Department of Economics

Working Paper No. 0311

http://nt2.fas.nus.edu.sg/ecs/pub/wp/wp0311.pdf

Competing Payment Schemes

Graeme Guthrie
Victoria University of Wellington

Julian Wright
National University of Singapore

September 9, 2003

Abstract: This paper presents a model of competing payment schemes. Unlike previous work on generic twosided markets, the model allows for the fact that in a payment system users on one side of the market (merchants) compete to attract users on the other side (consumers who may use cards for purchases). It analyzes how competition between card associations and between merchants affects the choice of interchange fees, and thus the structure of fees charged to cardholders and merchants. Implications for other two-sided markets are discussed.

 Corresponding author: Julian Wright, Department of Economics, Faculty of Arts and Social Sciences, National University of Singapore, AS2 Level 6, 1 Arts Link, Singapore 117570. jwright@nus.edu.sg. We thank seminar participants at the University of Auckland and the University of Melbourne for helpful comments, especially Matthew Ryan. Any errors are our own. Views expressed herein are those of the authors and do not necessarily reflect the views of the Department of Economics, National University of Singapore.

1 Introduction

Recently several models of payment schemes have been developed in order to analyze the optimal structure of fees in debit and credit card schemes. They ask the question: How much is charged to cardholders versus merchants for card transactions? Policymakers in a number of jurisdictions have been concerned that merchants pay too much to accept credit card transactions, costs that in their view are ultimately covered by consumers who pay by other means (see Chang and Evans, 2000, and Chakravorti and Shah, 2003).

The optimal structure of cardholder and merchant fees was first addressed by Baxter (1983), who viewed a payment transaction as a *joint* service consumed by cardholders and merchants. He emphasized the importance of a structure of fees for which cards are used whenever their joint transactional benefits exceed their joint costs. Baxter's was a normative analysis. Schmalensee (2002) and Rochet and Tirole (2002) have since provided positive analyses, explaining what determines the structure of fees set by a payment card association (such as MasterCard or Visa). While Schmalensee focuses on the trade-off between attracting cardholders and attracting merchants, Rochet and Tirole analyze a corner solution in which all merchants accept cards in equilibrium. A key feature of Rochet and Tirole's model is that they derive consumer and merchant demand from first principles, allowing for the fact merchants compete to attract customers (some of whom may be cardholders). Wright (2003a) combines elements of both approaches to explore the sources of divergence between the privately and socially optimal fee structures. However, all of these authors assume there is just *one* payment system that is choosing its price structure.[1] This paper relaxes this assumption.

In considering how competition between payment schemes determines the structure of fees between cardholders and merchants, this paper also falls within the recent literature on two-sided markets. These markets have the property that there are two types of agents that wish to use a common platform, and the benefits of each side depend on how many users there are on the other side of the network. Rochet and Tirole (2003) provide a general model of platform competition in a two-sided market. Examples they give include Adobe Acrobat (Acrobat Reader and Writer), payment cards (cardholders and merchants), platforms (hardware/console and software providers), real estate (home buyers and sellers), shopping malls (shoppers and retailers), and Yellow Pages (readers and advertisers). In contrast to their model of a single payment scheme (Rochet and Tirole, 2002), they treat users on both sides as end users, abstracting from the fact that an important feature of some of these markets is that one side competes amongst itself to sell to the other side (e.g. merchants, software providers, home sellers, retailers and advertisers).[2] In contrast, we allow for competition between merchants (sellers)

In addition to allowing for competition between merchants and competition between schemes, we also allow for consumers to make separate card-holding and card-usage decision. Our main analysis is presented in terms of a model of two competing card associations, such as MasterCard and Visa. For such schemes the structure of cardholder and merchant fees is determined by the level of the interchange fee, a fee set collectively by the members of the associations. The interchange fee is paid by the merchant's bank to the cardholder's bank on each card transaction. A higher interchange fee results in higher merchant fees and lower card fees. We model how the two types of competition alter the equilibrium interchange fee, and compare this to the socially optimal level. To obtain sharp results, we focus on the benchmark case

[1] This is also true of other recent models in the literature including Chakravorti and To (2000), Gans and King (2003), Schwartz and Vincent (2002), Wright (2003b), and Wright (2003c). Rochet and Tirole (2002) do discuss the impact of scheme competition in their model, but they do not find the resulting equilibrium.

[2] Other papers that model two-sided markets also abstract from the strategic interaction between sellers. These include Armstrong (2002), Caillaud and Jullien (2001, 2003), Parker and Van Alstyne (2000), and Schiff (2003).

in which consumers view the two types of cards as providing identical transactional benefits for making purchases, and merchants view the two types of cards as providing identical transactional benefits for receiving payment.

To compare results from our model with those of Rochet and Tirole (2002), we start with the benchmark case in which there is only a single card scheme, before allowing for competition between schemes in which consumers choose whether to hold none, one or both cards, and merchants choose whether to accept none, one or both cards. To see how merchant competition matters, in each case we start with a model in which the merchant is a monopolist, and so does not use card acceptance as a strategic instrument of competition. We then consider what happens when merchants compete in a Hotelling fashion.

A key determinant of the competitive fee structure (and interchange fee) is the extent to which consumers hold one or two cards when making purchases. With both forms of merchant behavior considered, when consumers only choose to hold one card, competition between card schemes does not result in lower interchange fees. In this case, by attracting cardholders, card schemes have a monopoly over access to these cardholders. This leads competing card schemes to care only about the surplus they can offer to cardholders, leaving no surplus to merchants (the case of a competitive bottleneck).

This competitive bottleneck outcome can be undermined provided some consumers hold both cards. Then the unique equilibrium in our model involves competing card schemes seeking to attract merchants exclusively, by offering maximal incentive for merchants to accept their cards. In the case of monopolistic merchants, this implies maximizing the expected surplus offered to merchants. In the case of competing merchants, this involves maximizing the expected joint surplus of consumers and merchants, given that competing merchants take into account the benefits that their customers get from being able to use cards. In either case, interchange fees are lower to the extent there is competition between schemes, but higher to the extent there is competition between merchants. When the two effects are combined, so there is competition between schemes and between merchants, the equilibrium interchange fee remains inefficiently low, but only to the extent that issuers and acquirers obtain positive margins.

The rest of the paper proceeds as follows. In Section 2 we present our model, starting with the case of a single card scheme. Section 3 considers the case of two identical competing payment card associations. Section 4 considers several extensions and implications of our analysis. Section 4.1 shows that our results also hold for the case of competing proprietary schemes, like American Express and Discover card, that set their cardholder and merchant fees directly. Section 4.2 considers how our results extend to allow for merchant heterogeneity. Section 4.3 explains how the existence of some cash-constrained consumers alters the equilibrium fee structure. Implications for policy and for the analysis of other two-sided markets are discussed in Sections 4.4 and 4.5. Finally, Section 5 provides some concluding remarks.

2 A single card scheme

We start by considering the simpler case of a single payment scheme where the only alternative to using cards is cash, before introducing competing payment schemes in the next section. This allows some of the analysis we will use in the next section to be first developed in a simpler setting. It also allows our model to be compared to the framework of Rochet and Tirole (2002) in the case in which there is only a single payment scheme. In addition to considering strategic merchants, as they do, we also allow for the case of monopolistic merchants, and for consumers to make separate joining and card usage decisions. Another difference between our model and theirs is that in our model consumers are assumed to receive their particular draw of transactional benefits from using cards once they have chosen which merchant to purchase from.

In the model of Rochet and Tirole, consumers get their draw of transactional benefits before they choose which merchant to purchase from. Clearly our timing assumption is made for modelling convenience. We think it is a reasonable modelling approach for two reasons. First, with our set-up, consumers still choose which merchant to purchase from (in the case of competing merchants) taking into account the expected benefits from using cards versus the alternative payment instrument. By accepting cards, merchants will raise consumers' expected benefit from purchasing from them, since consumers will gain the option of using cards for purchases. In fact, this section shows the timing assumption does not alter the equilibrium conditions under which merchants accept cards or consumers use cards: they are equivalent to the condition derived in Rochet and Tirole (2002) and Wright (2003a) for a Hotelling model of merchant competition, and to the condition derived in Baxter (1983) and Wright (2003c) in the case of monopolistic merchants. Thus, we do not think this particular timing assumption is driving the results we obtain. Additionally, the timing assumption can be motivated by the idea that consumers only learn of their particular need to use various types of payments once they are in the store.

A card association represents the joint interests of its members, who are issuers (banks and other financial institutions which specialize in servicing cardholders) and acquirers (banks and other financial institutions which specialize in servicing merchants). In such an open scheme, a card association sets an interchange fee a to maximize its members' collective profits.[3] The interchange fee is defined as an amount paid from acquirers to issuers per card transaction. In addition, we assume a cost of c_I per transaction of issuing and c_A per transaction of acquiring. A proprietary scheme incurs the cost $c_I + c_A$ of a card transaction. Competition between symmetric issuers and competition between symmetric acquirers then determines the equilibrium fee per transaction for using cards, f, and the equilibrium fee per transaction for accepting cards, m. In the case of a proprietary card scheme such as American Express, the scheme sets f and m directly to maximize its profits.

We follow Rochet and Tirole (2003, Section 6.2) and make the simplifying assumption that competition between symmetric issuers and between symmetric acquirers leaves some small constant equilibrium margin to issuers and to acquirers, so that

$$f(a) = c_I - a + \pi_I \qquad (1)$$

where π_I is some small constant profit margin.[4] Likewise, merchant fees are

$$m(a) = c_A + a + \pi_A, \qquad (2)$$

where π_A is some small constant profit margin. Intense intra-platform competition results in (as a first order approximation) the number of card transactions being taken as given, so that symmetric Hotelling or Salop style competition between banks would lead to the above equilibrium fees.

Card fees can be negative to reflect rebates and interest-free benefits offered to cardholders based on their card usage. Card fees decrease and merchant fees increase as the interchange fee is raised. The implication of the above assumption about bank competition is that the level of the interchange fee affects only the structure of fees, and not the overall level of fees: the sum of cardholder and merchant fees per-transaction, denoted by $l = f(a) + m(a)$, is independent of the interchange fee a (it equals

[3] In our set-up below this is also equivalent to assuming that the card association seeks to maximize the total number of card transactions. Notably, MasterCard and Visa (the card associations) obtain revenues from a small levy on each card transaction collected from their members, so they would seem to have the same incentives as their members.

[4] The form of card fees implies there is a one-to-one relationship between card fees and interchange fees.

$c_A + c_I + \pi_A + \pi_I$). As a result, a card association will maximize its members' profits by choosing its interchange fee to maximize its volume of card transactions.

As in Rochet and Tirole (2002), consumers get transactional or convenience benefits b_B from using cards as opposed to the alternative cash, and merchants get transactional or convenience benefits b_S from accepting cards relative to the alternative of accepting cash. The benefits b_B are drawn with a positive density $h(b_B)$ over the interval $[\underline{b}_B, \overline{b}_B]$. The hazard function $h(f)/(1 - H(f))$ is assumed to be increasing, where H denotes the cumulative distribution function corresponding to b_B. All merchants (sellers) are assumed to receive the same transactional benefits b_S from accepting cards (this assumption will be relaxed later). We will refer to $b_B - f$ as the 'surplus' to consumers from using cards, and $b_S - m$ as the 'surplus' to merchants from accepting cards. (Note, however, if merchants compete to attract cardholders, they will also profit from accepting cards through a business stealing effect.) We assume that

$$E(b_B) + b_S < l < \overline{b}_B + b_S \tag{3}$$

so as to rule out the possibility that there is no card use and to rule out the possibility that all consumers use cards.

It costs merchants d to produce each good, and all goods are valued at v by all consumers. There is a measure 1 of consumers who wish to buy from merchants. Consumers are assumed to each want to purchase one good. We also assume

$$v - d \geq l - \underline{b}_B - b_S + \frac{1 - H(\underline{b}_B)}{h(\underline{b}_B)},$$

which is used in Appendix B to show that even monopolistic merchants will set a price such that consumers who pay by cash will still want to purchase. Throughout, merchants are assumed to be unable to price discriminate depending on whether consumers use cards or not, so consumers will want to pay with the card if and only if $b_B \geq f$.[5] Using this property, we can define a number of important functions. The quasi-demand for card usage is defined as $D(f) = 1 - H(f)$, which is the proportion of consumers who want to use cards at the fee f. The average convenience benefit to those consumers using cards for a transaction is $\beta(f) = E[b_B | b_B \geq f]$, which is increasing in f. The expected surplus to a consumer (buyer) from being able to use their card at a merchant is

$$\phi_B(f) = D(f)\left(\beta(f) - f\right), \tag{4}$$

which is positive and decreasing in f. The expected surplus to a merchant (seller) from being able to accept cards is $D(f)(b_S - m)$, which, given that $f + m = l$, can be defined as

$$\phi_S(f) = D(f)\left(f + b_S - l\right). \tag{5}$$

Finally, the expected joint surplus to consumers and merchants from card usage is defined as

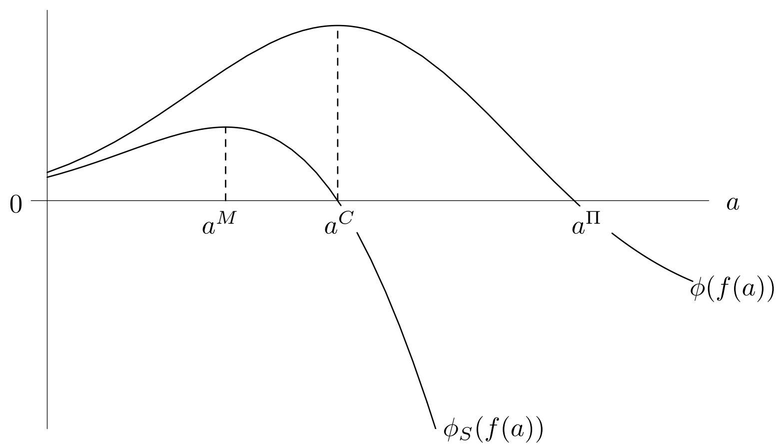
$$\phi(f) = D(f)(\beta(f) + b_S - l), \tag{6}$$

which equals $\phi_B(f) + \phi_S(f)$.

The following lemma summarizes some useful properties of these functions, and introduces three important levels of the interchange fee.

[5] This no price discrimination assumption can be motivated by the no-surcharge rules that card associations have adopted to prevent merchants from charging more to consumers for purchases made with cards. It can also be motivated by the observation of price coherence (Frankel, 1998) — that merchants are generally reluctant to set differential prices depending on the payment instrument used.

Figure 1: Important interchange fees



Notes. The interchange fee a^M maximizes $\phi_S(f(a))$, the expected surplus to a merchant from being able to accept cards, while a^C maximizes $\phi(f(a))$, the expected joint surplus to consumers and merchants from card usage. $\phi_S(f(a)) = 0$ at the interchange fee a^C, while $\phi(f(a)) = 0$ at a^Π.

Lemma 1

1. *There exists a unique interchange fee, denoted* a^C*, which maximizes* $\phi(f(a))$*. It equals*

$$a^C = b_S - c_A - \pi_A \tag{7}$$

and is the unique interchange fee which solves $\phi_S(f(a)) = 0$*.*

2. *There exists a unique interchange fee, denoted* a^M*, which maximizes* $\phi_S(f(a))$*.*

3. *There exists a unique interchange fee, denoted* a^Π*, which solves* $\phi(f(a)) = 0$*.*

4. *The interchange fees* a^C*,* a^M*, and* a^Π *satisfy* $a^M < a^C < a^\Pi$*.*

Proof. See Appendix A.1. ■

The results of this lemma are summarized in Figure 1.

Each consumer enjoys some intrinsic benefit, u, from holding a card, which is a random variable with positive density e over the interval $(-\infty, \overline{u}]$ for some $\overline{u} \geq 0$. We let $A(u) = 1 - E(u)$, where E is the cumulative distribution function for u. For some consumers $u < 0$, so that holding a card (which is not used) is inconvenient. This could also represent the case that there are some costs to issuers associated with managing a cardholder, if these costs are fully passed through to cardholders. For other consumers, cards offer more than just the ability to make transactions at merchants (for example, they may be used as security or to withdraw cash), in which case $u > 0$. The variable λ captures the measure of consumers who choose to hold a card.

The timing of the game is summarized as follows:

(i). The payment card association sets the level of its interchange fee a. Issuers and acquirers then set fees f and m to cardholders and merchants according to (1) and (2). Alternatively, a proprietary scheme sets f and m directly.

(ii). Consumers get their draw of u and decide whether or not to hold the card. Merchants decide whether or not to accept the card.

(iii). Merchants set their retail prices. If relevant, consumers decide which merchant to buy from.

(iv). Based on their individual realizations of b_B, consumers decide whether to use the card for payment (if they hold the card), or cash.

We first start with the case of monopolistic merchants.

2.1 Monopolistic merchants

By considering monopolistic merchants, we abstract from the business stealing motive that can influence a merchant's card acceptance strategy. This non-strategic approach is the setting that underlies the seminal analysis of Baxter (1983), in which he considered a merchant that accepted cards whenever the transactional benefits it obtained from doing so exceeded the merchant fee it was charged.

The following lemma describes the possible equilibria in stage (ii) of the game, and shows the Baxter condition for card acceptance also applies here.

Lemma 2 *Suppose a single card scheme has a card fee of f. If $\phi_S(f) < 0$, then monopolistic merchants reject the card and $\lambda(f) = A(0)$ consumers hold the card. If $\phi_S(f) \geq 0$, then monopolistic merchants accept the card, and $\lambda(f) = A(-\phi_B(f))$ consumers hold the card.*

Proof. We solve the game by working backwards through the stages. The stages are described in turn.

Stage (iv).

If consumers hold a card, they will use the card if $b_B \geq f$ and the merchant they buy from accepts the card and sets a price $p \leq v$. If consumers hold a card, they will use the card if $b_B \geq f + p - v$ and the merchant they buy from accepts the card and sets a price $p > v$. In any other circumstance, consumers will not use a card.

Stage (iii).

Provided it sets its price no higher than v, a merchant obtains profit of

$$\pi = p - d + \lambda(f)D(f)\,(b_S - m)\,I,$$

where I is an indicator variable which takes the value 1 if the merchant accepts the card and 0 otherwise. Alternatively, a merchant can set $p > v$, in which case it will only sell to cardholders who use cards, so that

$$\pi = \lambda(f)D(f + p - v)\,(p - d + b_S - m)\,I.$$

As we prove in Appendix B, it will not be profitable for a merchant to set a price above v, and exclude 'cash customers' (both those who do not hold a card and those who do not wish to use a card), provided the surplus from the good itself is sufficiently large. Instead, a monopolist will extract all the surplus from the cash customers by setting $p = v$, implying it earns a profit of

$$\pi = v - d + \lambda(f)\phi_S(f)I.$$

Stage (ii).

Given a merchant obtains a price of v regardless of whether it accepts cards or not, it will accept cards if and only if $\phi_S(f) \geq 0$. A consumer's benefit from holding the card is then $\phi_B(f)I + u$. It follows that consumers will hold a card if $u > 0$ and $\phi_S(f) < 0$ (using it only to withdraw cash and the like). They will also hold a card if $u > -\phi_B(f)$ and $\phi_S(f) \geq 0$. If neither condition applies, consumers will not hold a card. The equilibria in stage (ii) are thus those characterized by the lemma. ■

The card scheme's profit (that is, the total profit of the association's member banks) is zero if $\phi_S(f) < 0$, since merchants will not accept cards, and is equal to

$$\Pi(f) = (\pi_A + \pi_I)\, A\,(-\phi_B(f))\, D(f)$$

if $\phi_S(f) \geq 0$. Then we have

Proposition 1 *Facing monopolistic merchants, a single card scheme sets its interchange fee to solve $\phi_S(f) = 0$; that is, at $a = a^C$.*

Proof. A single card association maximizes $\Pi(f)$ by choosing f to maximize $A\,(-\phi_B(f))\, D(f)$ subject to the constraint $\phi_S(f) \geq 0$ which ensures merchants accept cards. The constraint is equivalent to $f + b_S \geq l$. Since $A(-\phi_B(f))D(f)$ is decreasing in f and the left hand side of the constraint is increasing in f, this implies the scheme will wish to set f as low as possible subject to the constraint. The constraint will be binding and the profit maximizing interchange fee solves $\phi_S(f) = 0$, which from Lemma 1 is precisely the interchange fee a^C defined in (7). ■

The single card scheme sets an interchange fee which leaves merchants with no surplus from accepting cards. This occurs when the fee charged to merchants equals the transactional benefits they obtain. This is also the (constrained) socially optimal level of the interchange fee, given monopolistic merchants. We have:

Proposition 2 *Facing monopolistic merchants, the welfare-maximizing interchange fee is $a = a^C$.*

Proof. See Appendix A.2. ■

Although the unconstrained socially optimal interchange fee is higher than a^C (it is $b_S - c_A + \pi_I$, so that consumers use cards whenever $b_B + b_S > c_I + c_A$), merchants will not accept cards if the interchange fee is set above a^C. The reason the interchange fee is too low here is that, facing stiff merchant resistance to accepting cards, the card scheme is unable to get merchants to absorb the issuers' and acquirers' margins. Instead, consumers who use cards will cover all of these margins, which leads to under-usage of cards. The fact that some consumers can still get utility from holding cards (even if they cannot use them) does not affect this conclusion, since these consumers still get the same utility from *holding* cards in the case they are able to use them.

