

**A Puzzle of Card Payment Pricing:**  
**Why Are Merchants Still Accepting Card Payments?**

Fumiko Hayashi<sup>1</sup>

First draft: December 28, 2004  
This version: September 6, 2005

**Preliminary: Please do not cite or quote without author's permission.**

**Abstract**

This paper presents models that explain why merchants accept payment cards even when the fees they face exceed the transactional benefits they receive from a card transaction. Such merchant behaviors can be explained by competition among merchants and/or the effectiveness of the merchant's card acceptance in shifting cardholders' demand for goods upward. The prevalent assumption used in payment card literature—merchants accept cards only when their transactional benefits are higher than the fees they pay—holds only for a monopoly merchant who faces an inelastic consumer demand. A card network that wants all merchants in a given industry to accept cards may set a lower merchant fee initially but in the long-run it will set the fee to the highest possible level, which may be higher than the sum of the merchant's transactional benefit and the merchant's initial margin without cards. Such merchant fees potentially create inequality between cardholders and non-cardholders. The paper also explores possible explanations for the recent gradual increases in merchant fees in the United States. Three possible explanations are 1) inflexible product price setting by merchants, 2) decreases (increases) in cardholder fees (rebates), and 3) increases in cardholding-customer proportion in a given industry.

---

<sup>1</sup>Fumiko Hayashi is a senior economist in the Payments System Research Department, Federal Reserve Bank of Kansas City, 925 Grand Boulevard, Kansas City, MO, 64198, e-mail: [fumiko-hayashi@kc.frb.org](mailto:fumiko-hayashi@kc.frb.org), Phone: (816) 881-6851. The author wishes to thank Antoine Martin, Mike Orlando, Richard Rosen, Marius Schwartz, Rick Sullivan, Zhu Wang, Julian Wright, and participants of IIOC 2005 for their comments, and Nathan Halmrast for outstanding research assistance. The views expressed in this article are those of the author and do not necessarily reflect those of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

## 1. Introduction

Credit and debit card payments are experiencing rapid growth. This rapid growth in card payments is attracting controversy and antitrust scrutiny in many countries. Recently, the European Commission and the Reserve Bank of Australia have issued their decisions on pricing of card payments.<sup>2</sup> Regulatory authorities in other countries, such as Mexico and the United Kingdom, are evaluating potential future regulations. At the center of those policymakers' attention are interchange fees, which are paid by the bank (called the acquirer) that processes the card transaction for the merchant to the bank (called the issuer) that has issued the payment card to the consumer. Interchange fees are typically set by the card network (or by its member financial institutions collectively) and in many instances they are considered by regulators to be too high.

Chart 1 shows the credit card interchange rates in Australia and the European Union (EU), before and after the reductions of interchange rates were forced. In these regions, the regulator's involvements on pricing of card payments lowered the credit card interchange rates significantly. The chart also includes the UK credit card interchange rates. In the United Kingdom, the antitrust authority has not decided whether it will regulate credit card interchange rates.<sup>3</sup> But if it does, then some industry observers anticipate the rates will be cut to 0.7 percent or lower.<sup>4</sup> Current U.S. credit card interchange rates are also included in the chart. As seen in the chart, the U.S. interchange rates are even higher than the rates in the three regions before the interchange rates were lowered.

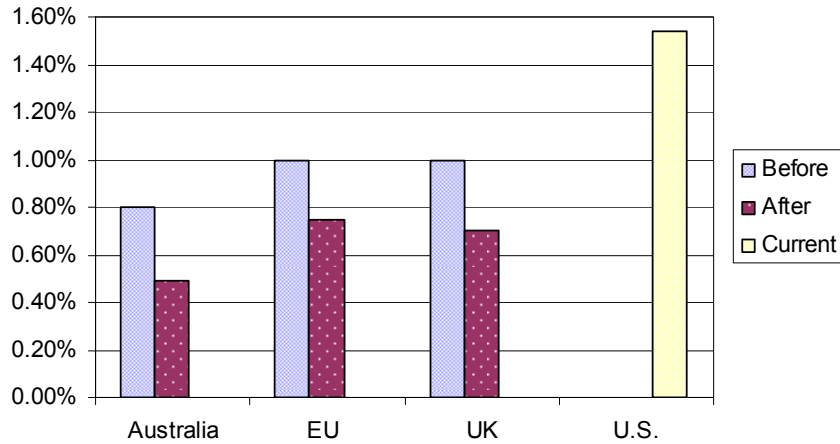
---

<sup>2</sup>While the Reserve Bank of Australia has regulated the credit card interchange rates, the European Commission has issued its decision to exempt Visa's interchange fees under European competition laws. Visa has voluntarily reduced its interchange rates for cross-border transactions within certain European Union countries.

<sup>3</sup>According to *The Herald*, November 11, 2004, a spokesman of the UK antitrust authority said "we would hope to issue a decision by the summer of next year."

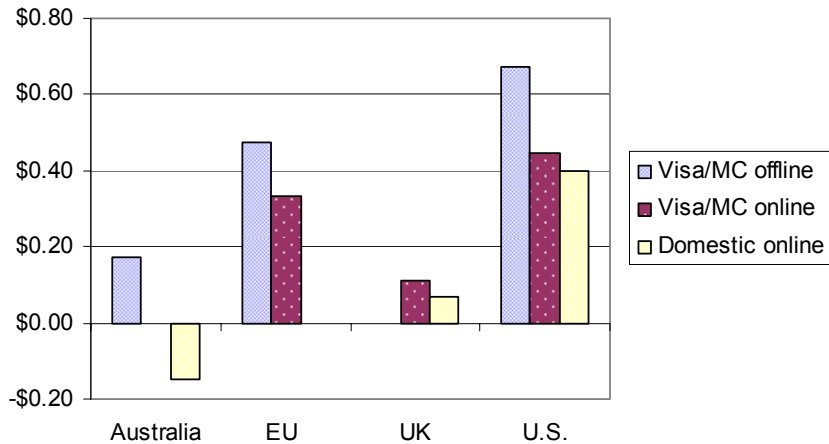
<sup>4</sup>Industry observers expect the interchange rates in the United Kingdom to be between 0.35 to 0.7 percent. See, for example, *The Times*, May 17, 2004 and November 11, 2004, and *Financial Times*, November 11, 2004.

Chart 1: Credit Card Interchange Rates in Selected Countries



Sources: Reserve Bank of Australia, Visa Europe, MasterCard International, and *American Banker*  
 Notes: “Before” = before the rate was forced to be lowered; “After” = after the rate was lowered; “Current” = as of November, 2004. In Australia, the regulation is effective for both Visa and MasterCard. The ‘before’ and ‘after’ rates are the average of Visa and MasterCard rates. In the EU, the European Commission made its decision on the Visa rate for cross-border transactions only. The ‘before’ rate is not publicly available, but the rate was estimated at about 1 percent according to the report “Credit Card Services” by the UK Monopolies and Mergers Commission and others. In the United Kingdom, the antitrust authority has not made a final decision on the credit card interchange rates as of November 2004. Therefore, the ‘after’ rate is an expectation by industry observers. They predict the regulated interchange rate will be between 0.35 to 0.7 percent. (See footnote 4.) The U.S. rate is the average of Visa and MasterCard rates.

Chart 2: Debit Card Interchange Fees for a \$50 Transaction in Selected Countries

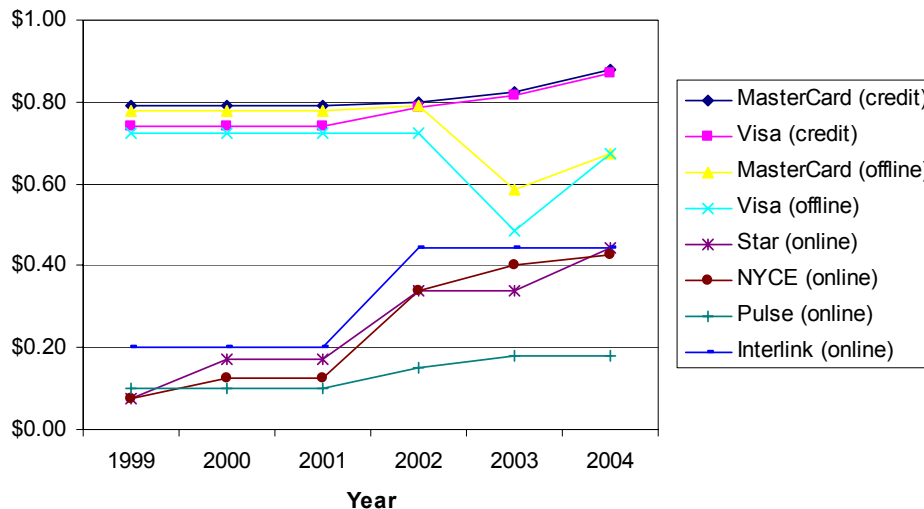


Sources: Reserve Bank of Australia, Visa Europe, MasterCard International, and *ATM & Debit News*  
 Notes: In Australia, there are no Visa/MasterCard online debit products. The domestic debit network’s interchange fees are paid by issuers to acquirers. For the EU, only cross-border debit products are listed. In the United Kingdom, the domestic debit network and MasterCard agreed in 2002 that the domestic products will be migrated into Maestro, the MasterCard online debit product by 2007. The rates shown were the 1998 rates. In the United States, there exist more than ten domestic online debit networks. The rate shown as domestic debit network is the weighted average of the top three domestic online debit interchange fees. To convert the currencies, the average exchange rates for the first three quarters of 2004 are used.

Chart 2 shows the debit card interchange fees in selected countries. Interchange fees for offline debit are higher than those for online debit.<sup>5</sup> In Australia, the online debit interchange fees go in the opposite direction—paid by the issuer to the acquirer. Although the chart does not include them, many countries have domestic debit schemes with zero interchange fees.<sup>6</sup> Similar to credit card interchange fees, debit card interchange fees in the United States are among the highest in the world.

In the United States, interchange fees for both credit and debit card transactions are among the highest in the world. Moreover, they have been increasing very rapidly for the past several years. Chart 3 shows the various interchange fees for a \$50 transaction at a retail store in the United States. Interchange fees for credit and online debit transactions have been increasing. Although the offline debit interchange rates were reduced after the settlement of a lawsuit by a group of merchants against the two offline debit networks in August 2003, they have increased since 2004.

Chart 3: Interchange Fees for a \$50 Transaction at Non-Supermarket: 1999-2004



Sources: *American Banker* and *ATM & Debit News*

<sup>5</sup>See Hayashi, Sullivan, and Weiner (2003) for the difference between online and offline debit.

<sup>6</sup>Those include Belgium, Canada, Denmark, Germany, Netherlands, and Switzerland.

Since interchange fees are a component of merchant fees charged by acquirers, increases in interchange fees result in increases in merchant fees.<sup>7</sup> More and more merchants have expressed their concerns with the fees and some even argue that the current fee level exceeds the benefits they receive in accepting cards. However, few merchants have stopped accepting card payments.

Two interesting questions arise from the experience of the U.S. payments industry. First, why do merchants keep accepting card payments even when, as they claim, merchant benefits in accepting cards are not higher than the fees they face? Merchants explain that the competitive pressure does not allow them to reject card payments and if they do so, they lose sales. Card networks, on the other hand, claimed that the benefits merchants receive today are greater than those in the past since, for example, card networks provide useful customer information that merchants can use to increase their profits. Therefore, the networks reason, merchants receive more benefits than fees (even though the fees today are higher than in the past), and if not they can drop their card acceptance. The networks also claimed that because of network externalities, payment cards generate more benefits since the number of merchants that accept payment cards has increased.

