rochet and Tirole (2002) in order to make it applicable to other contexts, like that of the Portuguese payment system. In particular, we allow for a possibility of a for-profit network (Rochet and Tirole 2002 considers only a not-for-profit association) and we consider two variants:

- In the first variant, that corresponds to the Portuguese debit card industry, banks compete on the two sides of the market for providing services both to cardholders (issuers) and merchants (acquirers).

[11]Two papers by Schwartz and Vincent (2000, 2002) relax this assumption in a particular context.

- In the second variant, that corresponds to the Portuguese credit card industry, banks only compete on the issuing side, while acquiring services are provided by a monopoly acquirer integrated with the network operator.

We study these two variants in Subsections 3.1 and 3.2. For expository purposes, we start by abstracting from competition with other networks. In Subsection 3.3, we introduce a second platform, and study the consequences of inter-network competition.

In Appendix 1, we give a more precise description of the Portuguese payment card industry, and explain the simplifications introduced in the model.

3.1 A Model of the Portuguese Debit Card Industry

Neglecting for the moment the competition with Redunicre, we model, in this subsection, the functioning of Rede Multibanco. It is illustrated by the following figure:

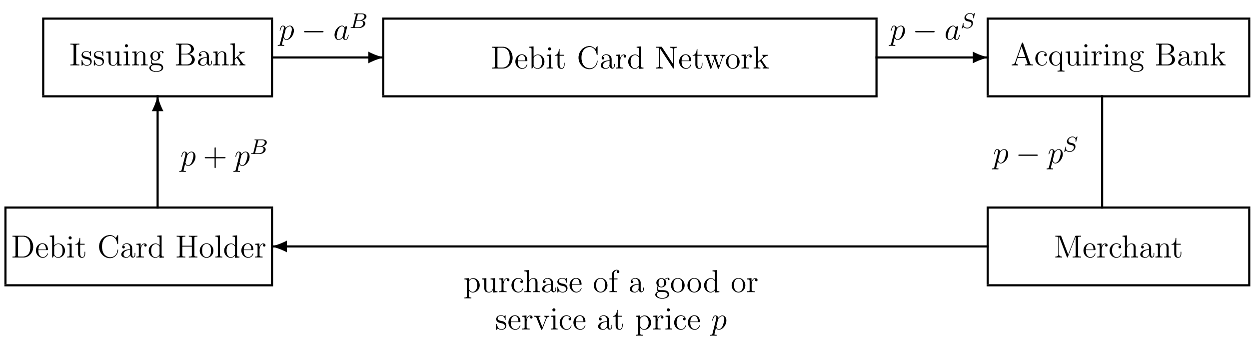


Figure 4: The flows of funds associated with a debit card payment.

When a debit card holder (who, for convenience, will be called the buyer) uses his debit card to purchase a good or service in a store at price p, the amount $p+p^B$ is debited from his account at the issuing bank, where p^B represents the cardholder fee.[12]

The issuing bank then transfers $p - a^B$ to the network, where a^B represents the interchange fee on the buyer's side. The network in turn transfers $p-a^S$ to the acquiring bank, where a^S represents the interchange fee on the seller's (i.e. merchant's) side. Finally, the account of the merchant in the acquiring bank is credited of the amount $p - p^S$, where p^S represents the merchant fee or discount.

[12] In practice, p^B is often zero or even negative (when cardholders receive frequent-flyer awards or cash-back bonuses). Also, p^B is often paid as an annual cardholder fee, independent of the number of transactions. This case is considered in Rochet-Tirole (2002), and leads to similar results.

When the network operator is operated on a not-for-profit basis, the two interchange fees coincide ($a^S = a^B = a$): This is the case studied by Rochet and Tirole (2002). We extend their analysis by allowing the network operator to keep a margin ($a^S - a^B$) on each transaction. For-profit systems may choose not to distribute substantial dividends and therefore behave as if they had a single interchange fee ($q^S \simeq a^B$); on the other hand, they may choose to make a profit on the platform ($a^S > a^B$).[13] We do not want to presume any thing on how the Portuguese systems behave and will behave in the future, and so consider both cases of a for-profit and not-for-profit system.

The timing of our model is as follows:

- The network owner (or owners) sets the two interchange fees a^B and a^S (with $a^S = a^B$ if it is not-for-profit).
- Banks compete for providing services to consumers and merchants. We do not make restrictive assumptions on the intensities of competition on these downstream markets, since they will turn out to be an important determinant of interchange fees. We just assume for tractability that all banks charge the same fees for service, p^B for issuers and p^S for acquirers. These fees are increasing functions of the net costs on each side of the market:

  $$p^B = f^B(c^B - a^B), \quad p^S = f^S(c^S + a^S),$$

  where f^B and f^S are increasing functions and c^B, c^S are the unit costs incurred by respectively issuers and acquirers for processing a card transaction. Perfect competition corresponds to the limit case where $f^B(c^B - a^B) = c^B - a^B$, and $f^S(c^S - a^S) = c^S - a^S$. In general, prices are strictly above costs: $f^B(c^B - a^B) > c^B - a^B$ and $f^S(c^S + a^S) > c^S + a^S$. As an illustration, we will often consider the case of *constant margins*: $f^B(c^B - a^B) = c^B - a^B + \pi^B$, $f^S(c^S + a^S) = c^S + a^S + \pi^S$, where π^B and π^S are strictly positive.[14]
- Merchants compete for attracting consumers, not only by setting retail prices p but also by deciding whether they accept debit card payments.
- Finally, consumers observe retail prices and which shops accept cards. They decide where to buy and which means of payment to use for their purchase.

[13]There are cases in which it could be optimal for the platform owners to choose a negative margin, i.e. to let the platform make a deficit, and therefore accept to re-inject cash periodically in the platform. We analyze these cases in Section 4.

[14]The assumption that markups π^B and π^S are fixed simplifies the analysis, but it is a strong one. One can for example assume that the acquiring and issuing sides are described as in an Hotelling/Lerner/Salop model of product differentiation.

For simplicity, we assume (as in Wright (2001) and Guthrie-Wright (2003)) that consumers are identical but that the convenience benefit[15] b^B for paying by card differs from one transaction to the other.[16] For expositional simplicity also, we assume that the cardholder chooses the retailer before knowing the specific convenience benefit b^B that he will experience when in the store. The expected net convenience benefit conditional on wanting to use the card in the store is

$$v^B(p^B) \equiv E[b^B - p^B | b^B > p^B].$$

The proportion of transactions with a gross convenience benefit greater than p^B is denoted $D^B(p^B)$ and interpreted as a "demand" function for card transactions on the buyers' side. For simplicity, we assume that merchants are identical:[17] the gross convenience benefit of a card payment for a merchant is denoted b^S.

3.2 A Model of the Portuguese Credit Card Industry

The main specificity of Redunicre, the dominant network in the Portuguese credit card industry,[18] is that acquiring activities are integrated within the system, so that banks compete only on the issuing side. This is illustrated by the following figure, which uses the same notation as Figure 4.

[15]Both convenience benefits and costs are measured with respect to an alternative payment means (cash or check) taken as benchmark.

[16]Rochet and Tirole (2002) assume instead that consumers are ex-ante heterogenous, which introduces additional complexities.

[17]The reader may wonder whether merchant inelastic demand is consistent with a fixed markup π^S. In the Hotelling model margins are compressed when the firms' (here, acquirers') marginal cost is close to the consumers' (here, the merchants') valuations. So consider the following variation on the Hotelling/Lerner/Salop model: *(i)* Fees $\{a^S, a^B\}$ are set; *(ii)* (Ex ante identical) merchants must pay a fixed cost C_0 in order to be able to use cards. They then learn their location in the product space; *(iii)* Acquirers set their prices. Then $\pi^S = t/n$ if t is the transportation cost, there are n acquirers and the circle has length 1. The average transportation cost is $t/2n$. Let b_0^S denote the gross convenience benefit. The merchants invest C_0 at stage *(ii)* if and only if $b_0^S + v^B(p^B) - \left[a^S + c^S + \pi^S + \frac{t}{2n}\right] \geq C_0$. Let $b^S \equiv b_0^S - \left(C_0 + \frac{t}{2n}\right)$. Then, we are back to the basic model. Furthermore, if C_0 is not too small, acquirer prices are not in the region where markups are compressed.

[18]We neglect for the moment the presence of a competing system (AMEX), that we introduce in Section 3.3.

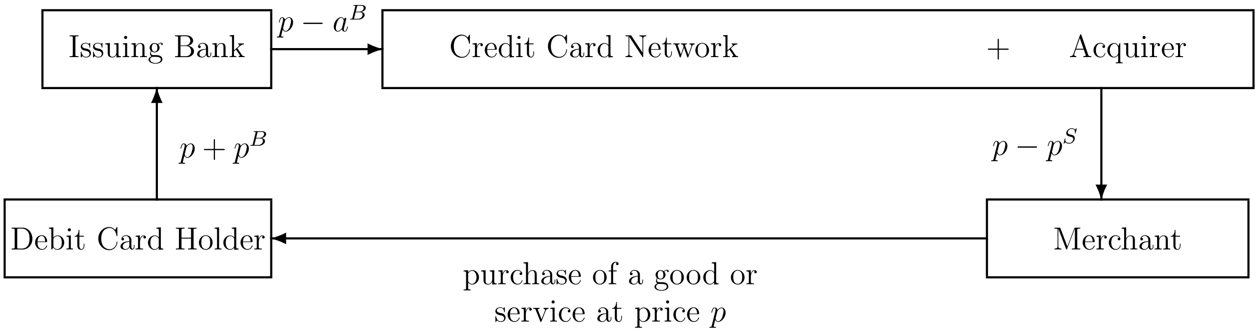


Figure 5: The flows of funds associated with a credit card payment.

The timing of our model is as follows:

- The network operator sets the interchange fee a^B and the merchant fee p^S.
- Issuing banks compete for services to consumers. The price resulting from competition on the issuing side is an increasing function of issuers' net cost $c^B - a^B$: $p^B = f^B(c^B - a^B)$. We will often consider, as an illustration, the case of constant margins: $p^B = c^B - a^B + \pi^B$.
- Merchants compete for attracting consumers, both by setting retail prices p but also by deciding whether they accept credit card payments.
- Finally, consumers observe retail prices and which shops accept credit cards. They decide where to buy and which means of payment to use for their purchase.

Like in the case of debit cards, we assume that consumers are identical but that the convenience benefit b^B for paying by card differs from one transaction to the other. Similarly, merchants are identical: the gross convenience benefit of a credit card payment for a merchant is denoted b^S.

3.3 The Impact of Network Competition

In this section, we extend our basic model of the payment card industry (under either variant: competing acquirers or integrated monopoly acquirer) by introducing two platforms (indexed by $i = 1, 2$). End users on both sides of the market then have the choice between using only one of the two platforms ("single-homing") or both ("multi-homing"). In Rochet and Tirole (2003) we analyze how network competition is affected by the intensity of consumer multi-homing, measured by a multi-homing index σ. In this report, we focus for simplicity on the two polar cases $\sigma = 0$ (single-homing) and $\sigma = 1$ (complete multi-homing).