In the limiting case of $\pi_A = \pi_I \to 0$, the profit maximizing interchange fee a^C equal $b_S - c_A$, which is also the Baxter interchange fee and the unconstrained socially optimal interchange fee.

2.2 Strategic merchants

In this section merchants may accept cards for strategic reasons, so as to attract customers from each other. Like Rochet and Tirole (2002), we model this by assuming there are two merchants who compete in a Hotelling fashion. In particular, consumers are uniformly distributed on the unit interval and the two merchants are located at either extreme. A consumer located at x faces linear transportation costs of tx from purchasing from merchant 1 and $t(1-x)$ from purchasing from merchant 2. These transportation costs can be summarized by the function $T_i(x) = tx\,(2-i) + t\,(1-x)\,(i-1)$, where $i = 1$ corresponds to merchant 1 and $i = 2$ corresponds to merchant 2. Note that the draws of u, x, and b_B are all assumed to be independent of one another.

The following lemma describes the possible equilibria in stage (ii) of the game.

Lemma 3 *Suppose a single card scheme has a card fee of f. If $\phi(f) < 0$, then competing merchants reject the card and $\lambda(f) = A(0)$ consumers hold the card. If $\phi(f) \geq 0$, then competing merchants accept the card and $\lambda(f) = A(-\phi_B(f))$ consumers hold the card.*

Proof. We solve the game by working backwards through time. The steps are described in turn.

Stage (iv).

Since consumers face the same price whether paying by card or by cash, they will only use the card if $b_B \geq f$ and the merchant they buy from accepts the card.

Stage (iii).

When deciding which merchant to buy from, consumers take into account their location in product space (their exogenous preference for the two merchants), whether or not they hold a card, the price charged by each merchant, and whether or not the merchants accept the card. A consumer located at x who buys from merchant i obtains indirect utility equal to

$$v_i = v - p_i - T_i(x)$$

if they do not hold a card, and equal to

$$v_i = v - p_i + \phi_B(f)I_i - T_i(x)$$

if they hold the card, where I_i is an indicator variable which takes the value 1 if merchant i accepts the card and 0 otherwise. The share of customers that firm 1 will attract is therefore

$$s_1 = \frac{1}{2} + \frac{1}{2t}\left(p_2 - p_1 + \lambda(f)\phi_B(f)\left(I_1 - I_2\right)\right), \tag{8}$$

where $\lambda(f)$ equals the measure of consumers holding cards given the card fee f. Firm 2 attracts a share $s_2 = 1 - s_1$ of consumers. Merchant i's profit is

$$\pi_i = s_i\left(p_i - d - \lambda(f)D(f)\left(m - b_S\right)I_i\right). \tag{9}$$

We solve for the Nash equilibrium in this stage by working out each merchant's profit-maximizing choice of prices given the share function (8). Solving these best responses simultaneously shows that the equilibrium prices are

$$p_1 = t + d + \lambda(f)D(f)I_1\left(m - b_S\right) + \frac{1}{3}\lambda(f)\phi(f)\left(I_1 - I_2\right), \tag{10}$$

$$p_2 = t + d + \lambda(f)D(f)I_2\left(m - b_S\right) + \frac{1}{3}\lambda(f)\phi(f)\left(I_2 - I_1\right). \tag{11}$$

Substituting (10) and (11) into (8) and (9) shows that firm i earns an equilibrium profit of

$$\pi_i = 2ts_i^2, \tag{12}$$

where

$$s_1 = \frac{1}{2} + \frac{1}{6t}\lambda(f)\phi(f)\left(I_1 - I_2\right) \tag{13}$$

and $s_2 = 1 - s_1$.

Stage (ii).

In this stage we determine equilibria in the subgame as simultaneous solutions of each party's best response, conditional on the card fee and merchant fee set by the card scheme in stage (i).

Merchants' best responses.

To work out merchants' optimal card acceptance policy we note that, regardless of what the other merchant does, each merchant will accept cards if doing so increases its equilibrium market share. Thus, each merchant will accept cards if doing so increases the function $\lambda(f)\phi(f)$, where the subscript for each merchant has been dropped since the function is the same for both merchants. Note that here merchants' card acceptance policy is determined by ϕ in the same way as it was determined by ϕ_S in Lemma 2. Thus, following the identical proof of Lemma 2 but replacing ϕ_S with ϕ, we end up with exactly the same results for merchant card acceptance, except with ϕ_S replaced by ϕ.

Consumers' best responses.

Their choice of card-holding depends on the benefits they get from holding a card, which depend on merchants' acceptance decisions. Given merchants are symmetric, the above result implies either both merchants will accept cards, if $\phi(f) \geq 0$, or both merchants will reject cards. In this case we can define $I = I_1 = I_2$. A consumer's additional benefit of holding the card is then $\phi_B(f)I + u$. If merchants reject the card (which happens when $\phi(f) < 0$), then consumers will hold the card if and only if $u \geq 0$. If merchants accept the card (which happens when $\phi(f) \geq 0$), then consumers will hold the card if and only if $u \geq -\phi_B(f)$. The equilibria in stage (ii) are thus those characterized by the lemma. ■

Notice from the proof above that card acceptance (that is, $I_i = 1$) does three things. First, it raises the demand faced by merchant i. It provides consumers with a valuable option from shopping at the merchant concerned, which is that they can use cards if doing so is convenient for them ($b_B > f$). The expected value of this option is measured by the term $\lambda(f)\phi_B(f)$ in the firm's market share equation (8). Second, and by symmetry, it lowers the demand faced by the rival firm. Thus, in this model card acceptance has a business stealing effect. Third, as the firm's profit equation (9) reveals, card acceptance changes the merchant's costs, increasing them if $m > b_S$ for the merchant concerned.

We are now ready to characterize equilibrium fee structures at stage (i) of the game. We consider the profit-maximizing interchange fee. The card scheme's profit is zero if $\phi(f) < 0$, since no merchants will accept cards, and equal to

$$\Pi(f) = (\pi_A + \pi_I)\, A(-\phi_B(f))D(f)$$

if $\phi(f) \geq 0$. We have:

Proposition 3 *Facing competing merchants, a single card scheme sets its interchange fee equal to a^{Π}, so that $\phi(f(a)) = 0$; that is, such that*

$$\beta(f(a^{\Pi})) + b_S = l. \tag{14}$$

Proof. Since $\pi_A + \pi_I > 0$, a single card association maximizes $\Pi(f)$ by maximizing $A(-\phi_B(f))D(f)$ subject to the constraint $\phi(f) \geq 0$ which ensures merchants accept cards. The constraint is equivalent to $\beta(f) + b_S \geq l$. Since $A(-\phi_B(f))D(f)$ is decreasing in f and $\beta(f)$ is increasing in f, this implies that the card scheme's profit is maximized by setting $f = f^{\Pi}$, where $\beta(f^{\Pi}) + b_S = l$ and $\phi(f^{\Pi}) = 0$. The existence and uniqueness of the corresponding interchange fee a^{Π} was proven in Lemma 1. ■

Condition (14) defines the merchant transactional benefits b_S, below which both merchants will not accept cards and above which both merchants will accept cards. Interestingly, this is the same equilibrium condition that Rochet and Tirole (2002) and Wright (2003a) obtained.[6] A single card scheme sets a

[6] To be precise, in Rochet and Tirole's model the term l on the right hand side is replaced by m. This difference only arises because in their model the consumers' fee f is effectively a fixed fee and so occurs regardless of the extent to which the card is used. Moreover, Rochet and Tirole also found another equilibrium was possible in which merchants with values of b_S slightly above this critical level both reject cards, although they rule out this second equilibrium. In their model,

fee structure which encourages maximal use of cards by consumers while ensuring merchants have just enough surplus so that they will accept cards. Since merchants compete amongst themselves, they take into account their customers' average surplus from using cards in deciding whether to accept cards or not. Thus, when the card scheme leaves merchants just indifferent between accepting cards or not, it implies that the expected joint surplus to consumers and merchants from the card scheme are exactly zero.

As in Rochet and Tirole (2002), the welfare-maximizing interchange fee is either the same as the privately chosen interchange fee set by a single scheme, or is lower.

Proposition 4 *Facing competing merchants, the welfare-maximizing interchange fee is*

$$a^W = \min\left\{b_S - c_A + \pi_I, a^\Pi\right\}. \tag{15}$$

Proof. See Appendix A.3. ■

The unconstrained socially optimal interchange fee is the one at which consumers face the joint costs of using cards less the transactional benefits that merchants obtain. This requires a card fee of $f = c_A + c_I - b_S$. Since at this card fee consumers face the full social costs and benefits of their card-holding decision, including the cost or benefit (u) simply from holding the card, there is no reason to distort card fees in this model to encourage or discourage additional card-holding. The only reason then to set a lower interchange fee is if this interchange fee implies merchants fees above which merchants would accept cards. It is never optimal to exclude merchants and so get no card transactions.

3 Competition between identical card schemes

We modify the model of Section 2 by assuming there are two competing identical card systems. Identical systems not only have the same costs and issuer and acquirer margins, they also provide the same benefits to cardholders and merchants. The only distinguishing feature of each card scheme is the fee structure it chooses. Specifically, each card association i sets an interchange fee denoted a^i.

Like the case with a single scheme, we follow Rochet and Tirole (2003, Section 6.2)[7] and assume that for competing card associations

$$f^i(a^i) = c_I - a^i + \pi_I \tag{16}$$

and

$$m^i(a^i) = c_A + a^i + \pi_A, \tag{17}$$

so that the interchange fee determines the structure but not the overall level of fees. Taking the limit of the equilibrium fee structure as π_I and π_A tend to zero allows us to capture the case of perfect intra-system competition.[8] As before, we define l as the sum of f and m, which, from the fact that the issuers and acquirers in each scheme are assumed to have the same costs and margins, does not depend on i. We also define $D^i = D\left(f^i\right)$, $\phi_B^i = \phi_B\left(f^i\right)$, $\phi_S^i = \phi_S\left(f^i\right)$, and $\phi^i = \phi\left(f^i\right)$.

each firm's acceptance decision is affected by what the other firm does. Due to our different timing assumption, each firm's acceptance decision is independent of what other firms are doing, and so we get a unique equilibrium.

[7] See also Hausman *et al.* (2003).

[8] Like Rochet and Tirole (2003) we do not deal with issues of duality, in which banks may be members of both card associations. See Hausman *et al.* (2003) for an analysis of duality. A further justification for this form of bank fees is that it can be used to recover the equilibrium fees that result from competition between two identical proprietary schemes, that set their fees directly to consumers and merchants. Section 4.1 carries out this exercise.

In stage (ii) consumers make their card-holding decision, which is now whether to hold none, one or both cards. Likewise, in stage (ii) merchants now have to decide whether to accept none, one or both cards. We let λ^i be the measure of consumers who hold card i only (singlehoming consumers), and λ^{12} be the measure of consumers who hold both cards (multihoming consumers).

We make use of two tie-breaking conventions. First, we assume if consumers are indifferent about holding one card or another, they will randomize to determine which card to hold. Second, we assume that if merchants are indifferent between accepting a card or not because they do not expect consumers to use the card, they will accept the card if doing so would increase their profits (or at least not decrease their profits) if some consumers did use the card.

The timing of the game is summarized as follows:

(i). Each payment card association sets the level of their interchange fee a^i. Issuers and acquirers then set fees f^i and m^i to cardholders and merchants according to (16) and (17). Alternatively, a proprietary scheme i sets f^i and m^i directly.

(ii). Consumers decide which cards to hold (neither, one or both). Merchants decide whether to accept cards (neither, one or both).

(iii). Merchants set their retail prices. If relevant, consumers decide which merchant to buy from.

(iv). Based on their individual realizations of b_B, consumers decide whether to use cards or cash for payment.

We start again with the case of monopolistic merchants, which, like the analysis of Rochet and Tirole (2003), allows competition in two-sided markets to be analyzed without considering the strategic interaction between sellers. Unlike Rochet and Tirole's analysis of this problem, we first consider the consumers' decisions about whether to hold none, one, or both cards, and assume all merchants get the same benefits from accepting cards.[9]

3.1 Monopolistic merchants

The following lemma describes the possible equilibria in stage (ii) of the game for any given card fees set by the schemes in stage (i).

Lemma 4 *Suppose two identical card schemes compete, with card schemes* 1 *and* 2 *having card fees* f^1 *and* f^2 *respectively.*

1. *If* $f^1 = f^2$ *then* $\phi_S^1 = \phi_S^2$ *and there are two cases to consider. If* $\phi_S^1 = \phi_S^2 \geq 0$*, then monopolistic merchants accept both cards and the singlehoming consumers will randomize over which card to hold. If* $\phi_S^1 = \phi_S^2 < 0$*, then monopolistic merchants reject both cards and there are no singlehoming consumers.*

2. *If* $f^i < f^j$ *then four equilibria are possible at stage (ii). If* $\phi_S^i, \phi_S^j < 0$*, then monopolistic merchants reject both cards and there are no singlehoming consumers. If* $\phi_S^i, \phi_S^j \geq 0$ *and* $(\lambda^i + \lambda^{12})\phi_S^i \geq \lambda^{12}\phi_S^j$*, then monopolistic merchants accept both cards and the singlehoming consumers will only hold card* i*. If* $\phi_S^i \geq 0$*,* $\phi_S^j < 0$*, then monopolistic merchants only accept card* i *and the singlehoming consumers*

[9] In Section 4.2, we consider a case in which all consumers hold both cards and merchants are heterogenous, a case closer to theirs.

will only hold card i. If $\phi_S^j \geq 0$ and $\phi_S^i < 0$ then monopolistic merchants only accept card j and the singlehoming consumers will only hold card j. (If $\overline{u} > 0$, then this equilibrium occurs more generally if $\phi_S^j \geq 0$ and $\phi_S^i < \phi_S^j$.)

Proof. We solve the game by working backwards through time.

Stage (iv).

To characterize consumers' usage of cards, it is useful to introduce another indicator variable, L_i^j, which captures the likelihood that, assuming they get a sufficiently high draw of b_B (that is $b_B \geq f^j$), consumers who hold both cards will prefer to use cards from scheme j at merchant i. For the case of monopolistic merchants, the subscript i is redundant. We retain it here since it will become relevant for the case of competing merchants considered in the next section. If $f^1 < f^2$ consumers will prefer to use card 1 if merchants accept both cards (or just card 1), and will prefer to use card 2 if this is the only card accepted. We thus define

$$L_i^1 = I_i^1 \text{ and } L_i^2 = I_i^2 \left(1 - I_i^1\right) \tag{18}$$

if $f^1 < f^2$. If $f^1 > f^2$ consumers will prefer to use card 2 if merchants accept both cards (or just card 2), and will prefer to use card 1 if this is the only card accepted. We thus define

$$L_i^1 = I_i^1 \left(1 - I_i^2\right) \text{ and } L_i^2 = I_i^2 \tag{19}$$

if $f^1 > f^2$. If $f^1 = f^2$ consumers will be indifferent about which card to use if merchants accept both, and so will randomize over card usage. If merchants only accept one card, then consumers prefer to use this card. We thus define

$$L_i^1 = I_i^1 \left(1 - \frac{1}{2} I_i^2\right) \text{ and } L_i^2 = I_i^2 \left(1 - \frac{1}{2} I_i^1\right) \tag{20}$$

if $f^1 = f^2$.

Provided merchants set their common price p no higher than v, consumers will only use a card if the transactional benefits of doing so are at least as high as the fee they face. If consumers only hold card i, they will only use the card if $b_B \geq f^i$ and the merchant they buy from accepts the card. The card use of consumers who hold both cards and draw $b_B \geq \min\{f^1, f^2\}$ is described by equations (18), (19) and (20). Alternatively, if merchants set their price above v, then if consumers only hold card i, they will only use the card if $b_B + v \geq f^i + p$ and the merchant they buy from accepts the card. The card use of consumers who hold both cards and draw $b_B \geq \min\{f^1, f^2\} + p - v$ is described by equations (18), (19) and (20).

Stage (iii).

Given there is a measure 1 of potential consumers, provided a merchant sets its price less than or equal to v in stage (iii), the merchant's profit is

$$\pi = p - d - \lambda^1 D^1 \left(m^1 - b_S\right) I^1 - \lambda^2 D^2 \left(m^2 - b_S\right) I^2 - \lambda^{12} \left(D^1 \left(m^1 - b_S\right) I^1 L^1 + D^2 \left(m^2 - b_S\right) I^2 L^2\right).$$

Alternatively, a merchant can set $p > v$, in which case the merchant will only sell to cardholders who use cards, so that

$$\begin{aligned} \pi \;=\; & \lambda^1 D\left(f^1 + p - v\right)\left(p - d + b_S - m^1\right) I^1 + \lambda^2 D\left(f^2 + p - v\right)\left(p - d + b_S - m^2\right) I^2 \\ & + \lambda^{12}\left(\left(f^1 + p - v\right)\left(p - d + b_S - m^1\right) I^1 L^1 + D\left(f^2 + p - v\right)\left(p - d + b_S - m^2\right) I^2 L^2\right). \end{aligned}$$

As we show in Appendix B, it will not be profitable for merchants to set a price above v, and exclude 'cash customers' (both those who do not hold a card and those who do not wish to use a card), provided the surplus from the good itself is sufficiently large. Instead, merchants will extract all the surplus from the cash customers by setting $p = v$, implying merchants earn a profit of

$$\pi = v - d + \Psi, \tag{21}$$

where

$$\Psi = \lambda^1 \phi_S^1 I^1 + \lambda^2 \phi_S^2 I^2 + \lambda^{12} \left(\phi_S^1 I^1 L^1 + \phi_S^2 I^2 L^2 \right). \tag{22}$$

Stage (ii).

In this stage we determine equilibria in the subgame as simultaneous solutions of each party's best response, conditional on the card fees and merchant fees set by the two card schemes in stage (i).

The merchant's best response.

To work out a merchant's optimal card acceptance policy we note that a merchant will accept cards if doing so increases the function Ψ.

We must consider two possibilities for consumers' card-holding. In the first possibility we consider, no consumers multihome, so that $\lambda^{12} = 0$ and the function Ψ is determined by the following table:

	$I^2 = 0$	$I^2 = 1$
$I^1 = 0$	0	$\lambda^2 \phi_S^2$
$I^1 = 1$	$\lambda^1 \phi_S^1$	$\lambda^1 \phi_S^1 + \lambda^2 \phi_S^2$

Recall that if a merchant is indifferent between accepting and rejecting card i because it does not expect consumers to use card i (so that accepting the card leaves the function Ψ unchanged), it will accept the card if doing so increases Ψ when consumers do use card i. This is true if and only if $\phi_S^i \geq 0$. Merchants therefore adopt the following policy: merchants reject both cards if $\phi_S^1, \phi_S^2 < 0$, accept both cards if $\phi_S^1, \phi_S^2 \geq 0$, and accept only card 1 (respectively, card 2) if $\phi_S^1 \geq 0 > \phi_S^2$ (respectively, $\phi_S^2 \geq 0 > \phi_S^1$). When all consumers hold at most one card, even though there are two card schemes, the condition that determines whether merchants accept cards is identical to the case with a single card scheme. Each individual merchant does not expect to be able to influence the number of cardholders of each type, and so it acts as though these are two segmented groups of consumers.

The second possibility is that some consumers hold both cards, so that $\lambda^{12} > 0$. Since L^1 and L^2 in equation (22) depend on the values of f^1 and f^2, we need to consider three different cases.

- If $f^1 = f^2$ then $\phi_S = \phi_S^1 = \phi_S^2$ and equation (20) implies that

$$\Psi = \phi_S \left(\lambda^1 I^1 + \lambda^2 I^2 + \lambda^{12} \left(I^1 + I^2 - I^1 I^2 \right) \right).$$

  In this case the function Ψ is determined by the following table:

	$I^2 = 0$	$I^2 = 1$
$I^1 = 0$	0	$\left(\lambda^2 + \lambda^{12} \right) \phi_S$
$I^1 = 1$	$\left(\lambda^1 + \lambda^{12} \right) \phi_S$	$\left(\lambda^1 + \lambda^2 + \lambda^{12} \right) \phi_S$

  Merchants' best response is to reject both cards if $\phi_S < 0$ and to accept both cards if $\phi_S \geq 0$.