The second question is why interchange fees have been gradually increasing despite the fact that the payments industry experienced technological advances, intensified competition, and strong merchant dissatisfaction. A variety of sources reported that technological advances have reduced the costs of processing card transactions.<sup>8</sup> Interchange fees can be used as a device for

---

<sup>7</sup>In the United States, acquirers completely pass through the interchange fee to merchants. Typically the interchange fee accounts for more than 75 percent of the merchant fee.

<sup>8</sup>See, for example, Evans and Schmalensee, *Paying with Plastic*, the MIT Press, 1999, p. 130, *ATM & Debit News*, February 12, 2004, and April 22, 2004, and “Murky Future for Interchange” *European Card Review*, September/October 2004.

shifting costs between issuers and acquirers.<sup>9</sup> If the interchange fee is used to cover all or part of the issuer's costs for providing card services, the reduction of some of their costs may reduce the fees. Intensified competition has been observed among card networks, issuers, acquirers, and processors, although some of them have greater market power than the others.<sup>10</sup>

The literature on payment cards has been growing, but only a few studies have analyzed merchant card acceptance explicitly. Most studies simply assume either that merchants accept cards when their transactional benefits from cards exceed the merchant fees or that merchants accept cards regardless of the level of merchant fees.<sup>11</sup> Therefore, one cannot use these models to answer the first question above. Some exceptions are Rochet and Tirole (2002), Chakravorti and To (2003), Guthrie and Wright (2003), and Wright (2003a, 2003b, 2004). Rochet and Tirole (2002), Guthrie and Wright (2003), and Wright (2003b, 2004) have found that if merchants compete against each other, they accept cards as long as the merchant fees do not exceed the sum of the merchant's transactional benefits and the cardholder's average net transactional benefits from cards.<sup>12</sup> Chakravorti and To (2003) focus on the credit card's revolving function and explain that competing merchants accept credit cards because it allows them to make sales to illiquid customers today rather than to wait for uncertain sales tomorrow. Their model also concludes that although each merchant chooses to accept cards, since all merchants accept cards in equilibrium each merchant's profit is lower.

---

<sup>9</sup>See, for example, Schmalensee (2002).

<sup>10</sup>In the ruling in the Department of Justice's antitrust case against Visa and MasterCard, the court found that these two card networks have market power in payment card markets. United States Court of Appeals for the Second Circuit, *United States of America v. Visa USA, Visa International Corp, and MasterCard*, August Term, 2002.

<sup>11</sup>Models with the assumption that merchants only accept cards when their transactional benefits from cards exceed the fees include Baxter (1983), Schmalensee (2002), and Bolt and Tieman (2003). In contrast, Frankel (1998), Gans and King (2002), Katz (2001), and Schwartz and Vincent (2004) assume that merchants accept cards regardless of the level of merchant fees, although they either implicitly or explicitly assume that the merchant fees need to ensure a certain level of profit for the merchants.

Chakravorti and To (2003) explain the merchants' credit card acceptance but do not explain their debit card acceptance. The other models explain both credit and debit card acceptance by merchants even when merchant fees are higher than their transactional benefits, however, some of their results seem to contradict what has been seen in the United States. For example, in Wright (2003b), merchants can increase their profit margins by accepting cards, which contradicts what some of the merchants are saying. In contrast, in Rochet and Tirole (2002), merchants cannot increase their margins by accepting cards. Their model, however, expects that the maximum merchant fee decreases as competition among issuers intensifies.<sup>13</sup>

For the second question, traditional economic theories suggest that technological advances and competition will reduce prices.<sup>14</sup> Since the payment industry is characterized as a two-sided market, some of the traditional theories that analyze one-sided markets are not applicable.<sup>15</sup> Nevertheless, a majority of previous literature on payment card networks predicts that network competition will reduce or at least will not raise interchange fees and/or merchant fees.<sup>16</sup> This is so, partly because most of the studies assume that merchants accept cards only when the merchant fees do not exceed their transactional benefits, and partly because most of them do not model issuers' behavior when the issuers decide which networks they will join.

---

<sup>12</sup>Guthrie and Wright (2003) show another case. Even though merchants are competing against each other, if consumers hold more than one branded card and they receive identical benefits from those cards, merchants may not accept the cards whose merchant fee does not exceed their transactional benefits to the merchants.

<sup>13</sup>In the United States, while the top ten credit card issuers' market share in terms of managed receivable is high and has increased from 79.4 percent in 2000 to 85.7 percent in 2003, the top ten debit card issuers' market share is low and steady. According to *EFT Data Book* (various years), both the online and offline debit issuers' market shares either in terms of number of cards issued or in terms of transaction volume have been around 30 to 35 percent since 2000.

<sup>14</sup>The interchange fee is not a price per se, but since the interchange fee is a major component of the merchant fee, its increase (or decrease) greatly affects the increase (or decrease) in the merchant fee. See also footnote 7.

<sup>15</sup>See, for example, Evans (2001) for the difference between one-sided and two-sided markets.

<sup>16</sup>Guthrie and Wright (2003) is an exception. When consumers hold at most one branded card, network competition may increase interchange fees. Other studies that analyze competition among card networks include Manenti and Somma (2002), Bolt (2003), Rochet and Tirole (2003), and Chakravorti and Roson (2004).

To answer the first question, this paper analyzes the merchant incentives to accept payment cards—both credit and debit cards. Four different markets that are characterized by merchant competitiveness (monopoly or a Hotelling model of competition) and by price elasticity of the market aggregate consumer demand (inelastic or elastic) are considered in turn. In the model, there is one card network that determines both the merchant fees and cardholder fee.<sup>17</sup> The network is assumed to be conservative in the sense that it sets a merchant fee so that all of the merchants in a given industry accept cards. Merchants are assumed to decide their card acceptance and to determine product prices they sell. The model assumes that the merchant's price setting is completely flexible as long as the merchant sets the same price for card users and cash users.<sup>18</sup> The assumption of the completely flexible price setting by merchants will be relaxed when exploring answers for the second question.

The results suggest that only monopoly merchants who are facing an inelastic consumer demand do not accept cards if the fees exceed the merchant's transactional benefits. In the other three markets, merchants accept cards even when the fees exceed their transactional benefits. As previous studies found, competing merchants accept cards for strategic reasons. Merchants initially hope that their card acceptance can lure customers away from their rivals, but later on they accept cards to keep their current customers. Even monopoly merchants accept cards when their transactional benefits are lower than the fees they pay if they face an elastic consumer demand. They do so not because they have a strategic reason but because card acceptance shifts their cardholder customers' demand upward and thus brings in incremental sales.

The paper also analyzes the welfare of cardholders, non-cardholders, and merchants separately because policymakers may want to evaluate the card pricing from both efficiency and

---

<sup>17</sup>In most networks, merchant fees are determined by each individual acquirer and cardholder fees are determined by each individual issuer. However, the fees do not vary very much by acquirer or issuer.



equity points of view. In comparison with the equilibrium without cards, if the network charges the highest merchant fee then cardholders are better off (or at least indifferent), non-cardholders are worse off, and merchants are either better off or indifferent. The total of the consumers' and merchants' surplus depends on price elasticity of the market aggregate consumer demand. In markets where the aggregate consumer demand is inelastic, the total of the consumers' and merchants' surplus with and without cards are the same.

The analysis above, however, does not provide explanations why interchange fees and thus merchant fees have been gradually increasing. To explore possible explanations, the model relaxes some of the assumptions made above. The recent gradual increase in interchange fees in the United States may be explained by inflexibility in price setting by merchants, gradual decreases (increases) in cardholder fees (rebates), and/or increases in cardholding customer proportion in a given industry.

The rest of the paper is organized as follows. Section 2 develops the model. Section 3 presents the highest merchant fee and welfare analysis of four different cases in turn. Section 4 explores possible explanations for the gradual increases in interchange fees and Section 5 concludes.

## **2. The model**

For simplicity, only two payment instruments are available in the model. They are card and cash. Each industry is small enough so that the card acceptance by the merchants in the industry does not affect the consumer's cardholding decision and therefore a fixed ( $\alpha$ ) percent of the customers for the industry hold a card.

---

<sup>18</sup>Due to no-surcharge rules imposed by card networks.

A card transaction brings transactional benefits to both the card users and the merchants who accept those cards. One of the transactional benefits of cards is reducing transactional costs associated with cash transactions for both consumers and merchants. When consumers pay with cash, the consumers incur some transactional costs besides the price of goods or services they purchase and the merchants also incur some transactional costs in addition to the costs of selling goods or services. Generally, transactional costs with cash for consumers include costs of obtaining cash (such as ATM fees and time to go to the bank) and risks associated with cash (such as theft); those for merchants include handling cash transactions (such as costs of labor, armed car, and bank fees).

Some transactions, such as those involving reservations, require more costs to consumers and merchants if they are paid by cash. Consider a hotel reservation: although the reservation may not be impossible with cash, cash is inconvenient if a customer wants a late arrival guarantee as well as an ability to cancel the reservation without paying fees. The hotel may want some advance payments from the cash customer to keep the room until he or she arrives, but it may hesitate to do so since it needs to send the cash back on the chance that if the customer cancels before the date that a cancellation fee is assessed. So, either the cash customer cannot obtain a late arrival guarantee or the hotel may lose business should the cash customer not show up. Card payment can diminish such inconvenience or loss. At the time of reservation, a card customer gives the hotel his or her card number. If the card customer cancels the reservation early, the hotel does not charge anything to the customer's card account. The hotel can keep the room until he or she arrives, but in case the customer does not show up without canceling the reservation, the hotel can simply charge a cancellation fee to the account.

Although there may exist some other transactional benefits of cards, the model assumes that card transactions reduce both card users' and merchants' transactional costs to zero but create no other transactional benefits to either merchants or card users. Card transactions require per transaction fees to cardholders,  $f$ , (if rewards instead of fees, then  $f$  is negative) and to merchants,  $m$ . For consumers, the true cost of purchasing a good or service is  $p + t_c$  with cash and  $p + f$  with a card, where  $p$  is the product price charged by the merchant and  $t_c$  is the transactional cost for cash users. For merchants, the true cost of selling a good or service is  $d + t_m$  with cash and  $d + m$  with a card, where  $d$  is the cost of selling a product regardless of the payment methods used for the transaction and  $t_m$  is the merchant's transactional costs with cash.

Assume that transactional costs with cash for consumers and merchants do not vary by each individual consumer or merchant but vary by industry. Thus, consumers who purchase goods from the same industry incur the same transactional costs with cash, and merchants who are in the same industry incur the same transactional costs with cash. Some empirical evidence can justify this assumption. The consumer payment study by Dove Consulting (2001) showed that the consumers' payment mix varies by the type of store. Hayashi and Klee (2003) also found that a consumer's payment choice depends on the transaction characteristics, such as average transaction value and physical characteristics at the points of purchase. Although larger merchants in a given industry may pay lower merchant fees than their smaller counterparts, typically U.S. merchant fees vary by the type of industry.<sup>19</sup> This assumption is also adopted by

---

<sup>19</sup>Both credit card networks and debit card networks in the United States set different interchange fees according to industry types. Typically, debit card networks have two fees: one for supermarkets and the other for non-supermarkets. Visa and MasterCard credit card networks set different fees for different categories of industry, such as supermarkets, general retail markets, emerging retail markets, hotel and car rentals, automated fuel dispenser, and so on.

some of the previous studies. Wright (2004) assumes that card benefits vary by industry. Wright (2000) and Katz (2001) assume that without cards, some of the transactions extract the consumer's entire surplus. Such transactions may happen only in particular industries.

