Pricing Payment Cards

Özlem Bedre* Emilio Calvano†

December 2008‡

Abstract

In a payment card association such as Visa, each time a consumer pays by card, the bank of the merchant (*acquirer*) pays an *interchange fee* (IF) to the bank of the cardholder (*issuer*) to carry out the transaction. This paper studies the determinants of socially and privately optimal IFs in a card scheme where services are provided by a monopoly issuer and perfectly competitive acquirers to heterogeneous consumers and merchants. Different from the literature, we distinguish card membership from card usage decisions (and fees). In doing so, we reveal the implications of an asymmetry between consumers and merchants: the *card usage* decision at a point of sale is delegated to cardholders since merchants are not allowed to turn down cards once they are affiliated with a card network. We show that this asymmetry is sufficient to induce the card association to set a *higher* IF than the socially optimal IF, and thus to distort the structure of user fees by leading to too low card usage fees at the expense of too high merchant fees. Hence, cap regulations on IFs can improve the welfare. These qualitative results are robust to imperfect issuer competition, imperfect acquirer competition, and to other factors affecting final demands, such as elastic consumer participation or strategic card acceptance to attract consumers.

JEL Classification: G21; L42; L31

Keywords: Payment card associations; Interchange fees; Merchant fees

*Toulouse School of Economics, GREMAQ, and (visiting at) Simon Graduate School of Business, University of Rochester, obedre@yahoo.com.

†Toulouse School of Economics and Department of Economics, Harvard University, ecalvano@fas.harvard.edu.

‡The first version of this paper is released in April 2008.

1 Introduction

In a payment card association such as Visa, each time a consumer settles a purchase by card, the bank of the merchant (*acquirer*) pays an *interchange fee* (IF) to the bank of the cardholder (*issuer*) to carry out the transaction. In practice, the IF is either set bilaterally by the issuer and the acquirer, or multilaterally by the members (issuers and acquirers) of the card scheme, or by regulatory agencies.[1] In the last decades, interchange fees have attracted much attention of economists[2], mainly because policy makers are concerned that IFs result in higher merchant fees, and thus higher retail prices, without leading to proven efficiencies. Interchange fee arrangements have already been subject to cap regulations[3] or found anti-competitive.[4] Under the pressure of the European Commission, in January 2008 MasterCard stopped setting a multilateral IF for cross-border consumer card transactions in Europe.[5] The EC is now investigating Visa's IFs.

This paper compares the determinants of socially and privately optimal interchange fees. The literature mostly does not take into account the fact that consumers make two types of decisions: subscribing to a payment card platform (*membership*) or not, and every time they make a purchase choosing whether or not to use the card (*usage*). The situation is different for merchants. They make one type of decision: subscribing to a payment card platform or not. Once they become a member of a card platform, they are not allowed to refuse cards of this platform. Hence, it is the cardholder who decides whether the card is used at an affiliated merchant. We show that taking this asymmetry into account and considering non-linear card fees change results considerably: When both merchants and consumers are heterogeneous, the payment card platform sets a too high IF to subsidize card users at the expense of too high merchant fees.[6] Different from the literature, the upward distortion of the privately optimal IF does not depend on cost and/or demand specifications or parameters. Our paper unambiguously, predicts that cap regulations on IFs can improve the social welfare, and thus delivers clear policy implications. However, we do not find any support for widely used issuer cost-based cap

[1] Levels of IFs vary between 0.5% and 2.5% of the transaction value. See the European Commission (EC)'s Retail Banking Sector Inquiry (2007) and the Reserve Bank of Australia (RBA)'s report (2007).

[2] See Chakravorti and To (2003), and Evans and Schmalensee (2005b) for a review of the literature on IFs and their regulation. Rochet (2003) provides a synthesis of the theoretical literature on IFs. Weiner and Wright (2005) compare practices of the payment card industry across various countries.

[3] For instance, in Australia, Spain, Switzerland and Mexico.

[4] In 2005, the UK Office of Fair Trading found MasterCard's IF arrangements illegal and recently in New Zeland card associations have been alleged of price fixing in the setting of IFs. For a review of recent regulatory developments in the world, see the RBA's report (2007).

[5] The European Commission, COMP/34.579, December 2007.

[6] We assume that merchants are not allowed to surcharge card users. Card schemes mostly impose a No-Surcharge Rule preventing their affiliated merchants from surcharging card users.

regulation. In line with the literature, we indeed find that the socially optimal IF reflects two considerations: relative demand elasticities (marginal users) and relative net surpluses (average users). We furthermore show that regulating the IF is not enough to achieve full efficiency in the industry. The IF affects only the allocation of the total user price between consumers and merchants whereas efficiency requires also a lower total price level due to positive externalities between the two sides.

We separate card membership from card usage decisions by assuming that consumers learn their convenience benefits from card transactions after their cardholding decisions. These benefits depend on, for instance, their cash holdings, the transaction value, the distance to the closest ATM, and the availability of foreign currency at the point of sale. Consumers hold a card in order to secure the option of paying by card in the future. Accordingly, we refer to the expected usage value of a payment card as the *option value*. Card membership decisions depend on *average* fees and benefits, whereas card usage decisions are determined by *marginal (transaction)* fees and benefits. Lower (or negative) transaction fees (e.g., reward programs) attract new members through a higher option value and foster card usage of each member. On the other side, merchants make only membership (card acceptance) decisions by comparing their *average benefit* from card payments with the *average merchant fee* to be paid to the acquirer. Fixed merchant fees and benefits are therefore found to be redundant in the analysis.

We consider a single card association which sets an IF to maximize the total profits of its members (issuers and acquirers). In the benchmark analysis, we show that a monopoly issuer sets the card usage fee equal to its transaction cost, which is the cost of issuing minus the IF, since it could internalize *incremental* card usage surpluses of buyers through a fixed membership fee. Perfectly competitive acquirers pass their transaction cost, which is the cost of acquiring plus the IF, fully to merchants (sellers). We first illustrate the conflict between buyers' and sellers' interests on the level of IF; the average buyer prefers a high IF, whereas the average seller prefers a low IF. Through issuer profits, the card scheme could internalize incremental transaction surpluses of buyers, and thus sets the IF maximizing buyers' card usage surplus. The socially optimal IF is lower than the privately optimal IF since the former takes into account incremental transaction surpluses of buyers as well as those of sellers. Hence, in equilibrium cardholders pay *too little* and merchants pay *too much* per transaction compared to what would prevail with the socially optimal IF. We furthermore show that the equilibrium card usage fee is negative (justifying reward programs) when the card acceptance demand is very insensitive to merchant fees.

Our results are not simply due to the internalization of the buyer surplus. The issuer also internalizes *some of the* merchant surplus via the IF. By setting the card transaction fee at its transaction cost, the issuer passes on all interchange revenue to card users[7] since the issuer should give incentives to cardholders to use their cards. It is profitable to subsidize card users only if there is a means to recoup the card usage surplus; a fixed membership fee serves as this means.

Full internalization of the buyer surplus is not essential for our qualitative results. We extend the benchmark analysis to elastic consumer participation resulting from (1) heterogeneity of consumers in their membership benefits (or costs) prior to cardholding decisions and (2) price competition between differentiated issuers. Intuitively, an issuer sets its card usage fee at the transaction cost as long as it internalizes *incremental* card usage surpluses of its existing cardholders. This is the case whenever consumers' membership benefits are independent from their transaction benefits. We show that the argument remains valid even if competition dissipates some or all membership revenue. In both extensions, issuer profits are shown to be increasing in buyers' card usage surplus.[8] Therefore, the card association sets again the IF maximizing the buyer surplus from card usage.

The qualitative results are robust to strategic card acceptance as a quality investment and/or to steal business from a rival. In such situations, merchants would be willing to pay more, ceteris paribus, for card acceptance, since they could internalize some of the average buyer surplus from card usage. Card associations could exploit the decrease in the *merchant resistance* to merchant fees by further increasing the IF. The results are also robust to imperfect acquirer competition. For instance, a monopoly acquirer could internalize *some* of the merchant surplus from card usage even if non-linear merchant fees are available. The reason is that heterogeneous merchants make only membership (card acceptance) decisions by comparing their average benefit with the average merchant fee.

Payment card networks are characterized by two-sided membership (network) externalities between consumers and merchants[9], as well as one-sided usage externalities from consumers to

[7]By this means, card users internalize the externality they exert on the rest of the industry when paying by card.

[8]For issuer competition, this is the case under standard conditions on demands for differentiated issuers. We show that such conditions hold in common demand specifications, e.g., Dixit (1979).

[9]The the value of accepting (holding) a card depends on how many consumers (merchants) hold (accept) that card.

merchants.[10] Baxter (1983) shows that IFs help internalize externalities between consumers and merchants, so sustain efficient card usage when the industry is perfectly competitive. Assuming some market power on the issuer side, Rochet and Tirole (2002) propose the first positive analysis and show that the profit-maximizing IF is either equal to or higher than the welfare-maximizing IF. They illustrate a possible upward distortion of the privately optimal IF by formalizing the fact that competing merchants accept a cost-increasing card in order to attract consumers. The greater the competitive edge guaranteed by card acceptance, the more likely that a card network sets a too high IF resulting in too high merchant fees. Schmalensee (2002) notes that by assuming homogeneous merchants Rochet and Tirole ignore the trade-off between attracting card usage and card acceptance, which is identified by Schmalensee. Wright (2004) extends Rochet and Tirole's analysis by assuming merchants to be heterogeneous in their benefits from card payments. But then, Wright shows that, the relation between the socially and privately optimal IFs becomes complex and depends on demand and cost parameters, as well as the relative degree of competition among acquirers and issuers. Competition between card schemes does not necessarily reduce the equilibrium IF. It instead results in an IF favoring the side which subscribes to only one card scheme (the competitive bottleneck).[11]

Our analysis encompasses the literature. If we assume that merchants are homogeneous in their benefits from card transactions and set fixed card fees at zero, we obtain Guthrie and Wright (2003, Proposition 2), which shows that the socially and privately optimal IFs coincide and result in under-provision of card payment services. If we allow for fixed card fees, we show that both the regulator and the association sets Baxter's IF implementing the first best card usage volume. Focusing on heterogeneous merchants, if we set card membership fees and benefits at zero, we get the benchmark of Rochet and Tirole (2003), that is a monopoly platform charging consumers and merchants per transaction. If we furthermore allow for competition among merchants, we obtain Rochet and Tirole (2002). In Rochet and Tirole (2002, 2003), linear or non-linear card fees, per-transaction and/or fixed benefits, would give the same results because consumers get a card if and only if they will use it for future transactions. Different from our paper, the authors assume that consumers know their convenience benefits from paying by card before their cardholding decision, i.e., consumers make only one decision, that is cardholding.

[10] Every time a cardholder pays by card, the merchant receives some benefits from the card transaction and pays a merchant fee to the acquirer.

[11] See Guthrie and Wright (2003). The theory of IFs has many parallels with the growing literature on access charges and two sided markets. See, for instance, Armstrong (2002, 2006), Laffont et al.(2003), and Rochet and Tirole (2003, 2006).

The paper is organized as follows. Section 2 introduces our general setup. Section 3 illustrates preliminary observations explaining why we focus on average merchant fees and benefits without loss of generality. Section 4 presents the benchmark analysis. Section 5 computes the Lindahl and Ramsey fees and compares them with the equilibrium outcome of regulating only IF. Section 6 introduces elastic consumer participation and imperfect issuer competition. Section 7 compares our setup and results with the literature where we also discuss the robustness to imperfect acquirer competition. Section 8 concludes with some policy implications. All formal proofs are presented in the appendix.

2 A Model of the Payment Card Industry

A single payment card association (e.g., Visa) provides card payment services to card users, that are cardholders and merchants, through intermediaries: issuers (cardholders' banks) and acquirers (merchants' banks). In order to fit the payment card industry, we assume that the card association prohibits merchants from surcharging buyers paying by card (the so called No-Surcharge Rule).[12] We also assume that issuers have market power while acquirers are competitive[13]. One motivation for the highly-competitive-acquirers assumption is the merchants' low search and switching costs, and limited brand loyalty. Mostly, merchants have some staff to deal with different banks for different services. However, consumers face high search and switching costs, and thus prefer dealing with one bank for all banking services and have brand preferences.[14] At the end of our analysis, we show that our main qualitative results are robust to allowing for symmetric market structures on the two sides, e.g., a monopoly issuer and a monopoly acquirer.

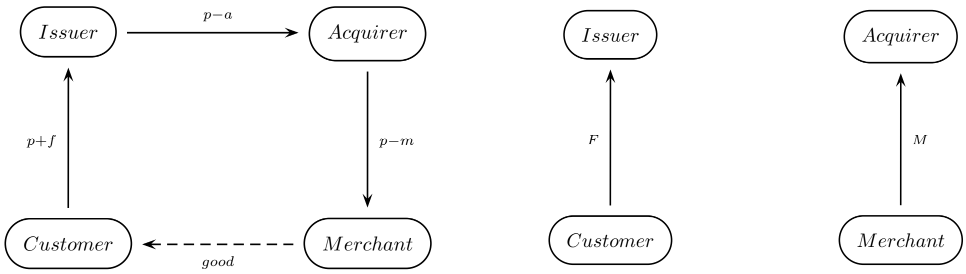
Consumption Surplus We consider a continuum (mass one) of consumers and a continuum (mass one) of locally monopoly merchants.[15] We assume that consumers are willing to purchase a unit good from each merchant. The consumption value of a unit good paid by cash is $v > 0$. A consumer gets $v - p$ from purchasing a unit good by cash at price p and the seller gets p from

[12] This is a common practice of payment schemes (e.g. Visa and MasterCard). Although surcharging is permitted in the UK and in Australia, it is uncommon mainly due to transaction costs of price discrimination among buyers using different forms of payment.

