Competing Payment Systems: Key Insights from the Academic Literature

Jean-Charles Rochet
Professor of Economics
Toulouse School of Economics and IDEI

Paper prepared for the Payments System Review Conference organized by the Reserve Bank of Australia and the Centre for Business and Public Policy, Melbourne Business School,
Sydney, 29 November 2007

Abstract

Most of the academic literature on retail payment systems focuses, for tractability reasons, on the case of a single payment system. However, a more realistic situation is one where several systems compete and where consumers have the choice between several means of payment (debit card, credit card, charge card, cheque, transfer…). This article summarizes the small literature that has modelled such a situation of competition between different systems. We derive some predictions about the competitive determination of user prices, consumer surplus and social welfare. We suggest some policy implications and also some possible directions for academic research.

1 Introduction

The fantastic development of payment card networks all over the world, together with the numerous interventions of courts of justice, regulators and competition authorities in their functioning have recently prompted the development of a sizeable academic literature on the topic. After the antitrust literature initiated by Baxter (1983),[1] several formal models of the payment card industry have been developed,[2] allowing a rigorous analysis of the impact of interchange fees on user fees and volumes in payment networks. For tractability reasons, this literature has focused on the case where a single payment system is available as an alternative to cash payments.

However in practice, consumers have the choice between several non-cash payment instruments (debit cards, credit-cards, charge cards, cheques, direct debit…) and several networks compete for providing each of these instruments. It is therefore important to extend the basic model developed in the academic literature described above, to a more realistic situation of competition between several payment systems. This is a delicate task that has only been undertaken in very few articles. The objective of this note is to summarize the key insights that can be obtained from these few articles, and to suggest some policy recommendations that can be derived from these key insights. I also indicate some directions of future research that would be useful to explore for providing better guidance to public decision makers.

The rest of this note is organized as follows: Section 2 recalls the analysis of a single payment system. Section 3 discusses the case of competing networks providing the same type of payment instrument. Section 4 considers the case where multiple payment methods are available. Finally section 5 examines several policy questions such as: is there a basis for price regulation in retail payment networks? Or do optimal interchange fee tend to zero when a payment system matures?

2 A single payment system

This section presents the economic analysis of the impact of interchange fees when a single payment system is available, as an alternative to cash payments. Baxter's seminal analysis is presented in 2.1: if banks and retailers were perfectly competitive, efficiency of cards usage could be obtained when the interchange fee is set at a certain level a_0 (which we call Baxter's level). In a perfectly competitive world, this outcome could also, in principle, be reached (independently of the level of the interchange fee) by allowing retailers to surcharge card payments since they would exactly pass to consumers the net costs they incur from cards payments (perfect surcharging). However in practice, banks and retailers are not perfectly competitive: banks charge more than their marginal costs, and retailers do not surcharge

[1] See in particular Carlton and Frankel (1995), Evans and Schamlensee (1995), Frankel (1998), Chang and Evans (2000), and Farrell (2006).

[2] See in particular Schmalensee (2002), Rochet and Tirole (2002, 2006 and 2007), Wright (2003, 2004), Gans and King (2003) and McAndrews and Wang (2006).

perfectly. The consequences of imperfect competition between banks and between retailers are analyzed in 2.2. Subsection 2.3 considers some extensions of the basic model and implications for price regulation.

2.1 Baxter's benchmark: perfectly competitive banks and retailers

Payment systems have the peculiarity that they provide a joint service to two users (the payor and the payee, whom I will call for simplicity the buyer and the seller). The socially efficient payment instrument is the one that minimizes the sum of the net costs[3] of all participants involved in the transaction. In the simple case where there is a unique card payment system[4] (providing the only alternative to cash payments), social efficiency is easily characterized by comparing the incremental costs (card vs cash) of the different participants. In a four party system these participants are the buyer (*B*), the seller (*S*), the issuer of the card (*I*) and the acquirer of the payment (*A*). Following the notation used by Rochet and Tirole (2002) let us denote by b_B and b_S the "benefits" of a card payment for the buyer and the seller (equal by definition to the differences between the cost of a cash payment and that of a card payment) and by c_I and c_A the marginal "costs" of a card payment (equal by definition to the differences between the cost of a card payment and that of a cash payment) for the issuer and the acquirer.