Merchants are assumed to set the same price for cash users and card users. The network imposes the no-surcharge rule, which prohibits merchants from charging a higher price to card users. Although the rule does not prohibit merchants from giving cash discounts, merchants likely choose to set the same prices since setting different prices may be costly.<sup>20</sup> In some countries, a surcharge has been allowed, but only a small percentage of merchants actually set different prices for card users and the rest of the consumers.<sup>21</sup>

Four different cases are analyzed and these four are characterized by consumer demand for goods—either elastic or inelastic—and merchant market competitiveness—either monopoly or competition according to the Hotelling model. Each industry faces either an elastic or inelastic aggregate consumer demand for goods. In the industry that faces an elastic demand, the consumer's quantity demanded depends on the consumer's true costs of purchasing goods or services rather than the price itself. Let us think about the hotel industry again. As explained above, making reservations with cash will likely be more costly for both customers and hotels than doing so with cards. Since cards reduce such transactional costs for customers, even when the price of the hotel stay does not change, there will either be more travelers or the current travelers will travel more. Industries that face inelastic demand, on the other hand, cannot shift industry aggregate demand by simply accepting payment cards. The consumer home appliances industry is an example of this. Each household needs at most a certain number of refrigerators or washers and dryers. Since these are large-dollar items, accepting cards reduces the consumer's

---

<sup>20</sup>Setting different prices may increase some administration costs or customer dissatisfaction.

<sup>21</sup>Those countries include Australia, Netherlands, and the United Kingdom.

risk or inconvenience associated with cash transactions. However, the consumers do not buy more refrigerators or washers and dryers because the merchants accept card payments. To simplify the analysis, each individual cardholder's and each individual non-cardholder's demand functions are assumed to be identical.<sup>22</sup>

In each industry, there is either a monopoly merchant or two merchants competing according to the Hotelling model. Merchants decide not only their card acceptance strategy but also their prices. In the next section, merchants are assumed to have complete flexibility in setting their product prices. Section 4 considers the case where merchants face some restrictions when they adjust their prices upward, since they may be afraid of antitrust scrutiny or customers' dissatisfaction toward their price increases.

The model assumes that there exists only one card network, which sets both merchant fees and the cardholder fee.<sup>23</sup> The network sets a unique merchant fee in a given industry and a universal cardholder fee. Although a variety of network objectives, such as maximizing the total members' profits, balancing the profits or costs between acquirers and issuers, or maximizing the total transaction volume, are suggested, since the paper focuses on the merchant card acceptance behavior, the network is assumed to have an incentive to set merchant fees as high as possible. The network is also assumed to be conservative when it sets merchant fees. It sets a merchant fee so that all of the merchants in a given industry accept cards.

It is assumed that consumers are either cardholders or non-cardholders. Non-cardholders need to pay with cash all the time. On the other hand, cardholders can choose their payment method—either cash or a card. Consumers observe all the decisions the merchants have made

---

<sup>22</sup>This assumption may not be realistic. The average income of credit card holders is higher than the average income of non-credit card holders. Thus demand functions are likely different.

<sup>23</sup>Closed system networks as well as open system networks can influence both merchant fees and cardholder fees. See, for example, Schwartz and Vincent (2004) and Wright (2003a).

(card acceptance and new price) before they determine their payment methods, the merchants from which they make purchases, and quantities (if demand is elastic).

Table 1: Timing of the Game

Period	Stage	Actions
0		Neither the monopoly merchant nor the duopoly merchants accept cards. They set the product price, $p^0$ so as to maximize their profits.
1	1	Given $p^0$ , the network sets the cardholder fee, $f$ , and the industry-specific merchant fee, $m$ .
	2	The merchant decides whether to accept cards or not, and determines the new price, $p^1$ , if it chooses to accept cards.
	3	The consumer decides from which merchant he or she makes purchases (if duopoly), which payment method he or she uses (if a cardholder), and how much he or she purchases (if demand is elastic).
N		Given $p^{N-1}$ , repeat the game in period 1.

The timing of the game is described in Table 1. We consider a multi-period game and in each period, there are three stages: at stage 1 the network sets a merchant fee and a cardholder fee, at stage 2 the merchants decide whether to accept cards and their prices, and at stage 3 the consumers decide from which merchant they make purchases, which payment method they use, and how much they purchase if demand is elastic.

### 3. Merchant acceptance and the highest possible merchant fees

#### 3.1 Monopoly merchant

##### 3.1.1. Inelastic demand

In period 0, for whatever reason the monopoly merchant does not accept cards. All of its customers (mass 1) pay with cash. Assume that each customer makes one transaction. The merchant profit in period 0 is:

$$\pi^0 = (p^0 - d - t_m), \quad (1)$$

where  $p^0$  is the product price the monopoly merchant charges in period 0. Assume that each consumer receives gross benefit,  $v$ , by purchasing one unit of the product. Since the monopoly merchant extracts the consumer's entire surplus, it sets the product price  $p^0$  so that:

$$p^0 = v - t_c. \quad (2)$$

Starting with stage 3 of period 1, a cardholder will use a card if  $t_c > f$ , and if the merchant accepts cards since the merchant sets a unique product price for both cash users and card users. That is, by using a card, the cardholder can save  $t_c - f > 0$  per transaction as opposed to paying with cash.

At stage 2 of period 1, given the merchant fee,  $m$ , charged by the network, the merchant decides whether to accept cards or not and determines the product price. The merchant's profit if it does not accept cards is the same as the profit in period 0. Two different levels of the merchant's profit from accepting cards are possible. One of them is the profit without changing the product price from  $p^0$ . Since the merchant has already extracted the cash users' entire surplus, raising the price from  $p^0$  implies that the merchant loses all cash customers. The other level of the merchant's profit is that of when the merchant extracts the card users' entire surplus. By doing so, the merchant raises the price to:

$$p^{1*} = v - f. \quad (3)$$

At this price, the merchant extracts the card users' entire surplus but loses the entire profit from cash users.

Given  $m$ , the merchant's maximum profit in period 1 is defined as:

$$\begin{aligned} \pi^c &= (v - t_c - d - t_m)(1 - \alpha) + (v - t_c - d - m)\alpha, & \text{if } (v - t_c - d - t_m)(1 - \alpha) \geq (t_c - f)\alpha, \\ &= (v - f - d - m)\alpha, & \text{otherwise.} \end{aligned} \quad (4)$$

From equations 1 and 4,  $\pi^c \geq \pi^0$  always holds as long as  $t_m \geq m$ . This implies that the merchant always accepts cards if the merchant fee is lower than the merchant transactional costs with cash.

From equation 4, the merchant may accept cards even when the merchant fee exceeds the merchant transactional costs with cash, if raising the price increases the merchant's profit.

At stage 1 of period 1, the network sets a cardholder fee,  $f$ , which should be lower than the transactional costs with cash for consumers,  $t_c$ . The highest possible merchant fee in period 1 is

$$\begin{aligned} \bar{m}^1 &= t_m, & \text{if } (v - t_c - d - t_m)(1 - \alpha) &\geq (t_c - f)\alpha, \\ &= t_m + t_c - f - \frac{1 - \alpha}{\alpha}(v - t_c - d - t_m), & \text{otherwise.} \end{aligned} \quad (5)$$

**Proposition 1 (Price and Welfare):** *Suppose the network sets  $f < t_c$  and  $m = \bar{m}^1$ . A monopoly merchant who faces an inelastic consumer demand accepts cards. Compared to the equilibrium without the cards,*

- a) *The merchant raises the product price if the cardholder's per transaction fee is low relative to his or her transaction benefits and/or the cardholder base is high enough;*
- b) *Non-cardholders' surplus is lower if the merchant raises the price, otherwise it is unchanged;<sup>24</sup>*
- c) *Cardholders' surplus is the same if the merchant raises the price, otherwise it is higher;*
- d) *Aggregate consumer surplus is unchanged if the merchant raises the price, otherwise it is higher;*
- e) *The merchant's surplus is the same.*



Suppose, in period 2, all exogenous variables, such as  $t_m$ ,  $t_c$ ,  $\alpha$ ,  $d$ , and  $v$ , are the same as in period 1. Can the network raise the merchant fee higher than the highest possible merchant fee in period 1, without changing the cardholder fee? Obviously, if the merchant did not raise the price in period 1, the network cannot raise the merchant fee, because the game in period 2 is exactly the same as the game in period 1. In this case, the highest possible merchant fee in period 2,  $\bar{m}^2$ , is the same as  $\bar{m}^1$ . Even if the merchant raised the price in period 1, the network cannot raise the merchant fee. Since the merchant can always go back to the original strategy (to not accept cards and to set the original price,  $p^0$ ), the new merchant fee needs to guarantee at least the same profit as the merchant's profit without accepting cards. Since the merchant has already extracted the card users' entire surplus in period 1, the merchant cannot raise its price anymore. The network set the merchant fee in period 1 so that the merchant profit with and without accepting cards breaks even and therefore the network cannot raise the merchant fee. Hence, the long-run equilibrium merchant fee will be reached from the first period the merchant accepts cards.

### 3.1.2 Elastic demand

In period 0, the monopoly merchant does not accept cards. All of its customers (mass 1) pay with cash. Assume that each individual customer's demand function  $D(p + t_c)$  is identical regardless of his or her cardholding. The merchant's profit function in period 0 is:

$$\pi^{nc}(p) = (p - d - t_m)D(p + t_c). \quad (6)$$

---

<sup>24</sup>If a consumer has no disutility from not purchasing any goods, non-cardholders' surplus is unchanged. Otherwise it is lower.

The monopoly merchant sets the product price so as to maximize its profit. The price in period 0 is defined as:

$$\begin{aligned} p^0 &= \operatorname{argmax} \pi^{nc}(p) \\ &= d + t_m - \frac{D(p^0 + t_c)}{D'(p^0 + t_c)} \end{aligned} \quad (7)$$

Starting with stage 3 of period 1, a cardholder will use the card if  $t_c > f$  and if the merchant accepts cards. At stage 2 of period 1, the merchant decides whether to accept cards; if it decides to accept, it adjusts the product price. The merchant's profit if it does not accept cards is the same as the profit in period 0. Given a merchant fee,  $m$ , the merchant's profit function if it accepts cards is defined as:

$$\pi^c(p; m) = (p - d - t_m)(1 - \alpha)D(p + t_c) + (p - d - m)\alpha D(p + f). \quad (8)$$

The merchant's profit maximizing price,  $p^*(m)$ , when it accept cards is defined as:

$$p^*(m) = d + t_m - \frac{(1 - \alpha)D(p^*(m) + t_c) + \alpha D(p^*(m) + f) - \alpha(m - t_m)D'(p^*(m) + f)}{(1 - \alpha)D'(p^*(m) + t_c) + \alpha D'(p^*(m) + f)}. \quad (9)$$

The merchant accepts cards if and only if  $\pi^c(p^*(m); m) \geq \pi^{nc}(p^0)$ .

At stage 1 of period 1, the network sets a cardholder fee,  $f$ , which should be lower than  $t_c$ . By assumption, the network will set the merchant fee as high as possible under the constraint of  $\pi^c(p^*(m); m) \geq \pi^{nc}(p^0)$ . Since general results are not available for the case of elastic consumer demand, the rest of this subsection restricts its attention to linear demand,

$D(x) = a - bx$ . Since  $\pi^c(p^*(m); m)$  is a monotone decreasing function of  $m$ , the highest merchant fee in period 1 set by the network,  $\bar{m}^1$ , satisfies  $\pi^c(p^*(\bar{m}^1); \bar{m}^1) = \pi^{nc}(p^0)$ . That is:

$$\bar{m}^1 = t_m + \left(\frac{2}{\alpha} - 1\right)(t_c - f) + \frac{2}{\alpha}A - \frac{2}{\alpha}\sqrt{(1 - \alpha)(t_c - f)^2 + 2(1 - \alpha)(t_c - f)A + A^2}, \quad (10)$$

where  $A = \frac{1}{2b} \{a - b(d + t_m + t_c)\}$ . Equation 10 yields the following proposition.