[13] To avoid dealing with conflicting interests within a bank, we assume that issuers are not in the acquiring business.

[14] See Evans and Schmalensee (1999), Rochet and Tirole (2002, 2005), and the EC's report (2007).

[15] In the extensions, we will discuss the robustness of our results to merchant competition.



(a) Card Transaction (b) Card Subscription/Acceptance

Fig. 1: Card Payments

this purchase.[16]

Card Usage Surplus Consumers (or buyers) get an additional payoff of $b_B - f$ when they pay by card, where b_B denotes the net per-transaction benefit[17] from paying by card rather than an alternative method, and f denotes the transaction fee to be paid to the issuer. Similarly, merchants (or sellers) get an additional payoff of $b_S - m$ when they are paid by card rather than other payment methods, where b_S denotes the per-transaction benefit received[18] and m denotes the merchant discount (or fee) to be paid to the acquirer. We do not restrict these quantities to be positive, we instead allow for negative benefits (distaste for card transactions) and negative fees (e.g. reward schemes like cash-back bonuses, frequent-flyer miles). For each card transaction, the issuer (the acquirer) incurs cost c_I (c_A). Let c denote the total cost of a card transaction, so $c = c_I + c_A$. The card association requires the acquirer to pay an interchange fee a per transaction to the issuer. The acquirer's (the issuer's) transaction cost is thus $c_I - a$ ($c_A + a$). Figure 1a summarizes the flow of transaction fees in a payment card association.

Card Membership Surplus Buyers and sellers get benefits from card membership which are not specific to transactions.[19] We denote these *fixed* benefits by B_B and B_S which buyers and sellers are aware of before their membership decisions. Let F and M denote the membership (fixed) fees to be paid respectively to the issuer and acquirer (Figure 1b). To simplify the notation, we assume that the fixed costs of issuing an extra card and acquiring an extra merchant are zero.

[16]We assume that unit cost of retailing is the same irrespective of the payment method and we set the retailing cost at zero to simplify the analysis.

[17]Such as foregoing the transaction costs of withdrawing cash from an ATM or converting foreign currency.

[18]Such as convenience benefits from guarantee on transactions, low transaction costs, and easy accounting.

[19]E.g. cardholders enjoy security of not carrying big amounts of cash, membership privileges (such as access to VIP), travel insurance, ATM services (such as account balance sheets, money transfers, etc.), social prestige (club effects); merchants benefit from safe transactions.

Consumers and merchants are heterogeneous both in their *usage* and *fixed* benefits from card payments. The benefits (b_B, b_S, B_B, B_S) are assumed to be independently distributed on some compact interval with smooth atomless cumulative distribution functions satisfying the Increasing Hazard Rate Property (IHRP).[20] Note that consumers' convenience benefits, (b_B, B_B), are independent of the market in which they purchase.

Timing

Stage i: The payment card association (alternatively a regulator) sets an interchange fee, a.

Stage ii: Banks observe a, simultaneously the issuer sets card fees (F, f)[21] and acquirers set symmetric merchant fees (M, m).

Stage iii: Merchants and consumers realize their membership benefits B_S and B_B. They then decide simultaneously whether to accept and hold the payment card, respectively, and which bank to patronize.

Stage iv: Merchants set retail prices. Merchants and consumers realize their transaction benefits b_S and b_B, consumers decide whether to buy and buyers decide whether to pay by card.

Consumers and merchants maximize their expected payoff. The card association sets the interchange fee to maximize the sum of the profits earned by its issuers and acquirers. The equilibrium is SPNE.

Consumption Surplus versus Transaction Surplus Let $G(b_B)$ and $g(b_B)$ denote respectively the cumulative distribution and density function of b_B. To simplify the benchmark analysis, we make the following assumption:

$$A1 : v \geq c - \underline{b_B} - \underline{b_S} + \frac{1 - G(\underline{b_B})}{g(\underline{b_B})}.$$

Guthrie and Wright (2003, Appendix B) show that under A1 monopoly merchants set $p = v$ regardless of whether they accept the card or not.[22] In other words, the utility of cash users,

[20]The IHRP leads to *log-concavity of demand functions* (for cardholding, for card usage, and for card acceptance), which is used to satisfy the second-order conditions of the optimization problems.

[21]Alternatively, issuer i sets (F_i, f_i) in the case of competing issuers.

[22]Note that this is different than the no-surcharge rule which prevents a merchant from price discriminating between card users and cash users.

v, is so high that merchants do not want to exclude cash users, so lose surplus that they could extract from cash users, by setting a price higher than v. A1 rules out the case where merchants accept cards to offer a better quality of services and recoup the extra consumer surplus from the quality improvement (card usage) by increasing retail prices. After solving the benchmark model, we show that relaxing A1 would reinforce our results.

3 Preliminary Observations

By A1, all merchants set $p = v$ and therefore all consumers purchase a unit good from each merchant.

If a merchant accepts cards, a proportion, α_B, of its transactions (to be determined in equilibrium) is settled by card. The net payoff of type B_S merchant from accepting cards is:

$$B_S - M + E\left[b_S - m\right]\alpha_B, \tag{1}$$

which is the sum of the membership and expected transaction surpluses when merchant fees are M, m. The number of merchants that join the payment card network is thus:

$$\alpha_S \equiv \Pr(B_S - M + E\left[b_S - m\right]\alpha_B \geq 0).$$

Note that α_S depends only on the average merchant benefit and fee, which are respectively:

$$\tilde{b}_S \equiv E\left[b_S\right] + \frac{B_S}{\alpha_B} \qquad \text{and} \qquad \tilde{m} \equiv m + \frac{M}{\alpha_B}.$$

We thus get $\alpha_S = \Pr(\tilde{b}_S \geq \tilde{m})$. There is a redundancy of merchant fees. By setting $M = 0$ and the transaction fee $m = \tilde{m}$, both the equilibrium card acceptance demand and the acquirer revenue[23] remain constant.[24] The card acceptance decision is sunk when b_S is learnt and thus cannot be affected by its realization. Only the *average* benefit known *before* the acceptance decision matters. For a given α_B, our framework is thus equivalent to a setup where merchants are heterogeneous in their *average* benefits prior to their card acceptance decisions. The same is not true on the consumer side, since consumers make *two* decisions: card membership and usage. Cardholding depends only on the average benefit and card fee, whereas card usage depends on

[23]This would still be the case if we assumed some market power on the acquiring side.

[24]Note that risk-adverse merchants would strictly prefer usage charges since they are borne only if a transaction occurs.

the transaction benefit and fee.

Without loss of generality we focus our attention on a simpler model where fixed merchant benefits and fees are set zero, and merchants are heterogeneous in their transaction (and thus average) benefit b_S which is realized *prior to* their card acceptance decisions. We assume that b_S is continuously distributed on some interval $[\underline{b_S}, \overline{b_S}]$ with CDF $K(b_S)$, PDF $k(b_S)$ and increasing hazard rate $k/(1-K)$.[25]

4 Benchmark Analysis

We now assume that consumers are homogeneous at the membership stage, i.e.,. we set $B_B = 0$. Either everyone or no one holds the card. We relax this simplifying assumption in Section 6 allowing for elastic cardholding.

4.1 Behavior of Consumers and Merchants

Cardholders pay by card if and only if their transaction benefit exceeds the usage fee. Thus, the quasi-demand for card usage is defined as

$$D_B(f) \equiv \Pr\left(b_B \geq f\right) = 1 - G(f),$$

which is the proportion of cardholders paying by card at fee f.

Given that all consumers hold the card, the merchant's surplus from accepting cards is

$$(b_S - m)D_B(f)$$

A merchant accepts cards whenever $b_S \geq m$.[26] The card acceptance demand is then[27]

$$D_S(m) \equiv 1 - K(m).$$

[25]Formally, for any pair of distributions $Z(b_S)$ and $F(B_S)$, one can always find a distribution $K(b_S)$ such that, for given m, f, the number of card users and card acceptance demand, and thus the equilibrium outcome coincide.

[26]Card acceptance is not affected by card usage, i.e., there is no externality imposed by consumers on merchant participation. We could restore this externality by allowing for fixed merchant fees, since the card usage demand then affects the average merchant fee, without changing our conclusions (see the discussion in the previous section).

[27]In the previous section, we define α_B and α_S as the proportions of respectively card users and card acceptance. In the current notation, α_i corresponds to D_i for $i = B, S$.

We define respectively buyers' and sellers' average surpluses from card usage:

$$v_B(f) \equiv E\left[b_B - f \mid b_B \geq f\right] \qquad \text{and} \qquad v_S(m) \equiv E\left[b_S - m \mid b_S \geq m\right].$$

The expected value or *option value* of holding a payment card is defined as

$$\Phi_B(f, m) \equiv v_B(f) D_B(f) D_S(m),$$

where $D_B(f)D_S(m)$ is the volume of card transactions at fees (f, m). The value of a card is increasing in expected usage at affiliated merchants, D_B, and in merchant participation, D_S. All consumers hold a card iff the option value of the card is higher than its fixed fee:

$$\Phi_B(f, m) \geq F.$$

4.2 Behavior of the Issuer and Acquirers

Taking the IF as given, perfectly competitive acquirers set a merchant fee equal to their transaction cost, $m^*(a) = a + c_A$. The issuer instead maximizes its profit under the consumers' participation constraint:

$$\max_{F,f} \left[(f + a - c_I) D_B(f) D_S(m) + F\right] \quad st.: \quad \Phi_B(f, m) \geq F. \tag{2}$$

The issuer internalizes buyers' card usage surpluses through $F = \Phi_B(f, m)$, so sets the usage fee at its transaction cost, $f^*(a) = c_I - a$. We thus have $f^*(a) + m^*(a) = c$ for any a.

4.3 Privately and Socially Optimal Interchange Fees

Given the equilibrium reactions of banks, we now establish three critical levels of IF: the buyers-optimal IF, a^B, which maximizes the buyer surplus from card usage (or the option value of the card), the sellers-optimal IF, a^S, maximizing the seller surplus from card usage, and a^V, which maximizes the volume of card transactions, .

Lemma 1 *Interchange fees* $\left(a^B, a^S, a^V\right)$ *exist uniquely and satisfy* $a^S < a^V < a^B$.

This lemma highlights the tension between consumers' and merchants' interests over the level of IF. An increase in the interchange fee has two effects. On one hand, it induces a higher merchant fee and thus lowers the number of shops where cards are welcome. On the other hand, it induces a lower card usage fee and thus increases the number of transactions per affiliated store. Under the IHRP, the average buyer surplus, v_B, is decreasing in card usage fee f, so increasing in IF. Symmetrically, the average seller surplus, v_S, is decreasing in merchant fee m, so in IF. Going above (below) the volume-maximizing IF increases the buyer (seller) surplus.

□ **Equilibrium Fees.**

Since perfectly competitive acquirers get zero, the association sets an IF to maximize the issuer's profits. Given $f^*(a) = c_I - a$ and $m^*(a) = c_A + a$, by setting an interchange fee the card association implicitly allocates the total cost of a transaction between the two sides of the market. The corresponding program is:

$$\max_{f,m} v_B(f) D_B(f) D_S(m) \qquad \textit{st.:}\ f + m = c. \tag{3}$$

The optimal allocation is such that the impact of a small variation of f on the option value of the card is equal to the impact of a small variation of m. The equilibrium IF thus implements the price structure that maximizes the option value. From Lemma 1 we know that a^B maximizes the card usage surplus of buyers. The card association therefore sets $a^* = a^B$ inducing the usage fee $f^B = c_I - a^B$ such that (see the proof of Lemma 1),

$$v_B(f^B) = \frac{c - f^B}{\eta_S(c - f^B)}, \tag{4}$$

where $\eta_S(m) = -\frac{mD'_S(m)}{D_S(m)}$ denotes the elasticity of the card acceptance demand, D_S. Observe that f^B is increasing in the elasticity of the merchant demand.

Corollary 1 *The equilibrium usage fee is negative if card acceptance demand is not very elastic (more precisely, if* $\eta_S < \frac{c}{v_B(0)}$*).*

When card acceptance decisions are not very sensitive to changes in the IF, in equilibrium, the card association subsidizes cardholders by setting an IF which induces negative transaction fees (e.g. reward programs like cash back bonuses and frequent flyer miles).