A card payment is socially efficient for a particular transaction if and only if the sum of users' benefits exceeds the sum of providers' costs:

$$b_B + b_S \geq c_I + c_A. \tag{1}$$

Baxter (1983) was the fist to emphasize that, even when there is perfect competition between banks (implying that user prices equal marginal costs on each side of the market) card usage can be socially inefficient. Indeed, provided that the seller accepts card payments, the buyer will choose to pay by card when his benefit exceeds the transaction fee p_B he faces. When issuers are perfectly competitive, this transaction fee is equal to the issuer's marginal cost c_I:

$$b_B \geq p_B = c_I. \tag{2}$$

Comparing with condition (1), Baxter concludes that social efficiency of card usage requires an interbank transfer (the interchange fee) from the acquirer to the issuer, equal to:

$$a_0 = b_S - c_A. \tag{3}$$

In such a way, the issuer's marginal cost becomes $c_I - a_0$ and the buyer chooses to pay by card whenever

$$b_B \geq c_I - a_0 = c_I + c_A - b_S,$$

[3] By net cost I refer to the difference between the cost and any benefit that might be associated with the user of a payment instrument. In all this article, "cost" means "net cost" and "benefit" means "negative cost".

[4] Following the seminal contribution of Baxter (1983), the economic analysis of this case (monopoly card scheme) is due to Rochet and Tirole (2002) and Wright (2003, 2004).

which is equivalent to the social efficiency condition (1). The interchange fee a_0 (which can be positive or negative, according to the values of b_S and c_A) restores efficiency of card usage.

Rochet and Tirole (2006c) and Farrell (2006) show that the same reasoning can be recast in terms of the usage externality that characterizes payment systems, due to the fact that, although the choice of the payment mean affects both users, only one of them (typically the buyer) makes this choice. By choosing to pay by cash (instead of card) the buyer increases the cost of the seller (or increases his benefit) by b_S, and reduces the cost of the acquirer by c_A. The usage externality is therefore perfectly internalized by the buyer whenever the buyer's fee is reduced by $b_S - c_A$.

Carlton and Frankel (1995) point out that, in Baxter's perfectly competitive world, perfect surcharging by sellers is an equivalent way to internalize the usage externality. Indeed if sellers charge different retail prices, p_{CASH} for cash payments and p_{CARD} for card payments, the buyer will choose a card payment whenever the gross benefit b_B of a card payment for the buyer exceeds the sum of the buyer fee p_B and the card surcharge $p_{\text{CARD}} - p_{\text{CASH}}$:

$$b_B \geq p_B + p_{\text{CARD}} - p_{\text{CASH}}.$$

When $p_B = c_I$ (competitive issuers) and $p_S = c_A$ (competitive acquirers), this condition is equivalent to the social efficiency condition : $b_B \geq c_I + c_A - b_S$, whenever

$$p_{\text{CARD}} - p_{\text{CASH}} = p_S - b_S,$$

which means that the card surcharge should equal the net cost inflicted to the merchant by the decision to pay by card. This condition, which is called "perfect surcharging", is obtained when there is perfect competition among retailers. It does not seem to be satisfied in practice (see below).

2.2 Introducing imperfect competition between banks and between retailers

Baxter's competitive benchmark is at odds with reality in several respects. In particular, banks and retailers are not perfectly competitive: they typically charge mark-ups over marginal costs, if only for covering their fixed costs. Moreover user prices do not typically react one for one to variations in interchange fees. The RBA reform in Australia is a good illustration of this fact. Although the reduction in interchange fees for credit cards was passed approximately one for one into merchant service charges by acquirers, the impact on cardholder fees was only a fraction of the reduction in issuer's revenue, while retail prices were not significantly affected [5](Chang et al. 2005). Moreover surcharging by Australian

[5] The RBA has a different view on this issue: it considers that Australian retailers did in fact pass on to consumers the reductions in the Merchant Service Charges but that these price changes were too small to be detected.(see Lowe 2006).

merchants, (which is possible since 2002) is far from perfect: very few retailers surcharge[6], and those who do sometimes surcharge more than their incremental cost of card payments[7].