**Proposition 2 (Merchant fee):** *The monopoly merchant who faces an elastic consumer demand accepts cards even if the merchant fee is higher than the merchant transactional benefit from cards.*

**Proof:** Clearly, the second and third terms of equation 10 are positive and the absolute value of the last term is smaller than the sum of the second and third terms. Thus,  $\bar{m}^1 > t_m$ . ■

Wright (2003b) analyzes merchant card acceptance in the industry where aggregate consumer demand is elastic and merchants are competing according to a Cournot model. The model also assumes that each consumer buys at most one unit in the industry. In such an industry, merchants will accept cards even when the merchant fee is higher than their transactional benefit from cards. In my model, each consumer has an elastic demand and buys as many units as he or she wants from the same merchant. Although a monopoly merchant does not have to compete, the merchant will accept cards even when the merchant fee is higher than their transactional benefit. This is because the merchant can make the cardholder's demand curve shift upward by accepting cards and the resulting incremental sales to cardholders bring more profit to the merchant.

Proposition 3 compares the product price with and without the cards and Proposition 4 summarizes the consumers' and merchant's welfare.

**Proposition 3 (Price):** *A monopoly merchant who faces a linear consumer demand will adjust the product price if it accepts cards.*

i) For  $t_m - (t_c - f) \leq m \leq \bar{m}^1$ , the product price with cards is higher.

ii) For  $m < t_m - (t_c - f)$  the product price with cards is lower.

**Proof:** From equations 7 and 9,

$$p^*(m) - p^0 = \frac{\alpha}{2} \{(m - t_m) + (t_c - f)\}. \quad (11)$$

By assumption,  $t_c > f$ . Obviously,  $p^*(m) > p^0$  when  $m \geq t_m - (t_c - f)$ , otherwise

$$p^*(m) \leq p^0. \quad \blacksquare$$

**Proposition 4 (Welfare):** *Suppose the network sets  $f < t_c$  and  $m = \bar{m}^1$ . A monopoly merchant who faces a linear demand accepts cards. Compared to the equilibrium without cards:*

- a) *Non-cardholders' surplus is lower;*
- b) *Cardholders' surplus is higher;*
- c) *Aggregate consumer surplus is higher;*
- d) *The merchant's surplus is the same.*

**Proof:** The consumers' surplus is measured by using the Marshallian demand curve. Non-cardholders' surplus depends only on the product price. The higher the price, the lower the non-cardholders' surplus is. From Proposition 3, when the network sets the highest merchant fees, the merchant sets a higher product price than that without cards. Therefore, non-cardholders are worse off.

Cardholders' surplus depends on the product price, the cardholder's per transaction fee, and the consumer's transactional costs with cash. The total costs for a cardholder when he or she uses a card,  $p^*(\bar{m}^1) + f$ , is always lower than the total costs for the cardholder when the merchant does not accept cards,  $p^0 + t_c$ . From equation 10:

$$\bar{m}^1 - t_m = t_c - f - \varepsilon, \quad (12)$$

for  $\varepsilon > 0$ . Equations 11 and 12 yield:

$$p^*(\bar{m}^1) = p^0 + \alpha(t_c - f) - \frac{\alpha\varepsilon}{2}. \quad (13)$$

From equation 13:

$$(p^*(\bar{m}^1) + f) - (p^0 + t_c) = -(1 - \alpha)(t_c - f) - \frac{\alpha\varepsilon}{2} < 0.$$

Denote  $v(\cdot)$  as a consumer's surplus. Aggregate consumers' surplus without cards,  $v(p^0 + t_c)$ , is lower than aggregate consumers' surplus when the merchant accepts cards and sets the product price at the highest level,  $(1 - \alpha)v(p^*(\bar{m}^1) + t_c) + \alpha v(p^*(\bar{m}^1) + f)$ .

$$\begin{aligned} & (1 - \alpha)v(p^*(\bar{m}^1) + t_c) + \alpha v(p^*(\bar{m}^1) + f) - v(p^0 + t_c) \\ &= \frac{\alpha}{2} \left\{ \varepsilon D(p^0 + t_c) + \frac{\alpha b \varepsilon}{4} + (1 - \alpha)b(t_c - f)^2 \right\} > 0 \end{aligned}$$

By assumption, the network sets the merchant fee so that the merchant's profit with cards is equal to the profit without cards. Therefore, the merchant's surplus is the same. ■

Propositions 3 and 4 imply that the merchant's card acceptance impacts the welfare of non-cardholders, cardholders, and the merchant differently. When the network sets the highest merchant fees, cardholders are better off, non-cardholders are worse off, and the monopoly merchant is indifferent. Even when the network sets the merchant fee lower than the merchant transactional benefit, non-cardholders are likely worse off. If cardholders' net transactional benefit from cards,  $t_c - f$ , is greater than the merchant's transactional benefit from cards,  $t_m$ , non-cardholders are always worse off. This result supports the findings of Schwartz and Vincent (2004). They compare the equilibrium with and without the no-surcharge rules and show that when a monopoly network charges a profit maximizing merchant fee from a monopoly merchant, cash users' (non-cardholders') surplus is lower under no-surcharge than under surcharge. Since cash users' surplus at the equilibrium without cards and that at the equilibrium with cards under

surcharge are the same, my results suggest that even when the network charges a merchant fee that is lower than the profit-maximizing level, cash users' surplus is lower under no-surcharge than under surcharge.

In this market, cardholders' welfare gain surpasses non-cardholders' welfare loss from card acceptance, and therefore consumers' aggregate surplus is always higher with cards than without cards. In contrast, in markets with a monopoly merchant who faces an inelastic demand, consumers' aggregate surplus is not necessarily higher with cards than without cards.

Now, let us consider the game in period 2. Suppose in period 1 the network sets the merchant fee,  $\bar{m}^1$ . Suppose also that all exogenous variables in period 2 are the same as those in period 1. Can the network raise the merchant fee higher than  $\bar{m}^1$ , without changing the cardholder fee? Suppose the network can raise the merchant fee from  $\bar{m}^1$  to  $\bar{m}^2$ . Since the merchant can always go back to the original strategy, the new merchant fee,  $\bar{m}^2$ , needs to satisfy the condition  $\pi^c(p^*(\bar{m}^2); \bar{m}^2) = \pi^{nc}(p^0)$ . Since  $\pi^c(p^*(m); m)$  is a monotone decreasing function of  $m$ ,  $\pi^c(p^*(\bar{m}^2); \bar{m}^2) < \pi^c(p^*(\bar{m}^1); \bar{m}^1)$ , but this contradicts  $\pi^c(p^*(\bar{m}^1); \bar{m}^1) = \pi^{nc}(p^0)$ . Therefore, the network cannot raise the merchant fee higher than  $\bar{m}^1$ . Similar to the case of the monopoly merchant who faces an inelastic demand, the long-run equilibrium will be reached from the first period the monopoly merchant accepts cards.

### 3.2 Duopoly merchants—Hotelling competition

Additional assumptions are made for this section. There are two merchants, Merchant A and Merchant B, who are competing with each other according to the Hotelling model. Consumers (mass 1) are uniformly distributed on the interval of  $[0, 1]$ , which is independent of their cardholding. Merchant A is located at point 0 and Merchant B is located at point 1. For the

consumer located at point  $x$ , where  $0 \leq x \leq 1$ , the transportation cost to Merchant A is  $tx$ , and the transportation cost to Merchant B is  $t(1-x)$ .

### 3.2.1 Inelastic demand

Before considering the equilibrium with cards, first we describe the equilibrium without cards (equilibrium in period 0). As is usual in the symmetric Hotelling model, the equilibrium prices,  $(p_A^0, p_B^0)$ , are the same for both merchants and are equal to the merchant's marginal cost,  $d + t_m$ , plus the transportation cost,  $t$ :

$$p_A^0 = p_B^0 = p^0 = d + t_m + t. \quad (14)$$

Each merchant's profit is equal to the margin times the market share:

$$\pi_A^0 = \pi_B^0 = \pi^0 = \frac{t}{2}. \quad (15)$$

Now, let us consider the equilibrium in period 1. At stage 3 of period 1, a non-cardholder chooses the merchant based on the prices. If  $t_c > f$ , a cardholder chooses the merchant based on the merchants' card acceptance and their prices. If the cardholder chooses the merchant who does not accept cards he or she uses cash, otherwise the cardholder pays with a card. If  $t_c \leq f$ , the cardholder acts as if he or she is a non-cardholder.

Suppose  $t_c > f$ . At stage 2 of period 1 the merchants decide whether to accept cards and determine the product prices. First, let us describe the cases where both merchants take the same card acceptance strategy. Suppose that the merchants accept cards. Given the other merchant's price  $p_j$ , merchant  $l$ 's profit function is defined as:

$$\pi_l^c(p_l; p_j) = (p_l - d - t_m)(1 - \alpha)\left(\frac{1}{2} + \frac{p_j - p_l}{2t}\right) + (p_l - d - m)\alpha\left(\frac{1}{2} + \frac{p_j - p_l}{2t}\right), \text{ for } l=A, B \quad (16)$$

Merchant  $l$ 's price in period 1 is:

$$p_l(l : \text{accept}; j : \text{accept}) = d + t_m + t + \alpha(m - t_m). \quad (17)$$

Merchant  $l$ 's profit is:

$$\pi_l(l : \text{accept}; j : \text{accept}) = \frac{t}{2}. \quad (18)$$

Suppose instead that the merchants do not accept cards. In this case, the equilibrium is the same as the equilibrium without cards.