□ **Optimal Regulation.**

The problem of a regulator is also stated as a cost allocation program similar to (3):

$$\max_{f,m} \left[v_B(f) + v_S(m)\right] D_B(f) D_S(m) \qquad st.:\ f + m = c \tag{5}$$

The regulator's problem with two-part tariff card fees would be equivalent to a Ramsey Planner's problem if card fees were linear. The optimal allocation satisfies:[28]

$$v_S(m) D_S(m) D_B'(f) = v_B(f) D_B(f) D_S'(m).$$

Increasing the usage fee discourages some consumers from using their cards and merchants lose $v_S D_S D_B'$ from such a reduction in card usage. Similarly, an increase in the merchant discount results in a reduction in the number of merchants accepting the card, which in turn decreases consumer surplus by $v_B D_B D_S'$. The optimal trade-off depends on how many end users are discouraged on one side, how much net surplus the other side loses due to this reduction in demand, and it balances the welfare losses of merchants and of consumers.

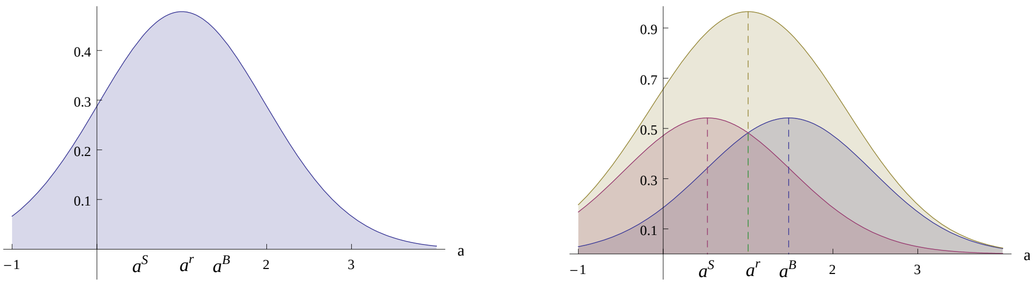
We now proceed to compare the regulator's choice with the choice of the association. Our formulation makes clear that the only difference is in the *allocation* of the total price c across the two sides of the market.

Proposition 1 *The privately optimal IF is higher than the socially optimal IF. Hence, in equilibrium, cardholders pay too little and merchants pay too much per transaction.*

The main conceptual difference between the privately and the socially optimal allocation is that the issuer does not take into account the average loss of sellers when lowering f at the expense of a higher m. This results in a price structure bias in favor of buyers. Equation (4) tells us that the distortion is more significant if the elasticity of merchant demand (merchant resistance) is lower. As in any two-sided market, merchant participation is necessary to realize transaction benefits. The size of the upward distortion is therefore limited by merchant participation. The easier to get merchants on board, the more significant the bias in favor of buyers.

Rochet and Tirole (2003, 2006b) derive the optimal pricing structure for a monopoly platform setting linear prices to both sides. As opposed to theirs, our equilibrium fees do not

[28] To obtain the optimality condition, we use $[v_B(f)D_B(f)]' = -D_B(f)$ and $[v_S(m)D_S(m)]' = -D_S(m)$, which arise from the definitions of $v_B(f)$ and $v_S(m)$.



(a) Volume of Transactions

(b) Sellers, Buyers and Socially Optimal IF

Fig. 2: Notable IFs

maximize the total volume of transactions. We thus cannot conclude that in equilibrium there is over-provision of card services just by noticing that the socially optimal IF is smaller than the privately optimal one. In fact, in equilibrium, there could be under-provision of card services, even though buyers get more favorable terms than what they would get if the industry were regulated. Figure 2 shows an example of under-provision in equilibrium when b_B and b_S are normally distributed around 0 and 2 with unitary variance and $c_I = c_A = 1/2$. There would be over-provision of payment card services in equilibrium, for instance, when seller demand is very insensitive to changes in IF. The volume-maximizing IF would then be very close to the buyers-optimal IF (since changing IF changes $D_B D_S$ nearly as in the same way as D_B) and thus the socially optimal IF would fall very below a^V, which together would result in an over-provision of card services in equilibrium. Denoting the elasticities of quasi-demands by $\eta_B = -\frac{fD'_B}{D_B}$ and $\eta_S = -\frac{mD'_S}{D_S}$, in line with the literature[29], we characterize the socially optimal price structure as:

$$\frac{f}{m} = \frac{\eta_B}{\eta_S} \div \frac{v_B}{v_S}.$$

The socially optimal allocation of the total price $f + m = c$ is achieved when relative user prices are equal to the ratio of the relative demand elasticities and the relative average surpluses.[30]

5 Efficient Fees

In this section we characterize the first best (Lindahl) fees and the second best (Ramsey) fees where we further require a balanced budget (BB) on both sides of the market.[31] A Ramsey

[29] Rochet and Tirole (2003, 2006).

[30] The same property holds for the optimal access charge between backbone operators or between telecom operators where the access charge allocates the total cost between two groups of users (consumers and web sites in backbone networks, call receivers and call senders in telecommunication networks). See Laffont et al. (2003).

[31] The rationale for the latter case comes from the problem of a regulator who can control end-user prices but cannot or does not want to run and/or subsidize operations, and therefore has to leave enough profits to keep the industry attractive for private investors.

planner can subsidize card users (i.e., set a usage fee lower than its associated cost) as long as the subsidy is financed by revenue from the same side (through the membership fee) or from the other side (through the IF). We will see that at the first and second best fees, the total user fee, $f+m$, is lower than the cost of a transaction, c. This observation shows that regulating only the IF is not enough to achieve efficiency.

5.1 Lindahl Fees.

Consider the problem of a public monopoly running the payment card industry in order to maximize the total welfare:

$$\begin{aligned} &\max_{F,f,m} W \equiv (f+m-c)D_B(f)D_S(m) + \left[v_B(f)+v_S(m)\right]D_B(f)D_S(m) \\ &st.: \quad \Phi_B(f,m) \geq F \end{aligned} \tag{6}$$

We ignore the participation constraint of consumers since it can always be satisfied by modifying appropriately the fixed fee, which has no impact on welfare. Let f^{FB}, m^{FB}, and p^{FB} denote respectively the first best levels of the card usage fee, the merchant fee and the total transaction price.

Proposition 2 *The first best total price (per transaction) is lower than the total cost of a transaction and equal to $c - v_B(f^{FB})$. The socially optimal allocation of such a price is achieved when*

$$v_B(f^{FB}) = v_S(m^{FB}),$$

that is, when the average buyer surplus is equal to the average seller surplus.

An extra card user (merchant) attracts an additional merchant (card user) which generates average surplus v_S (v_B). At the optimum, the two externalities must be equalized, so the total price is given by $f^{FB}+m^{FB} = c - v_S(m^{FB}) = c - v_B(f^{FB})$. Intuitively, each type of user is charged a price equal to the cost of a transaction minus a discount reflecting its positive externality on the other segment of the industry.

5.2 Ramsey Fees.

A Ramsey planner solves (6) subject to an additional constraint: $\Pi_A, \Pi_I \geq 0$, where Π_A and Π_I denote respectively acquirers' and the issuer's profits. The question is whether the first best fees can be implemented while providing the issuer and each acquirer with non-negative profits.

Proposition 3 *The second best total price is higher than the first best total price, but still lower than the total cost of a transaction. The second best fees are* $m^{SB} = c_A$ *and* $f^{SB} = c_I - v_S(c_A)$.

At the first best total price, if the merchant discount is strictly above the cost of acquiring, the usage fee would be so below the issuing cost (i.e., $f < c_I - v_B$) that taking all buyer surplus through the membership fee is not enough to guarantee a non-negative profit to the issuer. Hence, in equilibrium the budget balance condition of acquirers must be binding, $m^{SB} = c_A$. If $m^{FB} \neq c_A$, the planner would like to reduce m below c_A to equate the average surpluses of two sides, but then acquirers would not participate. Therefore, at the second best, the average surplus of buyers would be higher than the sellers', i.e.,. $v_B(c_I - v_S(c_A)) \geq v_S(c_A)$, which results in non-negative payoffs for the issuer.

Let us now summarize the key findings of the benchmark analysis. In the payment card industry, the first best efficiency requires a total user price lower than the cost of a transaction due to positive externalities between the two sides. A payment card association where card services are provided by a monopoly issuer and perfectly competitive acquirers distorts both the total user price *and* the allocation of the total price across the two sides of the market. Regulating the interchange fee corrects the latter distortion but not the former.[32] Therefore, a regulated industry is not fully efficient. The second best total price is higher than the first best, but still lower than the cost of a transaction. Below-cost card usage fees can be financed through membership revenue and thus do not necessarily trigger budget imbalances.

6 Extensions

6.1 Elastic Cardholding

In the benchmark model, the monopoly issuer, through the membership fee, fully internalizes the buyer surplus. This is the case because buyers are homogeneous at the membership stage (ex-ante). The objective of this section is to show that the issuer would set the fees maximizing buyers' card usage surplus even if it left some information rent to ex-ante heterogeneous consumers. In order to do so, we assume that holding a card provides extra benefits, B_B, which are not specific to transactions and are independently distributed across buyers' population.[33] We

[32]This is because the sum of the transaction fees (in subgame equilibrium) is independent of the IF.

[33]E.g., security of not carrying large amounts of cash, privileges for card members (such as access to VIP), travel insurance, ATM services (such as account balance sheets), social prestige (club effects).

allow for $B_B < 0$, i.e., intrinsic fixed costs of membership. Consumers are assumed to be heterogeneous at the membership stage such that B_B is distributed with a positive density $h(B_B)$ over its support $\left[\underline{B_B}, \overline{B_B}\right]$, and $H(B_B)$ refers to the corresponding CDF with an increasing hazard rate $\frac{h}{1-H}$.

□ **Behavior of Consumers and Merchants**

Under A1 all merchants set $p = v$ and only $D_S(m)$ of merchants accept the payment card. All consumers purchase a unit good from each merchant and only $D_B(f)$ of cardholders pay by card (whenever possible). At the membership stage a card is worth

$$B_B + \Phi_B(f, m),$$

that is, the sum of the membership benefits (or costs) and the option value of being able to pay by card. Type B_B gets a card if and only if the total benefits from cardholding exceed its price. The number of cardholders, which is denoted by λ, is then

$$\begin{aligned}\lambda(F - \Phi_B(f, m)) &= \Pr\left[B_B + \Phi_B(f, m) \geq F\right] \\ &= 1 - H\left(F - \Phi_B(f, m)\right),\end{aligned}$$

which is a continuous and differentiable function of card fees (F, f) and merchant discount m. In the benchmark model, consumers were ex-ante homogeneous, so the demand for cardholding was inelastic (λ was either 0 or 1). Here, consumers are ex-ante heterogeneous, so the cardholding is elastic.

□ **Behavior of the Issuer and Acquirers**

Perfectly competitive acquirers set $m^*(a) = c_A + a$. The issuer solves:

$$\max_{F,f} \left[(f + a - c_I)D_B(f)D_S(m) + F\right] \lambda(F - \Phi_B(f, m)) \tag{7}$$

The usual optimality conditions bring the equilibrium fees:

$$f^*(a) = c_I - a, \qquad F^*(a) = \frac{1 - H(F^*(a) - \Phi_B(a))}{h(F^*(a) - \Phi_B(a))}.$$ [34]

The fixed fee is characterized by the Lerner formula. The usage fee is set at the marginal cost of

[34] To simplify the expressions, we write $\Phi_B(a)$ instead of $\Phi_B(c_I - a, c_A + a)$.

issuing even though cardholding demand is elastic, i.e., the issuer cannot extract all surplus of buyers. The intuition is that, in equilibrium, the issuer internalizes the surplus of the "marginal buyer" rather than that of the "average buyer". Consider for instance usage fees f',f'' where $f' > f'' \geq c_I - a$. Lowering the usage fee from f' to f'' has two effects. First, cardholders save on the usage fee for inframarginal transactions. Second, the expected number of transactions, and thus the associated transaction surplus increases. The combination of these two effects increases the value of holding a payment card by

$$\Delta \equiv \Phi_B(f'', m) - \Phi_B(f', m)$$

Such incremental utility can in turn be captured by the issuer through a Δ-increase in the membership fee. Such an increase would leave the marginal buyer unchanged. These together increase the issuer's profit if

$$\Delta \geq (f' + a - c_I)D_B(f')D_S(m) - (f'' + a - c_I)D_B(f'')D_S(m),$$

or if

$$-\int_{f''}^{f'} \frac{\partial \Phi_B}{\partial f} df \geq \int_{f''}^{f'} \left[(f + a - c_I)D'_B(f) + D_B(f)\right] D_S(m)df,$$

which is the case since $\frac{\partial \Phi_B}{\partial f} = -D_B D_S$, $f'' \geq c_I - a$ and $D'_B < 0$. However, lowering the usage fee below $c_I - a$ is not profitable. Therefore, the issuer maximizes its profits by reducing the usage fee to its transaction cost and capturing incremental value of its card through a higher fixed fee.