Consider first the case of *single-homing* consumers: each consumer has at most one card. In this case, each system is effectively a monopolist on the merchant side. Indeed, the only way a merchant can benefit from a card payment by any given consumer is to accept the (unique) card owned by the consumer in question. Taking into account the increase in store attractiveness generated by card acceptance, a merchant will rationally accept card i if and only if the merchant fee p_i^S less or equal to the sum of his own convenience benefit b^S and the expected convenience benefit derived from card i by the buyer, i.e.

$$v^B(p_i^B) \equiv E[b^B - p_i^B | b^B > p_i^B].$$

Indeed, it is shown in Rochet and Tirole (2002), and also in Wright (2003c) and Guthrie and Wright (2003) that accepting card i provides the merchant with an increase in store attractiveness for which the merchant is willing to pay, besides the convenience benefit b^S, the increase $v^B(p_i^B)$ in the willingness of consumers to pay for his good or service. The merchant's net benefit of accepting card i is therefore equal to $b^S + v^B(p_i^B) - p_i^S$. In the single-homing case, card i is accepted by merchants if and only if this net benefit is non-negative:

$$p_i^S \leq b^S + v^B(p_i^B).$$

By contrast, in the polar case of complete *multi-homing* ($\sigma = 1$), the merchant knows that each consumer holds both cards in his wallet. Thus, by refusing one card, the merchant can still benefit from the convenience of a card payment by "steering" the consumer, i.e. by forcing him to pay with the other card. In the case (which we consider here for simplicity) where the two cards are perfect substitutes for both categories of users, merchants accept only the card(s) that provides(s) maximum total user surplus $[v^B(p_i^B) + b^S - p_i^S]D^B(p_i^B)$ among the two cards $i = 1, 2$.

4 The Determinants of Interchange Fees

In this section, we first determine the interchange fee that would be chosen by a social planner aiming at maximizing social welfare; we then use the benchmark model presented in Section 3 to characterize the interchange fee chosen by the system's owners as function of the structure of the industry and the intensity of competition both within and between networks. We conclude by suggesting criteria so as to determine whether interchange fees can be considered excessive.

4.1 Socially Optimal Interchange Fees

In our benchmark model, social welfare W is equal to the integral, for actual card transactions, of the difference between total user benefit $b^B + b^S$ and total provider cost $c^B + c^S$. In equilibrium, merchants always accept cards and buyers only use them if their convenience benefit for the transaction exceeds the (marginal) price p^B. Thus card transactions occur whenever $b^B > p^B$, and we can write social welfare as a function of p^B:

$$W(p^B) = \int_{p^B}^{\infty} (b^B + b^S - c^B - c^S) dH^B(b^B),$$

where $H^B(b^B) = 1 - D^B(b^B)$ represents the cumulative distribution function of the buyers' convenience benefit b^B. Notice that social welfare depends only on p^B, as long as p^S is such that merchants accept cards:

$$p^S \leq b^S + v^B(p^B).$$

Thus, if the social planner has no redistributive objectives, that is if it weighs equally the profits made by the banks and the platform(s) and the surplus obtained by final users, the socially optimal price structure is obtained when p^B maximizes W, i.e. for

$$p^B_W = c^B + c^S - b^S.$$

The merchant fee p^S only has to satisfy the inequality above. Thus there is only one socially optimal interchange fee on the buyer side, defined implicitly by

$$f^B(c^B - a^B_W) = p^B_W.$$

In the case of a constant margin π^B on the issuing side, i.e. when $f^B(c^B - a^B_W) = c^B - a^B_W + \pi^B$, the socially optimal interchange fee on the issuers' side is given explicitly by:

$$\begin{aligned} a^B_W &= b^S - c^S + \pi^B \\ &= a_0 + \pi^B. \end{aligned}$$

Notice that it exceeds Baxter's interchange fee a_0, because it internalizes the imperfectly competitive margin π^B on the issuers' side. Put differently, the interchange fee is raised so as to offset the demand contraction effect of market power on the issuing side.

By contrast, the merchants' inelastic demand creates an indeterminacy in the socially optimal interchange fee on the acquirers' side. This fee only has to satisfy the inequality:

$$f^S(c^S + a^S_W) \leq b^S + v^B(p^B_W).$$

In the case of a constant margin π^S on the acquirers' side, this inequality becomes:

$$a_W^S \leq b^S - c^S + v^B(p_W^B) - \pi^S.$$

Notice that it is compatible with budget balance of the platform only when

$$a_W^S \geq a_W^B,$$

i.e.:

$$v^B(p_W^B) \geq \pi^B + \pi^S.$$

This means that when market power in the banking industry is important (when $\pi^B + \pi^S$ is large), the socially optimal price structure is only attainable in a situation where the platform loses money.

4.2 Privately Optimal Interchange Fees

We now derive the interchange fee chosen by the network's owners as function of the structure of the industry and the intensity of competition both within and between networks. We start by a determination of the objective function of the platform.

4.2.1 The Objective Function of the Platform

The profit of the platform is given by

$$B^P = (a^S - a^B)D^B(p^B),$$

provided that merchants accept the card, i.e. provided that

$$p^S \leq b^S + v^B(p^B).$$

For simplicity we assume from now on that margins in downstream markets are constant :

$$p^B = c^B - a^B + \pi^B, \quad p^S = c^S + a^S + \pi^S.$$

We can express interchange fees as functions of final user prices:

$$a^B = c^B + \pi^B - p^B, \quad a^S = p^S - c^S - \pi^S,$$

so that the platform's profit can be written as:

$$B^P = (p^B + p^S - c^B - c^S - \pi^B - \pi^S)D^B(p^B).$$

To the extent that the platform may be owned by the main banks, its pricing policy (a^S and a^B, or equivalently p^S and p^B) may take into account not only the profit of the platform itself, but also those of the downstream banks (issuers and acquirers). To accommodate the variety of situations, we consider that the objective function of the platform is a weighted sum of the three profits:

$$B = B^P + y^A B^A + y^I B^I,$$

where B^A (resp. B^I) is the profit of acquirers (resp. issuers) and y^A (resp. y^I) measures the influence of acquirers (resp. issuers) on the board of directors of the system.

Let us provide a few illustrations of the flexibility of this objective function:

• *Not-for-profit platform controlled by issuers:* A not-for-profit platform controlled by the issuers maximizes B^I subject to the platform's break-even constraint $B^P = 0$ (and of course merchant acceptance). So $y^I = \infty$ and $y^A = 0$.

More generally, the objective function of a not-for-profit platform whose decision-making is determined through bargaining between issuers and acquirers is obtained by setting y^I/y^A equal to the relative weights of issuers and acquirers in the negotiation and either by setting both variables tend to infinity or by setting $B^P = 0$ right away.

• *For-profit platform:* Consider now a for-profit platform owned by issuers, acquirers and independent parties (the "investors"). Independent ownership or more generally a board of directors controlled by investors corresponds to $y^I = y^A = 0$. Next, suppose that acquirers control the board and own a fraction α of the platform's cash-flow rights. They then maximize $\{\alpha B^P + B^A\}$ and so $y^A = 1/\alpha$ and $y^I = 0$. And symmetrically an issuer-controlled platform maximizes $\{\beta B^P + B^I\}$ when issuers have a fraction β of the platform's cash-flow rights and so $y^I = 1/\beta$ and $y^A = 0$. Finally, one can envision a bargaining process among issuers, acquirers and investors, or a subset of these, yielding an objective function that is a weighted average of the three objective functions just derived. For example, if acquirers and issuers have cash-flow stakes α and β in the platform with $\alpha + \beta = 1$ and respective weights w^A and w^I ($w^A + w^I = 1$) in bargaining, and dividing by $(w^A\alpha + w^I B)$ in order to be consistent with the linear form posited above, then

$$B = \frac{w^A[B^A + \alpha B^P] + w^I(B^I + \beta B^P)}{w^A\alpha + w^I\beta},$$

and

$$y^A = \frac{w^A}{w^A\alpha + w^I\beta} \quad \text{and} \quad y^I = \frac{w^I}{w^A\alpha + w^I\beta}.$$

• *Coasian (efficient) bargaining*, in which all providers (issuers, platform, acquirers) reach an agreement that maximizes their total profit,[19] corresponds to $y^A = y^I = 1$.

[19] And possibly operate side transfers in order to reach a desired sharing of this profit.

The objective function becomes:

$$B = (p^B + p^S - c^B - c^S - (1 - y^A)\pi^S - (1 - y^I)\pi^B)D^B(p^B).$$

Due to possible differences between issuers and acquirers in ownership and cash-flow rights on the platform, the objective function of the platform may thus diverge from the total profit of the industry. Let us denote by $\pi = (1 - y^A)\pi^S + (1 - y^I)\pi^B$ the correcting term that appears in the above formula. By definition of y^A and y^I, we have that

$$\pi = \frac{(w^A - w^I)(\alpha\pi^B - \beta\pi^S)}{w^A\alpha + w^B\beta}.$$

When bargaining between the banks is efficient, we have seen that $y^A = y^I = 1$, and thus $\pi = 0$. This is also the case if issuers and acquirers have the same weight in bargaining ($w^A = w^I$). In the general case, π may be positive or negative, depending on whether acquirers or issuers have more weight ($w^A > w^I$ or the contrary) and whether profit margins (weighted by cash flow stakes) are bigger on the acquiring or issuing side ($\alpha\pi^B > \beta\pi^S$ or the contrary).

In the monopoly platform situation, p^S is set at the maximum level compatible with acceptance of the card by merchants,[20] i.e.:

$$p^S = b^S + v^B(p^B).$$

In this case, the objective function of the platform can be rewritten as

$$B = \int_{p^B}^{\infty} (b^B + b^S - c^B - c^S - (1 - y^A)\pi^S - (1 - y^I)\pi^B)dH^B(b^B).$$

For a for-profit platform, this expression is maximized when

$$p^B = p^B_m \equiv c^B + c^S + (1 - y^A)\pi^S + (1 - y^I)\pi^B - b^S.$$

Notice that this price structure $p^B_m, p^S_m = b^S + v^B(p^B_m)$ may imply that the platform loses money: $a^S_m < a^B_m$. In this case, the platform's owners have to inject cash on a regular basis. The above formulas give:

$$a^S_m - a^B_m = v^B(p^B_m) - y^I\pi^B - y^A\pi^S.$$

This is negative when margins π^B and π^S are high.

When the platform is not-for-profit, its objective function becomes

$$B = (y^A\pi^A + y^I\pi^I)D^B(p^B).$$

[20] This implies that total user surplus is driven down to zero.

Therefore it chooses the interchange fee that maximizes buyers' usage, i.e. the maximum interchange fee that is compatible with merchant acceptance.

We now examine in turn different possible configurations of the industry.

4.2.2 Monopoly Platform, Competition in Acquiring

Assuming away any ownership stake by independent investors, this corresponds to the model of Section 3.1. The Portuguese debit card industry can therefore be approximated by the bargaining solution derived above. The corresponding interchange fee is

$$a_m^B = c^B + \pi^B - p_m^B = b^S - c^S + y^I \pi^B + (y^A - 1)\pi^S.$$

We need to compare it with the socially optimal interchange fee

$$a_W^B = b^S - c^S + \pi^B.$$

We get:

$$a_m^B = a_W^B + (y^I - 1)\pi^B + (y^A - 1)\pi^S = a_W^B - \pi.$$

Thus, interchange fees for buyers are too high whenever π is negative.

We can also consider the case of a not-for-profit platform, i.e. add a zero-profit constraint:

$$p^B + p^S = c^B + c^S + \pi^B + \pi^S.$$

In this case, regardless of $\{y^I, y^A\}$, the platform chooses the maximum interchange fee that is compatible with merchant acceptance, leading to

$$p^B = p_A^B \equiv c^B + c^S - b^S - v^B(p^B).$$

The corresponding interchange fee is implicitly[21] determined by:

$$a_A^B = a_W^B + v^B(p_A^B) - \pi^B - \pi^S,$$

where "A" stands for "Association". Thus the interchange fee chosen by an association is too low (as compared with the social optimum) whenever the total profit margin $\pi^B + \pi^S$ exceeds net buyer surplus v^B.

[21] The determination of a_A^B is only implicit because p_A^B depends on a_A^B.