- If $f^1 < f^2$, then equation (18) implies that

$$\Psi = \lambda^1 \phi_S^1 I^1 + \lambda^2 \phi_S^2 I^2 + \lambda^{12} \left(\phi_S^1 I^1 + \phi_S^2 I^2 \left(1 - I^1 \right) \right).$$

In this case the function Ψ is determined by the following table:

	$I^2 = 0$	$I^2 = 1$
$I^1 = 0$	0	$\left(\lambda^2 + \lambda^{12}\right) \phi_S^2$
$I^1 = 1$	$\left(\lambda^1 + \lambda^{12}\right) \phi_S^1$	$\left(\lambda^1 + \lambda^{12}\right) \phi_S^1 + \lambda^2 \phi_S^2$

Merchants' best response is as follows: merchants reject both cards if $\phi_S^1, \phi_S^2 < 0$; they accept only card 1 if $\phi_S^2 < 0 \leq \phi_S^1$; they accept only card 2 if $\phi_S^2 \geq 0$ and $\lambda^{12}\phi_S^2 > (\lambda^1 + \lambda^{12})\phi_S^1$; and they accept both cards if $(\lambda^1 + \lambda^{12})\phi_S^1 \geq \lambda^{12}\phi_S^2 \geq 0$.[10]

- Exploiting the symmetry with the above case, merchants' best response is the same except the superscripts 1 and 2 are swapped.

Consumers' best responses.

At stage (ii), consumers decide which card(s) to hold, if any. Their choice of card-holding depends on the benefits they get from holding a card, which depend on merchants' acceptance decisions. A consumer's additional benefit of holding only card i is

$$\phi_B^i I^i + u,$$

while the additional benefit of holding two cards is

$$\phi_B^1 I^1 L^1 + \phi_B^2 I^2 L^2 + 2u.$$

If merchants reject both cards, then consumers with $u \geq 0$ will hold two cards, while consumers with $u < 0$ will hold no cards. If merchants accept only card i, then consumers will hold both cards if $u \geq 0$, will hold only card i if $-\phi_B^i \leq u < 0$ and will hold neither card if $u < -\phi_B^i$. If merchants accept both cards but card i has a lower card fee than the other card, then consumers will hold both cards if $u \geq 0$, will hold only card i if $-\phi_B^i \leq u < 0$ and will hold neither card if $u < -\phi_B^i$. If merchants accept both cards and both cards have the same card fees, then consumers will hold both cards if $u \geq 0$, will hold only a single card if $-\phi_B^1 = -\phi_B^2 \leq u < 0$ (in which case consumers will randomize over which card they will hold), and will hold neither card if $u < -\phi_B^1 = -\phi_B^2$. These results are summarized by the functions

$$\begin{aligned} \lambda^0(f^1, f^2) &= 1 - A(-\phi_B^1 L^1 - \phi_B^2 L^2), \\ \lambda^i(f^1, f^2) &= L^i \left(A(-\phi_B^i) - A(0)\right), \\ \lambda^{12}(f^1, f^2) &= A(0), \end{aligned}$$

which give the measure of consumers who hold neither card, just card i, or both cards respectively. Note if $\overline{u} = 0$ then $\lambda^{12} = 0$ and no consumers multihome.

Equilibria in the subgame.

Using the characterizations of consumers' and merchants' best responses, we can look for cases where both types of users have best responses to each other at stage (ii) — that is, we can look for possible equilibria in the subgame starting at stage (ii). There are three cases to consider based on the relative sizes of f^1 and f^2.

[10] Some of these results require a little thought. For example, consider the conditions required for accepting just card 1 to be optimal. If $\lambda^2 = 0$ then this decision is optimal because the merchant is indifferent between accepting only card 1 and accepting both cards (since consumers will never use card 2); the tie-breaking assumption means that the merchant will reject card 2 in this circumstance (since $\phi_S^2 < 0$). On the other hand, if $\lambda^2 > 0$, then accepting only card 1 is optimal because it leads to a higher value of Ψ than accepting both cards.

Case 1: $f^1 = f^2$**.** In this case $\phi_S^1 = \phi_S^2$. Then from above, an equilibrium in stage (ii) exists if $\phi_S^1 = \phi_S^2 \geq 0$. The merchant accepts both cards and the singlehoming consumers will randomize over which card to hold. An equilibrium also exists in stage (ii) if $\phi_S^1 = \phi_S^2 < 0$, in which case the merchant rejects both cards and there are no singlehoming consumers.

Case 2: $f^1 < f^2$**.** Then there are four possible equilibria at stage (ii). If $\phi_S^1, \phi_S^2 < 0$, there is an equilibrium in which the merchant rejects both cards and there are no singlehoming consumers. If $\phi_S^1, \phi_S^2 \geq 0$ and $(\lambda^1 + \lambda^{12})\phi_S^1 \geq \lambda^{12}\phi_S^2$ there is an equilibrium in which the merchant accepts both cards and the singlehoming consumers will only hold card 1. If $\phi_S^1 \geq 0$, $\phi_S^2 < 0$ there is an equilibrium in which the merchant only accepts card 1 and the singlehoming consumers will only hold card 1. If $\phi_S^2 \geq 0$ and $\phi_S^1 < 0$ there is an equilibrium in which the merchant only accepts card 2 and the singlehoming consumers will only hold card 2. (If $\overline{u} > 0$, then this equilibrium occurs more generally if $\phi_S^2 \geq 0$ and $\phi_S^1 < \phi_S^2$.)

Case 3: $f^1 > f^2$**.** By symmetry, this is the same as the above case except the superscripts 1 and 2 are swapped.

■

Note if there are some multihoming consumers (so $\lambda^{12} > 0$), there is the possibility of multiple equilibria in the stage (ii) subgame. This happens if $f^1 < f^2$, $\phi_S^1 \geq 0$, $\phi_S^2 \geq 0$, $\lambda^{12}\phi_S^2 > \lambda^{12}\phi_S^1$ and $(\lambda^1 + \lambda^{12})\phi_S^1 \geq \lambda^{12}\phi_S^2$ or if $f^1 > f^2$, $\phi_S^1 \geq 0$, $\phi_S^2 \geq 0$, $\lambda^{12}\phi_S^1 > \lambda^{12}\phi_S^2$ and $(\lambda^2 + \lambda^{12})\phi_S^2 \geq \lambda^{12}\phi_S^1$. The two equilibria are (a) merchants accept both cards and singlehoming consumers will only hold the card with lower card fees; and (b) merchants only accept the card with higher ϕ_S and the singlehoming consumers will only hold this card. If monopolistic merchants could choose one of these equilibria, they would choose the latter equilibrium. This maximizes their profits in the subgame. Where there are multiple equilibria, we select this equilibria in the subgame. An alternative possibility is to select the equilibria in the subgame in which merchants accept both cards. In this case, there is no (pure-strategy) equilibrium in the first stage of the game.[11]

We start our analysis of stage (i) equilibria by considering the special case in which consumers never want to hold both cards. This leads to the case of a 'competitive bottleneck'.

3.1.1 Consumers hold at most one card

With the assumption that consumers get no intrinsic benefit from holding cards ($\overline{u} = 0$), we can see from the characterization of consumers' best responses at stage (ii) of the game that consumers will not hold multiple cards. If merchants accept just one card, this is the card consumers will hold, while if merchants accept both cards, singlehoming consumers will decide which is the best card to hold. Then if both schemes set the same interchange fee (so $\phi_S \equiv \phi_S^1 = \phi_S^2$), merchants will accept both cards if $\phi_S \geq 0$ and neither otherwise. When merchants accept both cards, card-holding consumers randomize over which card to hold, and the members of such card schemes get (in aggregate) profits of

$$\Pi^1 = \Pi^2 = (\pi_I + \pi_A)\frac{A(-\phi_B(f))D(f)}{2}.$$

[11] This follows because any scheme i can attract all users by slightly undercutting the other scheme's interchange fee (and so card fee), provided $\phi_S^i \geq 0$. This would cause schemes to compete by setting high interchange fees to the point that $\phi_S^i = 0$, at which point the other scheme can attract merchants exclusively by setting a slightly lower interchange fee (higher card fee) given that some consumers will be holding both cards regardless of the fees.

If card schemes act to maximize their joint profits they will set their interchange fees so that $\phi_S^1 = \phi_S^2 = 0$, which leads to the highest level of $A(-\phi_B(f))D(f)$ such that merchants still accept cards. This is the interchange fee a^C defined in (7). We now show this is also the equilibrium outcome from competition between the two identical schemes.

Proposition 5 *If consumers get no intrinsic benefit from holding cards, the equilibrium interchange fee resulting from competition between identical card schemes facing monopolistic merchants equals a^C; that is, it solves $\phi_S(f(a)) = 0$. Merchants will (just) accept both cards and card-holding consumers will randomize over which card to hold. Each association shares in half the card transactions.*

Proof. The existence of a^C was proven in Lemma 1. The next step is to prove that this is an equilibrium using our analysis of equilibria in the subgame starting at stage (ii). From the analysis of consumers' best responses at stage (ii) of the game, with $\overline{u} = 0$, $\lambda^{12} = A(0) = 0$. No consumers will hold both cards. Note that at a^C, $\phi_S^1 = \phi_S^2 = 0$. If scheme 1 sets a card fee $f^1 < f(a^C)$, then $\phi_S^1 < \phi_S^2 = 0$, merchants will accept only card 2 and no consumers will hold card 1; scheme 1 will get no card transactions. If scheme 1 sets a card fee $f^1 > f(a^C)$ instead, then either $\phi_S^1 \geq 0$, in which case merchants accept both cards and no consumers hold card 1, or $\phi_S^1 < 0$, in which case merchants accept only card 2, and no consumers hold card 1; in either case, scheme 1 will get no card transactions. Thus, this is indeed an equilibrium.

This equilibrium is unique, since if any scheme i sets a fee structure such that $\phi_S^i > 0$, then the other scheme will always want to attract all consumers to hold its card by setting a lower card fee such that $\phi_S^j \geq 0$ and $\phi_S^j < \phi_S^i$. The optimal response of scheme i will be to match this fee structure. If any scheme i sets a fee structure such that $\phi_S^i < 0$, then merchants will reject its cards and the other scheme will always want to attract all consumers to hold it cards by setting a fee structure at which merchants will accept its cards (that is, with $\phi_S^j \geq 0$). The optimal response of scheme i will be to change its fee structure so that $\phi_S^i \geq 0$. Thus, the only equilibrium is one with $\phi_S^i = \phi_S^j = 0$. ■

Despite competition between identical schemes, they will each set their interchange fees as though they are a single scheme maximizing card transactions (and profits). When consumers hold only one card, the effect of competition between card schemes is to make it more attractive for each card scheme to lower card fees to attract exclusive cardholders to their network. Cardholders provide each card scheme with a bottleneck over a merchant's access to these cardholders. Since with no merchant heterogeneity a single scheme already sets the interchange fee to the point where merchants only just accept cards, there is no scope to further lower fees to cardholders by raising merchants' fees. Thus, despite competition between the schemes, their fee structure is unchanged from the case of a single scheme. Armstrong (2002) calls this situation a competitive bottleneck. Wright (2002) analyzes a similar competitive bottleneck that arises in mobile phone termination. The welfare implications of this equilibrium are exactly as characterized in Section 2.1. The socially optimal interchange fee is equal to a^C.

As we will see in the next section, this competitive bottleneck situation is quite a special one. It depends on the assumption that all consumers singlehome (hold just one card).

3.1.2 Some consumers hold both cards

Turning now to the case where some consumers get intrinsic benefits from holding cards, the implications of the resulting multihoming of consumers is dramatic. Rather than extracting all of the merchants' surplus, identical card schemes will compete by setting their interchange fee to maximize ϕ_S^i, the expected surplus to merchants from accepting cards. The (pure-strategy) equilibrium interchange fee is characterized in the following proposition.

Proposition 6 *If some consumers get intrinsic benefits from holding cards, the equilibrium interchange fee resulting from competition between identical card schemes facing monopolistic merchants equals a^M, which maximizes ϕ_S^i, the expected surplus to merchants from accepting cards. Merchants accept both cards and singlehoming consumers randomize over which card to hold. The measure of multihoming consumers is*

$$\lambda^M\left(a^M\right)=A\left(0\right),$$

the measure of consumers not holding cards is

$$\lambda^N\left(a^M\right)=1-A\left(-\phi_B\left(f\left(a^M\right)\right)\right),$$

and the measure of singlehoming consumers is

$$\lambda^S\left(a^M\right)=A\left(-\phi_B\left(f\left(a^M\right)\right)\right)-A\left(0\right).$$

Proof. The existence and uniqueness of a^M was proven in Lemma 1. From the analysis of consumers' best responses at stage (ii) of the game, $\lambda^{12}=A(0)>0$. Some consumers will hold both cards. We use this property and the analysis of equilibria in the stage (ii) subgame to show that a^M represents an equilibrium interchange fee.

Any scheme (say scheme 1) that sets a higher card fee f^1 (lower interchange fee), will result in $\phi_S^1<\phi_S^2$, so will imply an equilibrium at stage (ii) in which merchants accept both cards and the singlehoming consumers will only hold card 2. The measure of each type of consumer λ^N, λ^S and λ^M will not change since singlehoming consumers will get the same benefits from holding card 2, as previously they obtained from randomizing over which card to hold. Scheme 1 will get no card transactions.

Any scheme (say scheme 1) that sets a lower fee f^1 (higher interchange fee), will result in $\phi_S^1<\phi_S^2$, so will imply an equilibrium at stage (ii) in which merchants will only accept card 2 and the singlehoming consumers will only hold card 2. Again, the measure of each type of consumer will not change since singlehoming consumers will get the same benefits from holding card 2, as previously they obtained from randomizing over which card to hold. Again, scheme 1 will get no card transactions.

Thus, scheme 1 does strictly worse by setting a higher or lower interchange fee than that which maximizes ϕ_S, proving that this is an equilibrium.

It remains to prove that this equilibrium is unique. Suppose that it is not. Then there exists some other equilibrium in which one scheme (say scheme 1) sets an interchange fee such that $\phi_S^1<\phi_S^{\max}$. It is straightforward to show that scheme 2's best response is to set a different interchange fee so that $\phi_S^2>\phi_S^1$, in which case it attracts all card transactions. Thus, there can be no other equilibrium. ■

In equilibrium, both schemes set their interchange fee at the level a^M, and merchants accept both schemes' cards. This maximizes the expected surplus to merchants from accepting cards. Competition drives cards schemes to offer the maximal profit to merchants from card acceptance, in an attempt to have their card accepted exclusively, thus obtaining all card transactions. In equilibrium each card scheme shares equally in the card transactions.

The equilibrium interchange fee a^M can be compared to the interchange fee that maximizes the schemes' joint profits when merchants are monopolists — the interchange fee a^C defined in (7).

Proposition 7 *When some consumers get intrinsic benefits from holding cards, the equilibrium interchange fee resulting from competition between identical card schemes facing monopolistic merchants leads to an interchange fee lower than that which maximizes the schemes' joint profit (or joint card transactions).*

Proof. The result $a^M < a^C$ was proven in Lemma 1. ■

Since a^C is also the constrained socially optimal interchange fee for the case with monopolistic merchants, the competitive interchange fee here must be below the socially optimal interchange fee, even in the limit case $\pi_A = \pi_I \to 0$.

Proposition 8 *When some consumers get intrinsic benefits from holding cards, the equilibrium interchange fee resulting from competition between identical card schemes facing monopolist merchants leads to an interchange fee lower than that which maximizes overall welfare.*

Proof. See Appendix A.4. ■

By taking into account only the interests of one side of the market (merchants), the interests of the other side (consumers) are ignored. This results in card fees that are set too high and merchant fees that are set too low.

3.2 Strategic merchants

We now move to the case of competing card schemes and competing merchants. The analysis follows directly from that of monopolistic merchants. The main difference is that competing merchants accept cards to attract business from each other, so that attracting merchants now involves taking into account the surplus offered to their customers as well.

We first characterize equilibria at the stage (ii) subgame in which consumers decide which card(s) to hold and merchants decide which card(s) to accept.

Lemma 5 *Suppose two identical card schemes compete, with card schemes* 1 *and* 2 *having card fees* f^1 *and* f^2 *respectively.*

1. *If* $f^1 = f^2$ *then* $\phi^1 = \phi^2$ *and there are two cases to consider. If* $\phi^1 = \phi^2 \geq 0$*, then competing merchants will accept both cards and the singlehoming consumers will randomize over which card to hold. If* $\phi^1 = \phi^2 < 0$*, then competing merchants reject both cards and there are no singlehoming consumers.*

2. *If* $f^i < f^j$ *then four equilibria are possible at stage (ii). If* $\phi^i, \phi^j < 0$*, then competing merchants reject both cards and there are no singlehoming consumers. If* $\phi^i, \phi^j \geq 0$ *and* $(\lambda^i + \lambda^{12})\phi^i \geq \lambda^{12}\phi^j$*, then competing merchants will accept both cards and the singlehoming consumers will only hold card* i*. If* $\phi^i \geq 0$*,* $\phi^j < 0$*, then competing merchants will only accept card* i *and the singlehoming consumers will only hold card* i*. If* $\phi^j \geq 0$ *and* $\phi^i < \phi^j$*, then competing merchants will only accept card* j *and the singlehoming consumers will only hold card* j*.*

Proof. We solve the game by working backwards through time.

Stage (iv).

As in Lemma 4, if consumers only hold card i, they will only use the card if $b_B \geq f^i$ and the merchant they buy from accepts the card. The card use of consumers who hold both cards and draw $b_B \geq \min\{f^1, f^2\}$ is described by equations (18), (19) and (20).

Stage (iii).

Consumers take into account the price charged by each merchant when deciding which merchant to buy from, as well as their location in product space (their exogenous preference for the two merchants), whether they hold each of the cards, and whether the merchants accept each of the cards. A consumer located at x who buys from merchant i obtains indirect utility equal to

$$v_i = v - p_i - T_i(x)$$

if they hold neither card, equal to

$$v_i = v - p_i + \phi_B^j I_i^j - T_i(x)$$

if they just hold card j, and equal to

$$v_i = v - p_i + \phi_B^1 I_i^1 L_i^1 + \phi_B^2 I_i^2 L_i^2 - T_i(x)$$

if they hold both cards. Then the share of consumers that firm 1 will attract is

$$\begin{aligned} s_1 &= \frac{1}{2} + \frac{1}{2t}\Big(p_2 - p_1 + \lambda^1 \phi_B^1 \left(I_1^1 - I_2^1\right) + \lambda^2 \phi_B^2 \left(I_1^2 - I_2^2\right) \\ &\qquad + \lambda^{12}\left(\phi_B^1 (I_1^1 L_1^1 - I_2^1 L_2^1) + \phi_B^2 (I_1^2 L_1^2 - I_2^2 L_2^2)\right)\Big), \end{aligned} \tag{23}$$

while firm 2 attracts $s_2 = 1 - s_1$ of the consumers.

The next step is to determine the merchants' equilibrium prices conditional on the cards held by consumers and the cards accepted by merchants. Merchant i's profit is

$$\begin{aligned} \pi_i &= s_i \Big(p_i - d - \lambda^1 D^1 \left(m^1 - b_S\right) I_i^1 - \lambda^2 D^2 \left(m^2 - b_S\right) I_i^2 \\ &\qquad - \lambda^{12}\left(D^1 \left(m^1 - b_S\right) I_i^1 L_i^1 + D^2 \left(m^2 - b_S\right) I_i^2 L_i^2\right)\Big). \end{aligned} \tag{24}$$

We solve for the Nash equilibrium in this stage by working out each merchant's profit-maximizing choice of prices given the share function (23). Solving these best responses simultaneously implies that the equilibrium price p_i is

$$\begin{aligned} p_i =& t + d + \lambda^1 D^1 I_i^1 (m^1 - b_S) + \lambda^2 D^2 I_i^2 (m^2 - b_S) + \lambda^{12}(D^1 I_i^1 L_i^1 (m^1 - b_S) + D^2 I_i^2 L_i^2 (m^2 - b_S)) \\ &+ \frac{1}{3}(\lambda^1 \phi^1 (I_i^1 - I_j^1) + \lambda^2 \phi^2 (I_i^2 - I_j^2) + \lambda^{12}(\phi^1 (I_i^1 L_i^1 - I_j^1 L_j^1) + \phi^2 (I_i^2 L_i^2 - I_j^2 L_j^2))). \end{aligned} \tag{25}$$

Substituting (25) into (24) and simplifying terms, it can be shown that firm i earns an equilibrium profit of $\pi_i = 2ts_i^2$, where

$$s_1 = \frac{1}{2} + \frac{1}{6t}\left(\lambda^1 \phi^1 \left(I_1^1 - I_2^1\right) + \lambda^2 \phi^2 \left(I_1^2 - I_2^2\right) + \lambda^{12}\left(\phi^1 \left(I_1^1 L_1^1 - I_2^1 L_2^1\right) + \phi^2 \left(I_1^2 L_1^2 - I_2^2 L_2^2\right)\right)\right) \tag{26}$$

and $s_2 = 1 - s_1$.

Stage (ii).

In this stage we determine equilibria in the subgame as simultaneous solutions of each party's best response, conditional on the card fees and merchant fees set by the two card schemes in stage (i).

Merchants' best responses.