□ **Privately and Socially Optimal Interchange Fees**

The following result is analogous to Lemma 1, the only difference is that the measure of buyer surplus is modified to account for membership benefits.[35]

Lemma 2 *There exists a unique buyers-optimal IF, $\widetilde{a}^B$, a unique sellers-optimal IF, $\widetilde{a}^S$, and a unique volume-maximizing IF, $\widetilde{a}^V$, such that $\widetilde{a}^S < \widetilde{a}^V < \widetilde{a}^B = a^B$.*

Note that the buyers-optimal IF is the same as the one of the benchmark model since the

[35]The new measure is given by

$$\{E\left[B_B \mid B_B \geq F^*(a) - \Phi_B(a)\right] + \Phi_B(a)\} \lambda(F^*(a) - \Phi_B(a)).$$

IF maximizing the option value of the card, a^B, also maximizes the participation of buyers, and thus the total buyer surplus. One can prove that the sellers-optimal and the volume-maximizing IFs are higher than their counterparts in the benchmark analysis. Increasing (decreasing) IF not only increases (decreases) the quasi-demand for card usage but also continuously increases (decreases) the number of cardholders. Therefore, the volume of card transactions is maximized at a higher IF when cardholding is elastic. Similarly, sellers prefer a higher IF because they take into account the possible loss in card usage due to a reduction in the number of cardholders when IF decreases, i.e., they resist less to an increase in IF.

As in Section 3, both the association's and the regulator's problems can be expressed as cost allocation problems. The card association maximizes the issuer's profits subject to the equilibrium reactions of the issuer and acquirers:

$$\max_{F,f,m} F\lambda(F-\Phi_B(f,m)) \qquad \textit{st.:} \quad \textit{i.}\ f+m=c \ \textit{ii.}\ F=\frac{1-H\left(F-\Phi_B(f,m)\right)}{h\left(F-\Phi_B(f,m)\right)} \tag{8}$$

A regulator maximizes instead the total welfare:

$$\max_{F,f,m}\left\{\left[v_B(f)+v_S(m)\right]D_B(f)D_S(m)+E\left[B_B \mid B_B\geq F-\Phi_B(f,m)\right]\right\}\lambda(F-\Phi_B(f,m)) \tag{9}$$

subject to the same set of constraints as problem (8).

Proposition 4 *When both the cardholding and the card usage demands are elastic,*

i. the privately optimal IF is equal to the buyers-optimal IF, and

ii. the privately optimal IF is higher than the socially optimal IF. Hence, in equilibrium, cardholders pay too little and merchants pay too much per transaction.

The intuition behind part (i) is parallel to our discussion of the equilibrium usage fee. The issuer wants to maximize buyers' surplus from card usage even though it cannot capture all surplus of buyers. It is sufficient that the issuer captures incremental card usage surpluses of its existing cardholder base. As noted in Section 4.3, a socially optimal allocation takes into account also the seller surplus, thus the second part of the proposition follows along the lines of Proposition 2.

6.2 Competing Issuers

In this section, we modify our benchmark setup by allowing for imperfect competition between two issuers, denoted by I_1 and I_2, which provide differentiated payment card services in the same card scheme and charge their customers two-part tariff card fees. Consumers have preferences both for payments made by card instead of other methods of payment and for the issuer itself (i.e., brand preferences). Brand preferences are due to, for instance, quantity discounts (e.g., family accounts), physical distance to a branch, or consumers' switching costs deriving from the level of informational and transaction costs of changing some banking products (e.g., current accounts),

Card i refers to the payment card issued by I_i, for $i = 1, 2$. We denote the net price of card i by t_i, which is the difference between its fixed fee and the option value of holding card i, i.e., $t_i = F_i - \Phi_B(f_i, m)$ where the merchant fee is m. The demand for holding card i is denoted by $Q_i(t_i, t_j)$, for $i \neq j$, $i = 1, 2$. We make the following assumptions on the demand functions:

$$A2 \ : \ \frac{\partial Q_i}{\partial t_i} < 0 \qquad A3 : \frac{\partial Q_i}{\partial t_j} > 0 \qquad A4 : \left|\frac{\partial Q_i}{\partial t_i}\right| > \frac{\partial Q_i}{\partial t_j}$$
$$A5 \ : \ \frac{\partial^2 InQ_i}{\partial t_i^2} \qquad A6 : \left|\frac{\partial^2 InQ_i}{\partial t_i^2}\right| > \left|\frac{\partial^2 InQ_i}{\partial t_i \partial t_j}\right|$$

A2 states that the demand for holding a card is decreasing in its net price. A3 ensures the substitutability between the card services provided by different issuers so that the demand for holding card i is increasing in the net price of card j. By A4, we furthermore assume that this substitution is imperfect, and thus the own price effect is greater than the cross price effect. By assuming that Q_i is log-concave in net price t_i, A5 ensures the concavity of the optimization problems. A6 states that own price effect on the slope of the log-demand is higher than the cross price effect. In the Appendix, we give examples of classic demand functions for differentiated products [such as Dixit (1979), Singh and Vives (1984), Shubik and Levitan (1980)] which satisfy all of our assumptions.

☐ **Behavior of the Issuers and Acquirers**

Perfectly competitive acquirers set $m^*(a) = c_A + a$. Taking the IF and card j's fees given, I_i's problem is to set (F_i, f_i) by

$$\max_{F_i, f_i} \left[(f_i + a - c_I)D_B(f_i)D_S(m) + F_i\right] Q_i \left(F_i - \Phi_B(f_i, m), F_j - \Phi_B(f_j, m)\right),$$

so I_i sets $f_i^*(a) = c_I - a$ to maximize the consumer surplus from its card services. The option value of the card is therefore equal to $\Phi_B(c_I - a, c_A + a) \equiv \Phi_B(a)$ regardless of the identity of the issuer. I_i sets F_i^*, which satisfies

$$\epsilon_i(F_i^*, F_j; a) = 1,^{36}$$

where $\epsilon_i \equiv -F_i \frac{\partial Q_i / \partial F_i}{Q_i}$ refers to the elasticity of I_i's demand with respect to its fixed fee, F_i. Using the log-concavity of demand (A5) we get $\frac{\partial \epsilon_i}{\partial F_i} > 0$. Whenever ϵ_i is greater (less) than 1, I_i has a strict incentive to lower (raise) its fixed fee until $\epsilon_i = 1$. The following lemma shows that the issuers charge cardholders higher fixed fees when the option value of the card is higher.

Lemma 3. *When two differentiated issuers of a payment card scheme are competing in two-part tariff card fees, the equilibrium fixed fees are increasing in the usage value (option value) of the cards,* $\frac{\partial F_i^*}{\partial \Phi_B} > 0$ *for* $i = 1, 2$.

□ Privately and Socially Optimal Interchange Fees

Privately Optimal Interchange Fee. The association's problem is to set the IF maximizing the sum of the issuers' profits $\Pi_1^* + \Pi_2^*$ where

$$\Pi_i^* = F_i^* Q_i \left(F_i^* - \Phi_B(a), F_j^* - \Phi_B(a) \right),$$

given that $\epsilon_i(F_i^*, F_j^*; a) = \epsilon_j(F_j^*, F_i^*; a) = 1$. Now, our claim is that the association sets $a^* = a^B$ maximizing the option value of the card, $\Phi_B(a)$. We prove the claim by showing that each issuer gains more when Φ_B increases. Applying the Envelope Theorem, we derive

$$\frac{\partial \Pi_i^*}{\partial \Phi_B} = F_i^* \left[-\frac{\partial Q_i}{\partial t_i} - \frac{\partial Q_i}{\partial t_j} + \frac{\partial Q_i}{t_j} \frac{\partial F_j^*}{\partial \Phi_B} \right],$$

[36] Observe that the optimality condition is indeed given by the Lerner formula:

$$markup_i = \frac{1}{\epsilon_i},$$

where the markup of each duopolist issuer is equal to 1 since there is no fixed cost in our setup. If instead each issuer paid fixed cost C_I per card, the solution to I_i's problem would be

$$markup_i \equiv \frac{F_i^* - C_I}{F_i^*} = \frac{1}{\epsilon_i},$$

whereas we simply assume that $C_I = 0$, so we have $markup_i = 1$.

which helps us to identify two types of effects on I_i's profit as the option value changes:

- **Demand Effect:** The direct effect of the option value on the issuer's demand and is composed of own and cross demand effects, where
  - *own demand effect,* which is the first term in the brackets, is positive because the demand decreases in the net price of the card (A2) increasing in the option value of the card.
  - *cross demand effect,* which is the second term in the brackets, is negative because the demand increases in the net price of the rival's card (A3) decreasing in the option value.

The net demand effect is positive since the positive own demand effect dominates the negative cross demand effect (A4).

- **Strategic Effect:** The effect of the option value on the issuer's demand through changing the fixed fee of the rival (the third term in the brackets). Since the rival's fixed fee increases in the option value (by Lemma 3), the strategic effect is positive.

Hence, we show that both the demand and strategic effects are positive. Each issuer's profit is therefore increasing in the option value, Φ_B, and thus the association sets $a^* = a^B$ to maximize Φ_B. This statement is true even if the demand functions are asymmetric (see the proof of Lemma 3).

Through the association setting the IF at a_B, the issuers achieves double goals. First, maximizing the option value of the card maximizes the willingness to pay of consumers maximizing the amount of fixed fees collected (demand effect). Second, increasing the option value of the card, softens price competition, since a higher option value, ceteris paribus, induces the rival to set a higher fixed fee (strategic effect).

Socially Optimal Interchange Fee. For a given IF, the social welfare is higher than in the case of a monopoly issuer, since allowing for issuer competition does not change transaction fees which are again equal to the transaction cost of issuing, but issuer competition reduces the monopoly fixed fee to duopoly fixed fees. Because fixed fees are lump-sum transfers between cardholders and the issuers, they do not affect the average surplus from card usage. However, fixed fees affect the total number of cardholders, and thus the total volume of card transactions. For a given IF, the social welfare in the case of issuer competition is therefore equal to

$$W(a) = [v_B(c_I - a) + v_S(c_A + a)] \, D_B(c_I - a) D_S(c_A + a) \, [Q_1(F_1^*, F_2^*, a) + Q_2(F_1^*, F_2^*, a)]$$

where $\epsilon_1(F_1^*, F_2^*, a) = \epsilon_2(F_1^*, F_2^*, a) = 1$. Let a^{rc} denote the socially optimal fee under issuer competition.

Lemma 4. *Suppose that two differentiated issuers of a payment card scheme are competing in two-part tariff card fees. Assuming that the demands for the issuers are symmetric, we have $a^r < a^{rc}$, where a^r is the socially optimal IF in the case of a monopoly issuer.*

Issuer competition increases the social welfare for a given IF. When cardholding demands are symmetric, issuer competition reduces the scope for inefficiency as the socially optimal IF becomes closer to the privately optimal IF. The social welfare is therefore higher with issuer competition than in the case of a monopoly issuer.

The following proposition shows that the qualitative results of the benchmark model are valid when there is imperfect issuer competition:

Proposition 5 *When two differentiated issuers of a payment card scheme are competing in two-part tariff card fees where the cardholding demands are symmetric and satisfy A2-A6,*

i. the privately optimal IF is equal to the buyers-optimal IF, which is equal to a^B, and

ii. the privately optimal IF is higher than the socially optimal IF. Hence, in equilibrium, cardholders pay too little and merchants pay too much per transaction.

7 Comparisons with the Literature

In this section we compare some particular cases of our framework with the literature and illustrate that our paper encompasses the existing work. In the first part, we obtain Rochet and Tirole (2003)'s benchmark by focusing on linear card fees and transaction benefits ($F = B_B = 0$). We furthermore emphasize the role of separating cardholding from card usage decisions in distinguishing linear card fees (and transaction benefits) from non-linear card fees (and fixed benefits). Next, we suppose that merchants are homogeneous ($\underline{b_S} = \overline{b_S}$) and obtain the benchmark of Guthrie and Wright (2003). We moreover show that allowing for fixed card fees results in efficient card usage at the Baxter (1983)'s IF, which is both privately and socially optimal in this case. We then relax assumption A1 to take into account strategic card acceptance to attract consumers. We show that this extension reinforces our results and argue furthermore that introducing competition among merchants, like in Rochet and Tirole (2002), would strengthen our conclusions. Finally, contrary to Wright (2004), we discuss why our results are not deriven by the asymmetry in the pass-through rates of the issuers and acquirers. We moreover show that even if the market structure is completely symmetric on both sides, e.g., monopoly issuer and monopoly acquirer, the association still sets a higher IF than the socially optimal IF.

7.1 Cardholding vs Card Usage Decisions and Fees

With perfectly competitive acquirers and the monopoly issuer, the payment card association wants to maximize the issuer's profits, so the issuer plays the role of the platform owner, and through setting an IF, it determines also the merchant fee. If we considered only linear user fees and transaction benefits ($F = B_B = 0$), we would be in the benchmark case of Rochet and Tirole (2003, Section 2), where a monopoly platform charges consumers and merchants per transaction. Their Proposition 1 shows that a monopoly platform sets a total price by the standard Lerner formula for elasticity equal to the sum of the elasticities of user demands, and for a given total price, the price structure is given by the ratio of two elasticities. A socially optimal IF would account for the relative demand elasticities as well as the relative average surpluses of users. The comparison between the privately and socially optimal IFs would then depend on the elasticities of user demands and how much each group of users values a transaction.