4.2.3 Competing Platforms, Competition in Acquiring

The analysis of platform competition admits several alternative hypotheses. We here assume "member duality". That is, all issuers and acquirers are members of the two platforms and therefore offer both card services to end-users. However, we assume away board duality in order to rule out collusive platform policies;[22] in particular, an issuer owning a cash flow and control right over a given platform owns no share in and has no control right over the other platform. And similarly on the acquirer side. Platforms offer perfectly substitutable cards and we look for a symmetric equilibrium.[23] The pricing strategy of the two platforms crucially depends on consumers' multi-homing index σ, i.e. the proportion of them who hold the other card and therefore can substitute when their preferred card is not accepted by the merchant. We analyze the two extreme cases $\sigma = 0$ (single-homing) and $\sigma = 1$ (complete multi-homing).

The analysis of competitive platforms is more complex than that of a monopoly platform because banks have two incarnations: as users of the networks and as owners. For example, a set of issuers may own a platform and use the rival platform if the latter offers a higher interchange fee. The following results have been proved only in four polar cases: *(i)* independent platforms, *(ii)* platforms controlled by issuers, *(iii)* platforms controlled by acquirers, *(iv)* Coasian outcome (maximization of total profit). We suspect that the results are more general.

a) Consumer single-homing:

In this case each platform is effectively a monopolist on the merchant side, since consumers have only one card. Thus platforms charge the maximum merchant fee that merchants are ready to accept, taking into account the increase in store attractiveness generated by card acceptance. Thus

$$p_i^S = b^S + v^B(p_i^B),$$

and the total surplus of final users is driven down to zero.

Competition between the platforms to attract issuers yields[24] $a^B = a^S$. This in turn implies:

$$a_{SH}^B = a_W^B + [v^B(p^B) - \pi^B - \pi^S].$$

[22] The case of perfect collusion between platforms is formally equivalent to the case of a monopoly platform studied above.

[23] In practice, the cards (MultiBanco vs Visa Electron in the debit case, or Visa-MasterCard vs American Express) are not perfect substitutes and the market share of the two networks are not equal; but our symmetric model is used as a benchmark.

[24] In the case of issuer-owner platforms (or acquirer-owned platform by symmetry, or else in the Coasian case), would the issuers want to subsidize the platform? Suppose that a given platform is owned by half of the issuers (the case for subsidies is even weaker if it is owned by fewer than half). Then it can be

Thus the outcome of platform competition with consumer multihoming is formally equivalent to that of a monopoly not-for-profit platform.

b) Consumer complete multi-homing:

Contrarily to the previous case, merchants can steer consumers, who have the two cards in their wallets. And so platforms compete to attract merchants. In the case of independent platforms, each platform chooses the price structure that maximizes total user surplus under the break-even constraint:

$$\begin{cases} \max_{p^B,p^S}(v^B(p^B)+b^S-p^S)D^B(p^B) \\ \text{s.t. } p^S \geq c^S+\pi^S+c^B+\pi^B-p^B. \end{cases}$$

After some computations, this yields

$$p^B = c^B + c^S - b^S + (\pi^B + \pi^S),$$

or

$$a^B_{MH} = a^B_W - (\pi^B + \pi^S).$$

Thus interchange fee received by issuers is too low, as compared with the social optimum.

4.2.4 Monopoly For-Profit Platform, Integrated Acquiring

This corresponds to the model of Section 3.2. In this case the profit of the acquiring activity is simply added to the profit of the platform ($y^A = 1$). Moreover, since the platform is owned by issuers, $y^I = 1$. That is, interchange fees maximize total industry profit. Because

$$p^S = b^S + v^B(p^B) \quad \text{(zero user surplus)},$$

card usage is efficient:

$$p^B = p^B_W = c^B + c^S - b^S.$$

However the social surplus is confiscated by the banks, and final users (consumers and merchants) get no surplus.

---

shown that subsidies are indeed suboptimal if

$$|(D^B)'(p^B)/D^B(p^B)| < 2/\pi^B,$$

where the factor "2" refers to the fact that the subsidy will be enjoyed not only by the owners but also by non-owners. For reasonable estimates of the relative margin $\frac{\pi^B}{p^B}$ (less than 20%), this condition holds if the elasticity of demand is less than 10, which is likely to be satisfied.

The intuition behind this surprising result is two-fold. First, issuers receive the profit of the integrated platform-acquirer entity. Thus they internalize the other profits in the production chain. This implies that there is a single monopoly marginalization and no "double or triple marginalization" that would raise prices above single-monopoly prices. Second, monopoly pricing in this benchmark induces no demand distortion; to be certain, cardholder demand is elastic and so might be inefficient low under monopoly pricing. But as long as consumers are aware of merchants' card acceptance policies, the merchants internalize the cardholders' surplus, which therefore can be appropriated by the platform by taxing merchants accordingly. This makes the platform internalize the cardholder surplus. Due to the two-sidedness of the market, it suffices here that one side of the market be demand-inelastic to avoid monopoly distortions.

The interchange fee received by issuers is the socially efficient one:

$$a_W^B \equiv c^B - p_W^B + \pi^B = b^S - c^S + \pi^B.$$

4.3 Competing Platforms, Integrated Acquiring

The formulas of section 4.2.4 are directly applicable, since they do not depend on y^A or y^I.

The next section summarizes the main implications of our results.

4.4 When Can Interchange Fees be Considered Excessive?

Our results are summarized in the following table, where we focus on the interchange fee received by issuers, a^B the level of total users surplus, and the profit of the banking sector, depending on the industry structure. We analyze in turn the debit (competition in acquiring) and credit (integrated acquiring) cases.

	Interchange Fee	User surplus	Industry's profit
Monopoly (or Collusion)	$a_m^B = a_W^B + (y^I - 1)\pi^B + (y^A - 1)\pi^S$	0	Maximum[25]
Monopoly (Not for Profit)	$a_A^B = a_W^B + v^B(p^B) - \pi^B - \pi^S$	0	Moderate[25]
Competition + Single-Homing	$a_{SH}^B = a_W^B + v^B(p^B) - \pi^B - \pi^S$	0	Moderate
Competition + Multi-Homing	$a_{MH}^B = a_W^B - \pi^B - \pi^S$	Maximum	Small[25]

Table 1: The debit card market
(The socially optimal interchange fee is $a_W^B = b^S - c^S + \pi^B$)

Table 1 shows that the privately optimal interchange fees typically differ from the socially optimal level a^B_W. However there is no systematic bias. Notice also that the interchange fee received by issuers impacts the volume of card transactions, and thus social welfare, whereas the interchange fee paid by acquirers only impacts the repartition of social welfare between user surplus and industry's profit.

	Interchange Fee	User surplus	Industry's profit
Monopoly (or Collusion)	$a_m = a^B_W$	0	Maximum
Monopoly (Not for Profit)	$a^B_A = a^B_W + v^B(p^B) - \pi^B - \pi^S$	0	Moderate
Competition + Single-Homing	$a^B_{SH} = a^B_W + v^B(p^B) - \pi^B - \pi^S$	0	Moderate
Competition + Multi-Homing	$a^B_{MH} = a^B_W - \pi^B - \pi^S$	Maximum	Small

Table 2: The credit card market

A few comments are in order:

- When there is a unique platform, (monopoly), or when several platforms collude, interchange fees received by issuers are not necessarily excessive: the difference between a^B and a^B_W may be positive or negative, depending on the market power of banks (the level of π^B and π^S), the governance structure of platforms (reflected by the values of y^I and y^A), and the average net surplus of buyers $v^B(p^B)$. When issuers and acquirers have the same weight in the decision made by the platforms, the interchange fee received by issuers is socially optimal. Otherwise , it may be biased.

- Social welfare (i.e. the sum of user surplus and industry's profit) is maximized whenever $a^B = a^B_W$. This occurs where there is a monopoly platform (or collusion between competing platforms) and either Coasian bargaining ($y^A = y^I = 1$) or integrated acquiring.

- However, when there is a unique platform (monopoly), or when several platforms collude, interchange fees paid by acquirers can be viewed as excessive,[26] since they

[25]By maximum industry's profit, we mean that the platform makes a positive profit that adds to the ones made by the banks themselves. By moderate industry's profit, we mean that the platform does not make a profit on its own. By small industry's profit, we refer to a situation where the platform does not make any profit and simultaneously, the volume of transactions is reduced.

[26]This would be the point of view of a regulator aiming at maximizing user surplus (instead of social welfare). The first line of Table 2 shows that when there is a monopoly platform (or when several platforms collude) user surplus is minimum, even though social welfare is maximized.

lead to high merchant discounts and a confiscation of total user surplus by the platform.

- Platform competition does not guarantee a positive user surplus. Inter-platform competition creates surplus for end-users only when the two platforms do not collude or when they are not-for-profit, and when consumers hold the two systems' cards (multi-homing).

- A monopoly platform with integrated acquiring selects the socially optimal interchange fee for issuers (but zero user surplus). The integration of acquiring has no impact in the other cases (not-for-profit monopoly or competing platforms).

5 Robustness of our Conclusions

For expository purposes, we have chosen as a benchmark the simplest possible model of the payment card industry. This section analyzes the robustness of our conclusions to the relaxation of several simplifying assumptions of the model.

5.1 Imperfectly Informed Consumers

In particular, we have assumed that consumers are perfectly informed about card acceptance policies by merchants, whereas Baxter (1983), in his early contribution, assumed on the contrary that merchants' acceptance decisions were unknown to consumers (or at least that they did not have any impact on consumers' choice of stores to patronize). A more general specification, that encompasses the two cases, introduces a new parameter τ, interpreted as the proportion (thus $0 \leq \tau \leq 1$) of consumers who are aware of merchants' card acceptance policy, or else the fraction of purchases made by a given consumer and for which he is aware of the card acceptance policies before choosing a store. Our benchmark model corresponds to $\tau = 1$, while Baxter's case corresponds to $\tau = 0$.

When $\tau < 1$, the maximum fee that merchants are ready to pay is reduced to $b^S + \tau v^B(p^B)$.[27] This implies that the formulas of tables 1 and 2 have to be modified. The socially optimal interchange fees are unaffected[28], at least when b^S is large enough. The interchange fees chosen by the platforms are reduced in two cases: Monopoly not-for-profit platform (second rows of tables 1 and 2) and competition between two platforms

[27] This functional form requires some further assumptions (see eg. Rochet-Tirole 2002). On the other hand, the comparative statics with respect to τ don't rely on it.

[28] This is clear for a_W^B, since p_W^B is not affected by the reduction of τ. This is also true for a_S^B if merchants still accept cards, i.e. when $p_W^S \leq b^S + \tau v^B(p_W^B)$.

when consumers single-home (third rows of tables 1 and 2). Generally speaking, merchant resistance in stronger and the likelihood that privately optimal interchange fees are excessive is reduced when $\tau < 1$. Consider for example the case of a monopoly (for profit) platform and Coasian bargaining. The objective function of the platform is:

$$B = (p^B + p^S - c^B - c^S)D^B(p^B).$$

Replacing p^S by the above formula, we get:

$$B = \int_{p^B}^{\infty} (p^B + b^S + \tau(b^B - p^B) - c^B - c^S)dH^B(b^B).$$

This is maximum for:

$$p^B = p_m^B \equiv c - b^S + \frac{(1-\tau)}{\eta},$$

where $\eta = -\frac{(D^B)'}{D^B}$ is the semi-elasticity of buyers' demand. Finally

$$p_m^S = b^S + \tau v^B(p_m^B).$$

Thus the price paid by cardholders is too high (as compared with the socially optimal price $p_W^B = c - b^S$).