To work out merchants' optimal card acceptance policy we note that, regardless of what the other merchant does, each merchant will accept cards if doing so increases its equilibrium market share. Thus, each merchant will accept cards if doing so increases the function

$$\Phi = \lambda^1 \phi^1 I^1 + \lambda^2 \phi^2 I^2 + \lambda^{12}\left(\phi^1 I^1 L^1 + \phi^2 I^2 L^2\right), \tag{27}$$

where the subscript for each merchant has been dropped since the function is the same for both merchants. Note that here merchants' card acceptance policy is determined by Φ in the same way as it was determined by Ψ in Lemma 4. Thus, following the identical proof of Lemma 4 but replacing ϕ_S^i with ϕ^i, we end up with exactly the same results, except with ϕ_S^i replaced by ϕ^i. ■

As was the case for monopolistic merchants, if there are some multihoming consumers, multiple equilibria in the stage (ii) subgame are possible. The analysis of these multiple equilibria parallels that with monopolistic merchants where ϕ_S^i is replaced by ϕ^i. With competing merchants, we select the equilibria in the subgame that consumers and merchants would choose if they got together to choose one. As before, an equilibrium in the full game does not exist if the other equilibrium is selected in the subgame.

One interesting implication of Lemma 5 is that individual merchants will not necessarily prefer the card scheme with the lowest merchant fee. Individual merchants also care about the fees consumers will face from using cards (if cards are beneficial for merchants, they will want consumers to use them more), and the benefits consumers obtain from using cards (reflecting the fact that this allows merchants to attract more customers by accepting cards). Both aspects are combined in the function ϕ.

We start our analysis of stage (i) equilibria by considering the special case in which consumers never want to hold both cards. Like the case with monopolistic merchants, this results in a competitive bottleneck. Despite competition between schemes, each scheme sets its interchange fee at the same level as that set by a single monopolist scheme.

3.2.1 Consumers hold at most one card

Given Lemma 5 parallels the results of Lemma 4 with ϕ_S^i replaced by ϕ^i, in the case consumers get no intrinsic benefit from holding cards ($\overline{u} = 0$), Proposition 5 becomes:

Proposition 9 *If consumers get no intrinsic benefit from holding cards, the equilibrium interchange fee resulting from competition between identical card schemes facing competing merchants equals a^{Π}; that is, it solves $\phi(f(a)) = 0$. Merchants will (just) accept both cards and card-holding consumers will randomize over which card to hold. Each association shares in half the card transactions.*

Proof. The existence of a^{Π} was proven in Lemma 1. The next step is to prove that this is an equilibrium using our analysis of equilibria in the subgame starting at stage (ii). From the analysis of consumers' best responses at stage (ii) of the game, with $\overline{u} = 0$, $\lambda^{12} = A(0) = 0$. No consumers will hold both cards. Note that at a^{Π}, $\phi^1 = \phi^2 = 0$. The rest of the proof follows identically to the proof of Proposition 5, except with ϕ_S replaced by ϕ everywhere. ■

Again we have the case of a competitive bottleneck. Thus, despite competition between the schemes, their fee structure is unchanged from the case of a single scheme facing competing merchants. The welfare implications of this equilibrium are exactly as characterized by Proposition 4. The socially optimal interchange fee is either equal to a^{Π} or is lower.

3.2.2 Some consumers hold both cards

In this case we allow some consumers to obtain positive intrinsic benefits from holding cards (so $\overline{u} > 0$). As with monopolistic merchants, the implication for card scheme competition of this multihoming is dramatic.

Proposition 10 *When some consumers get intrinsic benefits from holding cards, the equilibrium interchange fee resulting from competition between identical card schemes facing competing merchants equals*

a^C, which maximizes ϕ^i, the expected joint surplus to consumers and merchants from using cards. Merchants accept both cards and singlehoming consumers randomize over which card to hold. The measure of multihoming consumers is

$$\lambda^M\left(a^C\right) = A(0),$$

the measure of consumers not holding cards is

$$\lambda^N\left(a^C\right) = 1 - A\left(-\phi_B\left(l - b_S\right)\right),$$

and the measure of singlehoming consumers is

$$\lambda^S\left(a^C\right) = A\left(-\phi_B\left(l - b_S\right)\right) - A(0).$$

Proof. Lemma 1 proved there exists a unique interchange fee a^C which maximizes ϕ^i. From the analysis of consumers' best responses at stage (ii) of the game, $\lambda^{12} = A(0) > 0$. Some consumers will hold both cards. We use this property and the analysis of equilibria in the stage (ii) subgame to show that a^C represents an equilibrium interchange fee. The rest of the proof follows identically to the proof of Proposition 6, except with ϕ_S replaced by ϕ everywhere. ■

In equilibrium, both schemes set their interchange fee at the level a^C, and merchants will accept both schemes' cards. The equilibrium involves merchants controlling which card is used by choosing which card to accept. Competition drives cards schemes to offer maximal profit to merchants, in an attempt to have their cards accepted exclusively, thus obtaining all card transactions. Since merchants compete amongst themselves, they take into account consumers' average surplus from using cards, so that competition between card schemes to attract merchants results in card schemes maximizing the expected joint surplus of end users (the expected joint net transactional benefits to consumers and merchants from using cards). In equilibrium each card scheme shares equally in the card transactions.

These interchange fees can be compared to the interchange fee that maximizes the schemes' joint profits — the interchange fee a^Π defined in equation (14).

Proposition 11 *When some consumers get intrinsic benefits from holding cards, the equilibrium interchange fee resulting from competition between identical card schemes facing competing merchants leads to an interchange fee lower than that which maximizes the schemes' joint profit (or joint card transactions).*

Proof. The result $a^C < a^\Pi$ was proven in Lemma 1. ■

Thus, competition between card schemes results in an equilibrium in which card schemes are jointly worse off, as there are fewer total card transactions compared to the case without competition. A more interesting comparison is whether the competitive interchange fee is higher or lower than the welfare maximizing interchange fee.

Proposition 12 *When some consumers get intrinsic benefits from holding cards, the equilibrium interchange fee resulting from competition between identical card schemes facing competing merchants leads to an interchange fee lower than that which maximizes overall welfare.*

Proof. See Appendix A.5. ■

Competing card schemes end up setting an inefficiently low interchange fee. They charge merchants too little, and cardholders too much. Despite the fact competing schemes maximize the expected surplus of end users, the competing card schemes set card fees too high from an efficiency perspective. They pass on their margins $\pi_A + \pi_I$ to cardholders, resulting in a distortion in the fees faced by consumers.

Table 1: Summary of results

	Single scheme	Competing Schemes		Social Planner	Relationships
		$\overline{u} \leq 0$	$\overline{u} > 0$		
Monopoly	$a^* = a^C$	$a_-^{**} = a^C$	$a_+^{**} = a^M$	$a^W = a^C$	$a_+^{**} < a^W = a^* = a_-^{**}$
Hotelling	$a^* = a^\Pi$	$a_-^{**} = a^\Pi$	$a_+^{**} = a^C$	a^W	$a_+^{**} < a^W \leq a^* = a_-^{**}$

Notes. a^* is the interchange fee which a single card scheme would choose, a_-^{**} is the equilibrium interchange fee resulting from competition between identical schemes when consumers hold at most one card, a_+^{**} is the equilibrium interchange fee resulting from competition between identical schemes when some consumers hold both cards, and a^W is the welfare-maximizing interchange fee.

When interchange fees are increased by the margins $\pi_A + \pi_I$, consumers will use cards if and only if the joint surplus exceeds the joint cost (efficient usage), while merchants will still accept cards (provided $\beta(f) - f \geq \pi_A + \pi_I$). Only in the limit as $\pi_A = \pi_I \to 0$ will equilibrium fees be set at the socially optimal level. In this case $a^C = b_S - c_A$ and $m = b_S$. This is the interchange fee that Baxter found was optimal — it ensures cards are only used when the joint surplus exceeds joint cost.

3.3 Summarizing results

Our results are summarized in Table 1. With both forms of merchant behavior considered, when consumers only hold one card, competition between card schemes does not result in lower interchange fees. In this case, by attracting cardholders, card schemes have a monopoly over access to these cardholders. This leads competing card schemes to care only about the surplus they can offer to cardholders, leaving no surplus to merchants (the case of a competitive bottleneck). This competitive bottleneck outcome can be undermined provided some consumers hold both cards. Then the unique equilibrium involves competing card schemes seeking to attract merchants exclusively by offering maximal incentive for merchants to accept their cards. In the case of a monopoly merchant, this implies maximizing the expected surplus offered to merchants. In the case of competing merchants, this involves maximizing the expected joint surplus of consumers and merchants, given that competing merchants take into account the benefit that their customers get from being able to use cards. In either case, interchange fees are lower as a result of competition between the card schemes.[12]

Interchange fees are also lower when there is a monopoly merchant rather than competing merchants. This is true both for the case without scheme competition (because $a^C < a^\Pi$) and for the case with scheme competition (because $a^M < a^C$). When merchants compete amongst themselves, card schemes will charge more to merchants and less to cardholders. Competition between merchants makes them less resistant to accepting cards, as card acceptance becomes a strategic device to attract customers. In this case, the business stealing effect can be responsible for inefficiently high merchant fees (and low card fees), while competition between schemes can be responsible for inefficiently low merchant fees (and high card fees). When the two effects are combined, so there is competition between schemes and between merchants, the equilibrium interchange fee turns out to be the same as the case without competition between schemes and without competition between merchants. The two types of competition exactly offset, and the resulting interchange fee is still inefficiently low, but only to the extent that issuers and

[12] Proposition 7 implies $a^M < a^C$ for the monopoly merchant case, and Proposition 11 implies $a^C < a^\Pi$ for the competing merchants case.

acquirers obtain positive margins.[13]

4 Extensions and implications

In Section 4.1 we consider what our analysis implies about competition between proprietary schemes such as American Express and Discover Card. Section 4.2 discusses some implications of merchant heterogeneity. Section 4.3 examines how the existence of cash-constrained consumers affects results. Some policy implications are discussed in Section 4.4, while Section 4.5 considers what our analysis suggests about competition in some other two-sided markets.

4.1 Competing proprietary schemes

In Section 3 we considered the case of competition between card associations, each of which sets an interchange fee to achieve its desired fee structure. It is also straightforward to determine what two competing proprietary schemes, which set the fees f and m directly, will do. Note that the profit of proprietary card scheme i is

$$\Pi^i = (f^i + m^i - c_A - c_I)T^i, \tag{28}$$

where T^i is the number of card transactions on system i.

In the cases we have considered, the results easily extend to two competing proprietary schemes. In each case, the analysis from stage (ii) onwards is the same as before since this assumed the fees f^i and m^i were taken as given. At stage (i) any equilibrium involving competition between identical proprietary schemes will involve the sum of their fees $f+m$ being driven down to cost c_A+c_I. Out of all fee structures that satisfy these constraints, only the structure of fees characterized in Section 3 will be equilibria. The equilibrium fees in the case of competing proprietary schemes then correspond to those implied by the interchange fees resulting from Section 3 in the limit as $\pi_A = \pi_I \to 0$. We therefore get

Proposition 13 *If consumers get no intrinsic benefit from holding cards (consumers singlehome), the equilibrium fees resulting from competition between identical proprietary schemes facing monopolistic merchants are characterized by $f^* = c_A + c_I - b_S$ and $m^* = b_S$.*

Proposition 14 *If some consumers get intrinsic benefits from holding cards (consumers multihome), the equilibrium fees resulting from competition between identical proprietary schemes facing monopolistic merchants are characterized by*

$$h(f^*)(f^* + b_S - c_A - c_I) = D(f^*)$$

and $m^* = c_A + c_I - f^*$.

Proposition 15 *If consumers get no intrinsic benefit from holding cards (consumers singlehome), the equilibrium fees resulting from competition between identical proprietary schemes facing competing merchants are characterized by $\beta(f) = c_A + c_I - b_S$ and $m^* = c_A + c_I - f^*$.*

Proposition 16 *If some consumers get intrinsic benefits from holding cards (consumers multihome), the equilibrium fees resulting from competition between identical proprietary schemes facing competing merchants are characterized by $f^* = c_A + c_I - b_S$ and $m^* = b_S$.*

[13] While this exact offset is likely to be a specific feature of our model, the presence of the two offsetting effects should be more general.

The first and fourth case correspond exactly to the socially optimal set of fees. In the second case, card fees are set too high and merchant fees too low. In the third case, card fees are set too low and merchant fees too high. These results are no different from the case of competing identical card associations in the limit as intra-platform competition drives issuing and acquiring margins to zero.

4.2 Merchant heterogeneity

In this section we follow Wright (2003a) and consider the case in which there are many industries, each of which has a different value of b_S (in some industries, being able to accept cards is more useful than in others). The random variable b_S is drawn with a positive density $g(b_S)$ over the interval $[\underline{b}_S, \overline{b}_S]$, and a cumulative distribution denoted G. Apart from this form of heterogeneity, industries are all alike. The particular draws of b_S are assumed to be unobserved by the card schemes, and are assumed to be independent of all the other draws. Consumers are exogenously matched to all the different industries, and so without loss of generality they buy one good from each industry. The merchants' (sellers') "quasi-demand" function which measures the proportion of merchants with transactional benefits above some level b_S is denoted $S(b_S) = 1 - G(b_S)$. The timing is the same as in the benchmark model.

We start by considering the case with a single card scheme and a monopoly merchant in each industry.[14] From Lemma 2, a monopolist merchant will reject the card if $\phi_S(f) < 0$ and will accept a card if $\phi_S(f) \geq 0$. Recall $\phi_S(f) = D(f)\,(f + b_S - l)$, so that merchants with $b_S \geq l - f$ will accept cards and others will not. There are therefore $S\,(l - f) = 1 - G\,(l - f)$ merchants that accept cards.[15] A consumer's benefit from holding the card is then $\phi_B(f)S\,(l - f) + u$, since a consumer gets an expected surplus of $\phi_B(f)$ from using their card at each merchant, and there are $S\,(l - f)$ merchants that accept cards. The term u is the consumer's intrinsic benefit of holding a card. It follows that consumers will hold a card if $u > -\phi_B(f)S\,(l - f)$, and not otherwise.

The card scheme's profit (that is, the total profit of the association's member banks) is then equal to

$$\Pi(f) = (\pi_A + \pi_I)\, A\left(-\phi_B(f)S\,(l - f)\right) D(f)S\,(l - f)\,,$$

since there are $A\left(-\phi_B(f)S\,(l - f)\right)$ consumers holding a card and each of them will use their card a proportion $D(f)$ of the time at a measure $S\,(l - f)$ of merchants. Then a single card scheme sets an interchange fee (or equivalently a card fee) to maximize $\Pi(f)$. This fee does not afford any useful interpretation.

As the above case demonstrates, it is difficult to obtain general results on the equilibrium interchange fee under merchant heterogeneity. This difficulty is heightened when there is also competition between payment schemes. One case which does afford a general interpretation under scheme competition is that in which all consumers hold at most one card. In this case we continue to get the competitive bottleneck outcome found without merchant heterogeneity. Schemes will compete by seeking to attract cardholders exclusively, and access to these cardholders becomes the bottleneck for merchants. However, unlike the case with homogenous merchants, the competitive bottleneck now results in competing schemes setting interchange fees that are too high, even from their own perspective.

We first characterize the equilibrium at stage (ii) in the game.

[14] In each case below, we note how the analysis changes when merchants compete in a Hotelling fashion.

[15] For the case with competing merchants in each industry, the condition for merchant acceptance $\phi_S(f) \geq 0$ is replaced by the condition $\phi(f) \geq 0$. This implies f is replaced by $\beta(f)$ in the function G below, and the function S below will change as a result. Other than this one modification, the analysis is unchanged in what follows.

Lemma 6 *Suppose two identical card schemes compete, with card schemes* 1 *and* 2 *having card fees* f^1 *and* f^2 *respectively. Suppose consumers can hold at most one card. Then* $S^i = 1 - G\left(l - f^i\right)$ *monopolistic merchants accept card* i*, and consumers hold card* i *if* $\phi_B^i S^i > \phi_B^j S^j$ *and* $u \geq -\phi_B^i S^i$*, will randomize over which card to hold if* $\phi_B^i S^i = \phi_B^j S^j$ *and* $u \geq -\phi_B^i S^i$*, and will not hold card* i *otherwise.*

Proof. Working back to stage (ii), a merchant's profit from accepting cards is identical to the case of a single merchant, for a given number of cardholders and given fees (that is, the analysis is identical to Lemma 4). Given no consumers hold both cards, merchants adopt the following policy: the merchant rejects both cards if $\phi_S^1, \phi_S^2 < 0$, it accepts both cards if $\phi_S^1, \phi_S^2 \geq 0$, and it accepts only card 1 (respectively, card 2) if $\phi_S^1 \geq 0 > \phi_S^2$ (respectively, $\phi_S^2 \geq 0 > \phi_S^1$). This implies merchants with $b_S \geq l - f^i$ will accept card i, and others will not accept card i, so that there are $S^i = 1 - G\left(l - f^i\right)$ merchants that accept card i. A consumer's additional benefit of holding just card i is

$$\phi_B^i S^i + u.$$

As a result, consumers will all hold the card with highest value of $\phi_B^i S^i$ provided $u \geq -\phi_B^i S^i$, will randomize over which card to hold if $\phi_B^i S^i = \phi_B^j S^j$ and $u \geq -\phi_B^i S^i$, and will not hold card i otherwise. ■

When consumers hold at most one card, competing schemes will seek to have their card held exclusively by consumers. This gives them a bottleneck over access to these consumers by merchants. If merchants want their customers to be able to use cards, then from the perspective of the merchant each card scheme (with its exclusive set of cardholders) is like a monopoly supplier. Then if both schemes set the same interchange fee (so $\phi_S \equiv \phi_S^1 = \phi_S^2$), the merchants accept both cards if $\phi_S \geq 0$ and neither otherwise. When the merchant accepts both cards, card-holding consumers randomize over which card to hold, and the members of such card schemes get (in aggregate) profits of

$$\Pi^1 = \Pi^2 = (\pi_I + \pi_A)\frac{A(-\phi_B(f)S\left(l-f\right))D(f)S\left(l-f\right)}{2}.$$

If card schemes act to maximize their joint profits, they will set their interchange fees so as to maximize $A(-\phi_B(f)S(l-f))D(f)S(l-f)$. This implies the same interchange fee as we found for the case with a single scheme. The following proposition characterizes the equilibrium interchange fees arising from competition between schemes.

Proposition 17 *If consumers hold at most one card, the equilibrium interchange fee resulting from competition between identical card schemes facing monopolistic merchants is that which maximizes consumers' expected surplus from holding a card* $\phi_B(f)S\left(l-f\right)$*. At these interchange fees, a merchant that accepts one card will accept both cards and card-holding consumers will randomize over which card to hold. Each association shares in half the card transactions.*

Proof. If any scheme sets a lower or higher card fee f, it will no longer get any cardholders and therefore any card transactions. This equilibrium is unique since if any scheme i sets a fee structure such that $\phi_B(f)S\left(l-f\right)$ is not maximized, then a single scheme can always adjust f to offer more to consumers from holding its card, which will result in its card being held exclusively. ■

The expected surplus from holding a card is the product of $\phi_B(f)$, the expected surplus a cardholder obtains from being able to use the card for a purchase, and $S(l-f)$, the proportion of merchants that accept cards. The expected surplus a cardholder obtains from being able to use their card for a purchase can be further broken down into the product of $\beta(f) - f$, the surplus they obtain from using a card for a purchase, and $D(f)$, the proportion of times the consumer would want to use a card for a purchase. Thus, the expected surplus which the card schemes will attempt to maximize can be expressed as the product of the surplus from using a card for a purchase and $D(f)S(l-f)$, the number of transactions using cards. Since a single scheme (or two schemes acting jointly) will maximize member profits by maximizing the number of transactions using cards, competing schemes will set a higher interchange fee (and lower card fee) if doing so raises the surplus to consumers from using cards for a purchase. We thus get that[16]

Proposition 18 *If $\beta(f) - f$ is decreasing in f, then the equilibrium interchange fee resulting from competition between identical card schemes when consumers hold at most one card cannot be below the joint profit maximizing interchange fee.*

Proof. Suppose both schemes set interchange fees at a^*, the equilibrium interchange fee resulting from competition between identical schemes, which maximizes $\phi_B(f)S(l-f)$. This interchange fee also maximizes $A(-\phi_B(f)S(l-f))\,\phi_B(f)S(l-f)$, since A is a monotonically decreasing function. This can be rewritten as $\theta(f) = (\beta(f) - f)\,T(f)$, where $T(f) = A(-(\beta(f) - f)\,D(f)S(l-f))\,D(f)S(l-f)$ is proportional to the joint profit of the two schemes. It follows that

$$\theta(f) \leq \theta(f^*) \quad \forall\, f \leq f^*.$$

Since $\beta(f) - f$ is decreasing in f for all f, then

$$\beta(f) - f \geq \beta(f^*) - f^* \quad \forall\, f \leq f^*.$$

Since $\theta(f)$ and $\beta(f) - f$ are both nonnegative for all f, it follows that

$$T(f) = \frac{\theta(f)}{\beta(f) - f} \leq \frac{\theta(f^*)}{\beta(f^*) - f^*} = T(f^*) \quad \forall\, f \leq f^*.$$

Therefore $\arg\max_f T(f) \geq f^*$ and the joint profit maximizing interchange fee cannot be above the equilibrium interchange fee a^*. ■

The intuition behind the result is simple. Holding constant the fraction of cardholders on each scheme, card schemes will maximize their respective profits by acting as though they do not face any competition. The market is effectively segmented. At this point, a small change in a scheme's interchange fee will have only a second order impact on the number of card transactions. However, a small increase in a scheme's interchange fee will have a first order impact on the average surplus consumers get from using cards. If the average surplus to those using cards increases when the interchange fee is increased above a^*, each card scheme will set interchange fees too high in an attempt to get consumers to switch to holding their card exclusively, an effect which ends up reducing the total number of card transactions and their members' profits.[17]

We can contrast the above case in which consumers hold at most one card to the case in which all consumers (that is, measure 1) hold both cards (which would arise if $u > 0$ for all consumers). We continue to assume all merchants are monopolists. Given consumers hold both cards, we only need to solve at stage (ii) for merchants' optimal decision for card acceptance. From Lemma 4, it follows that

Lemma 7 *Suppose two identical card schemes compete, with card schemes* 1 *and* 2 *having card fees f^1 and f^2 respectively. Suppose consumers hold both cards.*

[16] An earlier version of this paper (Guthrie and Wright, 2003) demonstrates the same result holds for the case in which merchants compete in a Hotelling fashion.