In Rochet and Tirole (2002, 2003), consumers know their convenience benefits from paying by card before their cardholding decision, so considering linear or non-linear card fees, per-transaction and/or fixed benefits would give the same results in their analysis. This is because consumers get a card if and only if they will use it for future transactions. This timing implicitly assumes that consumers make only one decision, whether to hold the card or not, by comparing their average benefit with the average card fee. In our model, however, convenience benefits from a card transaction are realized after the cardholding decision. We therefore allow the possibility of having cardholders who do not use the card once they realize that they have higher benefits from paying by cash at the point of sale. Consumers get the card in order to secure the *option* of paying by card in the future. Our timing is more realistic since it is true that consumers learn most of their transaction benefits from card usage when they are at the store (such as the availability of cash holdings, the transaction value, or the distance to the closest ATM) or after they decide to hold the card (such as discounts on products sold by affiliates). Furthermore, it allows us to capture the fact that consumers make two distinct decisions: cardholding and card usage. This timing is first used by Guthrie and Wright (2003), who restrict the analysis to linear fees, and by doing so, their benchmark results are the same as Rochet and Tirole (2002, 2003).

The main contribution of our paper is separating card membership from card usage decisions (and fees). Card membership decisions depend on *average* fees and benefits, whereas card usage decisions are determined by *transaction* fees and benefits. Fixed merchant fees and benefits are found to be redundant since merchants make only membership (card acceptance) decisions by comparing their average benefit from card payments with the average merchant fee. We show that this asymmetry between consumers and merchants induces issuers to subsidize card users because they could recoup card usage surplus of buyers through fixed fees. Therefore, issuers set card usage fees at their transaction cost passing all interchange revenue (some of the merchant surplus) to card users. The card scheme then sets the IF maximizing the card usage surplus of buyers.

7.2 Homogeneous Merchants

Merchants of the same sector with a similar scale (volume of sales) have more or less the same benefits from accepting a payment card, so they can be regarded as homogeneous in their card acceptance benefits. Acquirers are mostly aware of this fact and set the same fees for merchants of the same type.[37] On the other hand, there are some merchants whose category is not very well known or defined so that acquirers could not price discriminate across unknown types, and thus set the same merchant fee for different types of merchants. This case corresponds to our heterogeneous merchants analysis with unknown types.

Suppose now that all merchants receive the same convenience benefit, b_S, per card transaction. All merchants accept cards if and only if $b_S \geq m$. Perfectly competitive acquirers set $m^*(a) = c_A + a$. In this case, Baxter (1983) shows that setting an IF equal to $b_S - c_A$, which we call Baxter's IF, implements efficient card usage if issuers are also perfectly competitive setting $f^*(a) = c_I + a$. Intuitively, the first best could be implemented through the usage fee that induces buyers to internalize the externality they impose to the rest of the economy while paying by card, i.e., $f^{fb} = c - b_S$. His analysis is restricted to be normative since perfectly competitive

[37] In the EC's MasterCard case (COMP/34.579), the association of European hotels, restaurants, and bars (HOTREC) stated that their members paid the highest merchant fees compared to other sectors.

banks have no preferences over the level of IF. Going beyond Baxter, we assume imperfectly competitive issuers, and thus the privately optimal IF is well-defined in our analysis.

When issuers have market power and card fees are linear, Guthrie and Wright (2003, Proposition 2) show that the socially optimal IF results in under-provision of card payment services. The reason is the following. The regulator would like to set an IF above Baxter's IF to induce the optimal card usage in the presence of an issuer markup. But then merchants would not participate (as $m > b_S$). At the second best, the regulator sets Baxter's IF, which is also the privately optimal IF and results in under-provision of card services. Next proposition shows that allowing for fixed card fees prevents inefficient provision of card services by eliminating issuer markups:

Proposition 6 *When merchants are homogeneous, the privately and the socially optimal IFs always coincide. Furthermore,*

i. If imperfectly competitive issuers can charge only linear usage fees, there is under-provision of card payment services.

ii. If membership (fixed) fees are also available, there is socially optimal provision of card payment services.

Intuitively, since issuers could internalize incremental card usage surpluses of buyers through fixed fees, they set the usage fees at their transaction costs, $c_I + a$. Baxter's IF then implements the first best transaction volume.

7.3 Strategic Card Acceptance

By assuming monopoly merchants, we abstract away from business stealing effects of accepting payment cards. Rochet and Tirole (2002) are the first who analyze such effects in a model where merchants accept the card to attract customers from rival merchants who do not accept the card. For a given retail price, card acceptance increases the quality of merchant services associated with the option to pay by card. Consumers are ready to pay higher retail prices for the improved quality as long as they observe the quality.[38] Rochet and Tirole show that when merchants are competing à la Hotelling, they internalize the average surplus of consumers from card usage, $v_B(f)$, so merchants accept cards if and only if $b_S + v_B(f) \geq m$. In other words,

[38] The authors assume that only a proportion, α, of consumers observe which store accepts cards before choosing a store to shop. Here, we consider simply their extreme case of $\alpha = 1$, which is sufficient to make our point.

merchants pay $m - b_S$ to accept cards since they could recoup v_B through charging higher retail prices for their improved quality of services.

It is important to note that we do not need merchant competition to make this argument. A monopoly merchant would also be willing to incur a cost per card transaction, to offer a better quality of services to its customers (who value the option of paying by card), since it could then internalize *some*[39] of the average card usage surplus of buyers by charging higher retail prices.

We make assumption A1 to rule out card acceptance aiming to improve quality. Recall that A1 ensures a high enough consumption value by cash, v, so that merchants who accept cards do not want to exclude cash users by setting a price higher than v. In our setup, merchants accept cards only to enjoy convenience benefits from card payments, and thus they accept cards if and only if $b_S \geq m$. Once we relax A1, a merchant accepting cards might be willing to charge a price higher than v (exclude cash users, sell only to card users) since by increasing its price, it could internalize some of the buyer surplus from card usage. Anticipating this extra revenue from card users, a merchant might accept cost increasing cards. For instance, consider simply the case of homogeneous merchants and suppose that a merchant accepting cards prefers to set $p^* > v$, i.e., it gains more from setting $p = p^*$ than $p = v$. If the merchant sets p^*, only card users buy its product and the merchant gets[40]

$$\Pi_S^* = (p^* + b_S - m)D_B(f + p^* - v),$$

If the merchant sets $p = v$, all consumers buys its product and the merchant gets

$$\Pi_S = v + (b_S - m)D_B(f)$$

We assume that $\Pi_S^* > \Pi_S$, and thus the merchant prefers to set $p^* > v$. Since $D_B(f) >$

[39] Unlike Hotelling competition, total demand is decreasing in retail price. This is why the monopolist merchant could internalize *some of* the (not *all*) average card usage surplus.

[40] A monopolist merchant accepting cards sets its price by

$$\max_p (p + b_S - m)D_B(f + p - v) \ st.: \ p \geq v$$

The solution to the unconstrained problem is implicitly given by

$$p^* = m - b_S + -\frac{D_B(f + p^* - v)}{D_B'(f + p^* - v)}$$

The merchant's optimal price is p^* if it satisfies the constraint, i.e., $p^* > v$. Otherwise the merchant sets its price equal to v. We suppose here that the constraint is not binding in equilibrium.

$D_B(f + p^* - v)$ for $p^* > v$, our assumption ($\Pi_S^* > \Pi_S$) implies also that

$$V(p^*, f) \equiv p^* - \frac{v}{D_B(f + p^* - v)} > 0$$

where $V(p^*, f)$ is a positive function referring to the merchant's extra surplus from increasing its quality (so its retail price) through accepting cards. Putting it differently $V(p^*, f)$ refers to some of the average card usage surplus of buyers. The IHRP implies that p^* is decreasing in f (see the previous footnote). Using this together with the monotonicity of $D_B(.)$, we get that $V(p^*, f)$ is decreasing in f.

If the merchant does not accept cards, it gets $\Pi_S = v$. A merchant thus accepts cards whenever

$$\Pi_S^* = (p^* + b_S - m)D_B(f + p^* - v) \geq v \qquad or$$
$$b_S + V(p^*, f) \geq m$$

Anticipating extra surplus $V(p^*, f)$ from card users, the merchant is willing to pay more than its convenience benefit to be able accept cards, i.e., it resists less to an increase in m when it expects to get a higher surplus after accepting cards. Furthermore, the reduction in its resistance, $V(p^*, f)$, decreases in card usage fee f, so increases in the IF. When the association raises the IF, the merchant fee increases, which decreases the participation of merchants. Conversely, the increase in the IF decreases the card usage fee increasing $V(p^*, f)$. This in turn increases merchant participation. The latter effect does not exist in our original setup under A1. Hence, merchants would resist less to an increase in the IF if we relaxed A1, in which case the privately optimal IF would be even higher than what we found. Hence, relaxing A1 would reinforce our results: cardholders would pay even less and merchants would pay even more. The same conclusions would hold if we allowed business stealing effects by introducing competition among merchants, since such a modification in our setup would again weaken the resistance of merchants to an increase in IF [see Rochet and Tirole (2002, 2006a)]. For the case of heterogeneous merchants, we could make a similar argument for the marginal merchant: relaxing A1 would make the marginal merchant less resistant to an increase in the IF, and thus the association sets a higher IF.

7.4 Imperfect Acquirer Competition

Unlike most of the literature, we do not assume fixed margins for banks, but instead let the equilibrium margins be endogenous. Perfectly competitive acquirers have zero margins, whereas imperfectly competitive issuers' margins are determined by their equilibrium pricing strategies.[41] Wright (2004, Proposition 1) shows that the privately optimal IF is higher (lower) than the volume-maximizing IF if and only if the pass-through of costs to user fees is higher (lower) on the acquiring side than the issuing side when evaluated at the volume-maximizing IF. In our setup, perfectly competitive acquirers pass costs fully to user fees and, due to two-part tariff card fees, issuers pass costs fully to user fees, too. Different from Wright, we show (by Lemma 1, 2 and Proposition 5) that the privately optimal IF is *always* higher than the volume maximizing IF. Our result is driven by the fact that imperfectly competitive issuers are able to internalize incremental card usage surpluses of buyers through fixed fees, so the privately optimal IF aims to maximize buyers' card usage surplus by restricting the volume.

In the payment card industry, merchants make only membership (card acceptance) decisions by comparing their average benefit with the average merchant fee. Considering either linear or non-linear merchant fees therefore delivers the same results (see Section 3). If we assumed imperfect competition on the acquiring side, acquirers would always (irrespective of having linear or non-linear merchant fees) put some margin over their net cost when setting an average merchant fee since they could internalize some of the (not all) incremental merchant surpluses. This is because merchants make only membership decisions. The following lemma shows this claim for a monopoly acquirer.

Lemma 5 *Suppose that a monopoly acquirer charges merchants non-linear fees. The acquirer profits depend only on the average merchant fee and at optimum the acquirer puts some margin over its net cost when setting the average merchant fee.*

In particular, a monopoly acquirer would be able to internalize *some* of the merchant surplus while the monopoly issuer captures *incremental* card usage surpluses of buyers through fixed fees. Intuitively, by keeping the average merchant fee constant, if the transactional merchant fee is reduced, the increase in the merchant surplus cannot be appropriated by the acquirer since the amount of merchants accepting cards depend on the average merchant fee, and thus would

[41] Fixed margins for banks is a conventional (simplifying) assumption of the literature to get clear results when merchants are assumed to be heterogeneous (i.e., when banks' profits are not monotonic in IF) [see Rochet (2003), Guthrie and Wright (2003)].

be the same. However, by keeping the average card fee constant and reducing the transactional card fee, the issuer could internalize the increase in the card usage surplus of consumers, since the card usage increases.This asymmetry does not result from our assumptions but from the asymmetry between merchants and consumers: once a merchant becomes an affiliate of the card network, it is the cardholder who decides whether to pay by card (see our discussion in Section 3). We thus focus on linear merchant fees without loss of generality.

When the monopoly acquirer internalizes some of the merchant surplus, there would be a conflict between the acquirer and the issuer (due to conflicting interests of sellers and buyers) when setting an IF. Since the association maximizes the sum of the issuer and acquirer profits, the privately optimal IF would be *higher* than the socially optimal IF because the former accounts for incremental card usage surpluses of buyers and only some of the merchant surplus, whereas the latter accounts for incremental card usage surpluses of both buyers and sellers. We therefore obtain the following proposition.

Proposition 7 *When there is a monopoly acquirer and a monopoly issuer, the payment card association sets a higher IF than the socially optimal level. Hence, in equilibrium, cardholders pay too little and merchants pay too much per transaction.*

8 Policy Implications and Concluding Remarks

We show that a payment card association sets a higher interchange fee (IF) than the socially optimal IF, and thus distorts the structure of user fees by inducing too low card usage fees at the expense of too high merchant fees. By distinguishing card membership from card usage (transaction) decisions (and fees), we illustrate that the upward distortion of the equilibrium IF is mainly due to the fact that cardholders are the ones determining card usage once a merchant becomes a member of the card platform. This structural asymmetry between consumers and merchants induces issuers to subsidize card users, since they could recoup *incremental* card usage surpluses of buyers through fixed fees. Therefore, issuers set card usage fees at their transaction cost passing all interchange revenue (i.e., some of the merchant surplus) to card users. The card scheme then sets the IF maximizing the card usage surplus of buyers. The socially optimal IF is lower than the privately optimal IF because the former accounts for any change *both* in buyers' and sellers' card usage surpluses. We therefore show that there are efficiency gains from a cap regulation on the IF. However, we do not find any reason to apply widely used cost-based regulation, which sets a cap on IF reflecting issuers' (weighted or simple) average cost (such as transaction authorization, processing, fraud prevention). In line with the existing literature, we obtain a simple characterization of a socially optimal IF which reflects two considerations: relative demand elasticities (marginal users) and relative net surpluses (average users).