5.2 Non-Observable Merchant Heterogeneity

Another simplifying assumption was that the merchants' convenience benefit b^S for a card payment was identical accross merchants. A straightforward reinterpretation of our model accommodates observable heterogeneity, since platforms can select interchange fees that depend on observable characteristics of merchants. For example, supermarkets typically face lower interchange fees than other retail stores. However, the empirical observation of actual refusal by merchants suggests that unobserved merchant heterogeneity may be a relevant factor. However, if unobservable heterogeneity across merchants is introduced, the situation is more complex, since different merchants are then confronted with non-differentiated merchant discounts, which generates a (finite) elasticity of merchant demand for cards. In this case, our conclusions have to be amended. Rochet and Tirole (2003) show the socially optimal price structure becomes relatively complex, as it depends on demand elasticities and average surpluses on both sides of the market. In the case of a monopoly, not-for-profit platform, Rochet (2003) shows that the platform selects an interchange fee that may be too high or too low, as compared with the socially optimal one. In other words, there is not systematic bias. More specifically, the privately optimal interchange fee (the one chosen by the platform) is too low (as compared with the social optimum) when the average net surplus of sellers $v^S(p^S) = E[b^S - p^S | b^S > p^S]$ is smaller than the average net surplus of buyers,[29] $v^B(p^B)$. A more general analysis can be found in Rochet (2003), which confirms that there is no systematic bias between socially optimal and privately optimal interchange fees.

5.3 Fixed Fees

We have assumed that buyers pay per-transaction fees, whereas in practice they often pay fixed (yearly) fees. In Rochet and Tirole (2002), we study the impact of such fixed fees in a model where consumers differ ex-ante in their convenience benefit b^B for a card payment (non-observable consumer heterogeneity). The results are similar to those of the present model, the main difference being that net buyer surplus $v^B(p^B)$ has to be replaced by gross buyer surplus $\beta^B(p^B) = v^B(p^B) + p^B$. This increases the likelihood that interchange fees received by issuers are excessive.

5.4 How Proprietary Systems Perform the Balancing Act

In proprietary (three-party) systems (such as American Express in most countries or Discover), the network is the only intermediary in the payment card transaction between a cardholder and a merchant. The flow of funds is described in Figure 6.

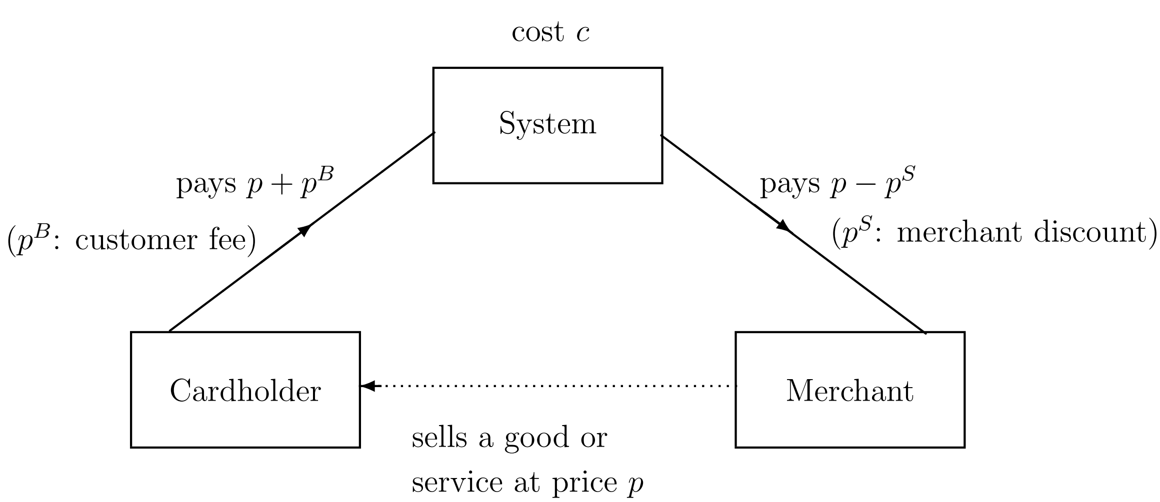


Figure 6: Flow of payments in a card transaction within a proprietary system.

[29]This condition is very rough, and does not provide a simple empirical test of excessively high interchange fees, since these surpluses are endogenous. However it is clearly satisfied when merchants are identical, since then $v^S = 0$.

The levels of the customer fee p^B and the merchant discount p^S are chosen by the managers of the proprietary system so as to maximize its profit. We can easily adapt our model of Section 3 to study the determinants of p^B and p^S, the prices faced by final users.

Consider for example the case of a monopoly system, and suppose that only a fraction $\tau < 1$ of the buyers are aware of merchants' acceptance decisions. The maximum fee that merchants are ready to accept becomes

$$p^S = b^S + \tau v^B(p^B).$$

The profit of the platform is

$$B = (p^B + p^S - c^B - c^S)D^B(p^B).$$

Replacing p^S by the above formula, we get:

$$B = \int_{p^B}^{\infty} (p^B + b^S + \tau(b^B - p^B) - c^B - c^S)dH^B(b^B).$$

This is maximum for:

$$p^B = p_m^B \equiv c - b^S + \frac{(1-\tau)}{\eta},$$

where $\eta = -\frac{(D^B)'}{D^B}$ is the semi-elasticity of buyers' demand. Finally

$$p_m^S = b^S + \tau v^B(p_m^B).$$

Thus the price paid by cardholders is too high (as compared with the socially optimal price $p_W^B = c - b^S$). Notice that the analysis is identical to the case of an open platform that maximizes the industry's profit (Coasian bargaining, i.e. $y^I = y^A = 1$).

In the polar case of competition between proprietary platforms offering perfectly substitutable cards we obtain instead:[30]

$$p_c^B = p_m^B,$$

and

$$p_c^S = c - p_c^B.$$

Thus perfect competition between proprietary platforms leads to a positive user surplus (and zero profit for the platforms) but does not correct the price distortion for buyers. In the case (studied here) where merchants' demand is inelastic, prices are unambiguously too high for buyers. In a more general model with elasticities on both sides, the direction of the bias may be reversed.

[30] The result is the same whether consumers have a single card (single-homing) or both (multi-homing).

5.5 Elastic Consumer Demand

We have also assumed so far that consumers' demand functions were inelastic. This allowed us to focus on the choice of payment instruments, for a fixed volume of trade. This assumption is very strong, but it simplifies the analysis. Relaxing it would allow to look at the impact of changes in the prices of card payments on consumer demand. Some commentators have argued that high IFs resulted ultimately in a reduction of consumer demand, given that merchants are likely to pass through the increase in the merchants discounts into the retail prices charged to consumers. Even if a complete analysis would be necessary to fully assess this statement, we suspect that it is flawed. The reason is that it neglects the convenience benefits of card payments for final users, which correspond in large part to the cost savings associated to not having to use cash or checks. Once all costs and benefits are properly taken into account, empirical evidence shows that card payments are on average more efficient than paper instruments. Thus Consumer Price Indices are not the appropriate metric for assessing the impact of IFs on consumer demand. Indeed consumer demand is likely to react to changes in "net" retail price, i.e. the retail price net of the average cost of payment instruments. In our model, this net retail price varies one to one with total user surplus. Introducing elastic consumer demand would therefore not alter significantly our results.

6 Regulatory and Antitrust Policies for Payment Card Networks

Interchange Fees (IFs) differ across countries and across transaction types. In several regions of the world (European Union, Australia, Israel), the mode of determination of IFs has come under scrutiny by competition authorities, spurred by complaints formulated by large retailers' associations. The purpose of this section is to build on the results of our theoretical model, presented in Section 5, and to lay out, in a non-technical fashion, the basic economic principles for choosing an interchange fee. We also examine the case for a public regulation of interchange fees.

This section is structured as follows. Subsection 5.1 recalls the fundamental two-sidedness of payment activities and the need to balance the demands of the two sides of the market (cardholders and merchants). Subsection 5.2 studies how payment card associations perform this balancing act by setting the appropriate IFs. Subsection 5.3 examines whether there is a case for a regulation of IFs by competition authorities. Subsection 5.4 concludes by comparing the payment card industry with regulated industries.

6.1 Payment Systems Are Two-Sided Markets

As we saw in Section 2, the fundamental characteristic of payment card systems is that every card transaction necessarily involves two users: a cardholder and a merchant. Thus, it is appropriate to view payment card systems as providing interdependent services to cardholders and merchants. Cardholders benefit from their holding a card only if their cards are accepted by a wide range of merchants, and merchants benefit from the card only if a sufficient number of consumers use it. Therefore, a payment card network can function effectively only if sufficient numbers of both cardholders and merchants participate in the network. To this "membership externality" must be added a "usage externality". In particular, merchants do not benefit from their patrons' holding a card if the latter use it only sporadically.

Thus, it is crucial for payment networks to find an effective method for balancing the prices on the two sides of the market. Payment card systems are but one of many examples of such two-sided markets. Detailed analyses of other examples, such as the software industry, videogames, internet portals, medias, and shopping malls are provided by Armstrong (2004), Evans (2003), and Rochet-Tirole (2003) for example (see below for a few illustrations). In all these industries as well, the crucial challenge for the platforms is to get both sides of the market "on board" while making a profit overall.

Is It Justified to Charge Different Prices to Different Users?

The determination of prices as a function of demand elasticities is familiar from other industries. Consider for instance airfares: a flight between any two destinations involves a common cost associated with the operation of the plane. The low prices granted to the elastic side (students, elderly, week-end stay over passengers) and the high fares for inelastic customers (businessmen) allocate the common cost so as to keep "everyone on board", so to speak. Charging different prices to different users, while resulting from profit maximization, is not a distortion due to market power: it is a socially efficient way of recovering the common cost while providing services to a larger number of users than would be otherwise the case.

In the case of two-sided markets, such as payment networks, where there exist positive externalities between different categories of users, it may even happen that one side of the market is left entirely free of charge. For example, cardholders are often exempt of transaction fees. This happens when cardholders are highly resistant to transaction charges. Put differently, economists' concerns about monopoly distortions relate to the price *level* and not to the price *structure*.

Another example is the pricing policy of Adobe (and many other software programs) in the software industry. Adobe Inc. sells Acrobat Exchange, the software needed to transform electronic files into the PDF format. The economic value of this software comes in great part from the fact that any potential reader can download for free the complementary software, needed to read PDF files (Acrobat Reader) on Adobe's web site. This reader is a "damaged" version of Acrobat Exchange. This price structure results from the fact that readers have a lower willingness to pay for the software than writers.

Other examples involving substantial cross-subsidies between the two sides of the market include portals, free TV networks and low-price newspapers, in which "eyeballs" are highly subsidized by advertizers; and videogames, for which consoles are sold below cost to game players by platforms.

6.2 How Do Payment Card Associations Perform the Balancing Act?

A payment card association such as Visa has two key characteristics: It is *open* and *not-for-profit*. The two properties need not come together: An infrastructure or a piece of intellectual property that are licensed on a non-discriminatory basis are open, but often for-profit. Conversely, a number of for-profit associations are closed. A proprietary platform such as Amex is both closed (actually in an extreme form: it is mostly vertically integrated into acquiring and issuing and in a number of countries does not license other issuers and acquirers) and for-profit. The openess and budget balance properties of Visa in turn imply two important distinctions with, say, Amex:

As we have seen in Section 3, in an open system the network's influence on the determination of final user prices is only indirect: it operates through the setting of interchange fees between member banks. Final users' prices p^B and p^S are not fixed by the network but result from competition among member banks within the network in the issuing and acquiring markets.[31] By contrast, a proprietary system directly sets p^B and p^S and receives the associated revenues (see Figure 6).

Second, like for-profit systems, a cooperative system has to perform the balancing act between the two sides of the market, and allocates the total cost between them in a proper fashion. However, as we noted above, it can do so only indirectly through the level of the interchange fee (IF).

[31] Note that a change in a in a induces changes in p^B and p^S, although not necessarily one for one, except if downstream markets are perfectly competitive.