[17] The condition in Proposition 18 is satisfied if the consumer quasi-demand for card use is linear (b_B follows the uniform distribution). Then $\beta(f) - f = (\bar{b}_B - f)/2$, so that $\beta(f) - f$ is decreasing in f.

1. *If $f^1 = f^2$ then $\phi_S^1 = \phi_S^2$ and there are two cases to consider. If $\phi_S^1 = \phi_S^2 \geq 0$, then monopolistic merchants accept both cards. If $\phi_S^1 = \phi_S^2 < 0$, then monopolistic merchants reject both cards.*

2. *If $f^i < f^j$ then four equilibria are possible at stage (ii). If $\phi_S^i, \phi_S^j < 0$, then monopolistic merchants reject both cards. If $\phi_S^i, \phi_S^j \geq 0$ and $\phi_S^i \geq \phi_S^j$, then monopolistic merchants accept both cards (but only card i gets used). If $\phi_S^i \geq 0$, $\phi_S^j < 0$, then monopolistic merchants only accept card i. If $\phi_S^j \geq 0$ and $\phi_S^i < \phi_S^j$, then monopolistic merchants only accept card j.*

Proof. Given $u > 0$ for all consumers, all consumers will hold both cards to obtain the instrinsic benefits of the cards (as well as to obtain any usage benefits). Working back to stage (ii), a merchant's profit from accepting cards is identical to the case of a single merchant, for a given number of cardholders and given fees. The results then follow directly from applying Lemma 4 when all consumers hold both cards. ■

Turning now to competition between schemes at stage (i), identical card schemes will compete by setting their interchange fee to attract merchants. We consider two cases. With equal interchange fees, consumers randomize over which card to use, and each card scheme receives half the card transactions. Lemma 7 implies scheme profits are

$$\Pi^1 = \Pi^2 = (\pi_I + \pi_A)\frac{D(f)S\left(l - f\right)}{2}.$$

In the second case, scheme 1 sets a lower interchange fee.[18] Then $f^1 > f^2$, and $D(f^1) < D(f^2)$. Define

$$b(f^1, f^2) = \frac{D(f^2)f^2 - D(f^1)f^1}{D(f^2) - D(f^1)}. \tag{29}$$

For industries with $b_S > l - b(f^1, f^2)$, merchants will accept cards from both schemes (but only card 2 will be used). Only merchants with high transactional benefits of accepting cards will be willing to accept the more expensive card knowing consumers will always use it. Merchants with b_S between $l - f^1$ and $l - b(f^1, f^2)$ will only accept cards from scheme 1, the cheaper card to accept. Merchants with yet lower transactional benefits of accepting cards will not accept either card. Scheme 1's profit is thus

$$\Pi^1 = (\pi_I + \pi_A)D(f^1)\left(S\left(l - f^1\right) - S\left(l - b(f^1, f^2)\right)\right),$$

while scheme 2's profit is

$$\Pi^2 = (\pi_I + \pi_A)D(f^2)S\left(l - b(f^1, f^2)\right).$$

A symmetric equilibrium in which both schemes set the same interchange fee a^* (and therefore the same card fee f^*) requires that

$$\frac{D(f^*)S\left(l - f^*\right)}{2} \geq D(f^1)\left(S\left(l - f^1\right) - S\left(l - b(f^1, f^*)\right)\right)$$

for any fee $f^1 > f^*$, and

$$\frac{D(f^*)S\left(l - f^*\right)}{2} \geq D(f^1)S\left(l - b(f^1, f^*)\right)$$

for any fee $f^1 < f^*$. The nature of the equilibrium, if any, depends on the specific distributions on b_B and b_S. A case in which there is an explicit solution is the case in which b_B and b_S are distributed uniformly (so that quasi-demand functions are linear).

Proposition 19 *If consumers hold both cards and quasi-demand functions are linear, the equilibrium interchange fee resulting from competition between identical card schemes facing monopolistic merchants is for both schemes to set $a_U^C = \frac{1}{3}\left(\overline{b}_S - c_A - \pi_A\right) - \frac{2}{3}\left(\overline{b}_B - c_I - \pi_I\right)$.*

[18] The case in which it sets a higher interchange fee follows by symmetry.

Proof. A symmetric equilibrium in which both schemes set the same interchange fee a^* requires that

$$\frac{D(f^*)S\left(l-f^*\right)}{2} \geq D(f^1)\left(S\left(l-f^1\right)-S\left(l-b(f^1,f^*)\right)\right) \tag{30}$$

for any lower interchange fee (so $f^1 > f^*$) and

$$\frac{D(f^*)S\left(l-f^*\right)}{2} \geq D(f^1)S\left(l-b(f^1,f^*)\right) \tag{31}$$

for any higher interchange fee (so $f^1 < f^*$).

With the uniform distribution on b_B and b_S, profits at the proposed equilibrium are

$$\frac{D\left(f^*\right)}{2}\left(\frac{\overline{b}_S-\left(l-f^*\right)}{\overline{b}_S-\underline{b}_S}\right).$$

Now consider the limit of scheme 1's profits as its interchange fee approaches the proposed symmetric equilibrium at a^* from below. This implies $f^1 > f^2 = f^*$, so that

$$D(f^1)\left(S\left(l-f^1\right)-S\left(l-b(f^1,f^*)\right)\right) \rightarrow D(f^*)\left(\frac{\overline{b}_B-f^*}{\overline{b}_S-\underline{b}_S}\right).$$

Thus, (30) will be violated for a sufficiently small decrease in scheme 1's interchange fee (increase in card fee) if

$$f^* < \frac{1}{3}\left(2\overline{b}_B-\overline{b}_S+l\right).$$

Alternatively, consider the limit of scheme 1's profit as its interchange fee approaches the proposed symmetric equilibrium at a^* from above. This implies $f^1 < f^2 = f^*$, so that

$$D(f^1)S\left(l-b(f^1,f^*)\right) \rightarrow D(f^*)\left(\frac{\overline{b}_S-l+2f^*-\overline{b}_B}{\overline{b}_S-\underline{b}_S}\right).$$

Thus, (31) will be violated for a sufficiently small increase in scheme 1's interchange fee if

$$f^* > \frac{1}{3}\left(2\overline{b}_B-\overline{b}_S+l\right).$$

Thus, the only potential candidate for a symmetric equilibrium is when

$$f^* = \frac{1}{3}\left(2\overline{b}_B-\overline{b}_S+l\right),$$

which corresponds to the interchange fee

$$a_U^C = \frac{\left(\overline{b}_S-c_A-\pi_A\right)-2\left(\overline{b}_B-c_I-\pi_I\right)}{3}. \tag{32}$$

To confirm this is an equilibrium, note that if scheme 1 sets a lower interchange fee, then its card transactions

$$D(f^1)\left(S\left(l-f^1\right)-S\left(l-b(f^1,f^*)\right)\right)$$

will decrease since $D(f^1)$ will decrease and $S\left(l-f^1\right)-S\left(l-b(f^1,f^*)\right)$ is independent of f^1. Alternatively, if scheme 1 sets a higher interchange fee (so that $f^1 < f^*$), then its card transactions

$$D(f^1)S\left(l-b(f^1,f^*)\right)$$

will decrease as

$$\begin{aligned}\frac{d\left(D(f^1)S\left(l-b(f^1,f^*)\right)\right)}{da^1} &= \frac{\overline{b}_S-l+2f^1+f^*-2\overline{b}_B}{\left(\overline{b}_B-\underline{b}_B\right)\left(\overline{b}_S-\underline{b}_S\right)} \\ &< \frac{\overline{b}_S-l+3f^*-2\overline{b}_B}{\left(\overline{b}_B-\underline{b}_B\right)\left(\overline{b}_S-\underline{b}_S\right)} \\ &= 0.\end{aligned}$$

To see why there are no asymmetric equilibria, note that if scheme 1 sets a lower interchange fee than scheme 2, then, as shown above, its card transactions

$$D(f^1)\left(S\left(l-f^1\right)-S\left(l-b(f^1,f^2)\right)\right)$$

will decrease since $D(f^1)$ will decrease and $S\left(l-f^1\right)-S\left(l-b(f^1,f^2)\right)$ is independent of f^1. This is true regardless of firm 2's interchange fee. Thus, scheme 1 will always want to set its interchange fee as close as possible to scheme 2's interchange fee (that is, an infinitesimal amount less than a^2). Clearly there can be no asymmetric equilibrium. ■

The equilibrium above can be compared to the joint profit maximizing and socially optimal interchange fees. If the schemes maximize their joint profits given all consumers hold both cards, they will set a common interchange fee to maximize

$$\left(\pi_I+\pi_A\right)D(f)S\left(l-f\right),$$

while a social planner will set the interchange fee to maximize

$$\left(\beta_B(f)+\beta_S\left(l-f\right)-c_A-c_I\right)D(f)S\left(l-f\right).$$

Given the assumption of a linear quasi-demand functions, $\beta_B(f)+\beta_S\left(l-f\right)-c_A-c_I$ is a constant, and the profit maximizing and socially optimal interchange fee coincide. (Schmalensee, 2002 and Wright, 2003a obtain the same result in somewhat different settings.) The socially optimal interchange fee is

$$a_U^W=\frac{\left(\overline{b}_S-c_A-\pi_A\right)-\left(\overline{b}_B-c_I-\pi_I\right)}{2}.$$

Comparing this to the equilibrium interchange fee resulting from competition between schemes, it follows using (3) that:

Proposition 20 *If consumers hold both cards and quasi-demand functions are linear, the equilibrium interchange fee resulting from competition between identical card schemes facing monopolistic merchants leads to an interchange fee lower than that which maximizes the schemes' joint profit, and than that which maximizes welfare.*

Proof. The result follows since

$$\frac{\left(\overline{b}_S-c_A-\pi_A\right)-2\left(\overline{b}_B-c_I-\pi_I\right)}{3}-\frac{\left(\overline{b}_S-c_A-\pi_A\right)-\left(\overline{b}_B-c_I-\pi_I\right)}{2}=\frac{l-\overline{b}_B-\overline{b}_S}{6}<0$$

given (3). ■

Allowing for merchant heterogeneity does not change the type of results obtained earlier with a homogenous merchant. When consumers already hold both cards, competing schemes still set interchange fees too low, as they attempt to get merchants to accept their cards exclusively.

In an earlier version of the paper (see Guthrie and Wright, 2003) we showed a similar result holds for the case in which there are two merchants in each industry (with the same b_S) which compete in a Hotelling fashion. Given the linear quasi-demand and quasi-supply specification, the equilibrium interchange fee was found to equal

$$\frac{2\left(\overline{b}_S-c_A-\pi_A\right)-\left(\overline{b}_B-c_I-\pi_I\right)}{3}.$$

Competition between schemes again resulted in lower interchange fees, and in fact interchange fees that are less than the privately optimal and socially optimal interchange fee (see Propositions 12 and 13 in

Guthrie and Wright, 2003).[19] Thus, while the analysis is made more complicated by the introduction of unobservable merchant heterogeneity, the qualitative results obtained for homogenous merchants broadly carry over to this case.

4.3 Cash-constrained consumers

Another reason why merchants may accept credit cards is to avoid losing sales to cash-constrained consumers. In this section we show how this possibility results in higher equilibrium interchange fees, and so higher merchant fees and lower card fees. We do so in the context of monopolistic merchants. For such merchants, the existence of cash-constrained consumers raises the benefits they obtain from accepting cards. On the other hand, since cash-constrained consumers will have their surplus v extracted from monopolistic merchants, they do not have a greater incentive to hold or use cards. To help rebalance the incentives faced by merchants versus cardholders, a higher interchange fee is both privately and socially optimal. This is true regardless of whether card schemes compete or not.

Assume a proportion q of consumers have insufficient cash to purchase goods. If they hold a card and the merchant they wish to buy from accepts this card, then such consumers will still be able to purchase. We assume that $q < 1/(1+B)$, where $-B$ is the minimum value of the function $D(x)+(x+v-d+b_S-l)D'(x)$.[20] The next section considers the case of a single card scheme, while Section 4.3.2 examines the case of two competing schemes.

4.3.1 Single card scheme

The four stages of the game are unchanged, but now a proportion q of consumers have insufficient cash to purchase at stage (iv); their only method of purchasing the good is using a card. This raises the possibility that by setting a price $p < v$, a monopolist merchant might be able to increase demand for its product from customers who are anyway cash constrained. However, in Appendix B we use the upper bound on q to show that monopolistic merchants will still want to price to extract all the surplus from cash consumers, setting a price equal to v. Any higher price will exclude too many cash consumers, while any lower price will not fully exploit cash consumers. When the merchant sets $p = v$, its profit equals

$$\pi = (1-q)\left(v-d+\lambda(f)D(f)\left(b_S-m\right)I\right)+q\lambda(f)D(f)\left(v-d+b_S-m\right)I,$$

which can be written as

$$\pi = (1-q)\left(v-d\right)+\lambda(f)D(f)\left(b_S+q\left(v-d\right)-m\right)I.$$

The merchant will therefore accept cards if

$$\lambda(f)D(f)(b_S+q(v-d)-m) \geq 0.$$

Equivalently, the merchant will accept cards if $\hat{\phi}_S(f) \geq 0$, where $\hat{\phi}_S(f) = D(f)(\hat{b}_S - m)$ and $\hat{b}_S = b_S + q(v-d)$. Much of the analysis in Section 2.1 carries through with b_S replaced by $\hat{b}_S$ everywhere. In particular, because the merchant's choice of $p = v$ implies that consumers obtain no surplus from the good itself, their benefits of holding cards do not depend on whether they are cash-constrained or not. Our analysis begins with the analog to Lemma 2.

[19] Note that consistent with the case with homogenous merchants, the equilibrium interchange fee is higher with competing merchants compared to with monopolistic merchants, and is lower with competing schemes compared to with a single scheme.

[20] Assuming that the density function h is continuous on $[\underline{b}_B, \overline{b}_B]$ ensures that the function $D(x)+(x+v-d+b_S-l)D'(x)$ is bounded on $[\underline{b}_B, \overline{b}_B]$. As we showed in the proof of Lemma 8, the lower bound we impose on $v-d$ in Section 2 ensures that the function, and hence its lower bound, is negative.

Lemma 8 *Suppose some consumers are cash-constrained and a single card scheme has a card fee of f. If $\hat{\phi}_S(f) < 0$, then monopolistic merchants reject the card and $\lambda(f) = A(0)$ consumers hold the card. If $\hat{\phi}_S(f) \geq 0$, then monopolistic merchants accept the card, and $\lambda(f) = A(-\phi_B(f))$ consumers hold the card.*

Proof. We solve the game by working backwards through the stages. The stages are described in turn.

Stage (iv).

If the merchant sets a price $p \leq v$ and accepts the card, cardholding consumers will use the card if (i) they are cash-constrained and $b_B \geq f + p - v$, or (ii) they are not cash-constrained and $b_B \geq f$. If the merchant sets a price $p > v$ and accepts the card, cardholding consumers will use the card if $b_B \geq f + p - v$. In any other circumstance, consumers will not use a card.

Stage (iii).

Provided it sets its price no higher than v, a merchant obtains profit of

$$\pi = (1 - q + q\lambda(f)D(f + p - v)I)(p - d) + ((1 - q)D(f) + qD(f + p - v))\lambda(f)(b_S - m)I$$

Alternatively, a merchant can set $p > v$, in which case it will only sell to cardholders who use cards, so that

$$\pi = \lambda(f)D(f + p - v)(p + b_S - d - m)I.$$

As we prove in Lemma 12, it will not be profitable for a merchant to set a price different from v. Instead, a monopolist will extract all the surplus from the cash customers by setting $p = v$, implying it earns a profit of

$$\pi = (1 - q)(v - d) + \lambda(f)D(f)(b_S + q(v - d) - m)I = (1 - q)(v - d) + \lambda(f)\hat{\phi}_S(f)I.$$

Stage (ii).

Given a merchant obtains a price of v regardless of whether it accepts cards or not, it will accept cards if and only if $\hat{\phi}_S(f) \geq 0$. Since consumers gain benefit of

$$D(f)(\beta(f) - f)I + u$$

from holding a card, whether or not they are cash-constrained, a consumer's benefit from holding the card is $\phi_B(f)I + u$. It follows that consumers will hold a card if $u > 0$ and $\hat{\phi}_S(f) < 0$ (using it only to withdraw cash and the like). They will also hold a card if $u > -\phi_B(f)$ and $\hat{\phi}_S(f) \geq 0$. If neither condition applies, consumers will not hold a card. The equilibria in stage (ii) are thus those characterized by the lemma. ■

The next result, which is the analog of Proposition 1, shows that the scheme sets an interchange fee to drive the merchants' surplus to zero. This involves the interchange fee being set at $a^C + q(v - d)$, where a^C is the corresponding interchange fee without cash-constrained consumers.

Proposition 21 *If some consumers are cash-constrained, a single card scheme facing monopolistic merchants sets its interchange fee to solve $\hat{\phi}_S(f) = 0$; that is, at $a = \hat{a}^C \equiv a^C + q(v - d)$.*

Proof. A single card association maximizes $\Pi(f)$ by choosing f to maximize $A(-\phi_B(f))D(f)$ subject to the constraint $\hat{\phi}_S(f) \geq 0$ which ensures merchants accept cards. The constraint is equivalent to $f + \hat{b}_S \geq l$. Since $A(-\phi_B(f))D(f)$ is decreasing in f and the left hand side of the constraint is increasing in f, this implies the scheme will wish to set f as low as possible subject to the constraint. The constraint will be binding and the profit maximizing interchange fee solves $\hat{\phi}_S(f) = 0$, which is precisely the interchange fee $\hat{a}^C$ defined in the statement of the proposition. ■

This is also the welfare maximizing interchange fee.

Proposition 22 *If some consumers are cash-constrained, the welfare-maximizing interchange fee for a single card scheme facing monopolistic merchants is $a = \hat{a}^C$.*

Proof. See Appendix A.6. ■

The higher interchange fee is socially desirable since it signals to cash-constrained consumers that they should hold (and use) cards more often given the surplus their purchases create for merchants. Note the higher interchange fee resulting from the existence of cash-constrained consumers results in some excessive usage of cards (by those who are not cash constrained) but this is necessary to help offset the insufficient holding and usage of cards by those consumers who are cash-constrained.

4.3.2 Competing card schemes

Higher equilibrium interchange fees also result from the existence of cash-constrained consumers when there are competing schemes. As we prove in Appendix B, provided the proportion of cash-constrained consumers is not too high, monopolistic merchants will again set $p = v$. When $p = v$, the merchant's profit equals

$$\begin{aligned}\pi &= q\Big(\left(\lambda^1 + \lambda^{12}L^1\right) D(f^1)(v + b_S - d - m^1)I^1 + \left(\lambda^2 + \lambda^{12}L^2\right) D(f^2)(v + b_S - d - m^2)I^2 \Big) \\ &\quad + (1-q)\Big(v - d + \left(\lambda^1 + \lambda^{12}L^1\right) D(f^1)(b_S - m^1)I^1 + \left(\lambda^2 + \lambda^{12}L^2\right) D(f^2)(b_S - m^2)I^2 \Big),\end{aligned}$$

which can be rewritten as

$$\pi = q\,(v-d) + \hat{\Psi},$$

where

$$\hat{\Psi} = \lambda^1 \hat{\phi}_S^1 I^1 + \lambda^2 \hat{\phi}_S^2 I^2 + \lambda^{12}(\hat{\phi}_S^1 I^1 L^1 + \hat{\phi}_S^2 I^2 L^2)$$

and

$$\hat{\phi}_S^i = D(f^i)(b_S + q(v-d) - m^i) = D(f^i)(\hat{b}_S - m^i).$$

In particular, merchants' resistance to accepting cards is lowered as a result of the benefit cards provide in capturing the surplus of cash-constrained consumers. Other than the fact that b_S is replaced by $\hat{b}_S$, and ϕ_S^i is replaced by $\hat{\phi}_S^i$, the analysis of Section 3.1 still applies. As before, with all their surplus being extracted from the purchase of goods, cash-constrained consumers face the same decision about whether to hold a card or not as other consumers. The following analog of Lemma 4 applies.