Next, let us describe the cases where each merchant takes a different card acceptance strategy from its rival's. Suppose that Merchant A accepts cards and Merchant B does not. Each merchant's profit function depends on  $t_c - f$  and  $t$ . If the ratio of a cardholder's net benefit from a card transaction to the transportation costs,  $(t_c - f)/t$ , is large, the cardholder's merchant choice does not depend on the price difference but depends on the difference of merchant card acceptance. The analysis below focuses on  $(t_c - f)/t \leq 2$ .<sup>25</sup>

Merchant A's profit function is defined as:

$$\pi_A^c(p_A; p_B) = (p_A - d - t_m)(1 - \alpha)\left(\frac{1}{2} + \frac{p_B - p_A}{2t}\right) + (p_A - d - m)\alpha\left(\frac{1}{2} + \frac{p_B - p_A}{2t} + \frac{t_c - f}{2t}\right), \quad (19)$$

and Merchant B's profit function is:

$$\pi_B^{nc}(p_B; p_A) = (p_B - d - t_m)\left\{(1 - \alpha)\left(\frac{1}{2} + \frac{p_A - p_B}{2t}\right) + \alpha\left(\frac{1}{2} + \frac{p_A - p_B}{2t} - \frac{t_c - f}{2t}\right)\right\}. \quad (20)$$

The product price and profit for each of the merchants are the following:

$$p_A(\text{accept}; \text{reject}) = d + t_m + t + \frac{\alpha}{3}\{(t_c - f) + 2(m - t_m)\} \quad (21)$$

---

<sup>25</sup>For  $(t_c - f)/t > 2$ , the two merchants differentiate their card acceptance behavior even when the merchant fee is low. One of the merchants concentrates on either cardholder customers or non-cardholder customers and sets monopoly price. The other merchant serves the other group of customers and sets a higher price by taking advantage of its rival's monopoly price. More generally, if  $n$  merchants are competing against each other and as long as  $(t_c - f)/t \leq n$ , such differentiated card acceptance behavior is not profitable to the merchants.



$$p_B(\text{reject}; \text{accept}) = d + t_m + t - \frac{\alpha}{3} \{(t_c - f) - (m - t_m)\} \quad (22)$$

$$\pi_A(\text{accept}; \text{reject}) = \frac{1}{2t} \left[ \left\{ t + \frac{\alpha}{3}(t_c - f) - \frac{\alpha}{3}(m - t_m) \right\}^2 - \alpha(1 - \alpha)(t_c - f)(m - t_m) \right] \quad (23)$$

$$\pi_B(\text{reject}; \text{accept}) = \frac{1}{2t} \left\{ t - \frac{\alpha}{3}(t_c - f) + \frac{\alpha}{3}(m - t_m) \right\}^2 \quad (24)$$

Consider Merchant A's card acceptance behavior. Suppose that Merchant B accepts cards. Merchant A will accept cards if and only if  $\pi_A(\text{accept}; \text{accept}) \geq \pi_A(\text{reject}; \text{accept})$ . This is equivalent to the merchant fee being lower than  $m_1$ , where  $m_1$  is defined as:

$$m_1 = t_m + t_c - f \quad (25)$$

Suppose instead that Merchant B does not accept cards. Merchant A will accept cards if and only if  $\pi_A(\text{accept}; \text{reject}) \geq \pi_A(\text{reject}; \text{reject})$ . This is equivalent to  $m \leq m_2$ , where  $m_2$  is defined as:

$$m_2 = t_m + \frac{3}{\alpha}t + \frac{9 - 7\alpha}{2\alpha}(t_c - f) - \frac{3}{\alpha} \sqrt{\frac{(1 - \alpha)(9 - 5\alpha)}{4}(t_c - f)^2 + 3(1 - \alpha)(t_c - f) + t^2} \quad (26)$$

From equations 25 and 26,  $m_1 \geq m_2$  always holds.

Since Merchants A and B are identical, Merchant A's card acceptance behavior described above is also applicable to Merchant B. Thus, pure strategy Nash equilibria are defined as follows: When the merchant fee is higher than  $m_1$ , there is one Nash equilibrium—both merchants do not accept cards; when the merchant fee is between  $m_1$  and  $m_2$ , there are two Nash equilibria—both merchants do not accept cards, or both accept cards; when the merchant fee is lower than  $m_2$ , there is one Nash equilibrium—both merchants accept cards.

At stage 1 of period 1, the network sets a cardholder fee,  $f$ , such that  $f < t_c$ . It also sets the merchant fee,  $m$ , so that both merchants accept cards. In order to have both of the merchants accept cards, the network will set the merchant fee lower than either  $m_1$  or  $m_2$ . For  $m \leq m_2$ ,

both merchants always accept cards regardless of their predictions for their rival's card acceptance behavior. For  $m > m_2$ , the merchant's card acceptance depends on its prediction for its rival's behavior. Since neither merchant previously accepted cards, each merchant may not believe with certainty that its rival merchant will accept cards when "both do not accept cards" can be an equilibrium. But each may think that its rival will accept cards with some probability.

Suppose that Merchant A predicts that Merchant B will accept cards with a probability of  $x_B$ , and will not accept cards with a probability of  $1 - x_B$ . Suppose also that Merchant A maximizes its expected profit. Given  $m$ , Merchant A's expected profit from accepting cards is defined as:

$$\pi_A^e(\text{accept}; m) = x_B \pi_A(\text{accept}; \text{accept}) + (1 - x_B) \pi_A(\text{accept}; \text{reject}),$$

and Merchant A's expected profit from not accepting cards is defined as:

$$\pi_A^e(\text{reject}; m) = x_B \pi_A(\text{reject}; \text{accept}) + (1 - x_B) \pi_A(\text{reject}; \text{reject}).$$

Similarly, Merchant B's expected profit can be defined. The network will set  $m$  so that

$$\pi_l^e(\text{accept}; m) \geq \pi_l^e(\text{reject}; m) \text{ for } l = A, B. \quad (27)$$

This is equivalent to  $m \leq m(\underline{x})$ , where  $m(\underline{x})$  is defined as:

$$m(\underline{x}) = t_m + \frac{1}{2\alpha(1 - 2\underline{x})} [\{2\alpha(1 - 2\underline{x}) + 9(1 - \underline{x})(1 - \alpha)\}(t_c - f) + 6t - 3\sqrt{4t^2 + 12t(1 - \underline{x})(1 - \alpha)(t_c - f) + \{4\alpha(1 - 2\underline{x})(1 - \underline{x})(1 - \alpha) + 9(1 - \underline{x})^2(1 - \alpha)^2\}(t_c - f)^2}], \quad (28)$$

where  $\underline{x} = \min\{x_A, x_B\}$ . Clearly,  $m(\underline{x})$  is an increasing function and  $m(0) = m_2$  and  $m(1) = m_1$ .

Suppose instead that each merchant minimizes the loss from making a wrong prediction. As long as each merchant is unsure about its rival's card acceptance behavior, there are two possible losses for each merchant. Suppose Merchant A predicts that Merchant B will accept cards. Then Merchant A accepts cards. Suppose that Merchant B actually chooses not to accept

cards. In this case, the loss of Merchant A from making a wrong prediction regarding Merchant B's card acceptance choice,  $L_A(\text{accept}; m)$ , is defined as:

$$L_A(\text{accept}; m) = \frac{t}{2} - \frac{1}{2t} \left[ \left\{ t + \frac{\alpha}{3}(t_c - f) - \frac{\alpha}{3}(m - t_m) \right\}^2 - \alpha(1 - \alpha)(t_c - f)(m - t_m) \right]. \quad (29)$$

Suppose Merchant A predicts that Merchant B will not accept cards. Then Merchant A rejects cards. Suppose again that Merchant B's actual card acceptance choice is other than predicted. In this case, the loss of Merchant A from making a wrong prediction regarding Merchant B's card acceptance choice,  $L_A(\text{reject}; m)$ , is defined as:

$$L_A(\text{reject}; m) = \frac{t}{2} - \frac{1}{2t} \left\{ t - \frac{\alpha}{3}(t_c - f) + \frac{\alpha}{3}(m - t_m) \right\}^2. \quad (30)$$

Similarly, Merchant B's losses are defined. From equations 31 and 32,

$L_l(\text{accept}; m) \leq L_l(\text{reject}; m)$  for  $l = A, B$  always holds for  $m_2 < m \leq \tilde{m}$ , where  $\tilde{m}$  is defined as:

$$\tilde{m} = t_m + \frac{4t(t_c - f)}{4t + 3(1 - \alpha)(t_c - f)}, \quad (31)$$

and  $L_l(\text{accept}; m) > L_l(\text{reject}; m)$  for  $l = A, B$  always holds for  $\tilde{m} < m \leq m_1$ . This implies that if both merchants minimize their losses from making a wrong prediction, they will accept cards if the merchant fee is lower than  $\tilde{m}$ .

The highest merchant fee the network can set in period 1 depends on the merchant objectives—whether to maximize expected profit or to minimize possible loss—and whether the merchant belief regarding their rival's card acceptance is observable to the network. In the case where the merchants maximize their expected profits and each merchant's belief,  $(x_A, x_B)$ , is observable to the network, the network sets  $m$  as high as  $m(\underline{x})$  (case 1). In the case where the merchants maximize their expected profits but each merchant's belief,  $(x_A, x_B)$ , is unobservable

to the network, the conservative network will set  $m$  as high as  $m_2$  (case 2).<sup>26</sup> In the case where the merchants minimize their possible losses, the network sets  $m$  as high as  $\tilde{m}$  (case 3).

$$\begin{aligned}\bar{m}^1 &= m(x), & \text{if case 1,} \\ &= m_2, & \text{if case 2,} \\ &= \tilde{m}, & \text{if case 3.}\end{aligned}\tag{32}$$

**Proposition 5 (Merchant fee):** *In the industry whose aggregate consumer demand is inelastic, the duopoly merchants competing according to the Hotelling model accept cards even if the merchant fee is higher than the merchant transactional benefit from cards.*

Proposition 5 implies that the merchants' strategic motive to accept cards enables the network to raise the merchant fee higher than the fee for a monopoly merchant. While the network charges the merchant fee to some monopoly merchants who face an inelastic demand as high as  $t_m$ , the fee it charges duopoly merchants who face the same aggregate consumer demand as the monopoly merchant is higher than  $t_m$ .

**Proposition 6 (Welfare):** *Suppose the network sets  $f < t_c$  and  $m = \bar{m}^1$ . In the industry whose aggregate consumer demand is inelastic, the duopoly merchants who are competing in the Hotelling model accept cards. Compared to the equilibrium without cards:*

- a) *Non-cardholders' surplus is lower;*
- b) *Cardholders' surplus is higher;*
- c) *Aggregate consumers' surplus is higher or at least the same;*
- d) *The merchants' surplus is the same.*

---

<sup>26</sup> Notice that  $m(0)=m_2$ .

**Proof:** Since consumer demand is elastic, assume that each consumer receives gross benefit,  $v$ , by purchasing one unit of the product. The net utility is, therefore,  $v - p - t_c$  for a cash transaction and  $v - p - f$  for a card transaction.

From equation 17, the new price,  $p^1 = (p_A^1 = p_B^1)$ , is higher because the merchant fee is higher than the transactional costs with cash. Since the non-cardholders' surplus depends only on the product price, non-cardholders are worse off.

The total cost for a cardholder when he or she uses a card,  $p^1 + f$ , is always lower than the total cost for the cardholder when the merchants do not accept cards,  $p^0 + t_c$ . The merchant will set the price at the highest,  $\bar{p}^1$ , when the network sets  $m = m_1$ .<sup>27</sup> Even at this highest price, the card holders are better off. From equation 17,  $\bar{p}^1 = d + t_m + t + \alpha(t_c - f)$ , and this yields:

$$(\bar{p}^1 + f) - (p^0 + t_c) = -(1 - \alpha)(t_c - f) < 0$$

Aggregate consumers' surplus without cards is  $v - p^0 - t_c$ , and that with cards under the highest product price  $\bar{p}^1$  is  $(1 - \alpha)(v - \bar{p}^1 - t_c) + \alpha(v - \bar{p}^1 - f)$ .

$$\{(1 - \alpha)(v - \bar{p}^1 - t_c) + \alpha(v - \bar{p}^1 - f)\} - (v - p^0 - t_c) = 0.$$

Since  $(1 - \alpha)(v - p - t_c) + \alpha(v - p - f)$  is decreasing in  $p$ , for all  $p^0 < p < \bar{p}^1$ ,

$$\{(1 - \alpha)(v - p - t_c) + \alpha(v - p - f)\} - (v - p^0 - t_c) > 0$$

From equations 15 and 18, the merchants' profits are unchanged. ■

Let us turn to the game in period 2. Can the card network raise the merchant fee higher than the highest merchant fee in period 1? Suppose all exogenous variables are the same as in period 1, and the network does not change the cardholder fee,  $f$ . In contrast to monopoly

---

<sup>27</sup> This highest level will be set when the merchants maximize their expected profit with  $\underline{x}=1$ .

merchants, duopoly merchants cannot go back to the original strategy easily. Given the fact that the other merchant is now accepting cards, “continue to accept cards” is the better strategy than “reject cards,” as long as the merchant fee is lower than a certain level, which may be higher than the highest fee in period 1.