6.2.1 The Impact of Banning the No Discrimination Rule

As discussed in Section 2, it is often argued that merchants' charging different prices for cash and card transactions is an alternative mechanism to the IF for internalizing externalities between merchants and cardholders, and that the no-discrimination rule (NDR) should be banned. The NDR is a rule that in some countries prevents the merchants affiliated with payment card networks from charging different prices for customers who pay by cash (or check) and customers who pay by card.

Despite the caveats stated in Section 2, let us assume that payment-based price discrimination by merchants is costless. The IF becomes totally ineffective if all merchants charge different prices for cash and card payments.[32] Indeed, in the model, merchants fully pass their net cost for a card payment (merchant discount minus convenience benefit) through to the customer who pays by card. If cardholders' fees to issuers are exactly proportional to transactions (in particular there is no yearly fee) and if cardholders are perfectly informed before selecting a merchant of card and cash prices charged by competing merchants, the choice of the payment instrument is determined by the sum of the fees paid by cardholders and merchants, which implies that the IF no longer plays a role.

Furthermore, if the issuing and acquiring markets are perfectly competitive, the prices faces by merchants and cardholders are not distorted by market power and so the externality between merchants and cardholders is perfectly internalized, resulting in an efficient usage of cards, as suggested by Katz (2001).

However, when the issuing and/or acquiring markets are not fully competitive, the second leg of the reasoning no longer applies. For example, we show in Rochet and Tirole (2002) that when issuers have market power, banning the NDR results in a systematic underprovision of card services, while with the NDR in place, the IF chosen by an association can result in an efficient card usage.

In a different vein, Wright (2003a) shows that when merchants have market power and cardholders' payments to issuers are not proportional to transactions,so that consumers pay a non-negligible yearly fee, merchants may be able to extract consumer surplus from card usage, destroying incentives for holding payment cards.

6.2.2 Does the No-Discrimination Rule Generate Distortions in the Choice of Payment Instruments?

It is sometimes argued that cash (and check) users are penalized by the NDR, since undifferentiated retail prices incorporate average transaction costs of merchants (including

[32] Rochet-Tirole (2002). Gans and King (2001), (2003) show that this neutrality of IFs when the NDR is banned holds under very weak assumptions.

merchant discounts for card transactions). Thus, customers who pay by cash incur a fraction of the costs of card payments, which provides an additional incentive for them to switch to cards. However, this does not imply that IFs selected by card associations distort the choice of payment instruments by consumers in favour of cards and at the expense of other payment instruments like cash or check. For, this choice is also distorted by the markups charged by issuers and acquirers, if these intermediaries enjoy market power. In a situation of under-utilization of payment cards, policies (such as the NDR) that encourage card usage increase rather than decrease welfare.[33]

A second argument in favor of the NDR builds on the fact that, in most countries, the prices of these other instruments may not fully reflect their costs. In particular, in many countries, check users do not internalize the costs of these checks, because there are no check fees for the vast majority of users. As a consequence, surcharging card users, while check users do not pay for their costs, would certainly not go in the direction of economic efficiency but instead hinder the development of electronic transactions.[34]

6.3 Is There a Case for Regulating Interchange Fees?

Merchants have repeatedly lobbied competition authorities to regulate IFs. This is easy to understand: retailers have a vested interest in bringing IFs down,[35] since such reductions would be partly reflected in reduced merchant discounts. Of course, cardholders have a vested interest in high IFs since decreases in IFs would cause cardholder fees to rise. A cardholders' lobby, assuming that it existed, would have some vested interest to push for a regulation that would increase IFs and thereby raise cardholders' benefits.[36]

As a general rule, public policy should be guided by social interest rather than special interests, and should rest on a full-fledged economic analysis of market failures. This section examines the market failures that have been advanced in relation with the use of IFs within payment card associations and presents a skeptical view concerning the likely efficacy of public regulation of the associations' IFs.

[33] Provided that they do not "overshoot", that is that they do not create overconsumption.

[34] This argument applies to the present situation in most European countries, where the costs of checks are not reflected in their prices for users. A different reasoning would apply to countries like Canada, where check fees are high.

[35] This is only true up to a limit since a low interchange fee reduces the number of cardholders: A decrease in merchant discounts, compensated by an increase in cardholder fees, may be detrimental to the welfare of merchants themselves. Indeed, when the elasticity of cardholders demand is high, even a small increase in cardholder fees may result in a large decrease in cardholders' card usage, and thus in the economic value of cards for merchants.

[36] Again, up to a limit, since cardholders want their card to be accepted by merchants.

6.3.1 The Interchange Fee is not a Fee for Service

The proponents of a cost-based regulation of interchange fees implicitly refer to a vertical structure, considering that issuers provide an intermediation service to acquirers, who then supply the final service to merchants. This view is erroneous because it ignores the role of cardholders as consumers of the payment services, on an equal footing with merchants. It fails to recognize the two-sidedness of the payment card industry, and the resulting need for balancing the demands of the two types of users. Unlike a fee for service in a vertically organized industry, the interchange fee affects not only the marginal cost of merchants, but also the size of the cardholder clientele and the usage of cards.[37]

6.3.2 Concerns About Anti-Competitive Behaviour in Relation With Interchange Fees

Do we expect the IFs chosen by platforms to be set at the "wrong level"? The recent theoretical literature (see in particular Rochet and Tirole (2002), (2003), Wright (2001)) shows that, although the socially optimal and privately optimal levels of IFs both depend on the same factors (issuers' and acquirers' costs, issuers' and acquirers' margins, cardholders' and merchants' demand elasticities), they are not equal in general. However, given the profession's current state of knowledge, there is hard to predict that the IFs chosen by a platform are systematically too high or too low, as compared with socially optimal levels. That it is not in the interest of a platform to choose IFs that deviate markedly from social optima, comes from three factors. First, network externalities imply that weakening the other side of the business reduces the demand on one's own side. Second, competition within networks implies that a reduction in the issuers' marginal cost of doing business is partly or fully competed away in favour of cardholders. Finally, competition between networks implies that merchants and/or cardholders can switch providers when one network decides to increase its prices. Let us briefly examine these three factors

Network externalities: Even a monopoly issuer (respectively, a monopoly acquirer) would not benefit from a very high (respectively, very low) IF. A high IF would result in substantial merchant resistance and would induce many merchants to reject the card. Therefore, even though a very high IF would result in a low marginal cost to the issuer of offering payment services to cardholders, it does not benefit the issuer if the cardholders reduce the use of the card because many merchants do not accept it. Symmetrically, even a monopoly acquirer would exercise restraint in setting the IF since a very low IF would lead to a

[37] Moreover it must be kept in mind that, in the case of an association, the IF is not retained as profit, but rather goes toward lowering the net costs on the side of the system that receives the IF. Ultimately, issuers receiving interchange tend to pass on some of the reduction in net costs in the form of lower cardholder prices.

correspondingly high cardholder fee, and so would discourage consumers from holding and using the card. Thus network externalities by themselves induce *some* restraint.

Within Network Competition: An increase in the IF does not go just into the issuers' pockets. It reduces the marginal cost of all issuers, and correspondingly reduces the price charged to cardholders. Indeed, some standard models of competition used in economic theory[38] predict that the increase in the IF is fully passed through to cardholders. While there may not be a full pass-through to cardholders, it is reasonable to assume that much of the increase in the IF is competed away (passed through to consumers). In this case, issuers gain little directly from an increase in the IF and would lose if this increase in the IF induced a substantial fraction of merchants not to take the card. To sum up, the fact that an increase in the IF is partly competed away reinforces the benefits to issuers from exercising restraint.

Between Networks Competition: Competition between networks is an additional force that restrains payment associations in their choice of prices. Suppose that issuers control the association and therefore would like to levy a high interchange fee. When the association competes with other payment systems and when a substantial fraction of cardholders multi-home, though, merchants are tempted to reject the cards of payment systems that charge high merchant discounts (see Section 4). System competition combined with cardholder multi-homing therefore force issuer-controlled association to exercise further restraint in their choice of interchange fee.

Rochet and Tirole (2002) and Guthrie and Wright (2003) show that competition between not-for-profit networks does not necessarily lead to a higher social welfare than a monopoly network. The intuition goes as follows: First, a standard benefit of competition is to bring the price *level* down. However, this benefit of competition does not exist when systems are associations whose not-for-profit status prevents exercise of market power. Competition therefore alters the price *structure* (the allocation of cost between the two sides of the market) rather than the price level. As shown in these papers, competition may lead systems to leave a large rent to a side of the market that has a credible threat to sever its relationship with one of the systems. As this rent comes at the expense of the other side of the market, the balancing act is perturbed.

6.4 Comparison with Other Industries

To be certain, the fact that an association's members, whether predominantly issuers or acquirers, have the socially worthwhile objective to bring both sides of the market on

[38] These include the perfectly competitive model and the Hotelling model of product differentiation.

board does not mean that the association, whoever controls it, will pick just the socially optimal IF. But this feature is not specific to the payment card industry: no industry ever engenders the socially optimal decisions. There is no reason to believe that airlines select the perfect bundle of routes, frequencies and prices, that patent holders perfectly maximize social welfare in their licensing choices, or that biotechnology start-ups perform the socially optimal amount of R&D.

The standard approach to public intervention in industries involves two steps:

a) the theoretical identification of a serious market failure and the validation of its empirical relevance,

b) the identification of the least distortionary way of addressing the market failure and a check that the remedy will not be worse than the illness.

For example, the regulation of telecommunications, electricity and railroad industries has traditionally been based on a broad intellectual consensus that certain segments represent natural monopolies and provide their owners with incentives to charge largely inflated, distortionary prices (part a)). Concerning part b) there has been much debate as to the proper mode of regulation as well as again a broad consensus that regulation itself introduces non-negligible distortions. Yet, most economists feel strongly enough about part a) that they are willing to accept the need for regulatory intervention in those industries, in spite of the concomitant regulatory distortions.

Proponents of a regulation of the IF must first build a theoretical paradigm that gathers broad intellectual consensus and demonstrates a clear market failure, show that the resulting distortions have a clear sign and a sizeable impact on welfare, and propose a form of regulation that is consistent with the underlying theory and is better that non-intervention. So far, no such theoretical paradigm has been achieved. On the contrary, recent academic work concurs to establishing that there is no systematic bias in the IFs selected by cooperative networks: there is no a priori theoretical reason to think that privately optimal IFs are higher or lower than socially optimal ones. By contrast, misunderstanding the economics of the two-sided markets and imposing cost-based regulation could impose substantial distortions in the industry.

A cost-based regulation of the IF might be an unfortunate precedent for two-sided markets. The same logic would then imply that advertizers' fees paid to TV networks, newspapers and portals[39] should be regulated on a cost basis so as to stop the subsidization of eyeballs by advertizers; that videogame developers[3] be entitled to regulated royalties and development kits' prices, and to above-cost console pricing; that the Internet[3] should

[39]These examples are developed in several mini-case studies in Evans (2003) and Rochet and Tirole (2003).

be regulated so as to stop the subsidization of websites by dial-up customers through bill-and-keep; that software reader programs[3] be charged the same price as software writer programs (they cost the same); and that social gatherings should be regulated so as to prevent payments to or free entry for attractive participants (e.g., celebrities) while others pay for entry.[3] We do not think that these implications are intended by the proponents of cost-based IFs.

6.5 Should Networks be Treated as Essential Facilities?

First, recall the definition of an "essential facility" or "bottleneck": An essential facility is an essential input, which is controlled by a dominant firm, that has no objective reason to deny competitors access to the essential facility. To qualify as an "essential input", it does not suffice that having access to it reduces the competitors' cost of production. Rather, duplicating the input must be "prohibitively costly", whatever this means. An "objective reason" to deny access may refer to a lack of capacity to accommodate new users of the essential facility (and a high cost of expanding this capacity), or to a technological incompatibility (that is costly to resolve), or else to rivals' access threatening intellectual property protection.