We also show that regulating the IF is not enough to achieve full efficiency in the payment card industry, since efficiency requires each user fee be discounted by the positive externality of that user on the rest of the industry and one tool (IF) is not enough to achieve efficient usage on both sides. Intuitively, we suggest that if a card scheme charged its member banks fixed membership fees as well as transaction fees[42], the platform could induce both consumers and merchants to internalize their externalities, and thus improve efficiency. We leave the characterization of an efficient IF mechanism for future research.

The qualitative results are robust to imperfect issuer competition, imperfect acquirer competition, and to many factors affecting final demands, such as elastic cardholding and strategic card acceptance to attract consumers.

Our setup does not incorporate the implications of competition among card schemes or other payment methods. However, as long as consumers subscribe to only one platform, i.e., holding one type of card, and merchants subscribe to more than one card platform, competing card schemes would like to attract consumers (competitive bottlenecks), and thus favor more the consumer surplus than the merchant surplus. In this case, the upward distortion of equilibrium IFs would be greater than the case of a monopoly card scheme. A thorough analysis is needed to see which side is going to hold/accept one type of card in equilibrium. A marginal decrease from the card association's IF is found to be socially desirable, however, we are unable to determine how much the IF should be decreased by. Too stringent price caps could be worse than no cap regulation. Our setup inherits all the practical limitations of setting socially optimal prices that depend on hardly observable characteristics of supply and demand.

[42] In this case, different transaction fees could be set to issuers versus acquirers.

Appendix

A Benchmark Analysis

A.1 Proof of Lemma 1

We first show that $v_B'(f) < 0$ and $v_S'(m) < 0$ under the Increasing Hazard Rate Property (thereafter IHRP). Consider first $v_B(f)$. Using $D_B(f) = 1 - G(f)$ and integrating by parts give

$$v_B(f) \equiv \frac{\int_f^{\bar{b}_B} D_B(b_B)db_B}{D_B(f)}. \tag{10}$$

Define $H(f) \equiv \int_f^{\bar{b}_B} D_B(b_B)db_B$. Notice that the IHRP is equivalent to $D_B'/D_B \equiv H''/H'$ decreasing in f. Given that H''/H' is decreasing, $H(\bar{b}_B) = 0$ and $H(f)$ is strictly monotonic, we have H'/H is decreasing due to **Bagnoli and Bergstrom (1989, Lemma 1).**[43] Using (10), decreasing H'/H is equivalent to $v_B'(f) < 0$. Similarly, we can establish that $v_S'(m) < 0$. Inequalities $v_B'(f) < 0$ ($v_S'(m) < 0$) implies $v_B D_B' + D_B > 0$ ($v_S D_S' + D_S > 0$) since $v_B' \equiv -\frac{v_B D_B' + D_B}{D_B}$ $\left(v_S' \equiv -\frac{v_S D_S' + D_S}{D_S}\right)$.

Given the best responses of the issuer ($f^*(a) = c_I - a$) and acquirers ($m^*(a) = c_A + a$), the problem of setting an IF to maximize some objective function is equivalent to the problem of allocating total price $f + m = c$ between buyers and sellers to maximize the same objective. For the sake of exposition, we follow the latter strategy and first characterize usage fees f^B, f^V, f^S which respectively maximize the buyer surplus from card usage, the total transaction volume, and the seller surplus subject to the subgame perfection. We then argue that, for $i = B, V, S$, usage fee f^i can be induced by a unique IF, a^i, which is given by $a^i = c_I - f^i$.

Existence and uniqueness of f^B: The buyer surplus from card usage is $\Phi_B(f, m)$, f^B is thus a solution to:

$$\max_f \Phi_B = v_B(f)D_B(f)D_S(c - f),$$

[43] The Generalized Mean Value Theorem of calculus ensures, for every x, the existence of a $\xi \in (x, \bar{b}_B)$ such that

$$\frac{H'(x) - H'(\bar{b}_B)}{H(x) - H(\bar{b}_B)} = \frac{H''(\xi)}{H'(\xi)}$$

If H''/H' is decreasing, for any $x < \xi$, it should then be the case that

$$\frac{H'(x) - H'(\bar{b}_B)}{H(x) - H(\bar{b}_B)} < \frac{H''(x)}{H'(x)}$$

Since H is monotone and $H(\bar{b}_B) = 0$, it must then be that $H'(x)H(x) < 0$ whenever $x < \bar{b}_B$. Multiplying both sides of the latter inequality by $H'(x)H(x)$ gives $H''(x)H(x) < (H')^2 - H'(\bar{b}_B)H'(x)$ and thus that $H''(x)H(x) - (H')^2 < 0$, which is equivalent to H'/H decreasing.

which has an interior solution only if $f \leq \overline{b_B}$, because otherwise no one pays by card. The quasi-demand D_B is maximized and equal to 1 when $f \leq \underline{b_B}$ and the issuer does not gain from reducing f below $\underline{b_B}$. Without loss of generality, we thus restrict the domain of f to be $\left[\underline{b_B}, \overline{b_B}\right]$. By the Weierstrass Theorem, there exists a maximum of the continuous function $\Phi_B(.)$ on the compact interval $\left[\underline{b_B}, \overline{b_B}\right]$. The IHRP and $v_B' < 0$ imply respectively that D_S and $v_B D_B$ are log-concave, which makes Φ_B log-concave in f. Hence, f^B is characterized uniquely by the first order optimality condition:[44]

$$\left[D_B\left(D_S + v_B D_S'\right)\right]_{f=f^B} = 0 \tag{11}$$

The *existence and uniqueness of* f^S*:* Symmetrically, f^S is a solution to

$$\max_f v_S(c-f) D_B(f) D_S(c-f).$$

Similar to the previous proof, the Weierstrass Theorem guarantees the existence of a maximum on the compact interval $\left[\underline{b_B}, \overline{b_B}\right]$. The objective function is log-concave because $v_S D_S$ is log-concave by $v_S' < 0$ and D_B is log-concave by the IHRP. The first order condition determines uniquely f^S:[45]

$$\left[D_S\left(D_B + v_S D_B'\right)\right]_{f=f^S} = 0 \tag{12}$$

The existence and uniqueness of f^V*:* The total volume of transactions is given by $D_B(f)D_S(m)$, so f^V is a solution to

$$\max_f D_B(f) D_S(c-f)$$

Since D_B and D_S are log-concave in f and m, respectively, the volume of transactions is log-concave in f. There thus exists a unique usage fee f^V which maximizes the volume. f^V is implicitly given by the first order condition:

$$\left[D_B' D_S - D_S' D_B\right]_{f=f^V} = 0 \tag{13}$$

[44]We get the optimality condition by using the fact that $[v_B(f)D_B(f)]' = -D_B(f)$.

[45]We get the optimality condition by using the fact that $[v_S(m)D_S(m)]' = -D_S(m)$.

Now, our claim is $f^B < f^V$.Consider the derivative of the volume at f^B:

$$[D_B D_S]'_{f=f^B} = \left[D'_B D_S - D'_S D_B\right]_{f=f^B}$$

By using (11), we can re-write the latter as $\frac{D_S}{v_B}(v_B D'_B + D_B)$, which is positive since $v_B D'_B + D_B > 0$ from $v'_B = -\frac{v_B D'_B + D_B}{D_B} < 0$. We thus have $[D_B D_S]'_{f=f^B} > 0$. Using the IHRP and (13), we then get $f^B < f^V$, which implies $a^B > a^V$.

Symmetrically, by using the IHRP and $v'_S < 0$, it can be shown that $f^S > f^V$, so $a^V > a^S$. We thus conclude that $a^S < a^V < a^B$.

A.2 Proof of Corollary 1

Observe that $v_B + \frac{D_S}{D'_S}$ is a decreasing function of usage fee since $v'_B < 0$ [see the previous proof] and $\frac{D_S}{D'_S}$ is increasing in merchant fee $m = c - f$ by the IHRP. Using (11), we have $f^B < 0$ if and only if $\eta_S(c) < \frac{c}{v_B(0)}$.

□ **The concavity of the regulator's problem**

By plugging the constraint into (5), the regulator's problem is rewritten as

$$\max_f \left[v_B(f) + v_S(c-f)\right] D_B(f) D_S(c-f)$$

The first order optimality condition

$$FOC_r : -v_B D_B D'_S + v_S D'_B D_S = 0$$

determines f^r if the second order optimality condition

$$SOC_r : D_B D'_S + v_B D_B D''_S + v_S D''_B D_S + D'_B D_S < 0$$

holds at such a critical point. Using FOC_r, we rewrite SOC_r as

$$SOC_r : \frac{D_B D'_S}{D'_B}\left[D'_B + D''_B v_B\right] + \frac{D'_B D_S}{D'_S}\left[D'_S + D''_S v_S\right] < 0.$$

We have $D'_B + D''_B v_B < D'_B - D''_B \frac{D_B}{D'_B}$ since $v_B < -\frac{D_B}{D'_B}$ by $v'_B < 0$ and that $D'_B - D''_B \frac{D_B}{D'_B} < 0$ by the IHRP. We thus get $D'_B + D''_B v_B < 0$. Symmetrically, we get $D'_S + D''_S v_S < 0$ by using $v'_S < 0$ and the IHRP. Hence, SOC_r holds and f^r is characterized uniquely by FOC_r.

A.3 Proof of Proposition 1

By definition f^B maximizes the surplus of buyers $v_B D_B D_S$ and f^S maximizes the surplus of sellers $v_S D_B D_S$, given that $f + m = c$. Lemma 1 shows the existence and the uniqueness of f^B and f^S, and that $f^B < f^S$. Any solution to problem (5), f^r, necessarily lies in (f^B, f^S) by the revealed preference argument. The regulated IF, $a^r = c_I - f^r$, is lower than the equilibrium (privately optimal) IF, $a^B = c_I - f^B$, since $f^B < f^r$.

A.4 Proof of Proposition 2

We decompose the planner's problem into a price allocation and a total price setting problem. We have already characterized in Proposition 1 the optimal allocation of a total price $p = c$. We are thus left to generalize the optimal allocation of any total price p and characterize then the optimal p. Let $f(p)$ and $m(p)$ denote the respective fees which implement the optimal allocation of p between buyers and sellers.

The social planner first solves

$$\max_f \left[p - c + v_B(f) + v_S(p - f)\right] D_B(f) D_S(p - f),$$

which characterizes $f(p)$ and $m(p)$ as follows:

$$(p - c)\left[D'_B D_S - D_B D'_S\right] - v_B D_B D'_S + v_S D'_B D_S = 0. \tag{14}$$

The planner then determines the socially optimal total price by

$$\max_p W(p) = \left[p - c + v_B(f(p)) + v_S(p - f(p))\right] D_B(f(p)) D_S(p - f(p))$$

Using $[v_S D_S]' = -D_S$ and the Envelope Theorem, we get

$$W'(p) = (p - c + v_B) D_B D'_S$$

which is equal to zero whenever $p = c - v_B$. At such a critical point the SOC, $(p - c + v_B) D''_S D_B + D_B D'_S < 0$, is verified, hence $p^{FB} = c - v_B(f^{FB}) < c$. By substituting p^{FB} into equation (14), we characterize the first best usage fee f^{FB} by $v_B(f^{FB}) = v_S(c - v_B(f^{FB}) - f^{FB})$.

A.5 Proof of Proposition 3

We will first show that in equilibrium the budget balance condition for acquirers must be binding. At first best prices, buyers hold the card if and only if $F^{FB} \leq \Phi_B(f^{FB}, m^{FB})$. Thus, at FB fees the issuer gets at most:

$$\overline{\Pi_I} \equiv (f^{FB} + v_B(f^{FB}) - c_I)D_B(f^{FB})D_S(m^{FB})$$

The upper bound $\overline{\Pi_I}$ is achieved when $F = \Phi_B(f^{FB}, m^{FB})$. By definition of the first best fees (from Proposition 3), we have $f^{FB} + m^{FB} = c - v_B(f^{FB})$. Substituting this equality into $\overline{\Pi_I}$, we rewrite the upper bound as:

$$\overline{\Pi_I} = (c_A - m^{FB})D_B(f^{FB})D_S(m^{FB})$$

If $m^{FB} > c_A$, we would get $\Pi_I < 0$, so the issuer's budget balance condition would not be satisfied. Hence, in equilibrium acquirers' budget balance condition must be binding. After plugging this into the planner's problem, we get the second best transaction fees as $m^{SB} = c_A$, $f^{SB} = c_I - v_S(c_A)$.

B Extensions

B.1 Elastic Cardholding

□ Proof of Lemma 2

Define functional K as

$$-\frac{(\mathrm{HR}^{-1})'}{1 - (\mathrm{HR}^{-1})'}$$

where HR^{-1} is the inverse of the hazard rate and thus decreasing by the IHRP. Note that $0 < K(\cdot) < 1$.

