Must-Take Cards and the Tourist Test*

Jean-Charles Rochet† and Jean Tirole‡

June 29, 2006

Abstract

Antitrust authorities often argue that merchants cannot reasonably turn down payment cards and therefore are forced to accept unacceptably high merchant discounts. The paper attempts to shed light on this "must-take cards" view from two angles.

First, the paper gives some operational content to the notion of "must-take card" through the "tourist test" (would the merchant want to refuse a card payment when a non-repeat customer with enough cash in her pocket is about to pay at the cash register?) and analyzes its relevance as an indicator of excessive interchange fees.

Second, it identifies four key sources of potential social biases in the payment card associations' determination of interchange fees: internalization by merchants of a fraction of cardholder surplus, issuers' per-transaction markup, merchant heterogeneity, and extent of cardholder multi-homing. It compares the industry and social optima both in the short term (fixed number of issuers) and the long term (in which issuer offerings and entry respond to profitability).

Keywords: Card payment systems, interchange fee, internalization, multi-homing, tourist test.

*We are grateful to Luis Cabral and especially Marius Schwartz for useful comments.

†IDEI and GREMAQ (UMR 5604 CNRS), Toulouse University, France.

‡IDEI and GREMAQ (UMR 5604 CNRS), Toulouse, and MIT.

1 Introduction

In payment cards systems such as Visa or MasterCard, the interchange fee (IF) paid by the merchant's bank to the cardholder's bank allocates the total cost of the payment service between the two users, cardholder and merchant. In several regions of the world, courts of justice, competition authorities, and banking regulators have asserted that these IFs are set at unacceptably high levels. Merchants, the argument goes, accept to pay the resulting high merchant discount because they are concerned that turning down cards would impair their ability to attract customers; that is, cards are "must-take cards" (Vickers 2005).[1]

In the UK, the Office of Fair Trading, following a multi-year investigation of MasterCard's credit card IFs, has announced its intention to regulate down these IFs, as well as plans to investigate Visa's. Similarly, under the pressure of the European Commission, Visa International agreed in 2002 to reduce its cross-border interchange fees on credit and debit transactions within the European Union. In Australia, after the publication of an extensive study of debit and credit card schemes in 2000, the Reserve Bank of Australia mandated a sizeable reduction of credit-card IFs, and is considering doing the same (or perhaps even mandating a zero IF) for debit transactions. Other countries where similar decisions have been made (or are seriously considered) by courts of justice, competition authorities or banking regulators include Israel, Spain, Portugal, Belgium, and the Netherlands.[2]

This paper offers a model of the payment card industry[3] that is sufficiently rich to ac-

[1]The potential anticompetitive effects of IFs have been discussed in a sizeable antitrust literature, in particular Carlton and Frankel (1995), Evans and Schmalensee (1995), Frankel (1998), Chang and Evans (2000) and Balto (2000). This literature is surveyed in Schmalensee (2003).

[2]The regulation of IFs may operate less directly, as was the case in the 2003 WalMart settlement in which MasterCard and Visa agreed to pay \$3 billion to merchants and to stop tying their credit and off-line debit card (the merchants in particular complained that the implicit IF on off-line debit was too high).

[3]This model elaborates on previous theoretical analyses of the impact of IFs, in particular Rochet and Tirole (2002), Schmalensee (2002), and Wright (2003a, 2003b, 2004). This literature is surveyed in Rochet (2003).

count for the complex effects of IFs on volumes of card payments, banks' profits, consumer welfare, retailers' profits and retail prices, yet simple enough to assess their regulation.

First, the paper gives some operational content to the notion of "must-take card" through the "tourist test": would the merchant want to refuse a card payment when a non-repeat customer with enough cash in her pocket is about to pay at the cash register? The merchant discount passes the tourist test if and only if accepting the card does not increase the merchant's operating costs. The paper analyzes its relevance as an indicator of excessive interchange fees.

Second, it identifies four key sources of potential social biases in the payment card associations' determination of interchange fees: internalization by merchants of a fraction of cardholder surplus, issuers' per-transaction markup, merchant heterogeneity, and extent of cardholder multi-homing. It compares the industry and social optima both in the short term (fixed number of issuers) and the long term (in which issuer offerings and entry respond to profitability).

The paper is organized as follows: Section 2 assesses the impact of the pricing of payment cards services on card acceptance decisions by merchants, card usage decisions by consumers and efficiency. It also introduces the tourist test. Section 3 looks at whether the interchange fees maximizing short-term welfare and total user surplus (the latter equal to the former minus banks' profits), respectively, meet the tourist test, in the benchmark model where banks' margins are constant. Section 4 allows these margins to vary with the interchange fee. Section 5 performs the same exercise in the long term, in which issuer entry and offerings respond to industry profitability (and so welfare and total user surplus coincide). Section 6 shows that retailer heterogeneity makes the tourist test more likely to produce false positives. Section 7 compares privately and socially optimal interchange fees, first in the case of a monopoly platform, then when several platforms compete. Section 8 concludes.

2 The model

2.1 Must-take cards

Consumers buy one unit of a retail good each and must decide which store to patronize. Once in the store they must further select a payment method, provided that the retailer indeed offers a choice among payment means. The benchmark model has a single card payment system. We assume "price coherence"; that is, the retailers find it too costly or are not allowed to charge different prices for transactions settled by card and by cash. Whenever a transaction between a consumer (buyer) and a retailer (seller) is settled by card, the buyer pays[4] transaction fee p_B and the seller pays merchant discount p_S. There are no annual fees and all consumers have a card.

The consumer's convenience benefit of using a card is a random variable drawn from cumulative distribution function H:

$$H(b_B) = \Pr\left(\widetilde{b}_B \leq b_B\right), \tag{1}$$

and density $h(b_B) = H'(b_B)$. The distribution has a monotone hazard rate: $\dfrac{h(b_B)}{1 - H(b_B)}$ is increasing. Because the net benefit of paying by card is equal to the difference $\widetilde{b}_B - p_B$, a card payment is optimal for the consumer whenever $\widetilde{b}_B \geq p_B$. The proportion of card payments at a store that takes cards is denoted $D_B(p_B)$:

$$D_B(p_B) = \Pr\left(\widetilde{b}_B \geq p_B\right) = 1 - H(p_B). \tag{2}$$

$v_B(p_B)$ denotes the average net cardholder benefit per card payment:

$$\begin{aligned} v_B(p_B) &\equiv E\left[b_B - p_B | b_B \geq p_B\right] \\ &= \frac{\displaystyle\int_{p_B}^{\infty} (b_B - p_B)\, dH(b_B)}{\displaystyle\int_{p_B}^{\infty} dH(b_B)} > 0. \end{aligned}$$

[4]We allow p_B to be negative, in which case the cardholder receives a payment from his bank, in the form of interest-free period, cash back bonuses or air miles awarded to the buyer every time he uses his card.

From the monotone likelihood ratio property, v_B is a decreasing function[5] of p_B.

The merchant's convenience benefit, b_S, is assumed to be homogeneous in a first step.

We make the following assumption on retailers' card acceptance policies:

Assumption 1: *A retailer accepts the card if and only if :*

$$p_S \leq b_S + s(\alpha, p_B), \tag{3}$$

where $s(\alpha, p_B)$ represents the increase in the quality of service associated with the option to pay by card, as perceived by customers, and α is a parameter that characterizes the consumers' awareness of the retailer's card acceptance policy.

The function $s(\cdot, \cdot)$ is increasing in α and decreasing in p_B, and satisfies:[6] $s(0, p_B) \equiv 0$, $s(1, p_B) \equiv v_B(p_B)$, *and* $\dfrac{\partial^2 s}{\partial\alpha\partial p_B} < 0$.

Appendix 1 shows that Assumption 1 is satisfied in the (unique equilibrium of the) classic Hotelling-Lerner-Salop model of retailing, where the parameter α is interpreted as the probability that consumers are aware of a retailer's card acceptance policy and s is linear in α ($s = \alpha v_B(p_B)$).

To understand formula (3), note that when consumers are unaware of the retailer's card acceptance policy ($\alpha = 0$, and thus $s = 0$), accepting the card does not help the retailer attract customers. And so the retailer accepts the card if and only if this reduces his operating cost: $p_S \leq b_S$. When consumers know that the card is accepted ($\alpha = 1$), they expect to enjoy extra surplus $v_B(p_B)$ per card payment (recall that they do not know their convenience benefit before going to the store), and so the retailer can increase the retail price by the amount $v_B(p_B) D_B(p_B)$ while keeping sales constant. This price

[5]This is a consequence of a well-known property of log-concave distribution functions (see Prékopa 1973).

[6]The assumption that s is decreasing in p_B is actually redundant: it is a consequence of $s(0, p_B) \equiv 0$ and $\dfrac{\partial^2 s}{\partial\alpha\partial p_B} < 0$. We state it for the reader's convenience.

increase must exceed the operating cost increase in order for the retailer to accept the card.

Formula (3) reflects the idea that in general (that is, if $\alpha > 0$ and thus $s(\alpha, p_B) > 0$), retailers internalize some of the cardholders' usage surplus and are therefore willing to accept cards even if their net cost on card transactions is higher (i.e., $p_S > b_S$). They are willing to incur a cost $p_S - b_S$ (providing it is not too large) on each card transaction, in order to offer a better quality of service to their customers (who value the option of paying by card). The parameter α measures the extent to which card acceptance makes their store more attractive to the consumer.

This internalization of consumer surplus is unrelated to competition among retailers. Indeed, formula (3) also applies to a retail monopolist and has much broader generality than the Hotelling-Lerner-Salop competition model would lead us to believe.[7] Card acceptance increases both the retailers' cost (if $p_S > b_S$) and quality of service for the consumer. Provided that consumers attach the same value to the increase in the quality of service, regardless of their willingness to pay for the good sold by the retailer,[8] the retailer's card acceptance decision depends only on the sum of the merchant net convenience benefit and the perceived quality increase brought about by card acceptance.

Retail prices are not a good measure of consumer surplus since they don't take transaction costs into account. The relevant measure of "consumer surplus" ϕ, which we call

[7] Consider for example a retail monopolist. By accepting the card he can obtain profit $\pi_A = \max_p (p - \gamma - c)\left[\alpha D(p-u) + (1-\alpha) D(p)\right]$ where p is the retail price, γ the cost of the good sold by the retailer, $C = (p_S - b_S) D_B(p_B)$ the expected net transaction cost of card payments, $D(\cdot)$ the demand for the retail good and $u = \int_{p_B}^{\infty} (b_B - p_B)\, dH(b_B)$. This profit is clearly decreasing in C, and increasing in α and u. Thus it is greater than the profit obtained when rejecting cards, $\pi_R = \max(p-\gamma)D(p)$, whenever C is smaller than some function $\psi(\alpha, p_B)$. This condition is equivalent to $p_S < b_S + s(\alpha, p_B)$, where $s(\alpha, p_B) = \dfrac{\psi(\alpha, p_B)}{D_B(p_B)}$.

[8] Recall that all consumers' convenience benefit b_B is drawn from the same distribution and that consumers do not know their individual benefit until they go to the store.

total user surplus,

$$\phi \equiv (b_S - p_S) D_B (p_B) + \int_{p_B}^{\infty} (b_B - p_B) \, dH (b_B)$$
$$\equiv \int_{p_B}^{\infty} (b_B + b_S - p_B - p_S) \, dH (b_B) ,$$

represents the expectation of the total surplus (total benefit $b_B + b_S$ minus total price $p_B + p_S$) derived from card payments by the two categories of users.