The concept of essential facility has figured prominently in regulated environments[40] and in industries solely subject to competition authorities' oversight.[41]

The first step in thinking about whether access should be mandated consists in assessing whether the facility is indeed essential: Can competitors in the downstream market procure the input in an alternative way at a reasonable cost? An example of his kind of reasoning is provided by Advocate General Jacobs in the Oscar Bronner (1998) decision. The Oscar Bronner case involved a morning delivery service in Austria (the Mediaprint newspapers gave access to this service to Wirtschaftsblatt, an independent but non -competing newspaper, but denied access to Der Standard). The Advocate General looked at alternative delivery systems, and further argued that while Der Standard was too small to have its own delivery system, getting access to the Mediaprint system would reduce its incentives to build a rival delivery network in cooperation with other independent newspapers. The Advocate General thereby re-stated the standard trade-off between access policies and system competition.[42]

[40] Regulation of local loops in telecommunications, of transmission grids in electricity, or of rail tracks and stations in rail transportation.

[41] Classic examples include Terminal Railroads (1912), Commercial Solvents (1973), various Computer Reservation Systems cases, wholesale markets (Gamco 1953, New York Stock Exchange 1963), stadiums, ports, airports, tunnels, as well as recent European cases involving intellectual property (Magill 1995, Mylan 1999 and IMS 2001).

[42] This trade-off is for example encountered in the case of the local loop in telecoms: Should competition be created by a) local loop access for specific services, b) local loop resale to competitors, or c) alternative

Second, and most difficult, is the question of the access price level. Unlike their regulatory counterparts, competition authorities often refrain from specifying access prices and content themselves with calling for a "reasonable rate" or "adequate financial compensation". Whether access should be granted to new entrants on a non-discriminatory basis depends on the level and depreciation of the investments sunk by the incumbents. At one extreme of the spectrum lies intellectual property law that, except in exceptional circumstances, does not mandate any open access ("compulsory licensing") to the intellectual property. Patents, trade secrets and other pieces of IP often fulfill the criteria enunciated in the definition of essential facilities. Yet, there is widespread recognition —and the legal framework— in support of the idea that investments must be rewarded and that an open access policy would destroy the incentive to innovate. Likewise, it would be inefficient to let companies standing on the sideline receive cheap access to an essential facility once investment has been sunk and technological and demand uncertainty resolved.

At the other extreme of the spectrum lie those essential facilities for which little investment has to be amortized, because the essentiality results from the granting of legal monopoly (say, an airport) or from the quasi-impossibility of duplicating the facility or reaching a minimum scale to reap network externalities (perhaps the case of a wholesale market in a medium-size town), rather than from investment or innovation. Open access should then be mandated.

As a general rule, competition authorities should be wary and carefully weigh the pros and cons of becoming de facto regulators of access. Yet, if a network a) has little non-amortized investment that its owners should see rewarded, b) is very difficult to duplicate, and c) excludes entrants, a policy of mandatory access sounds reasonable.

7 The Implications of Duality

How does board duality —the fact that the same banks sit on the boards of the two systems— impact industry outcomes? This question arose in the early 2000's for payment card networks, but in the context of *not-for-profit* networks. The US Department of Justice filed a complaint against Visa and MasterCard seeking, among other things, to force banks sitting on the board of one of the two associations to have an exclusive relationship with that association; that is, DoJ attempted to eliminate board duality (that had existed since the 1970s). The complaint turned out to be moot for three reasons: First, MasterCard abandoned its not-for-profit status and became a for-profit entity. Second, Visa and MasterCard responded to the complaint by effectively dropping board duality. Finally, the Court judge did not rule against duality anyway.

local loops?

In the case of two for-profit companies in general industries, antitrust authorities are highly suspicious (and rightly so) of monopolization through overlapping boards. The elimination of competition raises price to end-users and inefficiently reduces volume in the industry.

The standard industrial organization foundations for this hostility to overlapping boards however must be amended in the case of payment systems, because of their two-sided nature. In a nutshell, cartelization indeed has the potential to raise the price level, that is the total price charged for a transaction to the acquirer and the issuer combined; but it often leads to a better price structure, that is to a more efficient allocation of charges between the two sides.

To illustrate this, consider the case in which cardholders have a single debit card (in technical terms, they "single-home"). The first point to note is that in the absence of overlapping boards, the competing systems would compete only for cardholders (through issuers): See Section 4. On the merchant (acquirers) side, the systems are monopolists anyway since they have exclusive access to the cardholders under single-homing. Thus, the system-competition game consists in attracting cardholders by all means and then levy monopoly prices on merchants for the unique access to these cardholders. This one-sided-competition/ one-sided-monopoly situation results in very low price to cardholders and very high merchant discounts. A monopoly platform —or two platforms colluding through overlapping boards— charges more balanced prices, resulting in a more efficient allocation of resources. Indeed, one can construct simple examples in which end users (that is, cardholders plus merchants) in aggregate gain nothing from competition, while competition eliminates platform profit. This overall decrease in welfare results from an excessive card usage by cardholders.

Turning to the polar case of widespread multi-homing (all cardholders have the two systems' cards), the nature of competition is rather different. Under cardholder multi-homing, no system offers a unique access to a given cardholder. Merchants then have a strong incentive to accept solely cards of the system that is most cost effective for them. Platforms now compete for merchants as well as for cardholders.

In the current state of our knowledge, however, we do not feel that competition authorities should refrain from investigating the impact of duality. There are several reasons for this:

• Board duality under *for-profit* platforms has other consequences besides price level and structure. In particular, systems may collude to reduce competition in investment.[43]

• Alternative institutional arrangements, such as a not-for-profit, open-access status for

[43] See Guerin-Calvert-Ordover (1998).

platforms, are feasible, that solve the monopoly-price-level problem without necessarily inducing the competitive distortions. As shown by Hausman et al (2003), board duality, without being a panacea, has some attractions in this not-for-profit context. First, overlapping boards do not raise the price level when the platforms break even; so the classic argument underlying section 1 of the Sherman Act and article 81 of the EC Treaty does not hold. Second, unlike with for-profit platforms, overlapping boards do not induce underinvestment: the break-even condition ensures that investment costs are passed through to the issuers and acquirers, and so the systems are induced to trade off the cost of investment and the improved service provided to issuers and acquirers (and ultimately to end-users); in contrast, overlapping boards have an incentive to under-invest as part of the cost of investment comes in the form of a reduction in dividends (part of the benefits are competed away through platform competition). Third, coordination between the platforms may reduce some forms of wasteful competition (on the price structure —see above—, incompatibility of technologies, wasteful advertising) while still allowing some benchmarking of the two managerial teams.

8 Foreclosure and Tie-Ins

Buyers use both credit and debit cards. The credit facility brings about substantial benefits to some consumers, for some types of purchases or at specific moments of time. In other circumstances, credit is not needed.

Payment card associations, Visa and MasterCard,[44] offer both debit and credit cards and, until recently, engaged in a tie-in on the merchant side through the honor-all-cards (HAC) rule. In a class action initiated by WalMart (and involving more than five million U.S. merchants), this rule has come under attack on the grounds that the credit and debit card markets are separate markets and that the associations lever their market power in the "credit card market" (the tying market) to exclude on-line debit cards and thereby monopolize the "debit card market" (the tied market). Visa and Master Card agreed to abandon their HAC rules and pay over $ 3 billion in damages to the merchants. The objective of this section is to analyze the impact of the HAC rule.[45]

We extend our model of the payment card industry to include simultaneously two types of transactions, debit and credit, which may be either unbundled or tied on the merchant side.

To understand the impact of a tie-in, let us return to the factors relevant to the

[44]MasterCard is now for profit, but it was until recently an association.

[45]Tying of credit and debit is (currently) not an issue for Portugal, since credit and debit cards are managed by different networks (Red Unicre for credit cards, Red Multibanco for debit cards). However, this report aims at giving a broad picture of current issues in payment cards networks all over the world.

determination of merchant and cardholders' surpluses in the case of a single card (or under unbundling when multiple cards co-exist). With a single card, an association's choice of interchange fee is constrained in two ways:

- Even if the association faces no competition from another system, it must get both sides (consumers, merchants) on board. The interchange fee must be high enough so as to induce consumers to use the card, but low enough so as not to meet merchant resistance.[46]
- When competing with other payment systems, the association is further constrained, as each system tries to de-stabilize its rivals' balancing act, when consumers hold multiple cards: Systems try to "steer" merchants by lowering the interchange fee so as to incentivize them to turn down the rival card.

A not-for-profit association cannot exercise its market power by inflating the overall price *level*. By contrast, the association has discretion in the allocation of cost between cardholders and merchants and, like ordinary firms, it may or may not get the price *structure* (i.e., the relative prices for different end users) "socially right". On the one hand, both a social planner and an association ought to design the price structure so as to account for the elasticities on both sides of the market and thereby get both sides on board. On the other hand, the literature has identified factors, such as downstream (issuer, acquirer) market power or merchants' competition for market share, that may tilt an association's (or, for that matter, a proprietary system's) price structure away, in either direction, from the socially optimal one.

System competition is one such factor. Leaving aside the standard benefits of competition on managerial incentives,[47] system competition has an ambiguous impact on welfare because it influences only the price *structure* and not the price *level* (that must track cost, due to the not-for-profit status).

As noted above, when consumers hold multiple cards (multi-home), system competition tilts the price structure toward lower merchant discounts and higher cardholder fees, because merchants then have an incentive to turn down the card that is most expensive to them.[48]

[46] Such inducements are indirect to the extent that an increase in the interchange fee is partly or fully passed through by issuing banks to cardholders and by acquirers to merchants.

[47] E.g., through the owners' ability to benchmark their management's performance.

[48] This is most clearly the case when cardholders are unaware of the merchants' card acceptance policies before purchasing and so merchants do not make their store more attractive by accepting a card (the case studied by Baxter (1983)). This is also true in our model, but then merchants have to internalize the average surplus obtained by consumers when they pay by cards. The criterion that merchants adopt for accepting a card is then total user surplus.

Interestingly, steering is effective even when consumers are informed of the merchants' card acceptance policies and so are willing to pay more if the merchant accepts the card (as is assumed in the most of the treatment below). In the latter case, the card that is cheaper for the merchants is also less attractive for the cardholders (since the interchange fee is lower, cardholder benefits are lower — or their fees higher). When merchants are homogenous (or more generally when merchant heterogeneity is observable), competition reduces the interchange fee. In the polar case of a monopoly system, the association, in order to maximize volume, chooses an interchange fee equal to the highest value that merchants will bear. In the opposite polar case of perfect system competition, the equilibrium interchange fee is the (lower) one that maximizes total user surplus. Intuitively, merchants prefer the card that gives them the highest sum of their own surplus (convenience benefit minus merchant discount) and of the cardholder's average surplus (convenience benefit minus cardholder fee), since they internalize the latter when trying to attract customers. Under monopoly, the system tries to please consumers in order to maximize volume; under competition, merchants have an important say on the usage of a specific card, because rejecting a card is much less costly to them if cardholders multi-home; and so merchants receive a better deal (and cardholders a worse deal).

System competition focuses an association's attention on the system's own elasticities rather than on the socially more relevant end-user elasticities. And so, whether competition improves social welfare depends on the price-structure bias of a monopoly system. If the merchant discount (or the interchange fee) is initially too high, then competition forces it down and may improve welfare. By contrast, competition reduces welfare if the merchant discount (or the interchange fee) was initially too low.