In order to capture these features, Rochet and Tirole (2002) have developed the first fully-fledged model of an imperfectly competitive payment card industry, allowing a comparison between privately optimal and socially optimal interchange fees. They assume that issuers have market power while acquirers are perfectly competitive. Retailers are also imperfectly competitive: their decision to accept cards or not takes into account the impact of this decision on the attractiveness of their store to the customer. Rochet and Tirole (2002) show that socially optimal interchange fees are very difficult to determine, since their influence on user fees and retail prices is quite complex in an imperfectly competitive world. For example, imperfect competition between issuers implies that cardholder fees p_B are higher than issuers net marginal cost:

$$p_B = c_I - a + m, \tag{4}$$

where a is the level of the interchange fee while m denotes the profit margin of the issuer[8] Social efficiency of card usage is obtained where buyers receive the correct price signal

$$p_B = c_I + c_A - b_S. \tag{5}$$

Comparing with (4), we see that the interchange fee a^* that internalizes perfectly the usage externality is higher than Baxter's interchange fee a_0:

$$a^* = b_S - c_A + m > a_0 = b_S - c_A.$$

This is because social efficiency requires that the buyer does not internalize the issuers' margin m. However this condition implies that sellers pay more than their direct benefit b_S from card payments. This because the merchant service charge p_S cannot be lower than acquirer's total cost $c_A + a^*$:

$$p_S \geq c_A + a^* > c_A + a_0 = b_S. \tag{6}$$

Rochet and Tirole (2002) show that, in spite of condition (6), retailers may still be willing to accept card payments. This is because the option to pay by cards is attractive to their customers and may therefore increase the merchant's volume of sales for a given retail price. Rochet and Tirole (2002) consider a particular model of the retail sector (Hotelling model with full market coverage) where total retail demand is fixed. In this case, card acceptance only increases a retailer's volume of sales at the expense of his competitors. This is

[6] In a survey of Australian merchants commissioned by the RBA, it appears that"17% of very large merchants... surcharge" but that "surcharging by smaller firms is less common". However the RBA observes that the number of surcharging merchants increases over time : see Payment System Board Annual Report" (2007).

[7] See for example Payment System Board Annual Report" (2007), graph 12 shows that the Merchant Service Fees charged on Open systems cards were significantly below 1%, which is the average surcharge found in the survey of merchants cited in footnote 6.above.

[8] This margin m may be itself a function of the issuer net cost ($c_I - a$).

a pure business stealing effect: retailers' equilibrium profit is the same whenever they accept cards or not. However Wright (2003, 2004) shows that the result holds true in more general models of the retail sector (monopoly, Cournot competition,…). In all these models the maximum fee that retailers are ready to accept for card payments is $b_S + v_B$, where v_B represents cardholder surplus, equal to the increase in quality of service for customers, associated with the possibility to pay by cards. Since banks' profit typically increase with the level of the IF, a monopoly network will select the maximum IF a_m that is compatible with merchant acceptance:

$$c_A + a \leq b_S + v_B \quad \Leftrightarrow \quad a \leq a_m = b_S - c_A + v_B.$$

a_m is clearly higher than Baxter's IF $(a_0 = b_S - c_A)$ but may be higher or lower than $a^* = a_0 + m$. Thus there are two cases:

- if $m > v_B$ (high margins/low cardholder surplus), a^* is too high to be accepted by merchants, and the (second best) socially optimal IF is equal to the monopoly IF a_m,
- if $m < v_B$ (low margins/high cardholder surplus), a^* is attainable and the monopoly IF a_m is strictly higher than the socially optimal IF a^*.

Thus the basic conclusions to be drawn from the fully fledged economic model of a single payment system are the following:

- the socially optimal IF is higher than Baxter's IF level due to imperfect competition between issuers,
- the privately optimal IF either coincides with the socially optimal IF (when issuer margins are higher and/or cardholder surplus is low) or it is too high (when issuer margins are low and/or cardholder surplus is high).