2.2 The tourist test

Retailers often complain that they are "forced" to accept card transactions that increase their net costs. To understand this "must-take card" argument, one must distinguish between ex post and ex ante considerations. Once the customer has decided to buy from the retailer, it is in the latter's interest to "steer" the former to pay by cash or check instead of by card whenever $p_S > b_S$. But from an ex ante point of view, the retailer must also take into account the increase in store attractiveness brought about by the option of paying by card. Because retailers can always ex ante turn down cards, the "must-take card" argument refers to the ex post perspective.

Let us accordingly introduce the "tourist test": suppose the buyer in question is a tourist, who will never patronize the store again in the future and shows up at the cash register with ostensibly enough cash to pay the wares. The merchant discount passes the tourist test if the retailer is willing to allow this consumer to pay by card, or equivalently if accepting the card does not increase the retailer's operating costs: $p_S \leq b_S$.

Definition: *The merchant discount p_S passes the tourist test if and only if accepting the card does not increase the retailer's net operating costs: $p_S \leq b_S$.*

The attraction of the tourist test resides in the fact that the merchant pays no more than his convenience benefit from card payments. Capping the merchant discount at the merchant's convenience benefit prevents card payment systems from exploiting the internalization effect to force merchants to accept card payments that they do not want. Perhaps more importantly from the economic point of view, the absence of overpayment by the merchant suggests that the cardholder does not face excess incentives to use the card.

Whether the cap implied by the tourist test is reasonable, though, depends on whether cardholders are provided with the proper social incentive: The social optimum is reached only when the cardholders make the efficient decision with regards to the choice of payment method. As is usual, a "first-best rule" usually may no longer be adequate when the rest of the economy is already distorted. Potential existing distortions are:

(i) *cardholders' incentives are already distorted:* If merchants' fee equals their convenience benefit, the cardholder pays more than the net social cost of the card transaction (equal to the total cost of card payments, c_B+c_S, minus the merchant's benefit, b_S) whenever issuers (or acquirers for that matter) levy markups above cost. This suggests that cardholders may underconsume card payments if the merchant discount passes the tourist test.

(ii) *merchants are heterogenous:* When merchants differ in their convenience benefit (b_S), inframarginal merchants derive more benefit from card payments than marginal ones. The tourist test can be applied to each merchant (or at least to merchants who end up accepting the card), but cardholders, whose usage fee does not depend much on the merchant's identity, cannot be induced to exactly internalize the welfare of each merchant.

3 Welfare, total user surplus and the tourist test

We now model the payment card industry and investigate the impact of interchange fees on prices p_B and p_S, and ultimately on consumer surplus. Recall that, in a payment card association, the interchange fee (IF) a represents the amount paid[9] by the seller's bank

[9] Nothing prevents, both in our model and in reality, a from being negative, i.e., the IF from flowing from the issuer to the acquirer.

(the acquirer) to the buyer's bank (the issuer) for each card transaction. It reallocates the total cost[10] $c = c_B + c_S$ of processing the transaction between the two banks. The acquirers' net marginal cost becomes $c_S + a$ and the issuers' becomes $c_I \equiv c_B - a$. We simplify the analysis by assuming that acquirers are perfectly competitive (we later note how acquirer margins affect this analysis):

$$p_S = c_S + a. \tag{4}$$

By contrast, issuers may have (ex post) market power. In the benchmark model, we assume that issuers' margin m is constant:[11]

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

The interchange fee a passes the tourist test if and only if:

$$a \leq a^T \equiv b_S - c_S \quad \Leftrightarrow \quad p_B \geq p_B^T = c - b_S + m. \tag{5}$$

With an inelastic final demand, total user surplus is equal to:

$$\phi(p_B) \equiv \int_{p_B}^{\infty} (b_B + b_S - p_B - p_S)\, dH(b_B) = \int_{p_B}^{\infty} (b_B + b_S - c - m)\, dH(b_B).$$

Social welfare is equal (up to a constant) to the sum of total user surplus and banks' profit $\pi_B = mD_B(p_B)$:

$$W \equiv \phi + \pi_B = \int_{p_B}^{\infty} (b_B + b_S - c)\, dH(b_B).$$

Social welfare is a single-peaked function of p_B, and reaches its maximum at

$$p_B^W \equiv c - b_S. \tag{6}$$

[10]As in the rest of the paper, indices B refer to the buyer side, and indices S refer to the seller side. Thus c_B represents the marginal cost of the issuer (the buyer's bank) and c_S that of the acquirer (the seller's bank).

[11]In Section 4, we allow this margin to vary with a.

With perfectly competitive issuers ($m = 0$), the price p_B^W makes the consumer perfectly internalize the externality associated with the decision of paying by card. Indeed, the social cost of such a decision is not just the marginal cost c_B of the buyer's bank as it incorporates the externality exerted on the seller's side $c_S - b_S$.

The maximum interchange fee a^T that passes the tourist test thus corresponds to the socially optimal IF when banks are perfectly competitive, as was first pointed out by Baxter (1983). The merchant discount corresponding to this upper bound a^T (namely $p_S = c_S + a^T = b_S$) makes the retailer ex-post indifferent as to the buyer's choice of payment instrument (Farrell 2006).

When issuers have market power ($m > 0$), p_B^W still maximizes social welfare. The corresponding value of the interchange fee is then:

$$a^W = c_B - p_B^W + m = b_S - c_S + m > a^T.$$

By contrast, total user surplus ϕ is maximized[12] for a larger value p_B^{TUS}:

$$p_B^{TUS} = c - b_S + m. \tag{7}$$

p_B^{TUS} exceeds p_B^W because user surplus does not include the issuers' profit. Since issuers' profit $mD_B(p_B)$ decreases with p_B, a higher p_B (and thus a lower interchange fee) implies a lower expected profit for issuers and thus, around the social welfare optimum, a higher expected total user surplus.

The corresponding interchange fee a^{TUS} is given by:

$$a^{TUS} = c_B - p_B^{TUS} + m = b_S - c_S = a^T. \tag{8}$$

The behavior of functions W and ϕ is represented in Figure 1.

[12] There is a lower bound on cardholder fees, implied by Assumption 1: a retailer only accepts the card if $p_S \leq b_S + s(\alpha, p_B)$. Thus $p_B = c + m - p_S$ must at least be equal to p_B^m, defined implicitly by $c + m - p_B^m = b_S + s(\alpha, p_B^m)$. Since total user surplus ϕ and social welfare W are quasi-concave, whether this constraint binds or not does not modify the comparison between p_B^{TUS}, that maximizes ϕ, and p_B^W, that maximizes W.

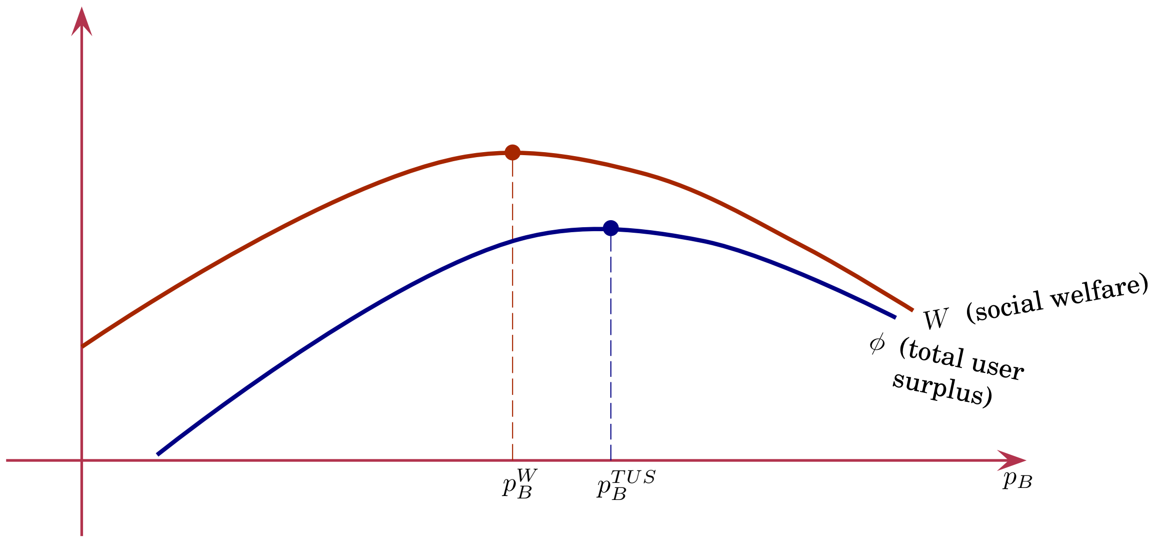

Figure 1: *Total user surplus ϕ, and social welfare W. The vertical difference between these functions represents the expected profit of issuers. It decreases with p_B. This explains why p_B^{TUS}, which maximizes ϕ, is to the right of p_B^W, which maximizes W.*

Proposition 1. *When issuers' margin is constant:*

i) the interchange fee a^W that maximizes social welfare offsets issuers' margin and thus is higher than the tourist test threshold a^T:

$$a^W = a^T + m.$$

ii) the interchange fee a^{TUS} that maximizes total user surplus is equal to a^T.

Thus in the benchmark model, the tourist test is a perfect test of interchange fees that are excessive from the point of view of total user surplus, but it is too restrictive from the point of view of social welfare.

Remark (acquirer margins): The analysis can be generalized to allow for acquirers' markups $m_S(p_S)$. The total margin m then equals the sum of acquirers' margin m_S and issuers' margin m_B. With constant markups (m_B and m_S invariant), the wedge between the welfare-optimal IF, which remains equal to $a^W = b_S - c_S + m_B$, and the TUS-maximizing IF, $a^{TUS} = b_S - (c_S + m_S)$, increases. Intuitively, the buyer internalizes his own cost, including m_B, and so in the TUS optimum, the internalization of the seller's surplus accounts for the cost, $c_S + m_S$, of acquiring. Note that $a^T = b_S - c_S - m_S$ still coincides with a^{TUS}. Thus Proposition 1 extends to a constant acquirer margin.

The result $a^{TUS} = a^T$, hence that the tourist test allows an exact detection of interchange fees that exceed the level that maximizes total user surplus ($p_S > b_S \iff a > a^T = a^{TUS}$) relies on three strong assumptions:

- issuers' margin is constant,
- the number of issuers is given,
- merchants are homogeneous.

We now relax each of these assumptions. Section 4 looks at variable issuers' margins. Section 5 considers the possibility of entry (or exit) of issuers. Section 6 allows for heterogenous merchants.

4 Variable issuers' margin

We now allow issuers' margin to vary with their net cost $c_I = c_B - a$. We model issuer competition in reduced form, denoting by $p_B(c_I)$ and $\pi_B(c_I)$ the issuers' price and profit as functions of c_I. We assume that p_B increases and π_B decreases with c_I.[13] Issuers' margin m is a function of p_B:[14]

$$m\left(p_B\left(c_I\right)\right) = p_B\left(c_I\right) - c_I. \tag{9}$$

There is *cost absorption* if $p'_B < 1$ (equivalently $m' < 0$) and *cost amplification* if $p'_B > 1$ (equivalently $m' > 0$). The benchmark case of Section 3 assumed *cost passthrough* ($p'_B = 1$, that is $m' = 0$).