The issue at stake, however, is not whether competition improves welfare, but whether, given competition, tying (the HAC rule) increases or decreases welfare. A first intuition might be that "bundling reduces competition, and so, if competition is socially desirable, bundling reduces welfare as well". This intuition turns out to be incorrect. We show that, when merchants are homogenous, *regardless of the desirability of system competition, the HAC rule always improves welfare.*

To see why this is so, suppose that consumers are informed of the merchants' card acceptance policies, and that a system issues both credit and debit cards, and faces more intense competition on one segment. To simplify the exposition, the system is a monopoly on credit cards and faces an on-line competitor for debit cards that is a perfect substitute. Then in the absence of the HAC rule, the outcome is the monopoly outcome for credit and the competitive outcome for debit. From our previous analysis, the interchange fee is higher on credit than on debit. This interchange fee structure is not the one predicated by the gross demand specificities of the two-sided debit and credit markets, but rather reflects the difference in the merchants' "bypass opportunities".

Suppose now that a merchant has to accept the system's two cards or none. Accepting the two cards can be an equilibrium behavior for merchants only if the total user surplus (which is also the merchants' reservation utility from accepting the system's cards) is greater than in the absence of the HAC rule, which is equal to the highest total user surplus that can be offered by the on-line network. The system gains *flexibility to rebalance its interchange fee structure* as the competitive constraint binds over the set of cards, rather than over each card. The system can therefore *increase volume* by raising the interchange fee on debit and lowering interchange fee on credit. Social welfare always increases.

There are a number of key differences with the rest of the literature on tying.[49] First, we analyze tying by an association. Hence, anticompetitive motives, like entry deterrence, cannot be associated with the standard purpose of raising the price level (an association can only affect the price structure).[50,51] Second, tying occurs in a two-sided market.[52] There is then a natural benefit of tying in terms of a greater flexibility to rebalance charges between the two sides.

9 Some Policy Implications for the Portuguese Payment Card Industry

Warning: The answers provided below have a very preliminary nature. Their validity may rely on specific factual knowledge. We will try to indicate where complementary information about the Portuguese payment card system would be useful in order to bring more refined answers.

The text in boxes refers to the questions posed by the Authority.

Debit card networks

[49]For recent views on tying in one-sided markets, see, e.g., Carlton-Waldman (2002), Choi-Stefanadis (2001), Evans-Salinger (2004), Nalebuff (2003, 2004), Rey et al (2003), and Tirole (2005).

[50]In Whinston (1990) for example, tying is entirely motivated by entry deterrence: By lowering its opportunity cost of selling in the competitive market (losing a sale in that market implies losing a sale in the monopoly market), the tie, provided that it is technologically irreversible, commits the tying firm to be aggressive and thereby may deter entry. In our model, like in Whinston's, tying may deter entry of a more efficient rival. But the consequences are different here because a not-for-profit firm cannot exercise its market power by raising the price level. So, for example, deterrence of entry of a slightly more efficient rival through tying still raises welfare in our model.

[51]We ignore "corporate governance" or "benchmarking" benefits of product market competition. But these benefits are to a large extent internalized by the tying firm.

[52]For overviews of the economics of two-sided markets, see Armstrong (2004) and Rochet-Tirole (2004).

- The two networks are owned by two financial institutions, SIBS and UNICRE that have almost the same shareholders (which include all the main banking and financial institutions developing activities in Portugal, encompassing also BCP, the issuer and acquirer for American Express) and are run by overlapping boards.

Question: What are the consequences of the shareholders overlapping between the two network competitors?

Institutional and theoretical background: The Visa-MasterCard long experiment with board duality discussed in Section 7 has only limited bearing on the Portuguese case to the extent that the two associations did not pay dividends to their owners and therefore there was no scope to raise price. By contrast, Multibanco and Redunicre are for-profit entities. Overlapping boards therefore have the potential to raise prices.

Section 7 noted that the analysis of the impact of overlapping boards is more complex in a two-sided-market context. Yet, we feel that duality ought to be investigated unless the platforms are turned into not for-profit platforms.

In some situations the issuing and the merchant bank are the same.

Question: what is the rationale for charging a merchant fee or merchant discount that includes a MIF as in the case where there is no identity between issuer and acquirer? What is the practice in other countries?

An "on-us" transaction is a transaction between a cardholder and a merchant, whose issuer and acquirer are the same entity. The interchange fee is then an internal transfer price from the acquiring division to the issuing division.

Let us first note that this situation is not specific to the payment system context: In telecom services (and more and more Internet services), the network terminating a call receives compensation —called the terminating charge— from the network issuing the call. With a small number of actors in the telecom industries, a non-negligible fraction of calls are then "on-net" (the equivalent of "on us"). That is, the termination charge washes out in the consolidated accounts of the firm.

In practice, acquirers are not prohibited from levying merchant discounts when they also serve the cardholder; and mobile telecom or Internet companies are not prohibited from charging receivers for calls or data when the caller or the sender is on net.

Nor should they be prohibited from doing so (even ignoring issues having to do with feasibility), as the issuers and acquirers would then have an incentive to disguise their affiliation through complex financial structures. An integrated acquirer does not enjoy a competitive advantage over other acquirers because he does not need to pay an interchange fee on the fraction of transactions coming from his issuing branch's cardholders. To see this, suppose that integrated acquirer A competes for a merchant with non-integrated acquirer B. Ignoring for a moment the impact on the issuing side, acquirer A would compare the total cost of servicing the merchant, which includes the interchange fee, to the price that he can obtain from the merchant (for example, the merchant discount charged by acquirer B if the two acquirers' services are similar); and so does acquirer B. Thus, ignoring the impact of competition on the acquiring side for A's issuing branch, there is competitive parity.

This reasoning, one will object, ignores the fact that A's issuing branch receives an interchange fee. Note, however, that A's issuing branch will receive this interchange fee regardless of whether A or B obtains the merchant's business. That is, A's issuing branch's profit on the cardholder, equal to the price paid by the cardholder minus the cost of servicing the cardholder (including the subsidy provided by the interchange fee) is not affected by which acquirer gets the merchant's business. The level-playing field is preserved.[53]

The theoretical point that the existence of "on-us" transactions does not destroy the level playing field and does not call for lower merchant discounts for such transactions receives indirect empirical support through the quasi-lack of integration between acquirers and issuers in the rather competitive US credit card system. Were integration to confer advantages to integrated firms and to destroy the level playing field, one would have expected a high degree of integration.

[53]If the issuing market is imperfectly competitive (and so issuers make profits) and if some merchants do not accept the card, the integrated acquirer has an incentive to be a bit more aggressive than its non-integrated rivals, because a "market expansion effect" on the merchant side increases the issuing branch's profit (in the jargon of economists, integration gives rise to a "Cournot effect"). While integrated acquirers may therefore gain market share at the expense of non-integrated ones, this often lowers prices and raises social welfare.

- The cards issued in both networks have almost the same characteristics except for two main differences:
  - The debit cards ELECTRON and MAESTRO are widely accepted in international networks, albeit MULTIBANCO is also accepted in some international networks.
  - The merchant discount differs from the REDE MULTIBANCO to the REDUNICRE, the latter being the highest.
  - There are no relevant differences in technological or qualitative terms between the cards.

Question: Does the similarity between the characteristics of debit cards from both networks and the significant differences between merchant charges confirm a lack of competition?

Our initial reaction to this question is that the differences in merchant charges per se may not signal a lack of competition and this for several reasons:

1) First, while the two cards may be perceived as identical by the merchants, they may differ in other respects that influence the pricing structure.

Suppose that one card, card A, offers much better international merchant acceptance than the other, card B. This difference in coverage induces a difference in demand — high-income, mobile cardholders will have a strong preference for card A—, as well as a difference in cost —the international interchange fee and the cost of servicing international transactions have no reason to coincide with their domestic counterparts—.

Consider for example the difference in clienteles. For many merchants, high-income, high-value-of-time consumers form an attractive clientele: They are "marquee buyers".[54] They are thus willing to pay high merchant discounts in order to please these customers through card acceptance.

Note: This analysis is incomplete. The only valid point is that demand and cost considerations may by themselves lead to diverging merchant discounts between Multibanco and Redunicre. To think about it further, we would need to know in particular: a) do the two systems indeed differ significantly in terms of international coverage?, b) how do the domestic and international interchange fees paid by a Portuguese acquirer differ?, c) do Portuguese issuers levy extra charges on cardholders for international transactions (either directly, or indirectly through a disadvantageous exchange rate)?

[54] To employ an expression in Rochet-Tirole (2003).

2) Even if the cards were similar on all accounts, the systems might be experimenting with "different business models". After all, divergences in business models are not uncommon, even though one would expect that they would be rarer in a mature industry.

3) Taking it for granted that there is a lack of system competition, one would have to ask why the overlapping boards select different merchant discounts for the two systems.[55]

Credit-card network

- BCP, the only acquirer for AMERICAN EXPRESS, is a major stockholder at UNICRE and SIBS.

Question: Given these two facts, can we consider network competition is possible in the credit card payment system? How?

BCP's stake in Unicre creates (asymmetric) board duality. As discussed above, there is a concern that this duality raises prices of Unicre directly, and more indirectly those of American Express.

The extent of this anti-competitive effect depends on how much influence BCP can exert on the board of Unicre: voting rights, ability to form coalitions (especially if other board members are divided),...

- UNICRE is the only acquirer for VISA, MASTERCARD, DINER'S CLUB INTERNATIONAL, JCB e TARJETA 6000.

Question: In a sole acquirer situation for both VISA and Mastercard credit card transactions can we still find incentives for network efficiency?

Does a sole acquirer situation provide insensitiveness to pressures over the merchant discount rates and extend that insensitiveness to the system, namely to the level of the interchange fees?

There are two questions here:

[55] One possible explanation for this was suggested to us by Luis Cabral: given that Portuguese banks encourage a gradual "migration" of users from MULTIBANCO to REDUNICRE, maintaining higher fees in the latter may be a hidden way to increase the effective average fee without having to increase individual nominal fees.

√ A monopoly in acquiring raises the usual worries about monopolies: high prices, high operating costs and low innovation. Thus, the cost of acquiring may be inflated by the lack of competition.

By contrast, it is not clear why the monopoly acquirer, as compared to an independent owner or a group of issuers or acquirers, would have less incentive to monitor and keep a lid on platform's cost.

√ Assuming away competition from Amex and from debit card networks, the integration between Unicre and Redunicre prevents a "chain of monopolies". Given the existence of monopolies in acquiring and at the system level, this integration may well be a good thing, due to the "Cournot or double marginalization effect": Suppose that Redunicre were to divest Unicre, but competition in acquiring did not develop. Then Unicre would charge a monopoly mark-up over its perceived cost, which itself would include a monopoly mark-up on interchange. Integration at least forces the two monopolies to internalize the negative impact of too high a mark-up on one activity on the other activity.

- In relation to fees, there is an indication of high interchange fees (estimated at 80% of the merchant discounts) in relation to the practice in other countries, and there is a practice of different merchant discounts (between 1% and 5%) for different economic activities and volumes.

Question: What factors could account for higher interchange fees in Portugal vis--vis other EU members? To what extent could these differences result from UNICRE being the sole acquirer for Visa and Mastercard credit card transactions?

It would be useful here to have more specific numbers (absolute values).

It is common practice outside Portugal to have differentiated merchant discounts for different economic activities and volumes. The reasons for this are many:

√ First, the costs of servicing different types of clients may not be the same for the acquirers.

√ Second, in order to get as many merchants on board as possible, merchants and therefore acquirers do price discriminate (this is the case for example in the US where the acquiring side is very competitive). For example, supermarkets in the US pay a substantially lower merchant discount as they have a lower willingness to pay for cards, which they can easily bypass (through installing ATMs or having their own store card).

It is worth emphasizing that there is nothing wrong with such discrimination. Getting everyone on board through differentiated tariffs ensures a widespread acceptance of cards.