If the merchants maximize their expected profits, regardless of whether each merchant’s belief on its rival’s probability of accepting cards is observable or unobservable to the network (cases 1 and 2), the merchant fee in period 2 is likely higher than the highest merchant fee in period 1. Since both merchants have accepted cards in period 1, each merchant may believe more strongly that its rival will continue to accept cards when “both merchants continue to accept cards” can be an equilibrium. Both  $x_A$  and  $x_B$  likely increase, and thus  $\underline{x}$  (minimum of  $x_A$  and  $x_B$ ) likely increases. Since  $m(\underline{x})$  is increasing in  $\underline{x}$ , the network can set a higher merchant fee with higher  $\underline{x}$ . Even in the case where  $x_A$  and  $x_B$  are unobservable to the network (case 2), the conservative network may not predict the lower of  $x_A$  and  $x_B$  being zero anymore given the fact that both merchants have accepted cards in period 1. Based on the new network’s prediction in  $x_A$  and  $x_B$ , the network can set the merchant fee higher than  $m_2$ .

If the merchants minimize their possible losses (case 3), the merchant fee in period 2 is likely the same as the highest merchant fee in period 1. Even though both  $x_A$  and  $x_B$  are now higher than before, as long as  $x_A$  and  $x_B$  are strictly smaller than 1, the merchant fee cannot be higher than  $\tilde{m}$ .

As more periods in which both merchants keep accepting cards have passed, each merchant’s belief in its rival’s probability of accepting cards will increase. In the long run, both  $x_A$  and  $x_B$  will converge to 1. Then the merchant fee will converge to  $\overline{m}^N$ , which is defined as:

$$\bar{m}^N = t_m + t_c - f, \quad (33)$$

regardless of the merchant objectives.

Notice that the long-run equilibrium merchant fee is the same as what Rochet and Tirole (2002) found. As explained above, in the short run the network may not set the long-run equilibrium merchant fee. This is because the conservative network wants to make sure that both of the merchants accept cards in the initial period, and by doing so it avoids setting a merchant fee which may lead to the equilibrium where both do not accept cards.

***Proposition 7 (Long-run equilibrium):*** *Compared to the equilibrium without cards, at the long run equilibrium with cards:*

- a) Non-cardholders are worse off;*
- b) Cardholders are better off;*
- c) Aggregate consumers' surplus is the same;*
- d) Merchants are indifferent.*

Merchant competition allows the network to set merchant fees higher. In industries whose aggregate consumer demand is inelastic, the network can charge monopoly merchants the merchant fees as high as the merchants' transactional benefits from cards or in some circumstances it can charge the fees higher than the merchants' transactional benefits. If the same industries are more competitive, the network can always charge the merchant fees higher than the merchants' transactional benefits. And in the long run it charges the fees that are equal to the merchants' transactional benefits plus card users' net transactional benefits. Moreover, in a competitive market where each merchant's initial profit margin,  $t$ , is smaller than the card user's net benefit,  $t_c - f$ , the long-run merchant fee exceeds the sum of the merchants' transactional benefit *and* the merchants' initial margin. In either monopoly or competitive markets, in the long

run, the sum of consumers' and merchants' surpluses with cards converges to the sum of their surpluses without cards.

### 3.2.2 Elastic demand

In order to make the model analytically solvable, the model is modified in the following way. Assume that a consumer buys at least one unit of product per trip. The consumer chooses a merchant based on the total cost—the sum of the product price, the transactional cost with the payment instrument he or she uses, and transportation cost. After the consumer decides from which merchant to purchase the products, the consumer determines the quantity based on the product price and the transactional cost.

Before considering the equilibrium with cards, first we describe the equilibrium without cards. This modified Hotelling model gives the equilibrium prices,  $(p_A^0, p_B^0)$ , which are the same for both merchants.

$$p_A^0 = p_B^0 = p^0 = d + t_m + \frac{tD(p^0 + t_c)}{D(p^0 + t_c) - tD'(p^0 + t_c)}. \quad (34)$$

Each merchant's profit is:

$$\pi_A^0 = \pi_B^0 = \pi^0 = \frac{1}{2} \frac{tD(p^0 + t_c)}{\{D(p^0 + t_c) - tD'(p^0 + t_c)\}}. \quad (35)$$

Now consider equilibrium with cards. In order to make the model analytically solvable, the merchant's decision-making is broken into two steps. First, the merchants decide whether to accept cards before determining their prices. Second, after observing each other's card acceptance strategy, they determine their prices. This assumption may make the highest merchant fee the network will charge slightly lower in each period, compared with the model where the merchants decide their card acceptance and prices at the same time.



Suppose  $t_c > f$ . At stage 3 of period 1, a cardholder chooses the merchant based on the merchants' card acceptance and prices. If the cardholder chooses the merchant who does not accept cards, he or she pays with cash, otherwise he or she pays with a card. Similar to the case where consumer demand is inelastic, merchant card acceptance and pricing behavior depend on the parameter values. Here, we analyze equilibrium only for  $(t_c - f)/t \leq 1$ .

At stage 2 of period 1, first let us suppose both merchants take the same card acceptance strategy. Given  $m$ , each merchant's profit when both merchants accept cards under price  $p^0$  is:

$$\pi_l(\text{accept}; \text{accept}) = (p^0 - d - t_m)\left(\frac{1-\alpha}{2}\right)D(p^0 + t_c) + (p^0 - d - m)\frac{\alpha}{2}D(p^0 + f), \text{ for } l = A, B, (36)$$

and when both merchants reject the cards under price  $p^0$  is:

$$\pi_l(\text{reject}; \text{reject}) = \frac{1}{2}(p^0 - d - t_m)D(p^0 + t_c), \text{ for } l = A, B. (37)$$

Suppose instead a merchant takes a different card acceptance strategy from its rival's. Consider the case where Merchant A accepts cards and Merchant B does not. Merchant A's profit is:

$$\pi_A(\text{accept}; \text{reject}) = (p^0 - d - t_m)\left(\frac{1-\alpha}{2}\right)D(p^0 + t_c) + (p^0 - d - m)\alpha\left(\frac{1}{2} + \frac{t_c - f}{2t}\right)D(p^0 + f), (38)$$

and Merchant B's profit is:

$$\pi_B(\text{reject}; \text{accept}) = (p^0 - d - t_m)\left(\frac{1}{2} - \frac{\alpha(t_c - f)}{2t}\right)D(p^0 + t_c). (39)$$

Suppose that Merchant B accepts cards. Merchant A accepts cards if and only if:

$$t_m + (p^0 - d - t_m)\left\{1 - \frac{t - (t_c - f)}{t} \frac{D(p^0 + t_c)}{D(p^0 + f)}\right\} \geq m. (40)$$

Suppose instead Merchant B does not accept cards. Merchant A accepts cards if and only if:

$$t_m + (p^0 - d - t_m)\left\{1 - \frac{t}{t + t_c - f} \frac{D(p^0 + t_c)}{D(p^0 + f)}\right\} \geq m. (41)$$

Define  $m_1$  and  $m_2$  that satisfy the equality in equations 40 and 41. Clearly,  $m_1$  is greater than  $m_2$ . When the merchant fee is higher than  $m_1$ , there is one Nash equilibrium; both merchants do not accept cards. When the merchant fee is between  $m_1$  and  $m_2$ , there are two Nash equilibria; both merchants do not accept cards, or both accept cards. When the merchant fee is lower than  $m_2$  there is one Nash equilibrium; both merchants accept cards.

As discussed in Subsection 3.2.1, the highest merchant fee the network can set in period 1 depends on the merchant objectives (whether maximizing the expected profit or minimizing possible losses) and whether each merchant belief is observable to the network. In any of the cases, the highest merchant fee in period 1 is at least as high as  $m_2$ .

**Proposition 8 (Merchant fee):** *In the industry whose aggregate consumer demand is elastic, the duopoly merchants competing in the (modified) Hotelling model accept cards even if the merchant fee is higher than the merchant transactional benefit from cards.*

After deciding card acceptance, both merchants set their prices. The profit maximizing prices are defined as:

$$p_A^1 = p_B^1 = p^1 = d + t_m + \frac{t\{(1-\alpha)D(t_c) + \alpha D(f)\} + \alpha(m-t_m)\{D(f) - tD'(f)\}}{(1-\alpha)\{D(t_c) - tD'(t_c)\} + \alpha\{D(f) - tD'(f)\}}, \quad (42)$$

where  $D(x) = D(p^1 + x)$  and  $D'(x) = D'(p^1 + x)$ . Since general results are not available for the case of elastic consumer demand, the rest of the subsection restricts its attention to linear demand.

From equations 34 and 42,

$$\left(1 + \frac{t^2 b^2}{AB}\right)(p^1 - p^0) = t^2 b^2 \frac{\alpha(t_c - f)}{AB} + (m - t_m) \frac{\alpha\{D(p^1 + f) + tb\}}{A},$$

where  $D(p) = a - bp$ ,  $A = (1-\alpha)\{D(p^1 + t_c) - tD'(p^1 + t_c)\} + \alpha\{D(p^1 + f) - tD'(p^1 + f)\}$ ,

and  $B = D(p^0 + t_c) - tD'(p^0 + t_c)$ . This provides the following proposition.

**Proposition 9 (Price):** *In the industry whose aggregate consumer demand is elastic, the duopoly merchants competing in the (modified) Hotelling model will adjust the product prices according to the merchant fee, if they accept cards.*

i) *For  $t_m \leq m \leq \bar{m}^1$ , the product price with cards is higher.*

ii) *For  $m < t_m$ , if  $t_m < (t_c - f)\delta$ , where  $\delta = \frac{b^2 t^2}{\{D(p^0 + t_c) + bt\} \{D(p^1 + f) + bt\}}$ , for any  $m > 0$*

*the product price with cards is higher; if  $t_m \geq (t_c - f)\delta$ , for  $m > t_m - (t_c - f)\delta$  the product price with cards is higher and for  $m < t_m - (t_c - f)\delta$  the product price with cards is lower.*

Similar to the monopoly merchant, the duopoly merchants will set product prices higher than the equilibrium prices without cards even when the merchant fee is lower than their transactional benefits. However, competition narrows the range of merchant fees that allows merchants to set the higher product prices. A monopoly merchant will set a higher product price when the merchant fee is higher than  $t_m - (t_c - f)$ . On the other hand, duopoly merchants will set higher product prices when the merchant fee is higher than  $t_m - (t_c - f)\delta$ , which is higher than  $t_m - (t_c - f)$ .

**Proposition 10 (Welfare):** *Suppose the network sets  $f < t_c$  and  $m = \bar{m}^1$ . In the industry whose aggregate consumer demand is elastic, the duopoly merchants who are competing in the Hotelling model accept cards. Compared to the equilibrium without cards:*

- a) *Non-cardholders' surplus is lower;*
- b) *Cardholders' surplus is higher;*
- c) *Aggregate consumers' surplus is higher;*
- d) *The merchants' surplus is higher or at least the same.*

**Proof:** See Appendix A.

From the results of 3.2.1, it is not hard to imagine that the network will raise the merchant fee in the periods after both merchants accept cards. In the long run, the merchant fee will converge to the highest possible level and the product prices will also converge accordingly. Under such merchant fee and product prices, the merchant's profit with cards becomes the same as the equilibrium profit without cards.