2.3 Extensions of the basic model and implications for price regulation

The results of Rochet and Tirole (2002) seem to suggest that a monopoly payment card system will never set IFs at a lower level than the social optimum. However, this feature is not robust: in general privately optimal IFs can be either too high or too low. For example, Wright (2003) considers the case where IFs cannot be differentiated perfectly for different types of retailers. In this case merchant acceptance of cards is elastic with respect to the MSC p_S, which increases merchant resistance.[9] The monopoly IF may be higher or lower than the socially optimal IF, which illustrates well that although a monopoly association will not in general select the socially optimal IF, there is no systematic bias: the monopoly IF can be

[9] In this case the socially optimal IF is characterized by the equality between cardholder surplus v_B and retailer surplus v_S (see Rochet 2003) for details).

lower or higher than the socially optimal IF, according to the value of some parameters (such as user surplus and elasticities of card payment volumes) that are very difficult to estimate empirically.

Gans and King (2003) show that when retailers surcharge perfectly, the level of IFs becomes neutral (in the sense that it does not impact card usage). In this case, a regulation of IFs has thus no impact on the volume of card transactions. Gans (2007) argues that this result is more general: he claims that in spite of the fact that relatively few Australian retailers do actually surcharge, the regulation of IFs by the RBA had essentially no impact.

Farrell (2006) argues that the target for a Competition Authority (as opposed to a regulator like the RBA) should be consumer surplus and not social welfare. In a first approximation (assuming that banks' margins do not vary too much with their costs), the level of the IF that maximizes consumer surplus is such that retailers are indifferent as to the choice of the payment instrument (card or cash) by the consumer.[10] This level is always below the monopoly IF a_m.

Finally, it is interesting to analyze the levels of user prices obtained when the monopoly platform is for profit and to compare these levels with the ones associated with a not-for-profit association, which we have just characterized. As can be expected the overall level of prices (in our notation $p_B + p_S$) is higher when the platform is for profit: the profit margin of the system has to be added to those of the banks. Moreover, the price structure is also unfavourable to merchants: when IFs can be differentiated across retailers' types, the monopoly platform selects the maximum MSC that is acceptable by merchants, like is the case for a not-for-profit platform. Therefore, a monopoly system extracts the maximum surplus from retailers, whether it is for-profit or not.

3. Competing Networks

The case of competing networks, providing substitutable payment card services to the two types of users, is more complex and has been only examined by very few papers: essentially Chakravorti and Roson (2006), Guthrie and Wright (2007) and Rochet and Tirole (2007). Chakravorti and Roson (2006) use the model designed by Rochet and Tirole (2003) for analysing platform competition in two-sided markets. They find that competition between for-profit card schemes unambiguously reduces the overall price of card payments ($p_B + p_S$ in our notation) but does not necessarily lead to a socially optimal price structure. This first insight is useful, but the model is not well adapted to study the determination of IFs in open payment schemes (it was designed for other industries). In particular it focuses on proprietary systems (that do not have explicit IFS) and also it does not capture the internalization by merchants of the quality of service provided by card payments to their customers. Chakravorti and Roson (2006) do no model explicitly the determination of retail prices. Thus they cannot

[10] Farrell (2006) calls this the Merchant Indifference Criterion.

address the question of the influence of card fees (both for cardholder and merchant) on the price and volume of retail transactions, and cannot deliver a complete welfare analysis.

The first fully-fledged model of a competitive payment industry is due to Guthrie and Wright (2007). It can be seen as the adaptation of Bertrand's model of perfectly competitive providers to the payment card industry. We summarize it below.

3.1 Guthrie and Wright's perfectly competitive benchmark

In this version of Guthrie and Wright's model, both issuers and acquirers are perfectly competitive, and two card schemes (that can be for-profit or not) provide perfectly substitutable payment card services. A first important consequence of the assumption that intersystem competition is perfect (cards are perfect substitutes both for consumers and retailers) is that the two card networks make a zero profit at equilibrium: the outcome of perfect intersystem competition is the same whether the card schemes are for-profit or not. A second important consequence of perfect intersystem competition is that, whenever the two card schemes are active at equilibrium, they both set the same interchange fee.[11]

The most important result obtained by Guthrie and Wright is that the forces of perfect competition are not enough in general to generate the socially optimal level of interchange fees. This comes from an impossibility to coordinate the two sides of the market, generating a multiplicity of competitive equilibria. Suppose for example that consumers decide to hold only one card (they "single-home"). In this case, retailers are better off accepting both cards (they "multi-home") since they would otherwise lose customers. For this reason, the equilibrium IF is biased in favour of cardholders: it corresponds to the maximum fee that merchants are ready to accept. It is in fact equal to the monopoly interchange fee a_m: inter-system competition is completely ineffective when consumers hold a single card. By contrast, if consumers decide to hold both cards (they "multi-home") retailers can threaten to reject the card that is the most expensive for them, which forces card schemes to pick the IF that is most favourable to retailers, namely Baxter's IF a_0.