Lemma 9 *Suppose some consumers are cash-constrained and two identical card schemes compete, with card schemes 1 and 2 having card fees f^1 and f^2 respectively.*

1. *If $f^1 = f^2$ then $\hat{\phi}_S^1 = \hat{\phi}_S^2$ and there are two cases to consider. If $\hat{\phi}_S^1 = \hat{\phi}_S^2 \geq 0$, then monopolistic merchants accept both cards and the singlehoming consumers will randomize over which card to hold. If $\hat{\phi}_S^1 = \hat{\phi}_S^2 < 0$, then monopolistic merchants reject both cards and there are no singlehoming consumers.*

2. *If $f^i < f^j$ then four equilibria are possible at stage (ii). If $\hat{\phi}_S^i, \hat{\phi}_S^j < 0$, then monopolistic merchants reject both cards and there are no singlehoming consumers. If $\hat{\phi}_S^i, \hat{\phi}_S^j \geq 0$ and $(\lambda^i + \lambda^{12})\hat{\phi}_S^i \geq \lambda^{12}\hat{\phi}_S^j$,*

then monopolistic merchants accept both cards and the singlehoming consumers will only hold card i. If $\hat{\phi}_S^i \geq 0$, $\hat{\phi}_S^j < 0$, then monopolistic merchants only accept card i and the singlehoming consumers will only hold card i. If $\hat{\phi}_S^j \geq 0$ and $\hat{\phi}_S^i < 0$ then monopolistic merchants only accept card j and the singlehoming consumers will only hold card j. (If $\overline{u} > 0$, then this equilibrium occurs more generally if $\hat{\phi}_S^j \geq 0$ and $\hat{\phi}_S^i < \hat{\phi}_S^j$.)

Proof. We solve the game by working backwards through time.

Stage (iv).

Suppose the merchant sets price $p \leq v$. Consumers who only hold card i and are cash-constrained will use the card if $b_B \geq f + p - v$ and the merchant they buy from accepts the card; if they are not cash-constrained, they will use the card if $b_B \geq f$ and the merchant they buy from accepts the card. The card use of cash-constrained consumers who hold both cards and draw $b_B \geq \min\{f^1 + p - v, f^2 + p - v\}$ is described by equations (18), (19) and (20); if they are not cash-constrained the corresponding condition is that $b_B \geq \min\{f^1, f^2\}$.

Alternatively, if merchants set their price above v, then if consumers only hold card i, they will only use the card if $b_B + v \geq f^i + p$ and the merchant they buy from accepts the card. The card use of consumers who hold both cards and draw $b_B \geq \min\{f^1, f^2\} + p - v$ is described by equations (18), (19) and (20).

Stage (iii).

As we show in Appendix B, merchants will extract all the surplus from the cash customers by setting $p = v$, implying merchants earn a profit of

$$\pi = q(v - d) + \hat{\Psi}. \tag{33}$$

Stage (ii).

In this stage we determine equilibria in the subgame as simultaneous solutions of each party's best response, conditional on the card fees and merchant fees set by the two card schemes in stage (i).

The merchant's best response.

To work out a merchant's optimal card acceptance policy we note that a merchant will accept cards if doing so increases the function $\hat{\Psi}$.

We must consider two possibilities for consumers' card-holding. In the first possibility we consider, no consumers multihome, so that $\lambda^{12} = 0$ and the function $\hat{\Psi}$ is determined by the following table:

	$I^2 = 0$	$I^2 = 1$
$I^1 = 0$	0	$\lambda^2 \hat{\phi}_S^2$
$I^1 = 1$	$\lambda^1 \hat{\phi}_S^1$	$\lambda^1 \hat{\phi}_S^1 + \lambda^2 \hat{\phi}_S^2$

Recall that if a merchant is indifferent between accepting and rejecting card i because it does not expect consumers to use card i (so that accepting the card leaves the function $\hat{\Psi}$ unchanged), it will accept the card if doing so increases $\hat{\Psi}$ when consumers do use card i. This is true if and only if $\hat{\phi}_S^i \geq 0$. Merchants therefore adopt the following policy: merchants reject both cards if $\hat{\phi}_S^1, \hat{\phi}_S^2 < 0$, accept both cards if $\hat{\phi}_S^1, \hat{\phi}_S^2 \geq 0$, and accept only card 1 (respectively, card 2) if $\hat{\phi}_S^1 \geq 0 > \hat{\phi}_S^2$ (respectively, $\hat{\phi}_S^2 \geq 0 > \hat{\phi}_S^1$).

The second possibility is that some consumers hold both cards, so that $\lambda^{12} > 0$. Since L^1 and L^2 in equation (22) depend on the values of f^1 and f^2, we need to consider three different cases.

- If $f^1 = f^2$ then $\hat{\phi}_S = \hat{\phi}_S^1 = \hat{\phi}_S^2$ and equation (20) implies that

$$\hat{\Psi} = \hat{\phi}_S \left(\lambda^1 I^1 + \lambda^2 I^2 + \lambda^{12} \left(I^1 + I^2 - I^1 I^2 \right) \right).$$

In this case the function $\hat{\Psi}$ is determined by the following table:

	$I^2 = 0$	$I^2 = 1$
$I^1 = 0$	0	$\left(\lambda^2 + \lambda^{12}\right) \hat{\phi}_S$
$I^1 = 1$	$\left(\lambda^1 + \lambda^{12}\right) \hat{\phi}_S$	$\left(\lambda^1 + \lambda^2 + \lambda^{12}\right) \hat{\phi}_S$

Merchants' best response is to reject both cards if $\hat{\phi}_S < 0$ and to accept both cards if $\hat{\phi}_S \geq 0$.

- If $f^1 < f^2$, then equation (18) implies that

$$\hat{\Psi} = \lambda^1 \hat{\phi}_S^1 I^1 + \lambda^2 \hat{\phi}_S^2 I^2 + \lambda^{12} \left(\hat{\phi}_S^1 I^1 + \hat{\phi}_S^2 I^2 \left(1 - I^1\right) \right).$$

In this case the function $\hat{\Psi}$ is determined by the following table:

	$I^2 = 0$	$I^2 = 1$
$I^1 = 0$	0	$\left(\lambda^2 + \lambda^{12}\right) \hat{\phi}_S^2$
$I^1 = 1$	$\left(\lambda^1 + \lambda^{12}\right) \hat{\phi}_S^1$	$\left(\lambda^1 + \lambda^{12}\right) \hat{\phi}_S^1 + \lambda^2 \hat{\phi}_S^2$

Merchants' best response is as follows: merchants reject both cards if $\hat{\phi}_S^1, \hat{\phi}_S^2 < 0$; they accept only card 1 if $\hat{\phi}_S^2 < 0 \leq \hat{\phi}_S^1$; they accept only card 2 if $\hat{\phi}_S^2 \geq 0$ and $\lambda^{12} \hat{\phi}_S^2 > (\lambda^1 + \lambda^{12}) \hat{\phi}_S^1$; and they accept both cards if $(\lambda^1 + \lambda^{12}) \hat{\phi}_S^1 \geq \lambda^{12} \hat{\phi}_S^2 \geq 0$.

- Exploiting the symmetry with the above case, merchants' best response is the same except the superscripts 1 and 2 are swapped.

Consumers' best responses.

At stage (ii), consumers decide which card(s) to hold, if any. Their choice of card-holding depends on the benefits they get from holding a card, which depend on merchants' acceptance decisions. Since consumers get no surplus from buying the good itself, cash-constrained consumers get the same benefits from cardholding as unconstrained consumers. A consumer's additional benefit of holding only card i is therefore

$$\phi_B^i I^i + u,$$

while the additional benefit of holding two cards is

$$\phi_B^1 I^1 L^1 + \phi_B^2 I^2 L^2 + 2u.$$

If merchants reject both cards, then consumers with $u \geq 0$ will hold two cards, while consumers with $u < 0$ will hold no cards. If merchants accept only card i, then consumers will hold both cards if $u \geq 0$, will hold only card i if $-\phi_B^i \leq u < 0$ and will hold neither card if $u < -\phi_B^i$. If merchants accept both cards but card i has a lower card fee than the other card, then consumers will hold both cards if $u \geq 0$, will hold only card i if $-\phi_B^i \leq u < 0$ and will hold neither card if $u < -\phi_B^i$. If merchants accept both cards and both cards have the same card fees, then consumers will hold both cards if $u \geq 0$, will hold only a single card if $-\phi_B^1 = -\phi_B^2 \leq u < 0$ (in which case consumers will randomize over which card they will hold), and will hold neither card if $u < -\phi_B^1 = -\phi_B^2$. These results are summarized by the functions

$$\begin{aligned}
\lambda^0(f^1, f^2) &= 1 - A(-\phi_B^1 L^1 - \phi_B^2 L^2), \\
\lambda^i(f^1, f^2) &= L^i \left(A(-\phi_B^i) - A(0) \right), \\
\lambda^{12}(f^1, f^2) &= A(0),
\end{aligned}$$

which give the measure of consumers who hold neither card, just card i, or both cards respectively. Note if $\overline{u} = 0$ then $\lambda^{12} = 0$ and no consumers multihome.

Equilibria in the subgame.

Using the characterizations of consumers' and merchants' best responses, we can look for cases where both types of users have best responses to each other at stage (ii) — that is, we can look for possible equilibria in the subgame starting at stage (ii). There are three cases to consider based on the relative sizes of f^1 and f^2.

Case 1: $f^1 = f^2$. In this case $\hat{\phi}_S^1 = \hat{\phi}_S^2$. Then from above, an equilibrium in stage (ii) exists if $\hat{\phi}_S^1 = \hat{\phi}_S^2 \geq 0$. The merchant accepts both cards and the singlehoming consumers will randomize over which card to hold. An equilibrium also exists in stage (ii) if $\hat{\phi}_S^1 = \hat{\phi}_S^2 < 0$, in which case the merchant rejects both cards and there are no singlehoming consumers.

Case 2: $f^1 < f^2$. Then there are four possible equilibria at stage (ii). If $\hat{\phi}_S^1, \hat{\phi}_S^2 < 0$, there is an equilibrium in which the merchant rejects both cards and there are no singlehoming consumers. If $\hat{\phi}_S^1, \hat{\phi}_S^2 \geq 0$ and $(\lambda^1 + \lambda^{12})\hat{\phi}_S^1 \geq \lambda^{12}\hat{\phi}_S^2$ there is an equilibrium in which the merchant accepts both cards and the singlehoming consumers will only hold card 1. If $\hat{\phi}_S^1 \geq 0$, $\hat{\phi}_S^2 < 0$ there is an equilibrium in which the merchant only accepts card 1 and the singlehoming consumers will only hold card 1. If $\hat{\phi}_S^2 \geq 0$ and $\hat{\phi}_S^1 < 0$ there is an equilibrium in which the merchant only accepts card 2 and the singlehoming consumers will only hold card 2. (If $\overline{u} > 0$, then this equilibrium occurs more generally if $\hat{\phi}_S^2 \geq 0$ and $\hat{\phi}_S^1 < \hat{\phi}_S^2$.)

Case 3: $f^1 > f^2$. By symmetry, this is the same as the above case except the superscripts 1 and 2 are swapped.

■

This implies that for the case in which consumers get no intrinsic benefits from holding cards, the equilibrium interchange fee is still the same as that set by a single monopoly scheme; that is, $a^C + q(v - d)$.

Proposition 23 *If some consumers are cash-constrained and no consumers get any intrinsic benefit from holding cards, then the equilibrium interchange fee resulting from competition between identical card schemes facing monopolistic merchants equals $\hat{a}^C$; that is, it solves $\hat{\phi}_S(f(a)) = 0$. Merchants will (just) accept both cards and card-holding consumers will randomize over which card to hold. Each association shares in half the card transactions.*

Proof. The existence of a^C was proven in Lemma 1, guaranteeing that $\hat{a}^C = a^C + q(v - d)$ also exists. The next step is to prove that this is an equilibrium using our analysis of equilibria in the subgame starting at stage (ii). From the analysis of consumers' best responses at stage (ii) of the game, with $\overline{u} = 0$, $\lambda^{12} = A(0) = 0$. No consumers will hold both cards. Note that at $\hat{a}^C$, $\hat{\phi}_S^1 = \hat{\phi}_S^2 = 0$. If scheme 1 sets a card fee $f^1 < f(\hat{a}^C)$, then $\hat{\phi}_S^1 < \hat{\phi}_S^2 = 0$, merchants will accept only card 2 and no consumers will hold card 1; scheme 1 will get no card transactions. If scheme 1 sets a card fee $f^1 > f(\hat{a}^C)$ instead, then either $\hat{\phi}_S^1 \geq 0$, in which case merchants accept both cards and no consumers hold card 1, or $\hat{\phi}_S^1 < 0$, in which case merchants accept only card 2, and no consumers hold card 1; in either case, scheme 1 will get no card transactions. Thus, this is indeed an equilibrium.

This equilibrium is unique, since if any scheme i sets a fee structure such that $\hat{\phi}_S^i > 0$, then the other scheme will always want to attract all consumers to hold its card by setting a lower card fee such that $\hat{\phi}_S^j \geq 0$ and $\hat{\phi}_S^j < \hat{\phi}_S^i$. The optimal response of scheme i will be to match this fee structure. If any scheme i sets a fee structure such that $\hat{\phi}_S^i < 0$, then merchants will reject its cards and the other scheme will always want to attract all consumers to hold it cards by setting a fee structure at which merchants will accept its cards (that is, with $\hat{\phi}_S^j \geq 0$). The optimal response of scheme i will be to change its fee structure so that $\hat{\phi}_S^i \geq 0$. Thus, the only equilibrium is one with $\hat{\phi}_S^i = \hat{\phi}_S^j = 0$. ■

Where some consumers get intrinsic benefits from holding cards, the equilibrium interchange fee now maximizes the modified surplus of merchants, $\hat{\phi}_S^i$.

Proposition 24 *If some consumers are cash-constrained and some consumers get intrinsic benefits from holding cards, then the equilibrium interchange fee resulting from competition between identical card schemes facing monopolistic merchants equals $\hat{a}^M$, which maximizes $\hat{\phi}_S^i$, the expected surplus to merchants from accepting cards. Merchants accept both cards and singlehoming consumers randomize over which card to hold. The measure of multihoming consumers is*

$$\lambda^M \left(\hat{a}^M\right) = A\left(0\right),$$

the measure of consumers not holding cards is

$$\lambda^N \left(\hat{a}^M\right) = 1 - A\left(-\phi_B\left(f\left(\hat{a}^M\right)\right)\right),$$

and the measure of singlehoming consumers is

$$\lambda^S \left(\hat{a}^M\right) = A\left(-\phi_B\left(f\left(\hat{a}^M\right)\right)\right) - A\left(0\right).$$

Proof. The existence and uniqueness of $\hat{a}^M$ can be proven in exactly the same way as the existence and uniqueness of a^M was proven in Lemma 1. From the analysis of consumers' best responses at stage (ii) of the game, $\lambda^{12} = A(0) > 0$. Some consumers will hold both cards. We use this property and the analysis of equilibria in the stage (ii) subgame to show that $\hat{a}^M$ represents an equilibrium interchange fee.

Any scheme (say scheme 1) that sets a higher card fee f^1 (lower interchange fee), will result in $\hat{\phi}_S^1 < \hat{\phi}_S^2$, so will imply an equilibrium at stage (ii) in which merchants accept both cards and the singlehoming consumers will only hold card 2. The measure of each type of consumer λ^N, λ^S and λ^M will not change since singlehoming consumers will get the same benefits from holding card 2, as previously they obtained from randomizing over which card to hold. Scheme 1 will get no card transactions.

Any scheme (say scheme 1) that sets a lower fee f^1 (higher interchange fee), will result in $\hat{\phi}_S^1 < \hat{\phi}_S^2$, so will imply an equilibrium at stage (ii) in which merchants will only accept card 2 and the singlehoming consumers will only hold card 2. Again, the measure of each type of consumer will not change since singlehoming consumers will get the same benefits from holding card 2, as previously they obtained from randomizing over which card to hold. Again, scheme 1 will get no card transactions.

Thus, scheme 1 does strictly worse by setting a higher or lower interchange fee than that which maximizes $\hat{\phi}_S$, proving that this is an equilibrium.

It remains to prove that this equilibrium is unique. Suppose that it is not. Then there exists some other equilibrium in which one scheme (say scheme 1) sets an interchange fee such that $\hat{\phi}_S^1 < \hat{\phi}_S^{\max}$. It is straightforward to show that scheme 2's best response is to set a different interchange fee so that $\hat{\phi}_S^2 > \hat{\phi}_S^1$, in which case it attracts all card transactions. Thus, there can be no other equilibrium. ■

The result is a lower equilibrium interchange fee than set by a single scheme:

Proposition 25 *If some consumers are cash-constrained and some consumers get intrinsic benefits from holding cards, the equilibrium interchange fee resulting from competition between identical card schemes facing monopolistic merchants leads to an interchange fee lower than that which maximizes the schemes' joint profit (or joint card transactions).*

Proof. The result $\hat{a}^M < \hat{a}^C$ follows immediately from replacing b_S with $\hat{b}_S$ everywhere in the proof that $a^M < a^C$ in Lemma 1. ■

and a lower interchange fee than that which maximizes overall welfare:

Proposition 26 *If some consumers are cash-constrained and some consumers get intrinsic benefits from holding cards, the equilibrium interchange fee resulting from competition between identical card schemes facing monopolist merchants leads to an interchange fee lower than that which maximizes overall welfare.*

Proof. See Appendix A.7. ■

However, compared to the case without cash-constrained consumers, the equilibrium interchange fee and the welfare maximizing interchange fee are higher.

4.4 Policy implications

Policymakers in some jurisdictions have claimed that competing credit card schemes (namely, MasterCard and Visa) set interchange fees too high.[21] Policymakers have also charged that there is a lack of competition between card schemes. For instance, the Reserve Bank of Australia (2002, p. 8) states "In Australia, credit card interchange fees are not determined by a competitive market." As a result, regulators in Australia and in Europe have required that card schemes lower their interchange fees, and the United Kingdom is investigating similar measures. In a generic two-sided market, there is no obvious link between greater inter-system competition and lower interchange fees (charging less to merchants and more to cardholders), or between greater inter-system competition and a more efficient *structure* of fees.[22] Our analysis provides a model which captures the specificities of the credit card market in which without scheme competition payment schemes may set interchange fees too high. In our model, provided some consumers hold both cards, competition between identical schemes will in fact lower interchange fees. This suggests that where there is a lack of scheme competition, regulation could improve matters. In light of our analysis we offer some guidance on this matter.

First, as our model demonstrates, competition may result in interchange fees that are too low, especially to the extent retailers do not accept cards for strategic reasons. Thus, it is not immediately obvious that the goal for regulators (in terms of setting interchange fees) should be to replicate the outcome of competition between schemes.

Where the lower interchange fees resulting from scheme competition do raise social welfare, it would seem natural to conduct some market analysis to ascertain whether card schemes have market power. In determining whether card schemes have market power, the policymaker would need to be careful to take into account the two-sided nature of the services being offered by card schemes. For instance, it would make no sense to conclude that high card fees imply market power without also considering the level of merchant fees. In a similar vein, our model provides no basis for the claim by policymakers that competition should drive interchange fees to cost, or the suggestion that cost-based interchange fees are efficient.[23]

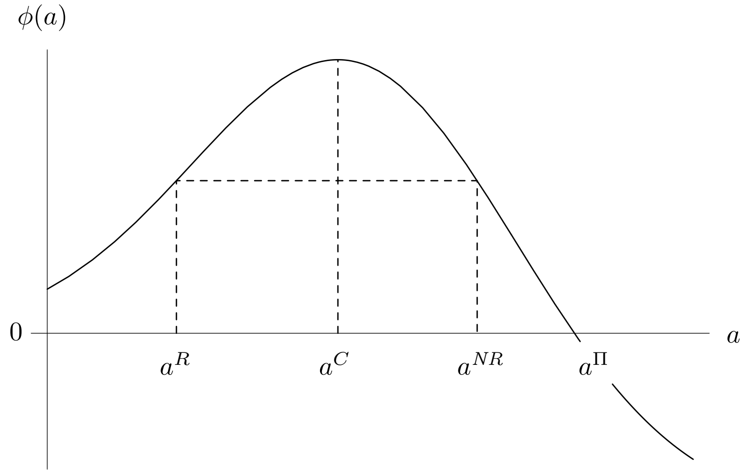
A related issue is that it is notable that in the U.S., where competition between credit cards is arguably quite strong (with DiscoverCard and American Express both competing for custom), interchange fees are

[21] See European Commission (2002), Office of Fair Trading (2003), and Reserve Bank of Australia (2002).

[22] For instance, Rochet and Tirole (2003) find in their model of generic competition in two-sided markets that, with linear demands, a monopoly scheme and competing schemes will set the same structure of fees (and so implicitly the same interchange fees). Moreover, the resulting structure of fees is socially optimal.