Existence and uniqueness of $\widetilde{f^B}$*:* $\widetilde{f^B}$ is a solution to:

$$\max_f \left[\int_{F(f)-\Phi_B(f)}^{\overline{B_B}} xh(x)dx + \Phi_B(f)\lambda(F(f) - \Phi_B(f)) \right]$$

$$\textit{where } \Phi_B(f) = v_B(f)D_B(f)D_S(c-f) \text{ and } F(f) = \frac{1 - H(F(f) - \Phi_B(f))}{h(F(f) - \Phi_B(f))}.$$

First, The Weierstrass Theorem guarantees the existence of $\widetilde{f^B}$ on $\left[\underline{b_B}, \overline{b_B}\right]$. By differentiating $F(f)$, we get $F' = K\Phi_B'$, which implies that $[F - \Phi_B]' = -(1-K)\Phi_B'$. The IHRP and $v_B' < 0$ imply respectively the log-concavity of D_S and $v_B D_B$, and thus Φ_B is log-concave. To satisfy the SOC of the problem, we furthermore assume that $[F - \Phi_B]$ is log-convex in f and that f^B maximizing the option value minimizes $[F - \Phi_B]$. We thus get that λ is log-concave in f. Hence, f^B necessarily maximizes the buyer surplus (gross of fixed fees).

The *existence and uniqueness of* $\widetilde{f^S}$*:* Taking as given the subgame perfection, $\widetilde{f^S}$ maximizes the surplus of sellers, so $\widetilde{f^S}$ is a solution to

$$\max_f v_S(c-f)D_B(f)D_S(c-f)\lambda(F(f) - \Phi_B(f))$$

The Weierstrass Theorem again guarantees the existence of $\widetilde{f^S}$ on $\left[\underline{b_B}, \overline{b_B}\right]$. Log-concavity of functions $v_S D_S$ (by $v_S' < 0$), D_B (by the IHRP), and λ, implies that $\widetilde{f^S}$ is uniquely determined by the first order optimality condition:

$$D_S(D_B + v_S D_B')\lambda + v_S(1-K)\Phi_B' h D_B D_S = 0 \tag{15}$$

The existence and uniqueness of $\widetilde{f^V}$*:* The total volume of transactions is given by $D_B D_S \lambda$, so $\widetilde{f^V}$ is a solution to

$$\max_f D_B(f)D_S(c-f)\lambda(F(f) - \Phi_B(f))$$

The Weierstrass Theorem guarantees the existence of $\widetilde{f^V}$ on $\left[\underline{b_B}, \overline{b_B}\right]$. Since quasi-demands D_B, D_S are log-concave in f (implied by the IHRP), and λ is a log-concave function of f, the volume of transactions $D_B D_S \lambda$ is log-concave in f. The unique usage fee, $\widetilde{f^V}$, is then implicitly given by the first order optimality condition:

$$\left(D_B' D_S - D_S' D_B\right)\lambda + (1-K)\Phi_B' h D_B D_S = 0 \tag{16}$$

Now, our claim is $f^B < \widetilde{f^V}$. By using the definition of f^B, i.e., $\Phi_B'(f^B) = 0$, we derive the volume of transactions at f^B as:

$$[D_B D_S \lambda]'_{f=f^B} = \frac{\lambda D_S}{v_B}\left(v_B D_B' + D_B\right),$$

which is positive since $v_B D'_B + D_B > 0$ from $v'_B < 0$. We thus have $[D_B D_S \lambda]'_{f=f^B} > 0$. The IHRP and (16) implies then that $f^B < \widetilde{f^V}$.

Symmetrically, by using the IHRP and $v'_S < 0$, it can be shown that $\widetilde{f^S} > \widetilde{f^V}$. We thus conclude that $f^B < \widetilde{f^V} < \widetilde{f^S}$. The IFs inducing these usage fees must then satisfy $\widetilde{a^S} < \widetilde{a^V} < a^B$.

□ Proof of Proposition 4

Part (i). The association's problem can be rewritten as

$$\max_f F(f)\lambda(F(f) - \Phi_B(f)),$$

$$where\ F(f) = \frac{1 - H(F(f) - \Phi_B(f))}{h(F(f) - \Phi_B(f))}$$

The Weierstrass Theorem guarantees the existence of a maximum on compact interval $\left[\underline{b_B}, \overline{b_B}\right]$. The first order optimality condition is

$$F'\lambda - Fh(F' - \Phi'_B) = 0,$$

Using $F' = K\Phi'_B$ (which we derive in the proof of Lemma 2), we rewrite the optimality condition as $[K\lambda + F(1-K)h]\,\Phi'_B = 0$. By using log-concavity of Φ_B, it can easily be shown that the second order condition holds. Therefore, the equilibrium usage fee is equal to f^B (i.e., the fee maximizing the option value of the card) (see the proof of Lemma 2).

Part (ii). By definition f^B maximizes the surplus of buyers (gross of fixed fees) and $\widetilde{f^S}$ maximizes the surplus of sellers. Lemma 2 shows the existence and the uniqueness of f^B and $\widetilde{f^S}$, and that $f^B < \widetilde{f^S}$. Usage fee $\widetilde{f^r}$, which maximizes the sum of buyer and seller surpluses, necessarily lies in the interval $(f^B, \widetilde{f^S})$. The regulated IF, $\widetilde{a^r} = c_I - \widetilde{f^r}$, must then be lower than the equilibrium (profit-maximizing) IF, $a^B = c_I - f^B$, since $f^B < \widetilde{f^r}$.

B.2 Competing Issuers

□ Examples of Demand Functions

The following examples of demand functions for differentiated products satisfy assumptions A2-A6.

(1) Linear symmetric demands of form, for $i = 1, 2, i \neq j$,

$$q_i = \frac{1}{1+\sigma} - \frac{1}{1-\sigma^2}p_i + \frac{\sigma}{1-\sigma^2}p_j$$

where q refers to demand, p refers to price, and σ measures the level of substitution between the firms (here, for imperfectly competitive issuers we have $\sigma \in (0, 1)$). These demands are driven from maximizing the following quasi-linear and quadratic utility function

$$U(q_i, q_j) = q_i + q_j - \sigma q_i q_j - \frac{1}{2}\left(q_i^2 + q_j^2\right)$$

subject to the budget balance condition, namely

$$p_i q_i + p_j q_j \leq I$$

(2) Dixit (1979)'s and Singh and Vives (1984)'s linear demand specification, for $i = 1, 2, i \neq j$,

$$q_i = a - bp_i + cp_j$$

where $a = \frac{\alpha(\beta-\gamma)}{\beta^2-\gamma^2}$, $b = \frac{\beta}{\beta^2-\gamma^2}$, $c = \frac{\gamma}{\beta^2-\gamma^2}$, and the substitution parameter is $\varphi = \frac{\gamma^2}{\beta^2}$, under the assumptions that $\beta > 0$, $\beta^2 > \gamma^2$, and $\varphi \in (0, 1)$ for imperfect substitutes.

(3) Shubik and Levitan (1980)'s demand functions of form, for $i = 1, 2, i \neq j$,

$$q_i = \frac{1}{2}\left[v - p_i(1+\mu) + \frac{\mu}{2}p_j\right]$$

where $v > 0$, μ is the substitution parameter and $\mu \in (0, \infty)$ for imperfect substitutes.

Special case: Hotelling Demand, for $i = 1, 2, i \neq j$,

$$q_i = \frac{p_j - p_i}{2t} + \frac{1}{2}$$

satisfies the assumptions except for A4 and A6 since the own price effect is equal to the cross price effect, that is

$$\left|\frac{\partial q_i}{\partial p_i}\right| = \frac{\partial q_i}{\partial p_j} \qquad \left|\frac{\partial^2 Inq_i}{\partial p_i^2}\right| = \frac{\partial^2 Inq_i}{\partial p_i \partial p_j}$$

which imply that the equilibrium fixed fees are independent of the option value, and thus independent of the IF. In this case, the issuers would not have any preferences over IF. Hence, the privately optimal IF is not well defined.

□ **Proof of Lemma 3**

Consider the FOC of I_i's problem:

$$FOC_i : Q_i\left(F_i - \Phi_B, F_j - \Phi_B\right) + F_i \frac{\partial Q_i}{\partial F_i} = 0$$

Solving FOC_i and FOC_j together gives us the equilibrium fees as functions of the option value, i.e., $F_i^*(\Phi_B)$ and $F_j^*(\Phi_B)$. The second-order condition holds by A5:

$$SOC_i : 2\frac{\partial Q_i}{\partial F_i} + F_i^* \frac{\partial^2 Q_i}{\partial F_i^2} < 0$$

The total derivation of FOC_i gives

$$SOC_i dF_i^* + \frac{\partial Q_i}{\partial Fj} dF_j^* - \left[\frac{\partial Q_i}{\partial F_i} + \frac{\partial Q_i}{\partial F_j}\right] d\Phi_B - F_i^* \left[\frac{\partial^2 Q_i}{\partial F_i^2} + \frac{\partial^2 Q_i}{\partial F_i \partial F_j}\right] d\Phi_B + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j} dF_j^* = 0$$

Rearranging the previous equation, we get

$$SOC_i dF_i^* + \left[\frac{\partial Q_i}{\partial Fj} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j}\right] dF_j^* = \left[\frac{\partial Q_i}{\partial F_i} + \frac{\partial Q_i}{\partial F_j} + F_i^* \frac{\partial^2 Q_i}{\partial F_i^2} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j}\right] d\Phi_B$$

Adding and subtracting $\frac{\partial Q_i}{\partial F_i}$ into the brackets on the right-hand side of the above equation, we re-write the equality as

$$SOC_i dF_i^* + \left[\frac{\partial Q_i}{\partial Fj} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j}\right] dF_j^* = \left[SOC_i + \frac{\partial Q_i}{\partial F_j} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j} - \frac{\partial Q_i}{\partial F_i}\right] d\Phi_B$$

Symmetrically, we take the total derivation of FOC_j and get

$$SOC_j dF_j^* + \left[\frac{\partial Q_j}{\partial Fi} + F_j^* \frac{\partial^2 Q_j}{\partial F_i \partial F_j}\right] dF_i^* = \left[SOC_j + \frac{\partial Q_j}{\partial F_i} + F_j^* \frac{\partial^2 Q_j}{\partial F_i \partial F_j} - \frac{\partial Q_j}{\partial Fj}\right] d\Phi_B$$

Solving together the last two equations gives us

$$\frac{\partial F_i^*}{\partial \Phi_B} = 1 - \frac{SOC_j \frac{\partial Q_i}{\partial F_i} - \frac{\partial Q_j}{\partial F_j}\left[\frac{\partial Q_i}{\partial F_j} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j}\right]}{SOC_i SOC_j - \left[\frac{\partial Q_i}{\partial F_j} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j}\right]\left[\frac{\partial Q_j}{\partial F_i} + F_j^* \frac{\partial^2 Q_j}{\partial F_i \partial F_j}\right]}$$

For the sake of exposition, we give the proof for *symmetric demands*. We can provide interested readers with the proof for asymmetric demand functions. If we assume that the demand functions are symmetric, the solution of the issuers' problems gives $F_i^* = F_j^*$. We then get

$$\frac{\partial F_j^*}{\partial \Phi_B} = \frac{\partial F_i^*}{\partial \Phi_B} = 1 - \frac{\partial Q_i/\partial Fi}{SOC_i + \frac{\partial Q_i}{\partial F_j} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j}}$$

If $\frac{\partial^2 InQ_i}{\partial t_i \partial t_j} < 0$, we have

$$\frac{(\partial^2 Q_i/\partial F_i \partial F_j)Q_i - (\partial Q_i/\partial F_i)(\partial Q_i/\partial F_j)}{Q_i^2} < 0$$

so that

$$\frac{\partial Q_i}{\partial F_j} - \frac{Q_i}{\partial Q_i/\partial Fi} \frac{\partial^2 Q_i}{\partial F_i \partial F_j} < 0$$

From FOC_i we have, $F_i^* = -\frac{Q_i}{\partial Q_i/\partial Fi}$, so we get

$$\frac{\partial Q_i}{\partial F_j} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j} < 0$$

Moreover, the log-concavity of Q_i (A5) implies that $SOC_i < \partial Q_i/\partial Fi$. Thus, we get $0 < \frac{\partial F_i^*}{\partial \Phi_B} < 1$.

If $\frac{\partial^2 InQ_i}{\partial t_i \partial t_j} > 0$, we have

$$\frac{\partial Q_i}{\partial F_j} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j} > 0.$$

A6 becomes $-\frac{\partial^2 InQ_i}{\partial t_i^2} > \frac{\partial^2 InQ_i}{\partial t_i \partial t_j}$, which implies that

$$-\left[\frac{\partial Q_i}{\partial F_i} + F_i^* \frac{\partial^2 Q_i}{\partial F_i^2}\right] > \frac{\partial Q_i}{\partial F_j} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j},$$

Using SOC_i, we get

$$\partial Q_i/\partial Fi > SOC_i + \frac{\partial Q_i}{\partial F_j} + F_i^* \frac{\partial^2 Q_i}{\partial F_i \partial F_j},$$

proving that $0 < \frac{\partial F_i^*}{\partial \Phi_B} < 1$.