More generally, Guthrie and Wright (2007) show that any IF in the interval (a_0, a_m) can be obtained in a competitive equilibrium between perfectly substitutable card schemes. Which particular level obtains depends on the "loyalty" of cardholders to their most preferred card, measured by the multi-homing index.[12].

3.2 Competing cards schemes with imperfectly competitive banks and retailers

As in the analysis of the monopoly case, the introduction of imperfect competition between banks and retailers complicates the picture substantially. The only attempts in this

[11] This is a good approximation of the pricing policy of MasterCard and Visa, that set very similar IFs in almost every country and sector where they are both active.

[12] The multi-homing index, defined by Rochet and Tirole (2003), measures the fraction of cardholders who switch to their second best card when their most preferred card is rejected by the retailer.

direction have been made so far by Guthrie and Wright (2003),[13] and by Rochet and Tirole (2007), who assume for simplicity that issuers' margin *m* is constant:

$$p_B = c_I - a + m.$$

As before, it is assumed that retailers compete for geographically dispersed consumers with a fixed total demand (Hotelling-Salop model). These assumptions allow to concentrate on the choice of the payment instrument. Consumers pick the retailer that offers them the lowest net price (sum of the retail price, transport costs and transaction costs associated with the choice of the payments instrument, card or cash). With these simplifying assumptions, social welfare is maximized for a value a_W of the interchange fee that is above Baxter's level a_0:

$$a_W = a_0 + m.$$

Thus in this world, competition between card schemes can lead to IFs that are too low with respect to the social optimum. For example when all consumers hold the two cards, we have seen that the equilibrium level of IF was Baxter's a_0, which is lower than a_W. However, if consumers' multi-homing index[14] is low (because most consumers insist on using their "preferred" payment card) the balance of bargaining power shifts toward consumers and the equilibrium IF can be higher than the socially optimal level a_W. Rochet and Tirole (2007) also analyze the case of heterogeneous merchants (as in Wright 2004). The fundamental result obtained by Guthrie and Wright (2007) remains valid: competitive IFs do not coincide in general with the social optimum but there is no systematic bias: the competitive level can be below or above the social optimum, depending on parameters (such as the elasticities of card payment volumes to cardholder and merchant fees) that are difficult to estimate econometrically.

4. Multiple payment methods

4.1 Credit versus debit

The main concern behind the RBA reform was not so much that there were too many card transactions overall but rather that the mix of debit and credit transactions was tilted toward credit transactions, that are supposed to be socially more costly. The only article so far that simultaneously models credit and debit card payments is Rochet and Tirole (2006b). However this article focuses for simplicity on the case where credit and debit cards are not substitutable. There are two types of consumption goods: "debit" goods that can be purchased by debit cards or in cash, and "credit" goods that can be purchased by credit cards or cheques. Rochet and Tirole (2006) assume that the credit card is offered by a monopoly scheme, while two competing schemes offer substitutable debit cards: an "on-line" (or EFTPOS) card, run by

[13] Guthrie and Wright (2003) is an extended version of Guthrie and Wright (2007). It also contains an analysis of the case where issuers have market power.

[14] This multi-homing index is difficult to measure empirically. Rysman (2006) and Snyder and Zinman (2007) provide empirical estimations of membership multi-homing (how many consumers hold two or several cards) and usage multi-homing (how many effectively use two or several cards).