[23] For instance, see Office of Fair Trading (2003, 3.12)

Figure 2: Effect of regulating one payment scheme



Notes. The equilibrium interchange fee in an unfettered market is a^C. The regulated interchange fee is a^R. The scheme that is not regulated will respond by setting its interchange fee at a^{NR}.

higher than in many other countries where competition between credit card schemes seems weaker. A reason for this could be that competition with other payment instruments (such as debit cards) may provide a stronger constraining force on interchange fees in other countries than it does in the U.S. In the U.S., on-line debit cards (which tend to have very low interchange fees) have not been as significant as in most other OECD countries. Thus, in deciding whether there is a lack of scheme competition, policy makers will likely need to consider types of payment instruments other than just credit cards.

Another issue in the context of interchange fee regulation is that such regulations are currently only being proposed for card associations (MasterCard and Visa) and not for the proprietary schemes (such as American Express) which set fees to cardholders and merchants directly. For instance, the Reserve Bank of Australia proposes regulations of the card associations' interchange fees which explicitly leave proprietary schemes free to set their fees. Since proprietary schemes, such as American Express, do not have to set interchange fees to achieve their desired price structure, any regulation of interchange fees could act as a potential handicap to card associations.

The consequences of any asymmetric approach to regulation can be examined in the context of our model, as illustrated in Figure 2. Suppose, for instance, there are two competing schemes, some consumers hold both cards, and merchants compete in a Hotelling fashion. (The monopoly case can be considered by replacing $\phi(f)$ with $\phi_S(f)$ in Figure 2.) Suppose one scheme has its interchange fee regulated below the privately set level, while the other remains free to set its interchange fee (or fee structure). Since competition between the schemes constrains the interchange fee to the competitive level a^C when merchants compete in a Hotelling fashion, the result of regulating one scheme and not the other is to constrain the interchange fee below a^C. This leads the rival scheme to take the whole market by setting a higher interchange fee. Even if merchants continue to accept cards from the regulated schemes, since consumers now face higher fees for using cards from the regulated scheme, the regulated scheme will not attract any card transactions. Merchants will continue to accept cards from the unregulated scheme even though the merchant fees for the regulated scheme are lower, since the more expensive cards allow them to attract customers from rivals who do not accept such cards. The scheme that is not regulated will respond by increasing its interchange fee above a^C, although not beyond the interchange fee set by a single card scheme a^{Π}, thereby attracting more card transactions and profit. Thus, if card associations have their interchange fees regulated below the competitive level a^C, the model predicts proprietary schemes will respond by setting *higher* merchant fees (and lower card fees), and will attract business away from card associations. Figure 2 illustrates this result by noting the interchange fee chosen by the scheme that is not regulated, denoted a^{NR}, for a particular interchange fee that is imposed on the regulated scheme, denoted a^R.

The above case provides some insight into the workings of competition between on-line debit cards offered directly by banks and off-line debit cards offered by the card associations. On-line debit cards offered by banks and off-line debit cards offered by the card associations are similar instruments both from the perspective of consumers and merchants.[24] Moreover, consumers who have access to debit cards offered by card associations usually also have access to an on-line debit card offered directly by their bank (this may also be their ATM card). The above results suggest competition between on-line debit and off-line debit will drive the interchange fees below the level that the schemes prefer. If card associations try to set their preferred pricing structure for off-line debit ignoring the existence of on-line debit (and thus set relatively high interchange fees), merchants will simply reject off-line debit knowing that such consumers will substitute by using on-line debit instead. One way for card associations to prevent this type of competition would be to tie the acceptance of their off-line debit cards to acceptance of credit cards, assuming the interchange fee for credit cards was not subject to the same kind of competitive pressure. This provides one interpretation of MasterCard and Visa's tying behavior, behavior that resulted in the 'Walmart case'.

In 2003 Walmart, together with a large number of other merchants, obtained a settlement of three billion U.S. dollars from MasterCard and Visa after Walmart alleged MasterCard and Visa tied off-line debit and credit cards together, and used market power in credit cards to set high merchant fees for off-line debit cards. If merchants accepted Visa credit cards under Visa's honor-all-cards rule, they also had to accept Visa off-line debit cards. Off-line debit cards have an interchange fee (and merchant fee) close to the levels used for credit cards, and according to Chakravorti and Shah (2003) are about three to five times more expensive for merchants to accept than on-line debit cards. The analysis provides one channel by which it is possible for the card associations' tying behavior to improve welfare. It does this by allowing the scheme to impose a different pricing structure for off-line debit cards from that chosen by banks for on-line debit, a pricing structure which could be more efficient given that competition between schemes can result in interchange fees being set too low. On the other hand, the tying behavior could also allow credit card schemes to set interchange fees for debit that are closer to the monopoly level a^{Π}, which could be excessively high.

4.5 Implications for other two-sided markets

One of the main motivations for this paper was to extend the existing literature on two-sided markets to the case in which one side of the market competes amongst itself to attract users on the other side. Many two-sided markets have this feature, including payment cards, shopping malls, Yellow Pages, and hardware/consoles and software providers. By examining the case with monopoly merchants and with competing merchants we are able to discern how allowing one type of user to compete amongst itself

[24] Typically, an on-line debit card transaction requires cardholders enter a pin number while an off-line debit card transaction requires cardholders sign to verify the transaction.

affects the equilibrium structure of fees. Competition between sellers generally increases the privately and socially optimal interchange fees (meaning it is optimal to charge more to merchants and less to consumers). When one type of user (sellers) competes amongst itself to attract the other type of user (buyers), the sellers tend to internalize the benefits of the buyers. This makes it more desirable to set fees which favor buyers, since by offering more surplus to buyers the schemes will find it easier to attract sellers. On the other hand, regardless of whether sellers compete, competition between schemes still lowers the relative charge to sellers when some buyers multihome, as schemes compete to be used exclusively by sellers.

Applying the same logic to other two-sided markets such as consumer directory services suggests that (i) competition between separate Yellow Pages (or other directories) should reduce the extent to which they will charge advertizers (and provide free services to consumers) and (ii) competition amongst advertizers should raise the charges they face. The fact that advertizers cover all of the costs of these services suggests that the second effect dominates. In a world of monopoly sellers and multihoming buyers, platforms such as these may charge the buyers and not the sellers for making use of their directories.

Another difference between payment cards and some other two-sided markets is the possibility for card schemes to set negative prices for card usage without inducing unbounded consumption.[25] Such pricing would be difficult to implement in the Yellow Pages business, for instance, although there may be limited opportunities for rebates to customers in other cases. Shopping malls sometimes offer prize draws and/or free-parking to consumers who make some specified level of purchases within the mall. To the extent that prices are constrained to be non-negative on one side of the market, the equilibrium fee structure could involve services being given free to this side of the market. In such cases, the fee structure may not respond to the nature of merchant or scheme competition over some range, at least until a positive fee becomes optimal.

An obvious feature of the equilibrium in our model is that it is sensitive to the ability of consumers and merchants to coordinate on cards from a particular scheme. We focused on the only pure-strategy equilibrium, in which schemes competed to attract merchants exclusively. The result of competition was relatively low merchant fees and high card fees. If, instead, consumers and merchants coordinate on the other equilibria in the stage (ii) subgame, in which merchants accept both cards and singlehoming consumers just hold the card with lower fees, then each scheme will have an incentive to set a slightly higher interchange fee than its rival (until merchants are left with no surplus, at which point either scheme can attract all merchants exclusively by offering a lower interchange fee). While there is no (pure-strategy) equilibrium in this case, it does raise the possibility that differences in users' beliefs could lead to divergent outcomes. This result could underlie the fact that sometimes quite different fee structures can emerge in apparently similar two-sided markets. For example, rental agencies (which help match tenants and landlords) typically charge landlords exclusively for the service, but in some cities such as Boston and New York the tenant typically pays the entire fee (see Evans, 2002).

5 Conclusions

This paper extended existing models of payment schemes to allow for competition between schemes. It examined how competition between identical schemes affects the choice of fee structure by card schemes, namely, how much to charge cardholders versus how much to charge merchants.

[25] Unbounded consumption is avoided with card transactions since consumers must still purchase the merchant's good to enjoy any frequent flyer miles, cash back or other inducements from the issuer.

There is a concern by policymakers that consumers face distorted incentives to use credit cards, as a result of low card fees (and other rebates) at the same time that merchant fees are set high. We addressed one implicit assumption behind the policymaker's concern — that a lack of system competition explains why MasterCard and Visa can set high interchange fees, and thus why schemes (including proprietary schemes) set high merchant fees and low card fees. We showed how scheme competition lowers interchange fees, although such competition can also result in interchange fees being set below the efficient level. When some consumers hold multiple cards, merchants tend to reject the more expensive card, causing schemes to compete by focusing on attracting merchants. Schemes do this by setting low interchange fees (charging too much to cardholders and too little to merchants).

While competition between schemes lowers equilibrium interchange fees, competition between merchants increases them. Surprisingly, the two types of competition exactly offset in our model. Taking into account both effects results in an equilibrium interchange fee that is identical to that chosen by a single scheme which faces a monopoly merchant. Moreover, this interchange fee maximizes the expected joint surplus of consumers and merchants from using cards. We also showed other scenarios for which equilibrium interchange fees are either too high or too low, sometimes both from the perspective of card schemes and of society. For the cases we examined, the results were robust to whether competition is between card associations or between proprietary schemes, and to whether there is merchant heterogeneity or not. Future research should consider alternative cases, such as that of imperfect competition between card schemes as well as between merchants.

References

Armstrong, M. (2002). "Competition in Two-Sided Markets," mimeo, Nuffield College, University of Oxford.

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

Caillaud, B. and B. Jullien (2001). "Competing Cybermediaries," *European Economic Review*, 45, 797–808.

Caillaud, B. and B. Jullien (2003). "Chicken & Egg: Competition Among Intermediation Service Providers," *RAND Journal of Economics*, 34(2), 309–328.

Chakravorti, S. and A. Shah (2003). "Underlying Incentives in Credit Card Networks," *The Antitrust Bulletin*, forthcoming.

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

European Commission (2002). "Case no. comp/29.373 - Visa International - multilateral interchange fee," *Official Journal of the European Communities*, 2002/L318/17, November 22.

Evans, D. S. (2003). "The Antitrust Economics of Two-Sided Markets," Yale Journal on Regulation, forthcoming.

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

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

Guthrie, G. and J. Wright (2003). "Competing Payment Schemes," Working Paper No. 245. Department of Economics, University of Auckland.

Hausman, J.A. and G.K. Leonard, and J. Tirole. (2003). "On Nonexclusive Membership in Competing Joint Ventures," *RAND Journal of Economics*, 34(1), 43–62.

Office of Fair Trading. (2003). "MasterCard Interchange Fees: Preliminary Conclusions," United Kingdom. February.

Parker, G. and M. Van Alstyne (2000). "Information Complements, Substitutes, and Strategic Product Design" Working Paper No. 299, March. Tulane University and University of Michigan.

Reserve Bank of Australia (2002). "Reform of credit card schemes in Australia. IV. Final reforms and regulation impact statement" Reserve Bank of Australia, Public Document, August.

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

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

Schiff, A. (2003). "Open and Closed Systems of Two Sided Networks," *Information, Economics and Policy*, forthcoming.

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

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

Wright, J. (2002). "Access pricing under competition: An application to cellular networks." *Journal of Industrial Economics*, 50, 289–315.

Wright, J. (2003a). "Determinants of Optimal Interchange Fees in Payment Systems," *Journal of Industrial Economics*, forthcoming.

Wright, J. (2003b). "Pricing in debit and credit card schemes," *Economics Letters*, 80, 305–309.

Wright, J. (2003c). "Optimal Card Payment Systems," *European Economic Review*, 47, 587–612.

A Proofs

A.1 Proof of Lemma 1

1. The derivative

$$\frac{d\phi}{da} = h(f)\,(b_S - l + f) = h(f)\,(b_S - m)$$

equals zero if and only if $b_S = m$ since h is positive over its support, $d\phi/da$ is positive for larger values of f and negative for smaller values of f. The unique interchange fee is

$$a^C = b_S - c_A - \pi_A.$$

At this interchange fee $\phi_S(f) = 0$.

2. The first order condition is

$$\frac{d\phi_S}{da} = -D(f) + h(f)(b_S - m) = 0$$

which, using the fact that $l = f + m$, implies

$$\frac{d\phi_S}{da} = -D(f) + h(f)(f + b_S - l) = 0.$$

The solution is

$$f = l - b_S + \frac{D(f)}{h(f)}.$$

Note the left hand side is increasing in f while the right hand side is decreasing in f given that the hazard function $h(f)/(1 - H(f))$ is increasing in f. Moreover, we have that at $\underline{b}_B$, the left hand side of the expression is less than $E(b_B)$ which is less than $l - b_S$ from (3), while the right hand side of the expression is greater than $l - b_S$. On the other hand, at $\overline{b}_B$ the left hand side of the expression is greater than $l - b_S$ from (3), while the right hand side of the expression is equal to $l - b_S$ since $D\left(\overline{b}_B\right) = 0$. It follows there must exist a unique solution to the first order condition, which we denote f^M. The corresponding interchange fee is a^M, which is characterized by

$$a^M = c_I + \pi_I - f^M. \tag{A-1}$$

Finally,

$$\begin{aligned} \left.\frac{d^2\phi_S}{da^2}\right|_{a=a^M} &= -h(f^M) - h'(f^M)(b_S - m) \\ &= -h(f^M) - \frac{h'(f^M)D(f^M)}{h(f^M)} \\ &< 0, \end{aligned}$$

where the final step follows from the assumption that the hazard function is increasing in f. Thus, a^M maximizes ϕ_S.

3. The condition $\phi(f) = 0$ is equivalent to $\beta(f) + b_S = l$. Note that $E(b_B) = \beta(\underline{b}_B) < \beta(f) < \beta(\overline{b}_B) = \overline{b}_B$ and so from (3) there is a sufficiently low interchange fee such that $\beta(f) + b_S > l$ and a sufficiently high interchange fee such that $\beta(f) + b_S < l$. It follows that there exists an interchange fee a^Π such that $\beta(f) + b_S = l$. Moreover since $\beta(f)$ is increasing in f, this interchange fee is unique.

4. It follows immediately from

$$f^M = \frac{D(f^M)}{h(f^M)} + l - b_S > l - b_S = f(a^C)$$

that $a^M < a^C$. Furthermore,

$$f(a^C) = l - b_S = \beta(f(a^\Pi)) > f(a^\Pi),$$

where we have used the fact that $\beta(f) > f$ to complete the last step. It follows that $a^C < a^\Pi$.

A.2 Proof of Proposition 2

If $\phi_S(f) \geq 0$, then the merchant accepts the card, $A(-\phi_B)$ consumers hold the card, and consumers who hold the card use it if $b_B \geq f$. Furthermore, because each transaction involving cards contributes $b_B + b_S - c_A - c_I$ to welfare, and because u and b_B are independently distributed, the total contribution of transactions involving cards to total welfare is

$$A(-\phi_B(f)) \int_f^{\bar{b}_B} (b_B + b_S - c_A - c_I) h(b_B) db_B.$$

However, consumers also derive benefits simply by holding cards: consumers for whom $u \geq -\phi_B(f)$ hold a card, receiving benefit u. Thus, card-holding contributes

$$\int_{-\phi_B(f)}^{\overline{u}} ue(u)du$$

to total welfare, where e is the density function of u. Adding the two welfare components together shows that when $\phi_S(f) \geq 0$ total welfare is[26]

$$W(f) = A(-\phi_B(f)) \int_f^{\bar{b}_B} (b_B + b_S - c_A - c_I) h(b_B) db_B + \int_{-\phi_B(f)}^{\overline{u}} ue(u)du. \tag{A-2}$$

If $\phi_S(f) < 0$ then the merchant rejects the card, and only consumers with $u \geq 0$ accept the card. In this case the only contribution to total welfare comes from the intrinsic benefits of card-holding, so that total welfare is

$$W(f) = W_0 \equiv \int_0^{\overline{u}} ue(u)du. \tag{A-3}$$

It is useful to note that when $a \leq a^C$ (so that the merchant accepts cards),

$$\begin{aligned}
W(f) &= A(-\phi_B(f)) \int_f^{\bar{b}_B} (b_B + b_S - c_A - c_I) h(b_B) db_B + \int_{-\phi_B(f)}^{0} ue(u)du + W_0 \\
&\geq A(-\phi_B(f)) \int_f^{\bar{b}_B} (b_B + b_S - c_A - c_I) h(b_B) db_B - \phi_B(f) \int_{-\phi_B(f)}^{0} e(u)du + W_0 \\
&= A(-\phi_B(f)) \int_f^{\bar{b}_B} (f + b_S - c_A - c_I) h(b_B) db_B + \phi_B(f) A(0) + W_0 \\
&= (f + b_S - c_A - c_I) D(f) A(-\phi_B(f)) + \phi_B(f) A(0) + W_0,
\end{aligned}$$

where we have used that

$$\phi_B(f) = \int_f^{\bar{b}_B} (b_B - f) h(b_B) db_B.$$

Further, differentiating W with respect to f in the region where $a \leq a^C$ gives

$$\begin{aligned}
W'(f) &= e(-\phi_B(f))\phi_B'(f) \int_f^{\bar{b}_B} (b_B + b_S - c_A - c_I) h(b_B) db_B \\
&\quad - A(-\phi_B(f))(f + b_S - c_A - c_I) h(f) - \phi_B(f)\phi_B'(f) e(-\phi_B(f)) \\
&= -e(-\phi_B(f)) D(f) \int_f^{\bar{b}_B} (f + b_S - c_A - c_I) h(b_B) db_B \\
&\quad - A(-\phi_B(f))(f + b_S - c_A - c_I) h(f) \\
&= (f + b_S - c_A - c_I) \left(-e(-\phi_B(f))(D(f))^2 - A(-\phi_B(f)) h(f) \right),
\end{aligned}$$

[26] Throughout the paper we ignore additional constant terms in the welfare function that can arise in models in which consumers incur transportation costs (for instance, in going to their preferred bank or merchant). These can be ignored because firms, issuers, acquirers, and card schemes are all symmetric.

where $\phi'_B(f)$ denotes the derivative of $\phi_B(f)$ with respect to f, which equals $-D(f)$. The term in large brackets in the expression above is always nonpositive. Since $a \leq a^C$ is equivalent to $f \geq l - b_S$, the other term in the expression for $W'(f)$ is non-negative. Therefore $W'(f) \leq 0$ for all $a \leq a^C$, and since

$$W(l - b_S) \geq (\pi_A + \pi_I)D(l - b_S)A(-\phi_B(l - b_S)) + \phi_B(l - b_S)A(0) + W_0 \geq W_0,$$

welfare is maximized by setting $a = a^C$.

A.3 Proof of Proposition 4

If $\phi(f) \geq 0$, then merchants accept the card, the proportion of consumers who hold the card is $A(-\phi_B)$ and consumers who hold the card use it if $b_B \geq f$. Using a similar argument to that in Appendix A.2, we can show that welfare equals the expression in equation (A-2) when $\phi(f) \geq 0$. If $\phi(f) < 0$ then merchants reject the card, only consumers with $u \geq 0$ accept the card, and welfare equals the expression in equation (A-3).

Using similar arguments to those in Appendix A.2, we can show that when $a \leq a^\Pi$ (so that merchants accept cards),

$$W(f) \geq (f + b_S - c_A - c_I)D(f)A(-\phi_B(f)) + \phi_B(f)A(0) + W_0$$

and

$$W'(f) = (f + b_S - c_A - c_I)\left(-e(-\phi_B(f))(D(f))^2 - A(-\phi_B(f))h(f)\right).$$

The term in large brackets in the expression above is always non-positive.

We consider two possibilities. Firstly, we show that if $a^\Pi > b_S - c_A + \pi_I$, then welfare is maximized by setting $a^W = b_S - c_A + \pi_I$. At this interchange fee $f(a^W) = c_A + c_I - b_S$, implying that $W'(f(a^W)) = 0$. Since $W'(f)$ is negative for lower interchange fees, and is positive for larger interchange fees, this a^W is a local welfare maximum. Since

$$W(f(a^W)) \geq \phi_B(f(a^W))A(0) + W_0 \geq W_0,$$

it is also a global maximum.

Secondly, we show that if $a^\Pi \leq b_S - c_A + \pi_I$, then welfare is maximized by setting $a^W = a^\Pi$. Since $f(a^\Pi) \geq c_A + c_I - b_S$, it follows that $W'(f) \leq 0$ for all $a \leq a^\Pi$, Since

$$W(f(a^\Pi)) \geq (f(a^\Pi) + b_S - c_A - c_I)D(f(a^\Pi))A(-\phi_B(f(a^\Pi))) + \phi_B(f(a^\Pi))A(0) + W_0 \geq W_0,$$

it follows that this a^W is a global welfare maximum.

A.4 Proof of Proposition 8

We begin by considering the welfare-maximization problem subject to the additional constraint that $f^1 = f^2$, and then show that relaxing this constraint does not raise welfare.