□ **Proof of Lemma 4**

When the demand functions Q_i, Q_j are symmetric, we have $F_i^* = F_j^*$ and $0 < \frac{\partial F_i^*}{\partial \Phi_B} = \frac{\partial F_j^*}{\partial \Phi_B} < 1$ from Lemma 3. Consider now the derivative of $Q_i(F_i^*, F_j^*, a)$ with respect to a:

$$Q_i'(a) = \left[\frac{\partial Q_i}{\partial F_i}\left(\frac{\partial F_i^*}{\partial \Phi_B} - 1\right) + \frac{\partial Q_i}{\partial F_j}\left(\frac{\partial F_j^*}{\partial \Phi_B} - 1\right)\right] \Phi_B'(a)$$

The first term inside the brackets represents the direct effect of the option value on Q_i, through changing F_i^*, and the second term represents the indirect effect of the option value on Q_i, through changing F_j^*. First, observe that $\Phi'_B(a^r) > 0$ because, from Lemma 1, the option value is concave in a, attains its maximum at $a = a^B$ and that $a^B > a^r$. Imperfect issuer competition (A3 and A4) then implies that the direct effect of the option value on Q_i dominates its indirect effect so that the term inside the brackets is positive. We therefore conclude that when two differentiated issuers are competing with symmetric demands, the demand for holding card i is increasing in interchange fee at $a = a^r$ so that $W'(a^r) > 0$. Since $W(a)$ is a log-concave function[46], we necessarily have $a^r < a^{rc} < a^B$.

□ **Proof of Proposition 5**

Following the lines of our benchmark analysis, we first define three important IF levels: the buyers-optimal IF, the sellers-optimal IF, and the volume maximizing IF,which we denote respectively by a^{Bc}, a^{Sc}, and a^{Vc}, where superscript c refers to issuer competition:

$$\begin{aligned} a^{Bc} &\equiv \arg\max_a v_B(c_I - a) D_B(c_I - a) D_S(c_A + a) \left[Q_1(F_1^*, F_2^*, a) + Q_2(F_1^*, F_2^*, a)\right] \\ a^{Sc} &\equiv \arg\max_a v_S(c_A + a) D_B(c_I - a) D_S(c_A + a) \left[Q_1(F_1^*, F_2^*, a) + Q_2(F_1^*, F_2^*, a)\right] \\ a^{Vc} &\equiv \arg\max_a D_B(c_I - a) D_S(c_A + a) \left[Q_1(F_1^*, F_2^*, a) + Q_2(F_1^*, F_2^*, a)\right] \end{aligned}$$

Part (i) Recall that we have, for $i = 1, 2$, $i \neq j$,

$$Q_i'(a) = \left[\frac{\partial Q_i}{\partial F_i}\left(\frac{\partial F_i^*}{\partial \Phi_B} - 1\right) + \frac{\partial Q_i}{\partial F_j}\left(\frac{\partial F_j^*}{\partial \Phi_B} - 1\right)\right]\Phi'_B(a)$$

Since $\Phi'_B(a^B) = 0$, it is straightforward that $Q_i'(a^B) = 0$. We then conclude that the IF maximizing the option value of the card also maximizes the card usage surplus when the issuers are imperfect competitors, i.e., $a^{Bc} = a^B$. Recall that the association sets $a^* = a^B$ because this maximizes both issuers' payoffs. Hence, the privately optimal IF coincides with the buyers-optimal IF.

Part (ii) Since the average surplus of buyers and the average surplus of sellers are decreasing in their own usage fees, i.e., $v'_B(f), v'_S(m) < 0$ (see the proof of Lemma 1), we have $a^{Sc} < a^{Vc} < a^{Bc}$. The regulator wants to maximize the sum of buyers' and sellers' surpluses, the socially

[46] In the benchmark analysis, we showed that $[v_B(c_I - a) + v_S(c_A + a)]\, D_B(c_I - a) D_S(c_A + a)$ is a log-concave function of a (see the Appendix). By assumption Q_i is log-concave. Hence, $W(a)$ is a log-concave function.

optimal IF is thus lower than the privately optimal one, i.e., $a^{rc} < a^{Bc}$.

C Comparisons with the Literature

C.1 Imperfect Acquirer Competition

□ **Proof of Lemma 5**

Under A1, all merchants set $p = v$, so all consumers purchase a unit good from each merchant and only $D_B(f)$ of cardholders would like to pay by card. Given that the fixed and transaction merchant fees are respectively M and m, a merchant of type b_S accepts card payments if and only if

$$(b_S - m)D_B(f) - M \geq 0.$$

The demand for card acceptance is then equal to $D_S(\widetilde{m})$ where $\widetilde{m} \equiv m + \frac{M}{D_B(f)}$ denotes the average merchant fee. A consumer holds a card if and only if $v_B(f)D_B(f)D_S(\widetilde{m}) \geq F$. Anticipating the reactions of merchants and consumers, the monopoly acquirer solves

$$\max_{M,m} \left[(m - c_A - a)\, D_B(f) + M\right] D_S(m + \frac{M}{D_B(f)}),$$

which is equivalent to

$$\max_{\widetilde{m}} \left[\widetilde{m} - c_A - a\right] D_S(\widetilde{m})D_B(f),$$

so the acquirer's profits depend only on the average merchant fee $\widetilde{m}$. At optimum the acquirer sets $\widetilde{m}^*$ where

$$\widetilde{m}^* = c_A + a - \frac{D_S(\widetilde{m}^*)}{D_S'(\widetilde{m}^*)},$$

i.e., puts a monopoly markup when setting the average merchant fee.

□ **Proof of Proposition 7**

At card fees (F, f) and merchant fee m, $D_B(f)$ of buyers would like to pay by card and $D_S(m)$ of merchants accept card payments. A consumer holds a card if and only if $\Phi_B(f, m) \geq F$. Anticipating the reactions of merchants and consumers, a monopoly issuer sets

$$f^* = c_I - a \qquad F^* = \Phi_B(f, m) = v_B(f)D_B(f)D_S(m),$$

A monopoly acquirer would set m^* where

$$m^* = c_A + a - \frac{D_S(m^*)}{D'_S(m^*)}.$$

Using the latter equality, we derive

$$\frac{\partial m^*}{\partial a} = \frac{1}{\left[-D_S(m^*)/D'_S(m^*)\right] + 1} > 0$$

Using the IHRP, one can show that $\frac{\partial m^*}{\partial a} < 1$.

The card association wants to maximize the sum of issuer and acquirer profits, that is

$$\max_a \left[-\frac{D_S(m^*)}{D'_S(m^*)} + v_B(f^*) \right] D_B(f^*) D_S(m^*),$$

We denote the association's IF by a^a. Let a^A denote the IF which maximizes the acquirer's profits only, i.e.,

$$a^A \equiv \arg\max_a \left[-\frac{D_S(m^*)}{D'_S(m^*)} D_B(f^*) D_S(m^*) \right],$$

which is then determined by the following FOC:

$$FOC_A : \frac{\partial m^*}{\partial a} \left[-\frac{D_S(m^*)}{D'_S(m^*)} \right]' D_B(f^*) D_S(m^*) + \left[-\frac{D_S(m^*)}{D'_S(m^*)} \right] \frac{\partial \left[D_B(f^*) D_S(m^*) \right]}{\partial a} = 0$$

Recall that a^S is the IF maximizing the merchant surplus, i.e., $a^S \equiv \arg\max_a v_S(m^*) D_B(f^*) D_S(m^*)$ where $v_S(.)$ is the average merchant surplus from card transactions. The FOC implicitly determines a^S:

$$FOC_S : \frac{\partial m^*}{\partial a} v'_S(m^*) D_B(f^*) D_S(m^*) + v_S(m^*) \frac{\partial \left[D_B(f^*) D_S(m^*) \right]}{\partial a} = 0$$

Using the IHRP, in Lemma 1, we have shown that $v'_S(m) < 0$, it is then straightforward to demonstrate that

$$\frac{\partial \left[D_B(f^*) D_S(m^*) \right]}{\partial a} \Big|_{a=a^S} > 0, \qquad (17)$$

i.e., $a^V > a^S$ where a^V is the volume maximizing IF.

Claim 1: $a^A > a^S$.

Proof of Claim 1. For any m, $v'_S(m) < 0$ implies that (see the proof of Lemma 1)

$$v_S(m) < -\frac{D_S(m)}{D'_S(m)} \tag{18}$$

By the IHRP, we furthermore have

$$\left[-\frac{D_S(m)}{D'_S(m)}\right]' < 0 \tag{19}$$

Inequalities (18) and (19) then imply that, for any m,

$$\left|v'_S(m)\right| > \left|\left[-\frac{D_S(m)}{D'_S(m)}\right]'\right|, \tag{20}$$

since otherwise we would get a contradiction with (18). Consider now FOC_A at $a = a^S$. Using inequalities (17) and (18), we obtain

$$FOC_A \underset{a=a^S}{|} > \left(\frac{\partial m^*}{\partial a}\left[-\frac{D_S(m^*)}{D'_S(m^*)}\right]' D_B(f^*)D_S(m^*) + v_S(m^*)\frac{\partial\left[D_B(f^*)D_S(m^*)\right]}{\partial a}\right)_{a=a^S}$$

Plugging FOC_S into the right-hand side of the inequality, we rewrite the latter inequality as

$$FOC_A \underset{a=a^S}{|} > \left(\frac{\partial m^*}{\partial a}D_B(f^*)D_S(m^*)\left(\left[-\frac{D_S(m^*)}{D'_S(m^*)}\right]' - v'_S(m^*)\right)\right)_{a=a^S}.$$

Using then inequality (20) brings us the result:

$$FOC_A \underset{a=a^S}{|} > 0,$$

we therefore prove that $a^A > a^S$.

Claim 2. The association sets a too high IF, i.e., $a^a > a^r$, where a^r is the socially optimal IF.

Proof of Claim 2. By definition

$$\begin{aligned} a^a &\equiv \arg\max_a \left[-\frac{D_S(m^*)}{D'_S(m^*)} + v_B(f^*)\right] D_B(f^*)D_S(m^*), \\ a^r &\equiv \arg\max_a \left[v_S(m^*) + v_B(f^*)\right] D_B(f^*)D_S(m^*). \end{aligned}$$

Revealed preference argument then proves the claim because we have $a^A > a^S$ (from Claim

1) where

$$a^A \equiv \arg\max_a \left[-\frac{D_S(m^*)}{D'_S(m^*)} D_B(f^*) D_S(m^*) \right], \quad a^S \equiv \arg\max_a v_S(m^*) D_B(f^*) D_S(m^*).$$

References

Armstrong, M. (2002): "The Theory of Access Pricing and Interconnection," *Handbook of Telecommunications Economics*, 1, 295–384.

——— (2006): "Competition in Two-Sided Markets," *Rand Journal of Economics*, 37(3), 668–681.

Armstrong, M., and J. Vickers (2006): "Competitive Nonlinear Pricing and Bundling," Discussion paper.

Baxter, W. (1983): "Bank Interchange of Transactional Paper: Legal and Economic Perspectives," *Journal of Law and Economics*, 26(3), 541–588.

Chakravorti, S. (2003): "Theory of Credit Card Networks: A Survey of the Literature," *Review of Network Economics*, 2(2), 50–68.

European Commission (2007): "Final Report on the Retail Banking Sector," Discussion paper.

Evans, D., and R. Schmalensee (2005a): *Paying With Plastic: The Digital Revolution In Buying And Borrowing*. MIT Press.

Evans, D., and R. Schmalensee (2005b): "The Economics of Interchange Fees and Their Regulation: An Overview," *a Conference Interchange Fees in Credit and Debit Card Industries: What Role for Public Authorities*, pp. 4–6.

Gans, J., and S. King (2003a): "Approaches to Regulating Interchange Fees in Payment Systems," *Review of Network Economics*, 2(2), 125–145.

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

Guthrie, G., and J. Wright (2003): "Competing Payment Schemes," Departmental Working Papers wp0311, National University of Singapore, Department of Economics.

——— (2007): "Competing Payment Schemes," *Journal of Industrial Economics*, 55(1), 37–67.

Laffont, J., S. Marcus, P. Rey, and J. Tirole (2003): "Internet Interconnection and the Off-Net-Cost Pricing Principle," *The RAND Journal of Economics*, 34(2), 370–390.

Reserve Bank of Australia (2007): "Reform of Australia's Payments System: Issues for the 2007/08 Review," Discussion paper.

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

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

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

Rochet, J., and J. Tirole (2006a): "Must-Take Cards and the Tourist Test," Discussion paper, Working Paper, IDEI.

——— (2006b): "Two-Sided Markets: A Progress Report," *The RAND Journal of Economics*, 37(3), 645–667.

Rochet, J., J. Tirole, and V. International (2002): *An Economic Analysis of the Determination of Interchange Fees in Payment Card Systems*. Visa International].

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

Weiner, S., and J. Wright (2005): "Interchange Fees in Various Countries: Developments and Determinants," *Review of Network Economics*, 4(4), 290–323.

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