[13] Revealed preference implies that these conditions are always satisfied for a monopoly issuer.

[14] This is more general than Rochet and Tirole (2003) where we assumed $p_B = f\left(c_I\right)$ with $0 \leq f' < 1$. $m(\cdot)$ is derived from $f(\cdot)$ by a simple change of variable: $m[f(c_I)] = f(c_I) - c_I$. Here we maintain the assumption that $f' \geq 0$ (p_B increases with c_I), but we do not require $f' < 1$. The assumption that $f' < 1$ implies that $m' < 0$ (cost absorption). The case $f' = 1$ corresponds to that of a constant margin. We also consider here the case where $f' > 1$ (e.g. Cournot oligopoly with isoelastic demand). In this cost amplification case m increases with p_B. Note however that $m' = \dfrac{f' - 1}{f'} < 1$.

For convenience, we take p_B (instead of a) as the variable of interest. By assumption, p_B is increasing in $c_I = c_B - a$, which implies that p_B is decreasing in a. We can thus reason on p_B, keeping in mind that an increase in the IF results in a decrease in p_B.

The total profit of the members of the association (that is of the issuers, since acquirers make no profit in our model) is thus:

$$\pi_B = m(p_B) D_B(p_B). \tag{10}$$

We assume that π_B is decreasing in p_B: Provided that merchants are kept on board, issuers prefer higher IFs.

The interchange fee a passes the tourist test if and only if

$$a \leq a^T \equiv b_S - c_S \quad \Longleftrightarrow \quad p_B \geq p_B^T = c - b_S + m\left(p_B^T\right).$$

p_B^T is uniquely defined since $m' < 1$ (see footnote 14). Similarly the value of the interchange fee that maximizes social welfare is then:

$$a^W = c_B - p_B^W + m\left(p_B^W\right),$$

where $p_B^W = c - b_S$. Thus

$$a^W = b_S - c_S + m\left(p_B^W\right) > a^T.$$

Finally, total user surplus

$$\phi(p_B) = \int_{p_B}^{\infty} \left[b_B + b_S - c - m(p_B)\right] dH(b_B)$$

is maximized for a larger value p_B^{TUS} that satisfies the first-order condition:

$$\phi'(p_B) = \left[p_B + b_S - c - m\right] D_B' - m' D_B = 0.$$

Thus

$$p_B^{TUS} = c - b_S + m + \frac{m' D_B}{D_B'}.$$

The corresponding interchange fee a^{TUS} is given by:

$$a^{TUS} = c_B - p_B^{TUS} + m\left(p_B^{TUS}\right) = a^T - \frac{m' D_B}{D_B'}\left(p_B^{TUS}\right).$$

Proposition 2. *The interchange fee a^{TUS} that maximizes total user surplus is higher than a^T in the cost amplification case ($m' > 0$) and lower than a^T in the cost absorption case ($m' < 0$). In both cases, the socially optimal interchange fee a^W exceeds a^T.*

Remark (acquirer margins): As in the case of constant margins, the analysis can be extended to allow for acquirer margins. The welfare optimal IF remains $a^W = b_S - c_S + m_B(p_B)$. By contrast, the formula of the TUS optimum is altered.[15]

Table 1 shows when the tourist test is likely to deliver false positives or negatives.

<table>
<tr><td>Interchange fee</td><td colspan="3">a^{TUS} a^T a^W</td></tr>
<tr><td>Tourist test</td><td colspan="2">PASSES</td><td>FAILS</td></tr>
<tr><td>Social welfare</td><td colspan="2">IF too low</td><td>IF too high</td></tr>
<tr><td>Total user surplus</td><td>IF too low</td><td colspan="2">IF too high</td></tr>
</table>

Cost absorption case

<table>
<tr><td>Interchange fee</td><td colspan="3">a^T a^{TUS} a^W</td></tr>
<tr><td>Tourist test</td><td>PASSES</td><td colspan="2">FAILS</td></tr>
<tr><td>Social welfare</td><td colspan="2">IF too low</td><td>IF too high</td></tr>
<tr><td>Total user surplus</td><td colspan="2">IF too low</td><td>IF too high</td></tr>
</table>

Cost amplification case

Table 1: Different thresholds for the interchange fee.

[15]Letting $m_S(p_S) = m_S(c_S + a) = m_S(c + m_B(p_B) - p_B)$ denote the acquirers' markup, then at the TUS optimum,

$$p_B + b_S - (c + m_B + m_S) = \frac{D_B}{D_B'}\left[m_B' - m_S'(1 - m_B')\right].$$

5 Entry: revisiting the notion of total user surplus

The previous section analyzed the interchange fees associated with two benchmarks, corresponding to the maximization of social welfare and to that of consumer surplus (TUS). Focusing on the narrow notion of consumer surplus is legitimate for a short-term analysis as long as the welfare of shareholders is weighted much less heavily than that of consumers. In the medium and long term, though, issuers respond to increased profitability by offering a wider variety of products or by reducing prices.

To illustrate the impact of entry, this section computes the "long-term total user surplus" first in the context of an homogeneous issuing industry in which issuers do not compete perfectly (so entry reduces price but does not increase variety), then in a context of monopolistic competition between differentiated issuers. While these environments are special, they in a sense are polar cases. The homogeneous-good case confers limited benefits on the entry mechanism: While entry benefits consumers through lower prices, the incentive to enter comes in large part from business stealing. Indeed, Mankiw and Whinston (1986) show that in an homogeneous-good, free-entry industry, the equilibrium number of firms under free entry exceeds the efficient number minus 1. Our monopolistic competition environment by contrast rules out any business stealing by assuming contestable niches, and thereby focuses on pure product creation. Thus the former (latter) environment is rather unfavorable (favorable) to a social accounting of the benefits from entry.

5.1 Homogeneous issuing industry

Let $c_I \equiv c_B - a = c - p_S$ denote the issuers' marginal cost, and N the number of issuers. The fixed cost of being in the issuing industry is $F > 0$.

Adapting our previous notation to account for the number N of issuers, we denote by $p_B = p_B(c_I, N)$ and $m = m(p_B, N)$ the equilibrium price and margin for a fixed number N

of issuers. We assume that $\frac{\partial m}{\partial N} < 0$ (more issuers imply smaller margins). The number of issuers is now endogenous, and given by the unique solution $N = N(p_B)$ to the zero-profit equation:

$$m(p_B, N)D_B(p_B) = NF. \tag{11}$$

Note that, as a consequence of our assumption that issuers' profit $m(p_B, N)D(p_B)$ decreases with p_B, the number of issuers $N(p_B)$ also decreases with p_B.

The long-term total user surplus (also equal to social welfare since issuers in the long term make no supra-normal profit) is equal to the sum of cardholder and merchant surpluses:

$$TUS^{LT} = \int_{p_B}^{\infty} [b_B + b_S - c - m(p_B, N(p_B))]\, dH(b_B).$$

The only difference with short-term total user surplus ϕ is that the margin $m(p_B)$ depends on p_B also through the number N of issuers. TUS^{LT} is maximized for

$$p_B^* = c - b_S + m + \frac{D_B}{D_B'} \left[\frac{\partial m}{\partial p_B} + \frac{\partial m}{\partial N} N' \right]. \tag{12}$$

Comparing with p_B^{TUS}, that maximizes short-term total user surplus ϕ, we see that p_B^* contains an additional term $\frac{D_B}{D_B'} \frac{\partial m}{\partial N} N'$, corresponding to the impact of p_B on the number of issuers, and thus indirectly on the level of issuers' margin. This additional term is negative since D_B', $\dfrac{\partial m}{\partial N}$ and N' are all negative. Thus

$$p_B^* < p_B^{TUS}.$$

The intuition for this result is that a lower cardholder fee p_B leads to the entry of more issuers and therefore increases competition, reduces issuers' margin and, ultimately, long-term total user surplus.

Note also that $p_B^* > p_B^T = c - b_S + m\left(p_B^T\right)$ if and only if there is long-term cost absorption:

$$\left[\frac{\partial m}{\partial p_B} + \frac{\partial m}{\partial N} N' \right] < 0.$$

Long-term cost absorption is less likely to prevail than short-term cost absorption $\left(\frac{\partial m}{\partial p_B} < 0\right)$, as the second term is positive.

To compare p_B^* and $p_B^W = c - b_S$ (that maximizes short-term welfare), differentiate the zero-profit condition that defines $N(p_B)$:

$$m(p_B, N(p_B))D(p_B) = N(p_B)F,$$

yielding:

$$\left(\frac{\partial m}{\partial p_B} + \frac{\partial m}{\partial N}N'\right) D_B + mD'_B = N'F. \tag{13}$$

By using formula (12) defining p_B^*, we get:

$$p_B^* - p_B^W = \frac{1}{D'_B}\left[mD'_B + D_B\left(\frac{\partial m}{\partial p_B} + \frac{\partial m}{\partial N}N'\right)\right].$$

where the right-hand side is computed at p_B^*. Thus

$$p_B^* - p_B^W = \frac{N'}{D'_B}F > 0.$$

We can now state our results in terms of the associated interchange fees a^*, a^{TUS} and a^W.

Proposition 3. *For an homogeneous issuing industry with free entry, the long-term total-user surplus maximizing interchange fee a^* lies in between a^{TUS}, that maximizes short-term total-user-surplus, and a^W, the short-term first best:*

$$a^{TUS} < a^* < a^W.$$

a^* passes the tourist test $\left(a^* \leq a^T\right)$ if there is long-term cost absorption ($m\left(p_B, N\left(p_B\right)\right)$ decreasing) and fails it in case of long-term cost amplification.

Cournot example: Consider Cournot competition with linear demand: $D_B\left(p_B\right) = 1 - p_B$. The industry is viable if the monopoly profit $\left(\frac{1-c_I}{2}\right)^2$ exceeds entry cost F. In this case the short-term equilibrium price and the short-term margin are:

$$p_B\left(c_I, N\right) = \frac{1 + Nc_I}{1+N} \text{ and } m\left(p_B, N\right) = p_B - c_I = p_B - \frac{(1+N)p_B - 1}{N} = \frac{1-p_B}{N}. \tag{14}$$

Thus there is short-term cost absorption $\left(\dfrac{\partial m}{\partial p_B} < 0\right)$, implying, by formula (8):

$$a^{TUS} = a^T - \left(\frac{\dfrac{\partial m}{\partial p_B} \cdot D_B}{D'_B}\right) \left(p_B^{TUS}\right) < a^T.$$

By contrast, there is long-term cost passthrough. Indeed, the zero-profit condition,

$$m\left(p_B, N\right)\left(1 - p_B\right) = NF,$$

yields :

$$m\left(p_B, N\left(p_B\right)\right) = \sqrt{F}.$$

Thus long-term total user surplus is maximized for $a^* = a^T = b_S - c_S < a^W = b_S - c_S + \sqrt{F}$.

5.2 Pure product variety

To illustrate the product-diversity argument, we build a stylized example that does not embody any business stealing effect. Consider a continuum of niche markets for cards, indexed by the fixed cost of entry F. All markets are identical but for the fixed cost of entry. Let $K(F)$ denote the cumulative distribution function of F.