#### **4. Possible explanations for gradual increases in merchant fees**

This section explores possible explanations why interchange fees and merchant fees have been gradually increasing in the last several years in the United States. Three possible explanations are discussed in turn. They are incompletely flexible price setting by merchants, gradual decreases (increases) in cardholder fees (rebates) by networks, and changes in market environments, such as increases in penetration rate of consumer cardholding and relative increases in transactional costs associated with payment instruments other than cards.

##### **4.1 Merchant inflexible price setting**

In Section 3, the model assumes that the merchant's price setting is completely flexible. Under this assumption, the long-run highest merchant fees can be set by the network relatively quickly. The network can set the highest fees to monopoly merchants from the first period they accepted cards. Although the network cannot set the long-run highest fees to competing merchants in earlier periods, once the merchants strongly believe that their rivals will continue to accept cards, the network can raise the fees to the highest level.

In reality, however, merchants may not be able to adjust their prices completely flexibly. Previous literature has found that the prices are rigid upward and/or downward by using micro or disaggregated data.<sup>28</sup> A large literature on prices has also explored reasons for the price rigidity.

---

<sup>28</sup>See, for example, Carlton (1986) and Kashyap (1995), Baharad and Eden (2003).

Among several reasons, the consumers' perception of price fairness is focused on by recent literature to explain that prices are rigid upward.<sup>29</sup> If a price increase is perceived to be unfair by consumers, it causes consumer dissatisfaction, which is costly to the merchant.

If merchants actually face some restrictions when they adjust prices upward, the network cannot set the highest fee found in the previous section in earlier periods of card acceptance. To see this, let us assume that merchants cannot increase their prices more than the growth rate of  $i$  in a period.

First, consider the fee for a monopoly merchant who faces an elastic demand. The long-run highest merchant fee,  $\bar{m}^N$ , is defined as in equation 10.<sup>30</sup> Given  $\bar{m}^N$ , the merchant profit maximizing price,  $p^*(\bar{m}^N)$ , satisfies  $\pi^c(p^*(\bar{m}^N); \bar{m}^N) = \pi^{nc}(p^0)$ . Suppose  $p^*(\bar{m}^N)$  is higher than the highest price the merchant can set in period 1, i.e.,  $p^0(1+i)$ . By definition,  $\pi^c(p^0(1+i); \bar{m}^N) < \pi^c(p^*(\bar{m}^N); \bar{m}^N)$ , and therefore  $\pi^c(p^0(1+i); \bar{m}^N) < \pi^{nc}(p^0)$ . This implies that the monopoly merchant will not accept cards if the network sets the merchant fee at  $m = \bar{m}^N$  in period 1. The network needs to set a lower merchant fee,  $\bar{m}^1$  in period 1 so that the merchant will accept cards.  $\bar{m}^1$  will be chosen to satisfy  $\pi^c(p^0(1+i); \bar{m}^1) = \pi^{nc}(p^0)$ .

Similarly, in period 2 the network may need to set the merchant fee,  $\bar{m}^2$ , which is lower than the long-run highest merchant fee. But the network can set the fee higher than  $\bar{m}^1$ , the fee in period 1. Suppose the merchant set its product price in period 1 at  $p^0(1+i)$ . Then the highest product price the merchant can set in period 2 is  $p^0(1+i)^2$ . If  $p^0(1+i)^2 < p^*(\bar{m})$ , the network cannot set the merchant fee as high as  $\bar{m}^N$ . At  $m = \bar{m}^1$ , however,  $\pi^c(p^0(1+i)^2; \bar{m}^1) > \pi^{nc}(p^0)$ ,

---

<sup>29</sup>See Kahneman, Knetsch, and Thaler (1986), Vaidyanathan and Aggarwal (2003), and Rotemberg (2004).

<sup>30</sup>Equation 10 defines the highest merchant fee in period 1. If the merchant's price setting is completely flexible, it is also the long-run highest merchant fee.

so the network can set the merchant fee higher than  $\bar{m}^1$  to satisfy  $\pi^c(p^0(1+i)^2; \bar{m}^2) = \pi^{nc}(p^0)$ .

In this way, the merchant fee is gradually increased to  $\bar{m}^N$ .

Second, consider the fee for competing merchants whose aggregate consumer demand is inelastic. In this case, even if the merchant price setting is completely flexible, it will take several periods for the network to raise the merchant fees to the long-run highest level. When the merchant price setting is not completely flexible, it may take even longer until the merchant fee reaches the highest level. The long-run highest merchant fee is reached when both of the merchants believe their rival will accept cards for sure (i.e.,  $x_A = x_B = 1$ ) if they can set their product prices completely flexibly. If the merchants cannot set their price freely, even when both merchants believe their rival will accept cards for sure, the merchant fee may not reach to the highest possible level.

To see this, let us suppose that in the 'n-1'th period after both merchants started accepting cards, each of them believes its rival will continue to accept cards for sure in the next period (period n). Suppose also that the highest product price in period n,  $\bar{p}^n (= p^{n-1}(1+i)$ , where  $p^{n-1}$  is the realized product price in period n-1) is lower than the profit maximizing price when the network sets the merchant fee at the long-run highest level,  $\bar{m}^N$ , which is defined in equation 33. From equation 17, the profit maximizing product price is:

$$p(\bar{m}^N) = d + t_m + t + \alpha(t_c - f). \quad (43)$$

The network needs to set the merchant fee so that  $\pi_l(\text{accept}; \text{accept}) \geq \pi_l(\text{reject}; \text{accept})$ , for

$l = A, B$ .<sup>31</sup> This is equivalent to the merchant fee being lower than  $\bar{m}^n$ , where  $\bar{m}^n$  is defined as:

$$\bar{m}^n = t_m + t_c - f - \frac{\{p(\bar{m}^N) - \bar{p}^n\}^2}{4\alpha t}. \quad (44)$$

Clearly from equations 33 and 44,  $\bar{m}^n$  is lower than the long-run highest merchant fee. As the product price gradually goes up, the merchant fee will be gradually raised to the highest level.

The merchant's inflexible price setting does not change the long-run highest merchant fees, although it prolongs the periods before the merchant fees reach the long-run highest levels. Therefore, in the long-run, product prices are not changed and neither is welfare. However, in the short-run, the merchant's inflexible price setting affects welfare, especially in competitive markets. Consider again the second situation discussed above, where the network sets the merchant fee at  $\bar{m}^n$  and competing merchants who face an inelastic demand accept cards and set the product price at  $\bar{p}^n$ . In this case, the equilibrium merchant profit is actually lower than the profit without cards at all.<sup>32</sup> Non-cardholders are worse off compared with the equilibrium without cards, but they are better off compared with the long-run equilibrium with cards. Cardholders are better off compared with either equilibrium.

#### **4.2 Lower cardholder fees and generous rebates**

Until this subsection, the model assumed that the network does not change the cardholder fee,  $f$ , over the periods. But this does not necessarily reflect the real world. In the United States, typically card users are not charged a per transaction fee for credit card transactions; rather many of them receive rewards and/or rebates from their card issuers. Recently, reward programs have been getting more popular and the rewards and/or rebates have been getting more generous.<sup>33</sup>

---

<sup>31</sup>To obtain  $\pi_l(\text{accept}; \text{accept})$  and  $\pi_l(\text{reject}; \text{accept})$ , see Appendix B.

<sup>32</sup>See Appendix B.

<sup>33</sup>It is reported by various sources. See, for example *American Banker*, May 11, 2005.

For debit card transactions, some issuers charge a per transaction fee to their cardholders but such issuers are in the minority.<sup>34</sup>

If the cardholder fee,  $f$ , has been reduced or the negative fee, such as rewards and rebates, has been increased, then the highest merchant fee will likely be increased. Consider the long-run equilibrium merchant fees. Clearly, for a monopoly merchant facing an elastic demand and for a competing merchant with an inelastic aggregate consumer demand, as the cardholder fee (rebate) decreases (increases), the long-run merchant fee increases. For a monopoly merchant facing an elastic demand, the long-run merchant fee is defined in equation 10. By taking derivative of  $\bar{m}^N$  with respect to  $f$ :

$$\frac{\partial \bar{m}^N}{\partial f} = \frac{2(1-\alpha)(t_c - f + A) - (2-\alpha)\sqrt{(1-\alpha)(t_c - f)^2 + 2(1-\alpha)(t_c - f)A + A^2}}{\alpha\sqrt{(1-\alpha)(t_c - f)^2 + 2(1-\alpha)(t_c - f)A + A^2}}, \quad (45)$$

where  $A = \frac{1}{2b} \{a - b(d + t_m + t_c)\}$ . Equation 45 is smaller than 0 but greater than  $-1$ .<sup>35</sup>

Therefore, a dollar decrease (increase) in a cardholder fee (rebate) increases the merchant fee by less than a dollar. For a competing merchant with an inelastic aggregate consumer demand, the long-run merchant fee is defined in equation 33. From this, a dollar decrease in a cardholder fee increases the merchant fee by a dollar. For a competing merchant with an elastic aggregate consumer demand, the long-run merchant fee is not formally analyzed in section 3. However, the results of the two cases (monopoly with an elastic demand and duopoly with an inelastic aggregate demand) suggest that a decrease in cardholder fee likely increases the long-run merchant fee for this type of merchant. For a monopoly merchant who faces an inelastic demand,

---

<sup>34</sup>According to the Board of Governors of the Federal Reserve System's report to the Congress on the disclosure of point-of-sale debit fees, 14 percent of depository institutions charge PIN fees for some of their customers. According to the same report, 13 percent of households reported that their depository institutions charge PIN fees.

<sup>35</sup>See Appendix C for the proof.



the long-run merchant fee depends on the condition, whether  $(v - t_c - d - t_m)(1 - \alpha) - (t_c - f)\alpha$  is positive or negative. If it is negative, a decrease in a cardholder fee increases the merchant fee. If it is positive, a small decrease in a cardholder fee may not increase the merchant fee. However, a large decrease will change the condition itself—from positive to negative—by doing so, a dollar increase in a cardholder fee increases the merchant fee by a dollar.

Since a decrease in the cardholder fee increases the long-run merchant fees, it will affect the long-run equilibrium product prices as well as welfare. Obviously, the higher the merchant fees, the higher the long-run product prices. Thus, in the long-run, non-cardholders are always worse off with a lower cardholder fee. Cardholders, on the other hand, are better off with a lower cardholder fee.<sup>36</sup> Even when a dollar decrease in the cardholder fee increases the merchant fees by a dollar, it does not increase the product prices by a dollar. A lower cardholder fee does not change the long-run merchants' surplus. It is the same as the merchants' surplus under the equilibrium without cards at all. Although in a market with an elastic demand a lower cardholder fee may generate some incremental sales to the merchants, it does not increase merchant profit.

### 4.3 Changes in market environment

The model has assumed that exogenous variables, such as  $\alpha$ ,  $t_m$ , and  $t_c$ , are constant over the periods. In reality, however, this may not be true and the changes in these variables affect the highest merchant fees the network can charge. If the merchant transactional benefits from cards,  $t_m$ , have increased over the periods, the highest merchant fees have been raised. The network also can raise the highest merchant fee, if the consumer transactional benefits from cards,  $t_c$ , have increased. The effect of an increase in  $\alpha$ , the proportion of customers that have a payment card in a given industry, varies by industry. In a competitive industry with an inelastic

aggregate demand, the long-run equilibrium merchant fee is not affected by  $\alpha$ . In the other industries, on the other hand, an increase in  $\alpha$  raises the long-run equilibrium merchant fees.