an association of banks, and an "off-line" (or scheme debit) card run by the same association that runs the credit card scheme. The objective of their article is to study the impact of a tie-in between the credit card and the "off-line" debit card. In the absence of such a tie-in (no Honour All Cards rule), interchange fees are determined as in Rochet and Tirole (2002) for the monopoly credit card industry and in Guthrie and Wright (2007) for the duopoly debit card industry. In particular, the interchange fee for credit is the highest IF a_m^C that is compatible with merchant acceptance. By contrast there are multiple equilibria in the debit card industry. To fix ideas, Rochet and Tirole (2006) assume that cardholders multi-home, which pins down the equilibrium IF to Baxter's level a_0^D. Consider now that happens if the credit card scheme imposes an Honour All Cards rule to retailers (i.e. they must accept the off-line debit card if they accept the credit card). By doing so, the credit card scheme is able to "rebalance IFs" i.e. to choose a combination of IFs for debit and credit that increases the total profit of banks on debit and credit transactions while remaining acceptable to retailers: credit IFs decrease while debit IFs increase. Because retailers internalize the surplus derived by their customers from the option to pay by card (debit and credit), they reject all combinations of IFs that decrease total user surplus. This means that the cards scheme that ties-in credit and debit cards cannot attract merchants unless it offers a combination of IFs that offers a higher total user surplus that the competing debit card.. Thus in this simple model, the tie-in of debit and credit is weakly beneficial to all parties: both total user (buyer + seller) surplus and issuers' benefit increase. It is true that the tie-in results in an increase in merchant fees for debit (and a decrease for credit) but these changes are passed on to consumers in the form of higher retail prices. Moreover merchant fees for credit decrease and the total volume of card transactions increases. The increase in retail prices for the debit good is more than compensated by a decrease in transaction costs for consumers: the "hedonic" price of the debit good (that incorporates transaction costs) actually decreases.

This surprising result (that a tie-in might benefit all parties, namely consumers, retailers and banks) is specific to two sided markets, and is only valid under some simplifying assumptions: homogenous retailers, no substitutability between debit and credit cards, and fixed retail demand. Without these assumptions, some retailers may be hurt by the HAC rule. However the "rebalancing" effect identified above remains robust. It illustrates a characteristic feature of two-sided markets: by bundling different goods and services on one side of the market, platforms can attract more users on the other side of the market, which typically generates a positive feedback for the "bundled side". A similar example is the bundling of advertisement and TV programs that allows TV channels to offer free TV to their viewers.

4.2 Further directions of research

There is a crucial need for developing models of retail payment systems where the substitutability between different payment means is explicitly captured. This is not an easy task, because the conditions characterizing efficient use of payment means are more complex

to obtain, and because multiple distortions exist, due to the interaction between several actors with different objectives. Consider for example the substitutability between credit and debit cards. For some transactions (such as "impulse purchases"), credit cards are the only possible means of payment for the consumers that do not have an easy access to other forms of credit. This means that the benefit derived by retailers from credit card acceptance is very high for these transactions, since they can lose the sale altogether if they do not accept credit cards. This explains that some retailers might be ready to accept very high fees for these transactions. By contrast, debit cards are perfect substitutes to credit cards for smaller transactions or for liquid consumers who have enough funds on their bank account. The main reason why some "convenience users" or “transactors” might still prefer to pay by credit cards is that they get "negative fees" in the form of cash back bonuses, air miles or other forms of rewards. From a social welfare point of view, it seems likely that credit card transactions of the first type (when the credit facility is really needed) are beneficial, while the second type of credit card payments (by convenience users) are inefficient. Since these two types of transaction are difficult to distinguish ex ante, social efficiency may only be attainable if an "incentive compatibility constraint", is imposed, requiring that the price of debit card payments is cannot be higher[15] than that of credit card payments:

$$p_B^D \leq p_B^C.$$

In the perfectly competitive benchmark, cardholder fees are equal to the net costs of issuers, namely

$$p_B^D = c_I^D - a^D, \;\; p_C^B = c_I^C - a^C,$$

Where, as before, c_I^k and a^k denote respectively the issuer cost and the interchange fee for cards of type k ($k = D$ for debit and k = C for credit). The above "incentive compatibility constraint" then amounts to introducing a cap on credit interchange fees:

$$p_B^D \leq p_B^C \;\; \Leftrightarrow \;\; a^C \leq a^D + c_I^C - c_I^D.$$

When the debit interchange fee a^D and the marginal cost c_I^D of debit payments for issuers are small, this cap is essentially linked to the issuer's cost c_I^C for credit transactions. Note that such a cap is similar to the cost-based regulations imposed to credit card interchange fees in some jurisdictions. However, the rationale for such a regulation is very different from the (wrong) one that is usually given, namely that IFs correspond to a fee for service paid by the acquirer to the issuer. Here the motivation for this cap is to provide the appropriate incentive for cardholders to use the socially efficient payment instrument according to the type of transaction they are involved in.