Each market is contestable (is an "auction market"). In equilibrium of a contestable market, there is a single firm, and this firm makes no profit; the markup $m\left(c_I, F\right)$ is the smallest solution of:

$$mD_B\left(c_I + m\right) = F.$$

As $\dfrac{d}{dm}(mD_B(c_I + m)) > 0$ in the relevant range, the contestable market example exhibits long-term cost amplification $\left(\dfrac{\partial m}{\partial c_I} > 0\right)$.

Let $F^*(c_I)$, a decreasing function, be defined by:

$$\max_m\{mD_B(c_I + m)\} = F^*(c_I).$$

It corresponds to the maximum fixed cost that an issuer can sustain. $K\left[F^*\left(c_I\right)\right]$ represents the mass of active issuers.

Note that

$$b_S - p_S = b_S - c_S - a = a^T - a.$$

Then

$$\begin{aligned} TUS^{LT} &= \int_0^{F^*(c_I)} \left[v_B\left(c_I + m\left(c_I, F\right)\right) + \left(b_S - p_S\right)\right] D_B\left(c_I + m\left(c_I, F\right)\right) dK(F) \\ &= \int_0^{F^*(c_I)} \left[v_B\left(c_I + m\left(c_I, F\right)\right) + \left(c_I - c_I^T\right)\right] D_B\left(c_I + m\left(c_I, F\right)\right) dK(F), \end{aligned}$$

where $c_I^T \equiv c_B - a^T = c - b_S$.

Using

$$v_B D_B = \int_{c_I + m(c_I, F)}^{\infty} \left[b_B - \left[c_I + m\left(c_I, F\right)\right]\right] dH\left(b_B\right),$$

we see that

$$TUS^{LT} = \int_0^{F^*(c_I)} \int_{c_I + m(c_I, F)}^{\infty} \left[b_B - m\left(c_I, F\right) - c_I^T\right] dH\left(b_B\right) dK(F).$$

Then at $c_I = c_I^T$:

$$\frac{dTUS^{LT}}{dc_I} = \frac{dF^*}{dc_I} v_B D_B k\left(F^*\right) - \int_0^{F^*(c_I)} D_B \frac{\partial m}{\partial c_I} dK(F) < 0$$

(where $F^* \equiv F^*\left(c_I\right)$).

Proposition 4. *The total-user-surplus maximizing interchange fee in the pure-product-variety model always exceeds the level given by the tourist test:* $a^* > a^T$.

6 Heterogenous retailers

To bring our benchmark model more in line with reality, we introduce unobservable heterogeneity in the retailers' convenience benefit b_S (heterogenous degrees of internalization α are examined in Appendix 5). Following Schmalensee (2002), we assume that sellers' convenience benefit is drawn from a continuous distribution with c.d.f. G:

$$G\left(b_S\right) = \Pr\left(\widetilde{b}_S \leq b_S\right).$$

Following Wright (2004), we assume that the retail sector consists in a continuum of retail subsectors,[16] each corresponding to a value of b_S. Merchants know their individual b_S before accepting cards. The buyers' distribution of convenience benefits, $H(b_B)$, is independent of the market in which they buy. Consumers buy one good in each of these markets and patronize the store that offers the best combination of retail price, transportation cost and quality of service (determined here by the retailer's decision of whether to accept cards). For given prices for card services (p_B and p_S), the equilibrium behavior of retailers is characterized by the same conditions as in Assumption 1 but now this behavior is conditional on the realization of b_S: A retailer in "subsector" b_S accepts card payments if and only if:

$$b_S \geq \widehat{b}_S \equiv p_S - s\left(\alpha, p_B\right). \tag{15}$$

The volume of card transactions is easily computed:

$$V = D_B\left(p_B\right) D_S\left(\widehat{b}_S\right), \tag{16}$$

where $D_S\left(\widehat{b}_S\right) = 1 - G\left(\widehat{b}_S\right)$ represents the "demand" for card transactions by retailers.[17]

Social welfare is maximized when two symmetric "Samuelson conditions" are satisfied:

$$p_B^* = c - E\left[b_S | b_S \geq \widehat{b}_S\right],$$

and

$$\widehat{b}_S = c - E\left[b_B | b_B \geq p_B^*\right].$$

These conditions imply that

$$v_B(p_B^*) = v_S(\widehat{b}_S) \equiv E\left[b_S - \widehat{b}_S | b_S \geq \widehat{b}_S\right],$$

[16] Wright (2004) builds a model of a payment card association with heterogenous merchants. He shows how the privately and socially optimal IFs depend on the elasticities on the two sides (merchants, cardholders) and argues that there is no systematic bias between the IF chosen by the association and the socially optimal IF.

[17] A similar multiplicative formula for the volume of card transactions (with $\alpha = 0$ and thus $\widehat{b}_S = p_S$) was first proposed by Schmalensee (2002) and later used in a more general context by Rochet and Tirole (2003).

where $v_S(\widehat{b}_S)$ denotes the average retailer surplus per card payment.

When merchants fully internalize cardholder surplus ($\alpha = 1$ and so the second Samuelson condition is met[18]) and $m = 0$ (perfect competition among issuers), this social optimum can be implemented by setting the interchange fee at the following level:

$$a^* = c_B - p_B^* = E\left[b_S | b_S \geq \widehat{b}_S\right] - c_S.$$

The corresponding merchant discount is

$$p_S = E\left[b_S | b_S \geq \widehat{b}_S\right].$$

In this case the *average* merchant (among those who accept cards) is ex post indifferent about the means of payment chosen by the consumer. This means that unless all retailers are identical, some of them would like to reject cards ex post. Efficiency cannot require that the tourist test be met by all participating merchants because cardholders must internalize the welfare of the *average* merchant and not of the marginal one, who, recall, values card payments less than the average merchant. Capping merchant discounts at the convenience benefit of the most reluctant merchants provides the cardholder with an incentive for underconsumption of card payments.

Proposition 5. *Assume that merchants differ in the net convenience benefit that they derive from card transactions, that all consumers are informed about merchants' card acceptance, and that issuers' margin is nil (perfect competition). Then the social optimum is characterized by the "balanced externality condition" $v_S = v_B$ and can be implemented by selecting a merchant discount equal to the average convenience benefit among the merchants who accept cards.*

Appendix 2 computes the interchange fees that maximize welfare and total user surplus when the issuers' markup m is constant. We here content ourselves with the expressions

[18]This condition is $\widehat{b}_S = c - E(b_B | b_B \geq p_B^*) = c - p_B - v_B(p_B^*)$. When $\alpha = 1$ and $p_B + p_S = c$, this condition coincides with the definition of $\widehat{b}_S$ (condition (15)), given that $s(1, p_B) = v_B(p_B)$.

and intuitions. The socially optimal prices $\left(p_B^W, p_S^W\right)$ are characterized by:

$$\frac{D_B'}{D_B}\left[v_S + [m-s]\right] = \frac{D_S'}{D_S}\left[1+\frac{\partial s}{\partial p_B}\right]\left[v_B + [m-s]\right], \tag{17}$$

where $D_B = D_B\left(p_B^W\right)$, $v_B = v_B\left(p_B^W\right)$ and $v_S = v_S\left(\widehat{b}_S^W\right)$.

The "optimal balancing" condition $\dfrac{D_B'}{D_B}v_S = \dfrac{D_S'}{D_S}v_B$ obtained in Rochet-Tirole (2003) must be amended to reflect internalization and market power. For an association, a decrease in price on one side must be offset by an equal increase of price on the other side. Losing, say, one seller implies a waste of surplus v_B on the buyer side, and conversely. This explains the basic optimal balancing formula. To understand how internalization and market power modify this formula, note first that an increase in p_S has two effects on seller demand as it leads to an equal[19] decrease in p_B and thus in a perceived increase in quality of service $|\dfrac{\partial s}{\partial p_B}|$. Thus everything is as if the net price increase on the seller side were $1+\frac{\partial s}{\partial p_B} < 1$, or, put differently, as if the elasticity on the seller side had been scaled down by this factor. Second, the surpluses v_j on side $j \in \{B, S\}$ created by increased demand on side i become $v_j + [m-s]$; the correction is equal to the markup m (the issuers' "surplus") minus the fraction of the buyers' surplus already internalized by the sellers. [The difference $m-s$ will play an important role in Section 7 as well.]

The buyer price p_B^{TUS} that maximizes total user surplus satisfies a formula similar to (17), except that it does not account for the issuers' margin m:

$$\frac{D_B'}{D_B}\left[v_S - s\right] = \frac{D_S'}{D_S}\left[1+\frac{\partial s}{\partial p_B}\right]\left[v_B - s\right]. \tag{18}$$

Note that for $m=0$ and $\alpha=1$, condition (17) (or (18)) yields the balanced externality condition, $v_B = v_S$, of Proposition 5.

[19]This is because this section assumes constant issuers' margin.

7 The privately optimal interchange fee

We now compute the IF that maximizes the issuers' profit, first when merchant benefits are homogeneous (Sections 7.1 and 7.2) and then when they are heterogenous (Section 7.3).

7.1 Monopoly platform

When b_S is the same for all merchants and there is no platform competition, the card association sets the IF at the maximum value that retailers accept. This is because issuers' profit $m(p_B)D_B(p_B)$ decreases with p_B.

The price p_B^m chosen by the monopoly association is given implicitly by:

$$p_S = b_S + s\left(\alpha, p_B\right).$$

Because $p_S = c - p_B + m\left(p_B\right)$ from formulas (4) and (9), we can rewrite this as:

$$b_S - c + p_B - m\left(p_B\right) + s\left(\alpha, p_B\right) = 0. \tag{19}$$

Using formula (19), the issuers' optimal buyer price can be rewritten as:

$$p_B^m = c - b_S + m\left(p_B^m\right) - s\left(\alpha, p_B^m\right). \tag{20}$$

Since $\dfrac{\partial^2 s}{\partial\alpha\partial p_B} < 0$ (by assumption 1), formula (20) shows that p_B^m increases with α.

Comparing (20) with formula (7) we see that p_B^m may be bigger or smaller than $p_B^W = c - b_S$, depending on issuer market power and on the value of α. The interchange fee chosen by the association is thus:

$$a^m = c_B - p_B^m + m(p_B^m) = b_S - c_S + s\left(\alpha, p_B^m\right). \tag{21}$$

Proposition 6. *i) A monopoly association selects the maximum interchange fee a^m that is accepted by retailers.*

ii) When $m\left(p_B^W\right) < s\left(\alpha, p_B^W\right)$ (a condition that is more likely to be satisfied when issuers' margin is small, merchant internalization is large and the net average cardholder benefit is large), a^m is larger than the socially optimal IF.

iii) When $m\left(p_B^W\right) \geq s\left(\alpha, p_B^W\right)$, the interchange fee a^m chosen by the association coincides with the (second best) socially optimal IF.