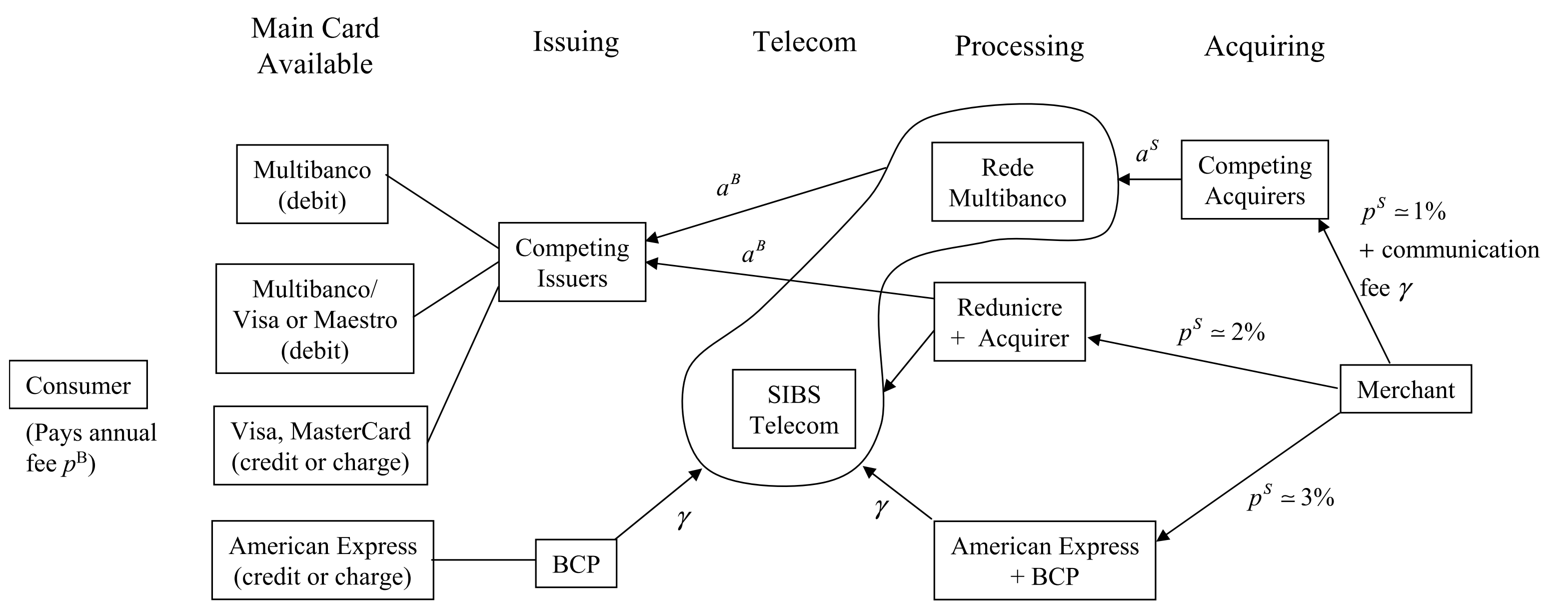
- SIBS telecommunication network and processing services are used by REDUNICRE and, most likely, by AMEX transactions.

Question: Is it efficient for the market to attribute to one organization the ownership of the entire telecommunication network and the sole responsibility over processing?

Are there incentives for network competition at the debit and credit card acquiring level given the close structural links between SIBS and UNICRE?

Having a single technical network by itself may not be an issue. However, SIBS could use its monopoly over processing to foreclose AMEX, or to provide access under unfavorable conditions. As long as this is not the case, there is no particular reason for competition authorities to be worried about this monopoly position of SIBS.

Appendix 1: The Portuguese Payment Card System



A Simplified Representation of the Portuguese Payment Card System

The models presented in the text separate the debit and credit markets, and neglect communication fees γ.

These two markets are represented together only to recall that BIBS Telecom provides processing services to the three networks (Multibanco, Redunicre, Amex BCP). This does not imply that the debit and credit markets should be considered as a single market.

The payment a^S from acquirers to Rede Multibanco roughly equals 70 % of the merchant fee p^S plus processing fees.

The payment a^B from the networks to the issuing bank equals a fixed proportion (70 % for Rede Multibanco, 80 % for Redunicre) of the merchant fee p^S minus processing fees.

REFERENCES

Armstrong, M. (2004), "Competition in Two-Sided Markets" mimeo, University College, London, UK.

Ausubel, L.M. (1991) "The Failure of Competition in the Credit Card Market," *American Economic Review*, 81: 50-81.

Baker, D.I. (1995), "Shared ATM Networks. The Antitrust Dimension", *Federal Reserve Bank of St. Louis Review*, 77: 5-17.

Balto, D.A. (2000) "The Problem of Interchange Fees: Costs without Benefits?", *European Competition Law Review*, 21: 215-224.

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

Baye, M., and J. Morgan (2001) "Information Gatekeepers on the Internet and the Competitiveness of Homogenous Product Markets", *American Economic Review*, 91: 454-474.

Caillaud, B. and B. Jullien (2003) "Chicken and Egg: Competition among Intermediation Service Providers", *Rand Journal of Economics*, 34(2): 309-328.

Carlton, D. W., and A. S. Frankel (1995) "The Antitrust Economics of Payment Card Networks," *Antitrust Law Journal*, 63 (2): 643-668.

Chakravorti, S. and W.R. Emmons (2001) "Who Pays for Credit Cards?" Federal Reserve Bank of Chicago Emerging Payments Occasional Paper Series, EPS-2001-1.

Chakravorti, S. and A. Shah (2001) "A Study of the Interrelated Bilateral Transactions in Credit Card Networks", Federal Reserve Bank of Chicago Emerging Payments Occasional Paper Series, EPS-2001-2.

Chakravorti, S. and T. To (2000) "Toward a Theory of Merchant Credit Card Acceptance", mimeo, Federal Reserve Bank of Chicago.

Chang, H.H. and S. Evand (2000), "The Competitive Effects of the Collective Setting of Interchange Fees by Payment Card Systems", *The Antitrust Bulletin*, Fall, 641-677.

Donze, J. and I. Dubec (2002) "Access Pricing in Shared ATM Networks", discussion paper, IDEI Center for Payment Cards.

Evans, D. (2003) "The Antitrust Economics of Two-Sided markets" *Yale Journal on Regulation*, Summer.

Evans, D. and M. Salinger (2004) "Why Do Firms Bundle and Tie? Evidence from Competitive Markets and Implications for Tying Law", forthcoming, *Yale Journal on Regulation*.

Evans, D. S., and R. L. Schmalensee (1993) *The Economics of the Payment Card Industry*, National Economic Research Associates.

—— (1995) "Economic Aspects of Payment Card Systems and Antitrust Policy Toward Joint Ventures", *Antitrust Journal Law*, 63: 861-901.

—— (2003) *Paying with Plastic: The Digital Revolution in Paying and Borrowing*, Cambridge, Ma: MIT Press, second edition.

Frankel, A. S. (1998) "Monopoly and Competition in the Supply and Exchange of Money," *Antitrust Law Journal*, 66 (2): 313-361.

Gans, J.S. and S.P. King (2001) "Regulating Interchange Fees in Payment Systems", Working Papers, University of Melbourne.

Gans, J.S. and S.P. King (2002) "A Theoretical Analysis of Credit Card Regulation", mimeo, University of Melbourne.

Gans, J.S. and S.P. King (2003) "The Neutrality of Interchange Fees in Payment Systems", *Topics in Economic Analysis and Policy*, Vol.3, No.1, Article 1.

Gilbert, R.J. (1990) "On the Delegation of Pricing Authority in Shared ATM Networks," mimeo, University of California Berkeley.

Guerin-Calvert, M. and J. Ordover (1998) "Maximizing Current and Future Network Competition in Payment Systems," paper presented at the 34th Annual Conference on Bank Structure and Competition.

Guthrie, G. and J. Wright (2003) "Competing Payment Schemes", mimeo, University of Auckland.

Hausman, J. G. Leonard and J. Tirole (2003) "On Non-exclusive Membership in Competing Joint Ventures", *Rand Journal of Economics*, 34(1): 43-62.

ITM (2000a), *Study Regarding the Effects of the Abolition of the Non-discrimination Rule in Sweden for European Commission Competition Directorate General*, Amsterdam.

Katz M. (1986) "Technology Adoption in the Presence of Network Externalities", *Journal of Political Economy*, 94: 822-841.

Katz M.L. (2001), "Reform of Credit Card Schemes in Australia" commissioned report for the Reserve Bank of Australia.

Katz, M. and C. Shapiro (1985) "Network Externalities, Competition, and Compatibility", *American Economic Review*, 75: 424-440.

Kim, J. (1998) "The Impact of Proprietary Positions and Equity Interest in the Pricing of Network ATM Services," mimeo, MIT.

Laffont, J.J., Rey, P. and J. Tirole (1998a) "Network Competition: I. Overview and Nondiscriminatory Pricing," *Rand Journal of Economics*, 29(1): 1-37.

—— (1998b) "Network Competition: II. Price Discrimination," *Rand Journal of Economics*, 29(1): 38-56.

Matutes, C. and J. Padilla (1994) "Shared ATM Networks and Banking Competition," *European Economic Review*, 38: 1113-1138.

McAndrews, J. (1996) "Retail Pricing of ATM Network Services", Working Paper 96-12, Federal Reserve Bank of Philadelphia.

McAndrews, J. (1997) "The Role of Direct Presentment in Non-Cash Payments", Research Paper, Federal Reserve Bank of New York.

McAndrews, J. and R. Rob (1996) "Shared Ownership and Pricing in a Network Switch," *International Journal of Industrial Organization*, 14: 727-745.

Nalebuff, B. (2004) "Bundling as an Entry Deterrent Device", *Quarterly Journal of Economics*, 119: 159-188.

Priest, G.L. (1977) "Cartels and Patent License Arrangements", *Journal of Law and Economics*, 20: 309-377.

Rochet, J.C. (2003) "The Theory of Interchange Fees: A Synthesis of Recent Contributions", *the Review of Network Economics*, 2(2), 97-124.

Rochet, J.C. and J. Tirole (2002) "Cooperation among Competitors: Some Economics of Payment Card Associations", *Rand Journal of Economics*, 33(4), 1-22.

Rochet, J.C. and J. Tirole (2003) "Platform Competition in Two-Sided Markets", *Journal of European Economic Association*, 1(4), 990-1029.

Rochet, J.C. and J. Tirole (2004) "Two-Sided Markets: An Overview", mimeo, IDEI.

Schmalensee, R. (2002a) "Payment Systems and Interchange Fees", *Journal of Industrial Economics*, 50: 103-122.

Schmalensee, R. (2002b) "Interchange Fees: A Review of the Literature", mimeo MIT.

Schwartz, M. and D.R. Vincent (2000) "The No-Surcharge Rule in Electronic Payments Markets: A Mitigation of Pricing Distortions", mimeo, Georgetown University.

Schwartz, M. and D.R. Vincent (2002) "Same Price, Cash or Credit: Vertical Control by Payment Networks", Georgetown University, Department of Economics Working Paper 02-01.

Small, J. and J. Wright (2000) "Decentralized Interchange Fees in Open Payment Networks: An Economic Analysis", mimeo, University of Auckland.

Tirole, J. (1988) *The Theory of Industrial Organization*, Cambridge, MIT Press.

Tirole, J. (2005) "The Analysis of Tying Cases: A Primer", *Competition Policy International*.

Wright J. (2000) "An Economic Analysis of a Card Payment Network", mimeo, NECG and University of Auckland.

Wright J. (2001) "The Determinants of Optimal Interchange Fees in Payment Systems", Department of Economics Working Paper No 220, University of Auckland.

Wright J. (2003a) "Optimal Card Payment Systems", *European Economic Review*, 47(4): 587-612.

Wright J. (2003b) "Pricing in Debit and Credit Card Schemes", *Economics Letters*, 80: 305-309.

Wright J. (2003c) "Why do Firms Accept Credit Cards?", mimeo, University of Auckland.