[15] Since these prices are often negative, this constraint means that rewards on debit cards should be at least equal to those on credit cards.

Of course this informal reasoning needs to be rigorously captured in a formal model where banks and retailers have some market power and behave strategically. Rochet (2007) contains some preliminary results along these lines.

5. Empirical findings, policy implications, and future research directions

Even though the economic analysis of competition in retail payment systems needs to be pursued further, we can already indicate some policy lessons that can be drawn from empirical findings. We also indicate directions for academic research that would need to be explored more systematically.

5.1 Retail payment systems are two-sided markets

Public interventions into payment card systems have often been prompted by the lobbying of retailers associations who wish to reduce their fees. Doing so, they fail to recognize that IFs contribute to the reduction of cardholder fees that encourage consumers to use card payments over less efficient payment means such as cash and cheques. However, recent empirical research (e.g. Zinman (2006), Ching and Hayashi (2007)…) clearly establishes that consumers react negatively to cardholder fees and tend to choose the payment instrument that minimizes their total transaction cost. Price structure really matters in retail payments, as has been long recognized by the card systems operators themselves. The balancing act that results from a careful reallocation of costs between the two sides of the market is a fundamental condition for the success of a retail payment system. Competition authorities and regulators should abandon one sided approaches to the retail payment industry.[16]

5.2 There is an asymmetry between the two sides of the market

The fact that retailers internalize some fraction of consumers' benefit (because the better quality of service offered to consumers by the option to pay by card makes their stores more attractive) implies that they are less resistant to high fees than cardholders. This is why the cost of payment instruments is often borne largely by merchants rather than consumers. But this is not necessarily bad for social welfare. A skewed price structure where one side of the market (here retailers) pay more than the other may be socially efficient, especially when banks have to recoup sizable fixed costs needed to maintain safe and efficient infrastructures (see Bolt and Tieman 2005).

[16] Wright (2006) contains an insightful description of the danger of using a one-sided logic in a two-sided market.

5.3 Card system operators and bank associations may have an interest in inflating credit IFs

Empirical evidence seems to suggest that higher credit IFs often result in higher profits for banks. This comes from the fact that price reactions to changes in IFs seem to be asymmetric. In the Australian case for example, reductions in credit IFs have been passed through almost one for one into merchant fees, but the corresponding increases in cardholder fees (and reductions in their rewards) have been significantly less than one for one. Even if the robustness of this observation needs to be checked carefully by empirical analysis of other systems and countries, it may explain why credit IFs are often much higher than debit IFs. As Rochet and Tirole (2002, 2006) have shown, this is not necessarily bad for social welfare, if the operating profits of banks allow them to cover the sizable fixed costs needed to increase the capacity and safety of their networks. On the other hand, IFs may be excessive if the issuers' profits are dissipated into wasteful marketing campaigns aimed at stealing business from their competitors.

5.4 Interchange fees are needed, even in mature payment card systems

The need to subsidize membership to internalize network externalities disappears when networks mature and cover a large fraction of potential users. However, payment networks are dominated by usage rather than classical network (membership) externalities. Even if all consumers hold debit cards, they need to be given incentives to use them. As we have seen, price elasticity of card usage by consumers seems to be much higher than that of card acceptance by merchants. It would be a mistake to mandate a cost-based cap on debit IFs, since it would completely overlook the two-sided nature of payment systems.

5.5 The substitutability between credit and debit cards needs to be understood better

Preliminary analysis of the substitutability between credit and debit cards (Rochet, 2007) seems to indicate a need for capping the difference between credit and debit IFs, in order to discourage the socially inefficient behaviour of "convenience users". However, it seems difficult to recommend a cost based regulation of credit IFs without a more complete understanding of this substitutability between credit and debit cards. In particular it would be important to understand whether there are indeed incentives for payment card networks to inflate the difference between credit and debit IFs. Generally speaking, the regulation of IFs is a very hazardous exercise, since socially optimal IFs depend in a complex fashion of parameters that are extremely difficult to estimate empirically. Even more importantly, the long term reactions of the payment industry to such a regulation are very difficult to predict.