Proof of Proposition 6

Part i) has already been noted. To establish parts ii) and iii), let us recall that social welfare is equal (up to a constant) to the sum of total user surplus and banks' profit:

$$W = \phi + mD_B = \int_{p_B}^{\infty} (b_B + b_S - c)\, dH(b_B),$$

which is maximum for $p_B = p_B^W$. Since W is quasi-concave in p_B, the socially optimal buyer price is equal to p_B^W when this is compatible with merchant acceptance, i.e., when $b_S - p_S^W + s\left(\alpha, p_B^W\right) > 0$, and to p_B^m otherwise. Now

$$p_S^W = c + m - p_B^W = m + b_S,$$

which establishes ii) and iii). ■

Proposition 6 extends an earlier result of Rochet and Tirole (2002) to the case of an arbitrary internalization coefficient[20] α. It shows that when there is a single association, when acquiring is perfectly competitive and when there is no unobservable heterogeneity among retailers, the association sets the highest possible IF a^m that retailers accept. a^m is always larger than the level a^{TUS} that maximizes total user surplus (and thus consumer surplus). However it is not necessarily larger than the socially optimal IF. If issuers' margin is large, or if retailers' acceptance of cards has a limited impact on their competitive position (for example if α is close to zero and/or the average benefit v_B of cardholders per card payment is small) the interchange fee that maximizes social welfare

[20] We also relax the assumption of "cost absorption" by issuers (see footnote 14).

is too large to be acceptable by retailers. The (second best) socially optimal IF then coincides with the privately optimal one.

7.2 Platform competition

We now extend our analysis to the competition between two card associations (indexed by $k = 1, 2$) in the context of the Hotelling-Lerner-Salop model of retailing: $s(\alpha, p_B) = \alpha v_B(p_B)$. Retailers and consumers are located uniformly on a circle, and α represents the probability that consumers are aware of a retailer's card acceptance policy (see Appendices 1 and 3 for details).

For simplicity we assume that the two cards are perfect substitutes for both buyers and sellers: However the two associations may set different IFs a_1 and a_2, in which case user prices (denoted p_B^k and p_S^k, $k = 1, 2$) also differ. Assume (without loss of generality) that $a_1 \leq a_2$ so that $p_B^1 \geq p_B^2$. With perfectly competitive acquirers, we have also:

$$p_S^1 = c_S + a_1 \leq p_S^2 = c_S + a_2. \tag{22}$$

Since cards are perfect substitutes and card 2 is more expensive, retailers would be inclined to accept only card 1 but, like in Section 2, they must take into account the impact of their acceptance decisions on consumers' patronage. We can characterize retailers' acceptance decisions by looking at the behavior of function ϕ_α:

$$\phi_\alpha(p_B) \equiv (b_S - p_S) D_B(p_B) + \alpha \int_{p_B}^{\infty} (b_B - p_B) \, dH(b_B).$$

This behavior is represented in Figure 2 below.

The function ϕ_α, which we call weighted total user surplus, coincides with total user surplus when $\alpha = 1$. As shown in Appendix 1, the optimal acceptance decision of a retailer is to accept the set of cards that maximizes the expected value of ϕ_α among consumers.

Note that when retailers accept both cards, then a consumer who holds both cards (we call such a consumer a multi-homer) uses only card 2 (since $p_B^2 \leq p_B^1$). Also, it is a dominated strategy for the associations to choose p_B^k in the decreasing part of ϕ_α. Thus we can assume without loss of generality that $\phi_\alpha(p_B^1) \geq \phi_\alpha(p_B^2)$. In turn, this implies that it is a dominated strategy for merchants to accept only card 2. If they only accept only one card, it will be card 1.

A complete analysis of platform competition lies outside the scope of this paper. We content ourselves with the analysis of two polar cases: Subsection 7.2.1 looks at the case of complete multi-homing, and Subsection 7.2.2 studies complete single-homing. Appendix 3 analyzes the retailers' acceptance decisions under partial multi-homing.

7.2.1 Complete multi-homing

We stick to the convention that $a_1 \leq a_2$ and therefore $p_B^1 \geq p_B^2$. Now, if all consumers have both cards, retailers accept both cards if and only if[21]

$$\phi_\alpha\left(p_B^2\right) \geq \max\left(0, \phi_\alpha\left(p_B^1\right)\right).$$

Since issuers' profit in network 2 is a decreasing function of p_B^2, network 2 wants to choose p_B^2 as small as possible, but it is constrained by the condition $\phi_\alpha(p_B^2) \geq \phi_\alpha(p_B^1)$. By symmetry the competition between networks results in (equal) prices, set to maximize $\phi_\alpha(p_B)$.[22]

Proposition 7. *In the case of complete multi-homing (all consumers have the two cards), both associations set the same interchange fee, a^{MH}, corresponding to the cardholder price p_B^{MH} that maximizes weighted total user surplus*[23] ϕ_α.

Furthermore

$$a^{MH} \leq a^{TUS} \text{ with strict inequality unless } \alpha = 1$$

[21] This condition is a particular case of the general condition (35) obtained in Proposition 11, Appendix 3. It corresponds to a proportion of multi-homers β_{12} equal to 1.

[22] This is the two-sided version of Bertrand's undercutting argument.

[23] Specifically,

$$a^{MH} = b_S - c_S - \frac{D_B}{D_B'}(p_B^{MH})\left[m'(p_B^{MH}) - (1-\alpha)\right]$$

where $p_B^{MH} \equiv \arg\max \phi_\alpha(p_B)$. The merchant discount is then given by $p_S^{MH} = c - p_B^{MH} + m(p_B^{MH})$.

and

$$a^{MH} \leq a^{T} \text{ if and only if } m' \leq 1-\alpha.$$

Thus the interchange fee set by competing networks is too low with respect to the value that maximizes total user surplus unless merchants fully internalize cardholder surplus. Furthermore, the interchange fee passes the tourist test provided that either there is cost absorption or passthrough by issuers ($m' \leq 0$) or merchants internalize only a small fraction of cardholder surplus (α small, given that $m' < 1$).

We now examine the polar case of complete single-homing.

7.2.2 Complete single-homing

If all consumers have a single card, card i is accepted if and only if $\phi_\alpha(p_B^i) \geq 0$. This implies that, as with a single network, card associations select the highest IF a^{SH} that retailers accept. The outcome of network competition is the same as the monopoly outcome characterized in Proposition 6:

- The price p_B^{SH} paid by cardholders is characterized implicitly by $\phi_\alpha\left(p_B^{SH}\right) = 0$, which gives:

$$p_B^{SH} = c - b_S + m\left(p_B^{SH}\right) - \alpha v_B\left(p_B^{SH}\right) = p_B^m.$$

- As illustrated by Figure 2, this price is lower than the value p_B^{TUS} that maximizes consumer surplus (or equivalently total user surplus) when issuers' margin decreases with p_B (cost absorption case) but the reverse may hold when issuers' margin increases with p_B (cost amplification case).

Proposition 8. *When there is complete single-homing, the interchange fee a^{SH} set by competing networks is equal to the monopoly interchange fee a^m. a^{SH} is thus larger than or equal to the socially optimal IF* $\min\left(a^m, a^W\right)$.

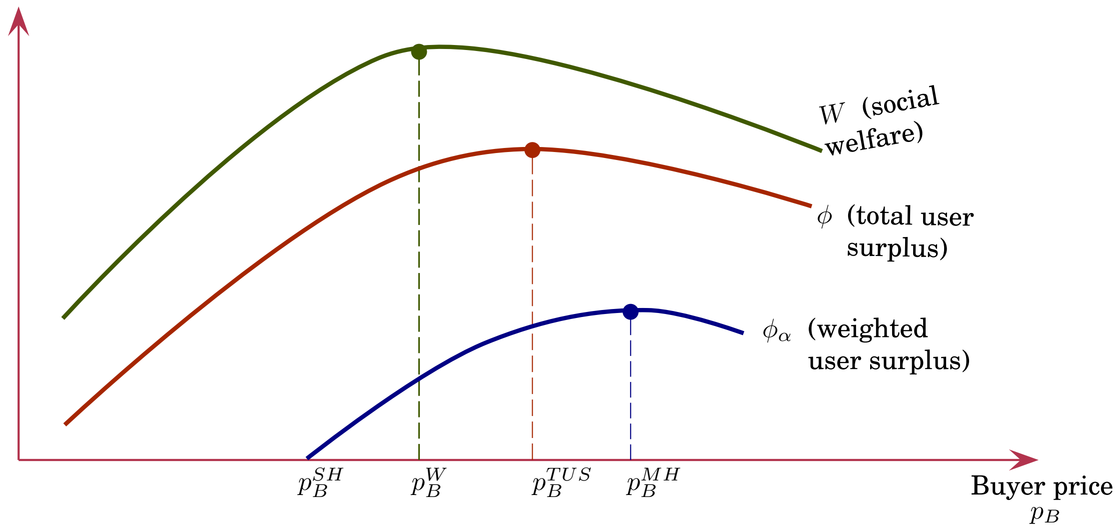

Figure 2: *Comparison of buyer prices (in the cost absorption case) when two card associations compete:*

- *when consumers single-home, the outcome is p_B^{SH} (i.e., the same as in the case of a single network),*
- *when consumers multi-home, the outcome is p_B^{MH}. It is larger than the price p_B^{TUS} that maximizes total user surplus in the cost absorption case (as represented here) but may be lower in the cost amplification case.*

We can now summarize the results for the homogeneous-merchant case.

In the case of perfect competition among issuers, the tourist test and the welfare and TUS thresholds coincide, as shown in Figure 3.

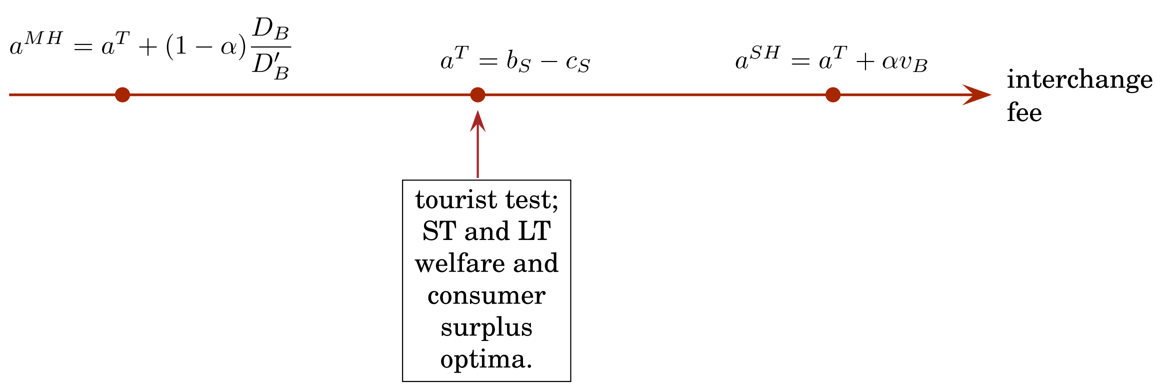

Figure 3: The perfect competition case

However, when issuers have some market power (which seems to be the belief of Competition Authorities in many regions of the world) the tourist test threshold can be too high or too low, according to whether cost absorption ($m' < 0$) or cost amplification prevails. This is represented in Figure 4.