Suppose $f^1 = f^2$. The first part of Lemma 4 describes the equilibria which can occur as functions of the common card fee f. If $\phi_S(f) < 0$, the monopolist merchant rejects both cards, only customers with $u \geq 0$ hold cards, and these customers hold both cards. Since no card transactions occur, the only contribution to welfare comes from the intrinsic benefits of holding cards. Thus, total welfare equals

$$W(f) = \int_0^{\overline{u}} 2ue(u)du = 2W_0,$$

where W_0 is defined in (A-3). If $\phi_S(f) \geq 0$, the monopolist merchant accepts both cards, customers hold both cards if $u \geq 0$, and hold one card if $-\phi_B(f) \leq u < 0$. The intrinsic benefit of holding cards thus contributes

$$\int_{-\phi_B(f)}^{0} ue(u)du + \int_{0}^{\overline{u}} 2ue(u)du = \int_{-\phi_B(f)}^{0} ue(u)du + 2W_0$$

to welfare. Customers who hold one or both cards use a card if $b_B \geq f$. Furthermore, because each transaction involving cards contributes $b_B+b_S-c_A-c_I$ to welfare, and because u and b_B are independently distributed, the total contribution of transactions involving cards to total welfare is

$$A(-\phi_B(f)) \int_{f}^{\bar{b}_B} (b_B + b_S - c_A - c_I)h(b_B)db_B.$$

Thus, total welfare equals

$$W(f) = A(-\phi_B(f)) \int_{f}^{\bar{b}_B} (b_B + b_S - c_A - c_I)h(b_B)db_B + \int_{-\phi_B(f)}^{0} ue(u)du + 2W_0. \tag{A-4}$$

Apart from the addition of a constant W_0, this is exactly the same welfare function as that in the proof of Proposition 2. It is therefore maximized by setting $a^1 = a^2 = a^C$.

Suppose that $f^1 \neq f^2$. From the second part of Lemma 4, either there are no singlehoming customers, or all singlehoming customers hold the same card. In the first case, welfare equals $2W_0$. In the second case, welfare equals $W(f^i)$, where W is defined by equation (A-4) and i is the card held. A necessary case for the second case to occur is that $\phi_S^i \geq 0$. It follows that relaxing the symmetry constraint cannot raise welfare, completing the proof that $a^1 = a^2 = a^C$ maximizes welfare.

A.5 Proof of Proposition 12

We begin by considering the welfare-maximization problem subject to the additional constraint that $f^1 = f^2$, and then show that relaxing this constraint does not raise welfare.

Suppose $f^1 = f^2$. The first part of Lemma 5 describes the equilibria which can occur as functions of the common card fee f. If $\phi(f) < 0$, the merchants reject both cards, only customers with $u \geq 0$ hold cards, and, since these customers hold both cards, total welfare equals $W(f) = 2W_0$. If $\phi(f) \geq 0$, the merchants accept both cards, customers hold both cards if $u \geq 0$, and hold one card if $-\phi_B(f) \leq u < 0$. Since customers who hold one or both cards use a card if $b_B \geq f$, we can use a similar argument to that in Appendix A.4 to show that total welfare equals the expression in equation (A-4). Apart from the addition of a constant W_0, this is exactly the same welfare function as that in the proof of Proposition 4. It is therefore maximized by setting $a^1 = a^2 = \min\{a^\Pi, b_S - c_A + \pi_I\}$.

Suppose that $f^1 \neq f^2$. From the second part of Lemma 5, either there are no singlehoming customers, or all singlehoming customers hold the same card. In the first case, welfare equals $2W_0$. In the second case, welfare equals $W(f^i)$, where W is defined by equation (A-4) and i is the card held. A necessary case for the second case to occur is that $\phi^i \geq 0$. It follows that relaxing the symmetry constraint cannot raise welfare, completing the proof that $a^1 = a^2 = \min\{a^\Pi, b_S - c_A + \pi_I\}$ maximizes welfare.

A.6 Proof of Proposition 22

The calculation of overall welfare differs slightly from the calculation in Proposition 2. There, the merchant's profit from selling the good could be ignored since its sales did not depend on the interchange fee. Now, however, the level of the interchange fee affects the level of sales to cash constrained consumers. We therefore include the merchant's profit from sales to cash-constrained customers.

If $\hat{\phi}_S(f) \geq 0$, then the merchant accepts the card, $A(-\phi_B)$ consumers hold the card, and consumers who hold the card use it if $b_B \geq f$. A total of $A(-\phi_B(f))qD(f)$ cash-constrained customers are able to buy the good, contributing

$$A(-\phi_B(f))qD(f)(v-d)$$

to overall welfare. Because each transaction involving cards contributes $b_B + b_S - c_A - c_I$ to welfare, and because u and b_B are independently distributed, the total contribution of transactions involving cards to total welfare is

$$A(-\phi_B(f)) \int_f^{\bar{b}_B} (b_B + b_S - c_A - c_I) h(b_B) db_B.$$

However, consumers also derive benefits simply by holding cards: consumers for whom $u \geq -\phi_B(f)$ hold a card, receiving benefit u. Thus, card-holding contributes

$$\int_{-\phi_B(f)}^{\overline{u}} u e(u) du$$

to total welfare, where e is the density function of u. Adding the three welfare components together shows that when $\hat{\phi}_S(f) \geq 0$ total welfare is

$$W(f) = A(-\phi_B(f)) \int_f^{\bar{b}_B} (b_B + \hat{b}_S - c_A - c_I) h(b_B) db_B + \int_{-\phi_B(f)}^{\overline{u}} u e(u) du,$$

where we have used the fact that $\hat{b}_S = b_S + q(v-d)$. If $\hat{\phi}_S(f) < 0$ then the merchant rejects the card, and only consumers with $u \geq 0$ accept the card. In this case, since cash-constrained consumers cannot buy the good, the only contribution to total welfare comes from the intrinsic benefits of card-holding, so that total welfare is

$$W(f) = \int_0^{\overline{u}} u e(u) du = W_0.$$

It is useful to note that when $a \leq \hat{a}^C$ (so that the merchant accepts cards),

$$\begin{aligned}
W(f) &= A(-\phi_B(f)) \int_f^{\bar{b}_B} (b_B + \hat{b}_S - c_A - c_I) h(b_B) db_B + \int_{-\phi_B(f)}^{0} u e(u) du + W_0 \\
&\geq A(-\phi_B(f)) \int_f^{\bar{b}_B} (b_B + \hat{b}_S - c_A - c_I) h(b_B) db_B - \phi_B(f) \int_{-\phi_B(f)}^{0} e(u) du + W_0 \\
&= A(-\phi_B(f)) \int_f^{\bar{b}_B} (f + \hat{b}_S - c_A - c_I) h(b_B) db_B + \phi_B(f) A(0) + W_0 \\
&= (f + \hat{b}_S - c_A - c_I) D(f) A(-\phi_B(f)) + \phi_B(f) A(0) + W_0,
\end{aligned}$$

where we have used that

$$\phi_B(f) = \int_f^{\bar{b}_B} (b_B - f) h(b_B) db_B.$$

Further, differentiating W with respect to f in the region where $a \leq \hat{a}^C$ gives

$$\begin{aligned}
W'(f) &= e(-\phi_B(f)) \phi_B'(f) \int_f^{\bar{b}_B} (b_B + \hat{b}_S - c_A - c_I) h(b_B) db_B \\
&\quad - A(-\phi_B(f))(f + \hat{b}_S - c_A - c_I) h(f) - \phi_B(f) \phi_B'(f) e(-\phi_B(f)) \\
&= -e(-\phi_B(f)) D(f) \int_f^{\bar{b}_B} (f + \hat{b}_S - c_A - c_I) h(b_B) db_B \\
&\quad - A(-\phi_B(f))(f + \hat{b}_S - c_A - c_I) h(f) \\
&= (f + \hat{b}_S - c_A - c_I) \left(-e(-\phi_B(f))(D(f))^2 - A(-\phi_B(f)) h(f) \right),
\end{aligned}$$

where $\phi'_B(f)$ denotes the derivative of $\phi_B(f)$ with respect to f, which equals $-D(f)$. The term in large brackets in the expression above is always nonpositive. Since $a \leq \hat{a}^C$ is equivalent to $f \geq l - \hat{b}_S$, the other term in the expression for $W'(f)$ is nonnegative, proving that $W'(f) \leq 0$ for all $a \leq \hat{a}^C$. Since

$$W(l - \hat{b}_S) \geq (\pi_A + \pi_I - q(v - d))D(l - \hat{b}_S)A(-\phi_B(l - \hat{b}_S)) + \phi_B(l - \hat{b}_S)A(0) + W_0 \geq W_0,$$

welfare is maximized by setting $a = \hat{a}^C$.

A.7 Proof of Proposition 26

We begin by considering the welfare-maximization problem subject to the additional constraint that $f^1 = f^2$, and then show that relaxing this constraint does not raise welfare.

Suppose $f^1 = f^2$. The first part of Lemma 9 describes the equilibria which can occur as functions of the common card fee f. If $\hat{\phi}_S(f) < 0$, the monopolist merchant rejects both cards, only customers with $u \geq 0$ hold cards, and these customers hold both cards. Since no card transactions occur, the only contribution to welfare comes from the intrinsic benefits of holding cards. Thus, total welfare equals

$$W(f) = \int_0^{\overline{u}} 2ue(u)du = 2W_0.$$

If $\hat{\phi}_S(f) \geq 0$, the monopolist merchant accepts both cards, customers hold both cards if $u \geq 0$, and hold one card if $-\phi_B(f) \leq u < 0$. The intrinsic benefit of holding cards thus contributes

$$\int_{-\phi_B(f)}^0 ue(u)du + \int_0^{\overline{u}} 2ue(u)du = \int_{-\phi_B(f)}^0 ue(u)du + 2W_0$$

to welfare. Customers who hold one or both cards use a card if $b_B \geq f$. Sales to cash-constrained consumers contribute

$$A(-\phi_B(f))qD(f)(v - d)$$

to the merchant's profit, and therefore to overall welfare. Furthermore, because each actual transaction involving cards contributes $b_B + b_S - c_A - c_I$ to welfare, and because u and b_B are independently distributed, the total contribution card transactions to total welfare is

$$A(-\phi_B(f)) \int_f^{\bar{b}_B} (b_B + b_S - c_A - c_I)h(b_B)db_B.$$

Combining these three components, total welfare equals

$$W(f) = A(-\phi_B(f)) \int_f^{\bar{b}_B} (b_B + \hat{b}_S - c_A - c_I)h(b_B)db_B + \int_{-\phi_B(f)}^0 ue(u)du + 2W_0. \tag{A-5}$$

Apart from the addition of a constant W_0, this is exactly the same welfare function as that in the proof of Proposition 22. It is therefore maximized by setting $a^1 = a^2 = \hat{a}^C$.

Suppose that $f^1 \neq f^2$. From the second part of Lemma 9, either there are no singlehoming customers, or all singlehoming customers hold the same card. In the first case, welfare equals $2W_0$. In the second case, welfare equals $W(f^i)$, where W is defined by equation (A-5) and i is the card held. A necessary case for the second case to occur is that $\hat{\phi}^i_S \geq 0$. It follows that relaxing the symmetry constraint cannot raise welfare, completing the proof that $a^1 = a^2 = \hat{a}^C$ maximizes welfare.

B Optimal monopoly pricing

In setting its single retail price p, a monopolistic merchant that wants to accept cards faces two alternatives. It can set $p = v$, so that cardholders who want to use cash will purchase, or it can set $p > v$, in which case it will only face demand from cardholders. If there is a single card scheme, the merchant's profit is

$$\pi = v - d + \lambda(f)D(f)\,(b_S - m) = v - d + \lambda(f)D(f)\,(f + b_S - l)$$

if it sets $p = v$, and

$$\pi = \lambda(f)D\,(f + p - v)\,(p - d + b_S - m) = \lambda(f)D\,(f + p - v)\,(f + p - d + b_S - l)$$

if it sets $p > v$. In the second case the marginal consumer equates the additional benefits of making a purchase with a card b_B with the additional cost $f + p - v$. The merchant will set $p = v$ if

$$v - d + \lambda(f)D(f)\,(f + b_S - l) > \lambda(f)D\,(f + p - v)\,(f + p - d + b_S - l)\,.$$

The following lemma shows that a sufficient condition for this result to hold is that

$$v - d \geq l - \underline{b}_B - b_S + \frac{D(\underline{b}_B)}{h(\underline{b}_B)}, \tag{B-1}$$

which we assumed in Section 2. This requires that the surplus of the good must be sufficiently large.[27]

Lemma 10 *Under the no-surcharge rule, if the surplus satisfies (B-1) then monopolistic merchants will set a price $p = v$ when there is a single payment scheme.*

Proof. Since the hazard function is increasing in f, the function $D(f)/h(f)$ must be decreasing in f. Therefore

$$\frac{D(f)}{h(f)} \leq \frac{D(\underline{b}_B)}{h(\underline{b}_B)} \leq v - d + \underline{b}_B + b_S - l \leq v - d + f + b_S - l,$$

with the inequalities being strict when $f > \underline{b}_B$. This implies that the function

$$D(f)(v - d + f + b_S - l)$$

is decreasing in f. In particular,

$$D(f + p - v)(p - d + f + b_S - l) < D(f)(v - d + f + b_S - l)$$

whenever $p > v$ (so that $f + p - v > f$). Therefore, if the merchant sets $p > v$, it earns profit of

$$\begin{aligned} \lambda(f)D(f + p - v)(p - d + f + b_S - l) &< \lambda(f)D(f)(v - d + f + b_S - l) \\ &= \lambda(f)D(f)(v - d) + \lambda(f)D(f)(f + b_S - l) \\ &\leq v - d + \lambda(f)D(f)(f + b_S - l). \end{aligned}$$

That is, its profit from setting $p > v$ is less than its profit from setting $p = v$. ■

The merchant's profit is a more complicated function of the retail price when there are competing card schemes, although the profit-maximizing choice remains $p = v$.

Lemma 11 *Under the no-surcharge rule, if the surplus satisfies (B-1) then monopolistic merchants will set a price $p = v$ when there are competing card schemes.*

[27] Wright (2003c) makes a similar assumption.

Proof. If the merchant sets $p = v$ then its profit is

$$\begin{aligned}\pi_{p=v} &= v - d + \lambda^1 D(f^1)\left(b_S - m^1\right) I^1 + \lambda^2 D(f^2)\left(b_S - m^2\right) I^2 \\ &\quad + \lambda^{12}\left(D\left(f^1\right)\left(b_S - m^1\right) I^1 L^1 + D\left(f^2\right)\left(b_S - m^2\right) I^2 L^2\right),\end{aligned}$$

while if it sets $p > v$, the merchant's profit is

$$\begin{aligned}\pi_{p>v} &= \lambda^1 D\left(f^1 + p - v\right)\left(p - d + b_S - m^1\right) I^1 + \lambda^2 D\left(f^2 + p - v\right)\left(p - d + b_S - m^2\right) I^2 \\ &\quad + \lambda^{12}\left(D(f^1 + p - v)\left(p - d + b_S - m^1\right) I^1 L^1 + D\left(f^2 + p - v\right)\left(p - d + b_S - m^2\right) I^2 L^2\right).\end{aligned}$$

Using identical arguments to those contained in the proof of Lemma 10, it follows that

$$D(f^i + p - v)(p - d + f^i + b_S - l) < D(f^i)(v - d + f^i + b_S - l), \quad i = 1,\ 2,$$

whenever $p > v$. This implies that

$$D(f^i + p - v)(p - d + b_S - m^i) < D(f^i)(v - d + b_S - m^i), \quad i = 1,\ 2,$$

whenever $p > v$, and the proof that $\pi_{p=v} > \pi_{p>v}$ follows almost immediately. ■

The analysis is slightly more complicated when some consumers are cash-constrained. However, provided we impose some technical conditions, monopolistic merchants will continue to set $p = v$. We start with the case of a single payment scheme.

Lemma 12 *If the density function h is continuous on $[\underline{b}_B, \overline{b}_B]$, the surplus satisfies (B-1), and the proportion of cash-constrained consumers q is less than $1/(1 + B)$, then monopolistic merchants will set $p = v$ when there is a single payment scheme.*

Proof. We need to consider three possibilities: $p < v$, $p > v$, and $p = v$.

First, if a monopolistic merchant sets $p < v$ and accepts cards then a cardholding customer who is not cash constrained will buy the good using a card provided that $v + b_B - p - f \geq v - p$, which holds if and only if $b_B \geq f$, and will buy it using cash otherwise. A cardholding customer who is cash-constrained will buy the good using a card provided that $v + b_B - p - f \geq 0$, which holds if and only if $b_B \geq f + p - v$, and will not buy it otherwise. Customers who do not hold cards will buy the good using cash unless they are cash-constrained. In total, the merchant's profit equals

$$\pi = \left(1 - q + q\lambda(f) D(f + p - v) I\right)(p - d) + \left((1 - q) D(f) + q D(f + p - v)\right)\lambda(f)(b_S - m) I$$

Differentiating with respect to p, and using the assumption that the function $D(x) + (x + v - d + b_S - l) D'(x)$ is bounded below by $-B$, gives

$$\begin{aligned}\frac{d\pi}{dp} &= 1 - q + q\lambda(f)\Big(D(f + p - v) + (f + p + b_S - d - l) D'(f + p - v)\Big) I \\ &\geq 1 - q + q\lambda(f)(-B) I \\ &= 1 - q - q\lambda(f) B I.\end{aligned}$$

The last line is positive provided that

$$q < \frac{1}{1 + \lambda(f) B I},$$

a condition which is guaranteed by our assumption that $q < 1/(1 + B)$. Thus, a monopolistic merchant will not set $p < v$.

Second, suppose a monopolistic merchant sets $p > v$ and accepts cards. Since it is never optimal to buy the good with cash, a cardholding customer will buy the good using a card provided that $v+b_B-p-f \geq 0$, which holds if and only if $b_B \geq f+p-v$, and will not buy it otherwise, whether or not they are cash constrained. Customers who do not hold cards will not buy the good. Therefore, the merchant's profit equals

$$\pi = \lambda(f)D(f+p-v)(p-d+b_S-m)I.$$

As we showed in the proof of Lemma 10, the merchant can always achieve a higher profit by setting $p = v$. ■

The analysis is further complicated by the presence of competing payment schemes.

Lemma 13 *If the density function h is continuous on $[\underline{b}_B, \overline{b}_B]$, the surplus satisfies (B-1), and the proportion of cash-constrained consumers q is less than $1/(1+B)$, then monopolistic merchants will set $p = v$ when there are competing payment schemes.*

Proof. Suppose the merchant sets $p \leq v$. Then its profit equals

$$\begin{aligned}
\pi &= q\Big(\left(\lambda^1+\lambda^{12}L^1\right) D(f^1+p-v)(p+b_S-d-m^1)I^1 \\
&\quad + \left(\lambda^2+\lambda^{12}L^2\right) D(f^2+p-v)(p+b_S-d-m^2)I^2\Big) \\
&\quad + (1-q)\left(p-d+\left(\lambda^1+\lambda^{12}L^1\right) D(f^1)(b_S-m^1)I^1+\left(\lambda^2+\lambda^{12}L^2\right) D(f^2)(b_S-m^2)I^2\right).
\end{aligned}$$

Differentiating with respect to p, and using the assumption that the function $D(x)+(x+v-d+b_S-l)D'(x)$ is bounded below by $-B$, gives

$$\begin{aligned}
\frac{d\pi}{dp} &= 1-q+q\left(\lambda^1+\lambda^{12}L^1\right)\Big(D'(f^1+p-v)(p+b_S-d-m^1)+D'(f^1+p-v)\Big)I^1 \\
&\quad + q\left(\lambda^2+\lambda^{12}L^2\right)\Big(D'(f^2+p-v)(p+b_S-d-m^2)+D'(f^2+p-v)\Big)I^2 \\
&\geq 1-q+q\left(\lambda^1+\lambda^{12}L^1\right)(-B)I^1+q\left(\lambda^2+\lambda^{12}L^2\right)(-B)I^2 \\
&= 1-q-qB\left(\lambda^1I^1+\lambda^2I^2+\lambda^{12}(I^1L^1+I^2L^2)\right).
\end{aligned}$$

The last line is positive provided that

$$q < \frac{1}{1+B(\lambda^1I^1+\lambda^2I^2+\lambda^{12}(I^1L^1+I^2L^2))},$$

a condition which is guaranteed by our assumption that $q < 1/(1+B)$. Thus, a monopolistic merchant will not set $p < v$. If a monopolistic merchant sets $p > v$, its profit is

$$\begin{aligned}
\pi &= \lambda^1 D\left(f^1+p-v\right)\left(p-d+b_S-m^1\right)I^1+\lambda^2 D\left(f^2+p-v\right)\left(p-d+b_S-m^2\right)I^2 \\
&\quad +\lambda^{12}\left(D(f^1+p-v)\left(p-d+b_S-m^1\right)I^1L^1+D\left(f^2+p-v\right)\left(p-d+b_S-m^2\right)I^2L^2\right).
\end{aligned}$$

Using identical arguments to those contained in the proof of Lemma 12, it follows that

$$D(f^i+p-v)(p-d+b_S-m^i) < D(f^i)(v-d+b_S-m^i), \quad i = 1,\ 2,$$

whenever $p > v$, so that a monopolistic merchant will never set $p > v$. It follows that $p = v$ is the profit-maximizing price. ■