References

Armstrong, M. (2006) "Competition in Two-Sided Markets", *Rand Journal of Economics,* 37(3), 668-691.

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

Bolt, W. and A. Tieman (2005), "Social Welfare and Cost Recovery in Two-Sided Markets", *IMF Working Paper*, n° 05/194.

Carlton, D. and A. Frankel (1995), "The Antitrust Economics of Credit Card Networks", *Antitrust Law Journal*, 63(3), 903-915.

Chakravorti, S. and R. Roson (2006), "Platform Competition in Two-Sided Markets: The Case of Payment Networks", *Review of Network Economics*, 5(1), 118-142.

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

Chang, H., D. Evans and D. Garcia-Schwarz (2005), "The Effect of Regulatory Intervention in Two-Sided Markets", *Review of Network Economics*, 4, 4, 328-358.

Ching A. and F. Hayashi (2007), "Payment Card Reward Programs and Consumer Payment Choice", working paper, Federal Reserve Bank of Kansas City.

Evans, D. (2002),"The Antitrust Economics of Two-Sided Markets", *Joint Centre for Regulatory Studies*.

Evans, D. and R. Schmalensee (2005), "The Economics of Interchange Fees and Their Regulation: An Overview", *MIT Sloan Working Paper*, Cambridge.

Farrell, J. (2006), "Efficiency and Competition between Payment Intruments", *Review of Network Economics,* 5(1), 26-44.

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

Gans, J. (2007), "Evaluating the Impact of the Payment System Reform", Submission to the *RBA Review of Payment System Reforms*.

Gans, J. and S. King (2003), "The Neutrality of Interchange Fees in Payment Systems", *Topics in Economic Analysis and Policy*, 3(1), Article 1.

Guthrie, G. and J. Wright (2003), "Competing Payment Schemes", working paper 0311, NUS.

Guthrie, G. and J. Wright (2007), "Competing Payment Schemes", *Journal of Industrial Economics*, LV, 37-67.

Lowe, P. (2006) "Opening Statement to House of Representatives Standing Committee on Economics, Finance and Public Administration" http://www.rba.gov.au/Speeches/2006/sp_ag_150506.html

McAndrews, J. and Z.Wang (2006)."Microfoundations of Two-Sided Markets: A Payment Card Example" Payments System Research Working Paper, Federal Reserve of Kansas City.

Payment System Board Annual Report" (2007) http://www.rba.gov.au/PublicationsAndResearch/PSBAnnualReports/2007/Pdf/card_pay_sys.pdf

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

Rochet, J.C. (2007), "Debit vs Credit: The Role of Interchange Fees", (in preparation).

Rochet, J.C. and J. Tirole (2002), "Cooperation among Competitors: The Economics of Payment Card Associations", *Rand Journal of Economics*, 33(4), 549-570.

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

Rochet, J.C. and J. Tirole (2006a), "Two-Sided Markets: A Progress Report", *Rand Journal of Economics*, 37(3), 645-667.

Rochet, J.C. and J. Tirole (2006b), "Tying in Two-Sided Markets and the Impact of the Honor All Cards Rule", mimeo TSE.

Rochet, J.C. and J. Tirole (2006c), "Externalities and Regulation in Card Payment Systems", *Review of Network Economics*, 5(1), 1-14.

Rochet, J.C. and J. Tirole (2007), "Must-Take Cards and the Tourist Test", mimeo, TSE.

Rysman, M. (2007), "An Empirical Analysis of Payment Card Usage".

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

Snyder, C. and J. Zinman (2007), "Consumer Homing on Payment Cards: From Theory to Measurement", working paper Dartmouth College, Hanover NH.

Suarez, J. (1994), "Closure Rules, Market Power and Risk-Taking in a Dynamic Model of Bank Behavior", LSE, FMG DP 196, London UK.

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. (2004), "One-Sided Logic in Two-Sided Markets", *Review of Network Economics*, 3, 1.

Zinman, J. (2007), "Debit or Credit", mimeo, Dartmouth College, Hanover NH.