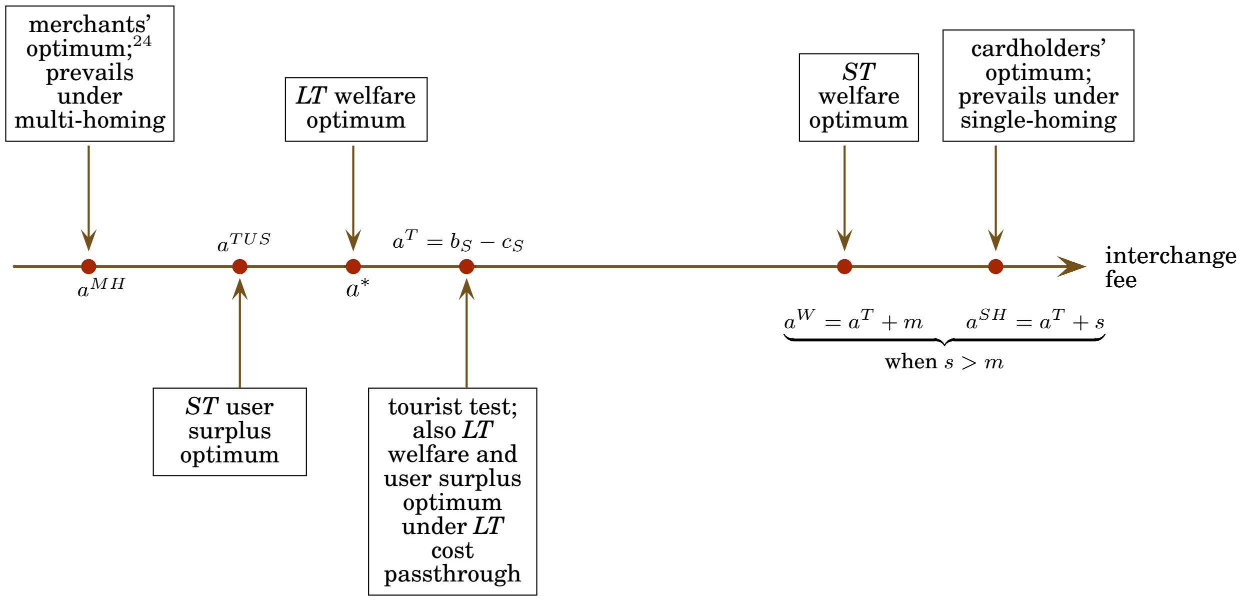

Figure 4a: Optimal values of the interchange from the points of view of merchants, cardholders and social welfare (cost absorption case)

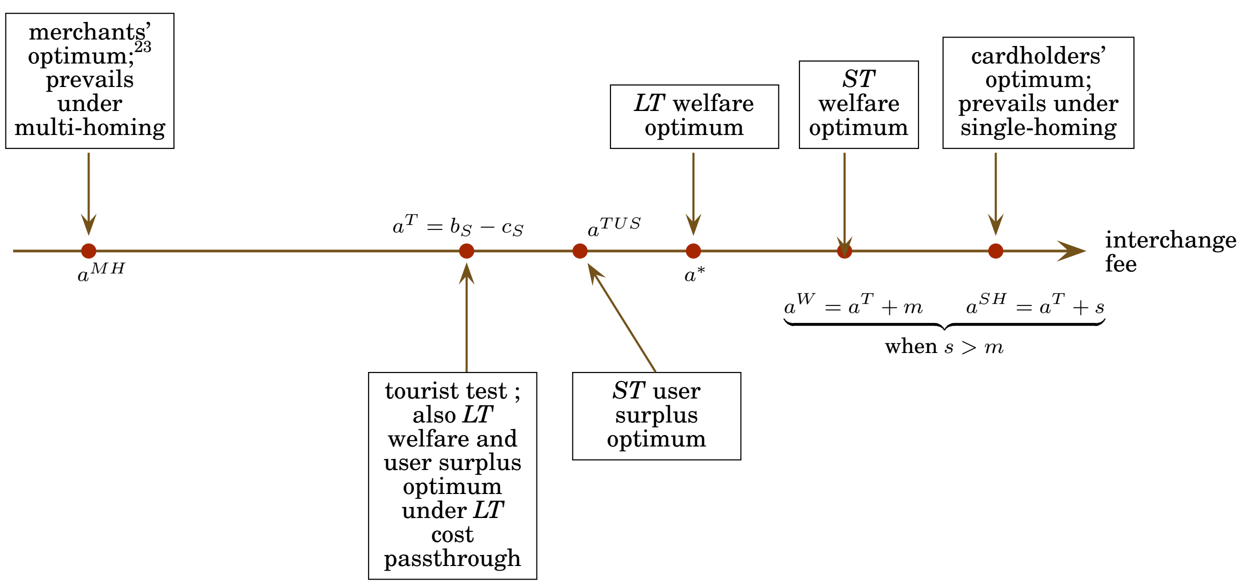

Figure 4b: Optimal values of the interchange fee from the points of view of merchants, cardholders and social welfare (cost amplification case)

[24]This assumes that retail profit decreases with retail cost, so that merchants' profit decreases with a.

7.3 Heterogenous merchants

Let us return to the extension of Section 6, in which merchants differ in their convenience benefit b_S, and compare the payment system's choice of interchange fee with the TUS and welfare benchmarks obtained in Section 6. The following result (whose proof can be found in Appendix 4) assumes that the issuers' margin m is constant, so that $p_B + p_S$ is itself a constant, and equal to $c + m$.

Proposition 9. *Assume that merchants differ in the net convenience benefit b_S that they derive from card transactions and that issuers' margin is constant (perfect cost passthrough).*
i) Provided that social welfare is quasi-concave in the buyer price, the IF selected by a monopoly association is lower than the socially optimal value if and only if average merchant surplus per card payment exceeds cardholders':

$$v_S\left(\widehat{b}_S^m\right) > v_B\left(p_B^m\right). \tag{23}$$

ii) This condition is stronger than the average tourist test ($p_S < E\left(b_S | b_S > \widehat{b}_S\right)$) when $\alpha < 1$, and equivalent to it when $\alpha = 1$.
iii) This condition is also necessary and sufficient for the privately optimal IF to be lower than the value that maximizes total user surplus when total user surplus is quasi-concave in the buyer price.

Proposition 9 shows that merchant heterogeneity has interesting consequences: It implies that the price structure chosen by a monopoly platform, in the absence of a regulation, is not systematically biased in favor of cardholders.[25] Intuitively, issuers want to maximize volume and do not account for buyer and seller surpluses. The bias induced by a private choice of IF depends on the relative surplus of the two groups of users. When

[25]Condition (23) is never satisfied when merchants are homogeneous. As established by Proposition 6, the monopoly IF is then higher than or equal to the socially optimal IF, which is consistent with Proposition 9.

the average net benefit of retailers from card payments v_S is greater than the average net benefit of consumers v_B, the IF chosen by a monopoly platform is too low, from both viewpoints of social welfare and total user surplus.

To prove Proposition 9 (see Appendix 4), one first demonstrates that in the absence of IF regulation, a monopoly network chooses the price structure (p_B^m, p_S^m) that maximizes volume $V = D_B D_S$ subject to $p_B + p_S$ being constant. It is characterized by the equality of the semi-elasticities, where the sellers' semi-elasticity must be reduced to account for their internalization of buyer surplus:

$$\frac{D_B'}{D_B}\left(p_B^m\right) = \frac{D_S'}{D_S}\left(\widehat{b}_S^m\right)\left[1 + \frac{\partial s}{\partial p_B}\right] \tag{24}$$

8 Conclusion

Merchants may accept the card even when their net benefit from doing so is negative. Accordingly, we introduced the "tourist test" (would the merchant want to refuse a card payment when a non-repeat customer with enough cash in her pocket is about to pay at the cash register?). The tourist test is attractive because it caps the merchant discount at a level at which the payment system cannot exploit the internalization effect to force merchants to accept a retail-cost-increasing means of payment. We analyzed its relevance as an indicator of "excessive" interchange fees. The relevant welfare benchmarks for this analysis are (a) short-term total welfare, (b) short-term consumer welfare, and (c) consumer welfare when issuer entry and offerings respond to industry profitability.

The paper's first contribution has been to assess whether the tourist test is a proper test for detecting excessive interchange fees. While merchants are not forced into transactions that they would not wish ex post when the IF passes the tourist test, we unveiled a number of reasons why the tourist test may yield false positives even if we focus on total user surplus (TUS), which does not account for issuers' markups. First, in the short run, the TUS-maximizing IF fails the tourist test if the issuing industry's prices exhibit cost amplification (conversely, cost absorption leads to false negatives). Second, in the long term, issuer markups translate into entry and thereby lower prices and increased variety, and so the short-term analysis yields TUS-maximizing IFs that are smaller than their long-term counterpart. Third, merchants are heterogenous, and an interchange fee that properly guides cardholders' decisions must reflect the average, not the marginal merchant benefit. This implies that the merchants who benefit least from the card, say the large retailers, are likely to fail the tourist test at the social optimum.

In the absence of platform competition or under cardholder single-homing and merchant homogeneity, the interchange fee chosen by issuers exceeds the short-term socially optimal level if and only if the fraction of cardholder benefits internalized by merchants (which depends on how knowledgeable about merchants' acceptance policies cardholders are) exceeds the issuers' per transaction markup. Under platform competition and multi-homing, the IF is smaller than the value that maximizes consumer surplus (cardholders' plus merchants' surplus), and a fortiori the value that maximizes social welfare.

References

[1] Balto, D. A.(2000) "The Problem of Interchange Fees: Costs without Benefits?" *European Competition Law Review*, 21(4): 215–224.

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

[3] Carlton, D. W., and A.S. Frankel (1995) "The Antitrust Economics of Credit Card Networks," *Antitrust Law Journal*, 63(2): 643–668.

[4] Chang, H. H., and D.S. Evans (2000) "The Competitive Effects of the Collective Setting of Interchange Fees by Payment Card Systems," *Antitrust Bulletin*, 45(3): 641–677.

[5] Evans, D. S., and R. Schmalensee (1995) "Economic Aspects of Payment Card Systems and Antitrust Policy toward Joint Ventures," *Antitrust Law Journal*, 63(3): 861–901.

[6] Farrell, J. (2006) "Efficiency and Competition between Payment Instruments", *Review of Network Economics*, 5(1): 26–44.

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

[8] Lal, R., and C. Matutes (1994) "Retail Pricing and Advertising Strategies," *Journal of Business*, 67: 345–370.

[9] Mankiw, G., and M. Whinston (1986) "Free Entry and Social Inefficiency," *Rand Journal of Economics*, 17(1): 48–58.

[10] Prékopa, A. (1973) "On Logarithmic Concave Measures and Functions," *Acta Scientiarum Mathematicarum*, 34: 339–343.

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

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

[13] —— (2003) "Platform Competition in Two-Sided Markets," *Journal of the European Economic Association*, 1(4): 990–1029.

[14] Schmalensee, R. (2002) "Payment Systems and Interchange Fees," *Journal of Industrial Economics*, 50(2): 103–122.

[15] —— (2003) "Interchange Fees: A Review of the Literature," in *The Payment Card Economics Review*, Vol. 1, Cambridge: payingwithplastic.org/National Economic Research Associates, 2003: 25–44.

[16] Vickers, J. (2005) "Public Policy and the Invisible Price: Competition Law, Regulation and the Interchange Fee", Proceedings of a conference on "Interchange Fees in Credit and Debit Card Industries" (Federal Reserve Bank of Kansas-City, Santa Fe, New Mexico, May 4-6, 2005) 231–247.

[17] Wright, J. (2003a) "Optimal Card Payment Systems," *European Economic Review*, 47: 587–612.

[18] —— (2003b) "Pricing in Debit and Credit Card Schemes," *Economics Letters*, 80: 305–309.

[19] —— (2004) "The Determinants of Optimal Interchange Fees in Payment Systems," *Journal of Industrial Economics*, 52(1): 1–26.

Appendix 1: The Hotelling-Lerner-Salop model satisfies Assumption 1

There is a continuum of consumers (of total mass normalized to one) with quasi-linear preferences. They spend their income on a composite good or "cash good" taken as a numeraire and on one unit of a "card good" sold by R retailers (being a "card good" means that consumers can pay by card as long as merchants accept it. "Cash goods" include leisure/work). The utility from purchasing the card good can differ across consumers, but is large enough, so that the aggregate demand for the card good is constant and equal to one.[26] To capture the intensity of competition in the (card good) retail sector, we use the Lerner-Salop model of product differentiation: Retailers and consumers are located uniformly on a circle of length normalized to one. The timing is as follows:

1. First, each consumer learns his preference across brands of the card good offered by the retailers, as well as the prices chosen by the retailers. Furthermore, he learns all stores' card acceptance policies with probability α, and does not learn any with probability $1-\alpha$. The consumer then chooses which store to patronize. The optimal choice minimizes the sum of three terms: the retail price p_R^j, the transportation cost $t\Delta^j$ incurred for going to the store (where Δ^j is the distance to the store and $t > 0$ is a given parameter), and the expected transaction cost associated with the payment mode (this term is detailed below).

2. Second, after choosing a store, the consumer learns his convenience benefit of using a card rather than cash in the particular instance, and chooses the payment mode among the ones accepted by the retailer. For simplicity, we restrict the analysis to two payment modes: card (if it is accepted by the retailer) and an alternative payment mode (cash or check). The relative cost $\widetilde{b}_B$ of this alternative payment mode for the consumer (also

[26] The analysis of the variant where the demand for the card good is elastic is more complex but gives similar results. It is available from the authors upon request.

equal to his convenience benefit for a card payment) is random, and drawn from another continuous distribution with cumulative distribution function H:

$$H(b_B) = \Pr\left(\widetilde{b}_B \leq b_B\right). \tag{25}$$

We adopt the convention that $\widetilde{b}_B$ is the convenience cost of a cash/check payment and 0 is that for a card payment.

As we noted, this convenience benefit $\widetilde{b}_B$ is observed by the consumer only once he is in the store. The net benefit of paying by card is thus equal to the difference $\widetilde{b}_B - p_B$. A card payment is optimal for the consumer whenever this net benefit is positive. The proportion of card payments is denoted $D_B(p_B)$:

$$D_B(p_B) = \Pr\left(\widetilde{b}_B > p_B\right) = 1 - H(p_B). \tag{26}$$

Retailers $j = 1, \cdots, R$ compete in two stages:[27]

1. First, they simultaneously decide whether to accept the card. We denote the decision of retailer j by a variable x^j equal to one if retailer j accepts the card, and zero if he does not.

2. Second, they simultaneously set their retail prices: p_R^j is chosen by retailer j so as to maximize his profit:

$$\pi^j = \left[p_R^j - \gamma - b_S - x^j (p_S - b_S) D_B(p_B)\right] y^j, \tag{27}$$

where γ is the cost of producing the card good, and b_S is the cost of the alternative payment mode for the seller (assumed for the moment to be constant across sellers). We adopt a convention similar to that for cardholders: b_S is the retailer's cost of a cash/card payment, while that for a card payment is normalized at 0. Thus $(p_S - b_S) D_B(p_B)$ represents the

[27] The timing here is irrelevant: the equilibrium would be the same if the first and second stages were simultaneous. This is because we assume that consumers' transactional benefits are drawn ex post.

expected net cost of card payments for the seller (incurred only when $x^j = 1$). Finally, y^j represents the market share of retailer j.

Retailer j's market share is a function of the retail prices set by retailer j and his neighbors $j-1$ and $j+1$, as well as the card acceptance decisions x^j, x^{j-1} and x^{j+1}:

$$y^j = \frac{1}{R} + \frac{1}{2t}\left[p_R^{j-1} + p_R^{j+1} - 2p_R^j - \alpha\left(x^{j-1} + x^{j+1} - 2x^j\right) s_B\left(p_B\right)\right] \tag{28}$$

where

$$s_B\left(p_B\right) \equiv \int_{p_B}^{\infty} \left(b_B - p_B\right) dH\left(b_B\right)$$

denotes the expected surplus that a buyer derives from the option of paying by card, and $\alpha \in [0,1]$ represents the proportion of consumers who are informed about the card acceptance decisions of retailers.[28] Our next proposition relates the retailers' acceptance decision to prices p_B and p_S and to the internalization parameter α.[29]

Proposition 10. *The retail sector equilibrium is unique and symmetric.*

- *retailers (all) accept the card if and only if:*

$$p_S \leq b_S + \alpha v_B\left(p_B\right) \tag{29}$$

  (otherwise none accepts the card). If this condition is satisfied, then:

- *retailer pass through card transaction costs (or benefits) into the retail price:*

$$p_R^* = \gamma + b_S + \frac{t}{R} - \left(b_S - p_S\right) D_B\left(p_B\right) \tag{30}$$

  (where $\gamma + b_S$ is the marginal cost of a cash transaction and $\frac{t}{R}$ the Hotelling mark-up), and the total profit of the retail sector is constant:

$$\pi = \frac{t}{R}, \tag{31}$$

[28] It may seem strange that when $\alpha < 1$ some consumers know the prices, but not the card acceptance decisions. Note however that a) they do not need to know all retail prices (as in Lal-Matutes 1994, it suffices that they know at least one price for the reasoning to apply), and b) whether they know some price or not, card acceptance helps attract consumers but in this latter case, the more general functional form posited in Assumption 1 is needed.

[29] Proposition 10 can easily be extended to the case where two card networks offer identical cards and set identical IFs.

- *consumers' total purchase cost is given by:*

$$\left[\gamma + E\left(\widetilde{b}_B\right) + b_S\right] + \frac{5t}{4R} - \phi_1, \qquad (32)$$

where ϕ_1 is obtained by taking $\alpha = 1$ in formula (29)*.*

Proof of Proposition 10

Suppose, first, that all retailers accept the card. In this case, formulas (27) and (28) show that the profit of retailer j is maximized when

$$0 = \frac{\partial \pi^j}{\partial p_R^j} = y^j - \left[p_R^j - \gamma - b_S - (p_S - b_S)\, D_B\left(p_B\right)\right]/t.$$

The equilibrium market shares are all equal $\left(y^j \equiv \frac{1}{R}\right)$, and so are retail prices:

$$p_R^j \equiv p_R^* = \left[\gamma + b_S + \frac{t}{R}\right] - (b_S - p_S)\, D_B\left(p_B\right)$$

which establishes formula (30). Consumer total purchase cost is then equal to the sum of the retail price p_R^*, the average transportation cost $\frac{t}{4R}$ and the expected transaction cost for the cardholder $E\left(\widetilde{b}_B\right) + \int_{p_B}^{\infty} (p_B - b_B)\, dH\left(b_B\right)$. Formulas (32) and (31) are then immediate.

Suppose now that retailer j considers rejecting the card. A new price equilibrium arises where all retailers except j increase their price and market share, whereas retailer j decreases his price and market share but also his cost (assuming $b_S < p_S$). It is easy to check that the net effect is to decrease retailer j's profit if and only if

$$\phi_\alpha \equiv (b_S - p_S)\, D_B\left(p_B\right) + \alpha \int_{p_B}^{\infty} (b_B - p_B)\, dH\left(b_B\right) \geq 0.$$

We call ϕ_α the weighted total user surplus. Intuitively, with a linear demand (stemming from the uniform distribution of consumers), retailer j can lower his price by $\alpha \int_{p_B}^{\infty}(b_B - p_B)dH(b_B)$ and keep the same market share as when he accepts the card. The first term in the latter inequality is (minus) the cost saving associated with rejecting the card.

The condition is thus equivalent to the assertion that accepting the card maximizes the perceived or weighted total user surplus, where only a fraction α of the buyer surplus from using the card is internalized. This condition ends the proof of formula (29) and, thus, of Proposition 10. The equilibrium is unique. ■

Finally, let us discuss our choice of the Lerner-Salop model for the description of retail demand. This model's linear demand allows a convenient aggregation of demands by consumers who are informed and uninformed about card acceptance policies. The retailers' card acceptance policy rule ($\phi_\alpha \geq 0$) holds for arbitrary demand functions when $\alpha = 0$ or 1. To see this, introduce the *perceived* and *real* hedonic prices:

$$\widehat{p}^j \equiv p^j - x^j \alpha v_B \left(p_B\right)$$

and

$$\widetilde{p}^j \equiv p^j - x^j v_B \left(p_B\right).$$

For arbitrary demand functions and $\alpha \in \{0, 1\}$, retailer j's profit is:

$$\left[\widehat{p}^j - \left[\gamma + b_S - x^j \phi_\alpha\right]\right] y^j \left(\widehat{p}^j, \widehat{p}^{-j}\right).$$

Thus when $\alpha \in \{0, 1\}$, retailer j accepts the card iff $\phi_\alpha \geq 0$, independently of the shape of the demand function y^j. That $\phi_\alpha \geq 0$ is also the exact criterion for acceptance when $0 < \alpha < 1$ by contrast hinges on the linearity of demand.

Appendix 2: Welfare and TUS optima under heterogenous merchants

Social welfare is given by:

$$W = \int_{p_S - s(\alpha, p_B)}^{\infty} \int_{p_B}^{\infty} \left[b_S + b_B - c\right] dH\left(b_B\right) dG\left(b_S\right) \tag{33}$$

and total user surplus has the same expression with $[b_S + b_B - c - m]$ as the integrand.

Social welfare is maximized for a value p_B^W of buyers' price such that:

$$\frac{\partial W}{\partial p_B}(p_B^W, \widehat{b}_S^W) = \frac{\partial W}{\partial \widehat{b}_S}(p_S^W, \widehat{b}_S^W)\left(1 + \frac{\partial s}{\partial p_B}\left(\alpha, p_B^W\right)\right). \tag{34}$$

Using formula (33) we see that

$$\frac{\partial W}{\partial p_B}(p_B^W, \widehat{b}_S^W) = D_B'(p_B^W)\int_{\widehat{b}_S^W}^{\infty}(p_B^W + b_S - c)dG(b_S).$$

Now by definition:

$$v_S(\widehat{b}_S^W)D_S(\widehat{b}_S^W) = \int_{\widehat{b}_S^W}^{\infty}(b_S - \widehat{b}_S^W)dG(b_S).$$

Thus we can write:

$$\frac{\partial W}{\partial p_B}(p_B^W, \widehat{b}_S^W) = D_B'(p_B^W)D_S(\widehat{b}_S^W)\left[v_S(\widehat{b}_S^W) + m - s\left(\alpha, p_B^W\right)\right].$$

Similarly

$$\frac{\partial W}{\partial \widehat{b}_S}(p_B^W, \widehat{b}_S^W) = D_S'(\widehat{b}_S^W)D_B(p_B^W)\left[v_B(p_B^W) - s\left(\alpha, p_B^W\right) + m\right].$$

Thus the first-order condition for welfare maximization can be rewritten as:

$$\frac{D_B'(p_B^W)}{D_B(p_B^W)}\left[v_S(\widehat{b}_S^W) + m - s\left(\alpha, p_B^W\right)\right] = \frac{D_S'(\widehat{b}_S^W)}{D_S(\widehat{b}_S^W)}\left[v_B\left(p_B^W\right) - s\left(\alpha, p_B^W\right) + m\right]\left(1 + \frac{\partial s}{\partial p_B}\left(\alpha, p_B^W\right)\right)$$

This is equivalent to formula (17).

Appendix 3: Retailers' acceptance decisions under partial multi-homing

For simplicity, this appendix takes as given the proportions of buyers who own the two cards. Specifically, let β_k $(k = 1, 2)$ denote the proportion of buyers who own only card k (the single-homers) and β_{12} denote the proportion of buyers who own both cards (the multi-homers). Like before, we assume for simplicity that all consumers have a card[30] (i.e., $\beta_1 + \beta_2 + \beta_{12} = 1$) and also that only a proportion $\alpha \in [0, 1]$ of consumers are

[30] This is not inconsistent with our model, which assumes that issuers do not charge fixed fees to cardholders.

aware of retailers' card acceptance policy before they select which store to patronize. The next proposition characterizes the equilibrium of the retail sector (both in terms of card acceptance and retail prices) as a function of payment card prices and parameters α, β_1, β_2 and β_{12}.

Proposition 11. *At the equilibrium of the retail sector:*

- *retailers accept both cards if and only if:*

$$\beta_1 \phi_\alpha\left(p_B^1\right) + \left(\beta_2 + \beta_{12}\right) \phi_\alpha\left(p_B^2\right) \geq \max\left[0, \left(\beta_1 + \beta_{12}\right) \phi_\alpha\left(p_B^1\right), \left(\beta_2 + \beta_{12}\right) \phi_\alpha\left(p_B^2\right)\right] \quad (35)$$

If this condition is satisfied then:

- *retail prices*[31] *are given by:*

$$p_R^* = \left[\gamma + b_S + \frac{t}{R}\right] - \beta_1\left(b_S - p_S^1\right) D_B\left(p_B^1\right) - \left(\beta_2 + \beta_{12}\right)\left(b_S - p_S^2\right) D_B\left(p_B^2\right),$$

- *aggregate demand for the card good equals*[32] $\beta_1 D\left(u_1^*\right) + \left(\beta_2 + \beta_{12}\right) D\left(u_2^*\right)$, *where*

$$u_k^* = p_R^* + \frac{t}{4R} + E(\widetilde{b}_B) - s_B\left(p_B^k\right), \qquad k = 1, 2.$$

- *consumer surplus equals* $\beta_1 S\left(u_1^*\right) + \left(\beta_2 + \beta_{12}\right) S\left(u_2^*\right)$,
- *finally, the total profit of the retail sector is*

$$\pi_R = \frac{t}{R}\left[\beta_1 D\left(u_1^*\right) + \left(\beta_2 + \beta_{12}\right) D\left(u_2^*\right)\right]$$

Proof of Proposition 11:

It proceeds similarly to that of Proposition 10. In equilibrium, retailers accept the (set of) cards that maximize the expectation of weighted user surplus ϕ_α over all buyers. Accepting only card $k(k = 1, 2)$ allows a fraction $(\beta_k + \beta_{12})$ of buyers to pay by card generating

[31] We assume that retailers find it too costly or are not allowed to charge different prices for transactions settled through different cards.

[32] For simplicity, we assume that the determinants of the choice of cards by consumers are independent from the parameters that determine the gross utility $\tilde{u}$ obtained by consumers when they consume the card good.

weighted user surplus $\phi_\alpha(p_B^k)$. The right-hand side of condition (35) corresponds to the maximum of three outcomes: accepting no card, accepting card 1 alone, and accepting card 2 alone. The left-hand side of condition (35) corresponds to the expectation of weighted user surplus when the merchant accepts both cards. In this case multi-homers prefer to use card 2 (since $p_B^2 \leq p_B^1$), which explains the fraction $(\beta_2 + \beta_{12})$ of buyers who use card 2. This establishes the first bullet point in Proposition 11. Retail prices (second bullet point) are given at equilibrium by the average unit cost faced by merchants (including the net cost of card payments) plus a constant margin $\frac{t}{R}$. The other bullet points are immediate. ■

Appendix 4: Proof of Proposition 9

Given our assumption that issuers' margin is constant, their profit is maximized for a value p_B^M of buyers' price that maximizes the volume of card transactions, given by:

$$V(p_B) = D_B(p_B)D_S(\widehat{b}_S(p_B)),$$

where $\widehat{b}_S(p_B) = c + m - p_B + s\,(\alpha, p_B)$.

Now p_B^m is given by the first-order condition:

$$\frac{V'}{V}(p_B^m) = \frac{D_B'}{D_B}(p_B^m) - \frac{D_S'}{D_S}(\widehat{b}_S^m)(1 + \frac{\partial s}{\partial p_B}\,(\alpha, p_B^m)) = 0,$$

which gives (24).

When social welfare is quasi-concave with respect to the buyer price, the privately optimal IF (associated with buyer price p_B^m) is excessively low whenever

$$\Delta \equiv \frac{\partial W}{\partial p_B}(p_B^m, \widehat{b}_S^m) - \left[1 + \frac{\partial s}{\partial p_B}\,(\alpha, p_B^m)\right] \frac{\partial W}{\partial \widehat{b}_S}(p_B^m, \widehat{b}_S^m) < 0.$$

Adapting the formulas obtained above, we see that

$$\begin{aligned} \Delta \quad &= \quad D_B'(p_B^m)D_S(\widehat{b}_S^m)\left[v_S(\widehat{b}_S^m) + m - s\,(\alpha, p_B^m)\right] \\ &\quad -[1 + \frac{\partial s}{\partial p_B}\,(\alpha, p_B^m)]D_S'(\widehat{b}_S^m)D_B(p_B^m)[v_B(p_B^m) - s\,(\alpha, p_B^m) + m]. \end{aligned}$$

Applying condition (24) allows us to simplify this expression:

$$\begin{aligned}\Delta &= D'_B(p_B^m)D_S(\widehat{b}_S^m)\left[v_S(\widehat{b}_S^m) + m - s(\alpha, p_B^m) - v_B(p_B^m) + s(\alpha, p_B^m) - m\right],\\ &= D'_B(p_B^m)D_S(\widehat{b}_S^m)\left[v_S(\widehat{b}_S^m) - v_B(p_B^m)\right].\end{aligned}$$

Thus $\Delta < 0 \Leftrightarrow v_S(\widehat{b}_S^m) > v_B(p_B^m)$, and the proof of Proposition 9i) is complete.

For establishing 9ii) it suffices to notice that $v_S\left(\widehat{b}_S^m\right) = E\left(b_S|\widehat{b}_S^m\right) - p_S^m + s(\alpha, p_B^m)$ and that $s(\alpha, p_B^m) \leq v_B(p_B^m)$, with strict inequality when $\alpha < 1$ and equality when $\alpha = 1$.

Part iii) results from the comparison of formulas (17) and (24) ■

Appendix 5: Heterogeneity in internalization

Let us investigate heterogeneity in α (b_S being the same for all merchants). Merchants accept cards if and only if their internalization parameter α exceeds a critical value $\alpha^*(p_B)$ defined implicitly by the relation:

$$b_S = p_S - s(\alpha, p_B) = c + m(p_B) - p_B - s(\alpha, p_B). \tag{36}$$

(We assume that there is cost absorption or cost passthrough ($m' \leq 0$). In this case, α^* is a decreasing function[33] of p_B. Thus increasing p_B gets more merchants on board. Social welfare has a simple expression:

$$W = \left[\int_{p_B}^{\infty} (b_B + b_S - c)dH(b_B)\right] \cdot \Pr(\alpha \geq \alpha^*(p_B)).$$

At the (first best) social optimum all merchants accept cards ($\alpha^* = \alpha_{\min}$, where $\alpha_{\min}$ is the minimum value of α among merchants) and:

$$p_B^W = c - b_S.$$

This can be implemented by setting the interchange fee at

$$a = b_S - c_S + m = a^W,$$

[33] Since $\alpha^* \leq 1$ and $v'_B \geq -1$, the assumption $m' \leq 0$ implies that the right-hand side of (36) decreases in p_B.

provided this is compatible with universal merchant acceptance:

$$s\left(\alpha_{\min}, p_B^W\right) \geq m\left(p_B^W\right).$$

If this condition is not satisfied, let $\psi(p_B) = \Pr(\alpha \geq \alpha^*(p_B))$ denote the proportion of merchants who accept cards, as a function of the cardholder price p_B. The socially optimal level of p_B, which we denote p_B^W, satisfies:

$$\frac{W'}{W}(p_B) = \frac{(p_B + b_S - c)D_B'}{\int_{p_B}^{\infty}(b_B + b_S - c)dH(b_B)} + \frac{\psi'}{\psi}(p_B) = 0.$$

By contrast, a monopoly network selects a cardholder price p_B that maximizes the profit of issuers:

$$\Pi = m(p_B)D_B(p_B)\psi(p_B).$$

Therefore p_B^m satisfies:

$$\frac{\Pi'}{\Pi} = \frac{m'}{m}(p_B) + \frac{D_B'}{D_B}(p_B) + \frac{\psi'}{\psi}(p_B) = 0.$$

Under monotone comparative statics (that is, if Π is log concave) we can compare p_B^m and p_B^W. Indeed if $m' \leq 0$ then, given that $D_B' \leq 0$, and

$$\frac{p_B + b_S - c}{\displaystyle\int_{p_B}^{\infty}(b_B + b_S - c)dH(b_B)} < \frac{1}{D_B},$$

we have:

$$\frac{\Pi'}{\Pi}(p_B^m) = 0 = \frac{W'}{W}(p_B^W) \geq \frac{D_B'}{D_B}(p_B^W) + \frac{\psi'}{\psi}(p_B^W) \geq \frac{\Pi'}{\Pi}(p_B^W),$$

which implies, by log concavity of Π, that

$$p_B^W \geq p_B^m.$$

Similarly total user surplus can be written as:

$$TUS\left(p_B\right) = \left[\int_{p_B}^{\infty}\left[b_B + b_S - c - m\left(p_B\right)\right]dH\left(b_B\right)\right]\psi\left(p_B\right).$$

The term between brackets is maximum for a value of (p_B) that is above p_B^T (this is because we have assumed cost absorption: see Table 1).

Since $\psi' > 0$, log concavity assumptions guarantee that this is a fortiori true of p_B^{TUS}. Therefore $a^{TUS} < a^T$.

Thus we have shown:

Proposition 12. *When the internalization parameter α is heterogenous across merchants and there is cost absorption or pass through $(m' \leq 0)$, the interchange fee selected by a monopoly association is (weakly) bigger than the socially optimal interchange fee.*[34]

Moreover, under the same assumptions, the tourist test threshold is above the level of the interchange fee that maximizes total user surplus.

[34]For analogous reasons, the buyer price p_B^{SH} that results from competition between identical platforms when cardholders single-home is greater than p_B^m. This contrasts with the case of homogeneous merchants, where $p_B^{SH} = p_B^m$.