The proportion of customers that have a payment card in a given industry,  $\alpha$ , may have increased as the penetration rate of cardholding among population has increased. In the United States, the penetration rate of credit card holding among consumers has increased for the last several years. According to the Federal Reserve Board's Survey of Consumer Finance (SCF), 72.7 percent of families had at least one bank-type credit card in 2001, which is up from 67.6 percent in 1998 and from 66.5 percent in 1995.<sup>37</sup> The debit card penetration rate is hard to obtain, but the proportion of families that hold at least one transaction account can be a proxy. According to the SCF, the proportion had been increased from 87 percent in 1995 to 90.5 percent in 1999, and to 90.9 percent in 2001.

Changes in  $t_m$  and  $t_c$  are hard to quantify. Reportedly, merchant transactional costs associated with payment instruments other than payment cards have been constant for the last several years.<sup>38</sup> However, as payment card networks claim, if payment cards generate other transactional benefits—not just saving transactional costs associated with cash or checks—and if such transactional benefits either to merchants or to consumers have increased, the long-run equilibrium merchant fees can be raised.

## 5. Conclusion

A large group of merchants accept cards even when their transactional benefit from cards is lower than the fee they pay, as long as a card user's net transactional benefits are positive. Only monopoly merchants who are facing an inelastic consumer demand may deny cards when

---

<sup>36</sup>An exception is the case of monopoly merchants facing an inelastic demand who serve only cardholders.

<sup>37</sup>Kennickell, Starr-McCluer, and Surette (2000) and Aizcorbe, Kennickell, and Moore (2003).

<sup>38</sup>See, for example, Food Marketing Institute (1994, 1998).

the fee exceeds its transactional benefit. This suggests that the analyses based on the assumption that merchants accept cards only when their transactional benefits from cards surpass their fees are potentially misleading.

Although whether the network has an incentive to charge the highest possible merchant fee or not is in question, if it does, it will charge the fee as high as the sum of the merchant's transactional benefit and part or all of the cardholder's net transactional benefits including rewards. The network may not charge the highest fees in earlier periods of card acceptance in a given industry but in the long run it will set the fee to the highest possible level.

Merchant competition allows the network to set higher merchant fees. The network can always set higher merchant fees in more competitive markets. Moreover, in competitive markets the merchant fees in the long run may exceed the sum of the merchant's initial margin and the merchant's transactional benefit.

The highest possible merchant fees may be more efficient than the case without cards, if the real cost of processing card transactions is lower than the sum of the merchant's and the consumer's transactional costs with the other payment instruments. However, these fees potentially create inequality among consumers. While debit cards are more accessible to all consumers, credit cards are not.<sup>39</sup> In the United States, about 75 percent of families are estimated to hold at least one credit card.<sup>40</sup> In order to have a credit card, a consumer needs to meet various criteria, such as income, credit history, and so on. These criteria are necessary for banks as well as for society to have safe and sound payment systems. However, sometimes credit cards are used simply as a transactional means rather than as a tool for borrowing.<sup>41</sup> If higher credit card

---

<sup>39</sup>According to the SCF (2001), about 13 percent of U.S. families did not have a checking account. About 20 percent of families that did not have a checking account did not have enough money or have credit problems.

<sup>40</sup>According to the SCF (2001).

<sup>41</sup>According to the SCF (2001), about 40 percent of cardholding families did not borrow on credit cards.

merchant fees imply higher product prices, those who cannot have a credit card need to pay higher prices. Cardholders also pay higher prices but they are compensated for that by receiving rewards.

The paper pointed out three possible explanations for the recent gradual increases in merchant fees in the United States: incompletely flexible price setting by merchants, gradual decreases (increases) in cardholder fees (rebates) by networks, and changes in market environments, such as increases in penetration rate of consumer cardholding. Inflexible price setting by merchants does not change the long-run equilibrium merchant fees, but the other two increase the merchant fees in the long run.

Since the model assumes a monopoly network, the network does not have to compete for merchants, for consumers, and/or for issuers. But if card networks are actually competing against each other, does the competition lower merchant fees? The answer requires a formal analysis; however, a pessimistic speculation would be that network competition may not reduce merchant fees but it may rather raise merchant fees at least in the current situation in the United States. One study reported that consumers are beginning to hold fewer cards, given universal acceptance of credit/debit cards by merchants, and that some consumer's card usage is heavily driven by reward programs.<sup>42</sup> In order to attract more cardholders, the networks or their issuers may want to increase cardholder rewards if it does not cause rejection of their cards by merchants. Since typical merchant fees consist of per transaction fees and negligible fixed fees, merchants accept as many networks' cards as they can. Merchants are generally hesitant to reject customers' payment choices. Although they may steer customers from one payment instrument to another, in the case that their customers are not willing to use the merchant's preferred method of payment, they accept other payments to satisfy their customers. This sort of merchant behavior

may allow competing networks to increase the cardholder rebates and thus to raise the merchant fees, as shown in section 4. As long as the merchant fee does not exceed the level that gives merchants negative profits, merchants may have no choice but to continue accepting cards.

---

<sup>42</sup>The 2004 Preferred Card Study conducted by Edger, Dunn & Company. See [www.edgardunn.com/eletter/2004-02/](http://www.edgardunn.com/eletter/2004-02/).

## References

- Aizcorbe, Ana, Arthur Kennickell, and Kevin Moore. 2003. "Recent Changes in U.S. Family Finances: Evidence from the 1998 and 2001 Survey of Consumer Finance," *Federal Reserve Bulletin*, 89, pp. 1-32.
- Baharad, Eyal, and Benjamin Eden. 2003. "Price Rigidity and Price Dispersion: Evidence from Micro Data," Department of Economics, Vanderbilt University, Working Papers 0321.
- Baxter, William F. 1983. "Bank Interchange of Transactional Paper: Legal Perspectives," *Journal of Law and Economics*, 26. pp. 541-588.
- Board of Governors of the Federal Reserve System. 2004. Report to the Congress on the Disclosure of Point-of-Sale Debit Fees.
- Bolt, Wilko. 2003. "Retail Payments in the Netherlands: Some Facts and Some Theory," De Nederlandsche Bank Research Memorandum WO number 722.
- Bolt, Wilko, and Alexander F. Tieman. 2003. "Pricing Debit Card Payment Services: An IO Approach," IMF Working Paper WP03/202.
- Carlton, Dennis. 1986. "The Rigidity of Prices," *American Economic Review*, 76 (4), pp. 637-658.
- Chakravorti, Sujit, and Roberto Roson. 2004. "Platform Competition in Two-Sided Markets: The Case of Payment Networks," mimeo.
- Chakravorti, Sujit, and Ted To. 2003. "A Theory of Credit Cards," working paper, Federal Reserve Bank of Chicago.
- Dove Consulting. 2001. *2001 Study of Consumer Payment Preferences*.
- Evans, David. 2002. "The Antitrust Economics of Two-Sided Markets," AEI-Brookings Joint Center working paper.
- Evans, David, and Richard Schmalensee. 1999. *Paying with Plastic*. The MIT Press.
- Frankel, Alan. 1998. "Monopoly and Competition in the Supply and Exchange of Money," *Antitrust Law Journal*, 66 (2), pp. 313-361.

- Food Marketing Institute. 1994. *Benchmarking Comparative Payment Methods: Costs and Case Studies*.
- Food Marketing Institute. 1998. *EPS Costs: A Retailers Guide to Electronic Payment Systems Costs*.
- Gans, Joshua S., and Stephen P. King. 2002. "A Theoretical Analysis of Credit Card Regulation," Melbourne Business School Working Paper No. 2002-11.
- Giesen, Lauri. 2004. "Murky Future for Interchange," *European Card Review*, September/October 2004.
- Guthrie, Graeme, and Julian Wright. 2003. "Competing Payment Schemes," National University of Singapore Department of Economics Working Paper No. 0311.
- Hayashi, Fumiko, and Elizabeth Klee. 2003. "Technology Adoption and Consumer Payments: Evidence from Survey Data," *Review of Network Economics*, 2 (2), pp. 175-190.
- Hayashi, Fumiko, Richard Sullivan, and Stuart E. Weiner. 2003. *A Guide to the ATM and Debit Card Industry*. Kansas City. Federal Reserve Bank of Kansas City.
- Kahneman, Daniel, Jack Knetsch, and Richard Thaler. 1986. "Fairness as a Constraint on Profit Seeking Entitlements in the Market," *American Economic Review*, 76 (4), pp. 728-741.
- Kashyap, Anil. 1995. "Sticky Prices: New Evidence from Retail Catalogs," *Quarterly Journal of Economics*, 110 (1), pp. 245-274.
- Katz, Michael. 2001. "Reform of Credit Card Schemes in Australia II," Commissioned Report to the Reserve Bank of Australia.
- Kennickell, Arthur, Martha Starr-McCluer, and Brian Surette. 2000. "Recent Changes in U.S. Family Finances: Results from the 1998 Survey of Consumer Finances," *Federal Reserve Bulletin*, 86, pp. 1-29.
- Maneti, Fabio M., and Ernesto Somma. 2002. "Plastic Clashes: Competition among Closed and Open Systems in the Credit Card Industry," mimeo.
- Rochet, Jean-Charles, and Jean Tirole. 2002. "Cooperation among Competitors: Some Economics of Payment Card Associations," *Rand Journal of Economics*, 33 (4), pp. 549-570.
- Rochet, Jean-Charles, and Jean Tirole. 2003. "Platform Competition in Two-Sided Markets," *Journal of European Economic Association*, 1 (4), pp. 990-1029.

- Rotemberg, Julio J. 2004. "Fair Pricing," NBER Working Paper Series, 10915.
- Schmalensee, Richard. 2002. "Payment Systems and Interchange Fees," *The Journal of Industrial Economics*, 50 (2), pp. 103-122.
- Schwartz, Marius, and Daniel Vincent. 2004. "The No Surcharge Rule and Card User Rebates: Vertical Control by a Payment Network," working paper.
- United States Court of Appeals for the Second Circuit, United States of America v. Visa USA, Visa International Corp, and MasterCard, August Term, 2002.
- Vaidyanathan, Rajiv, and Praveen Aggarwal. 2003. "Who is the Fairest of Them All?" *Journal of Business Research*, 56 (6), 453-463.
- Wright, Julian. 2000. "An Economic Analysis of a Card Payment Network," mimeo, NECG and University of Auckland.
- Wright, Julian. 2003a. "Optimal Card Payment Systems," *European Economic Review*, 47 (4), pp. 587-612.
- Wright, Julian. 2003b. "Why Do Firms Accept Credit Cards?," mimeo.
- Wright, Julian. 2004. "The Determinant of Optimal Interchange Fees in Payment Systems," *The Journal of Industrial Economics*, 52 (1), pp. 1-26.



## Appendix A: Proof of Proposition 10

(a) is obvious from proposition 9. (b) It is equivalent to show  $p^1 + f < p^0 + t_c$  at  $m = m_1$ .

$$\begin{aligned} \left(1 + \frac{t^2 b^2}{AB}\right) \{(p^1 + f) - (p^0 - t_c)\} &= -\frac{t_c - f}{AB} \left[ (1 - \alpha) t^2 b^2 + \frac{D(p^0 + t_c) + tb}{(t + t_c - f) D(p^0 + f)} \right. \\ &\quad \left. \{(1 - \alpha) t D(p^0 + t_c) (D(p^1 + f) + tb) + (t_c - f) D(p^0 + f) D(p^1 + t_c) \right. \\ &\quad \left. + \alpha b (t_c - f) (t + t_c - f) D(p^0 + f) + tb (t_c - f) D(p^1 + f) + t^2 b^2 (t_c - f) \} \right] < 0 \end{aligned}$$

(c) To show this, first we show that aggregate quantities demand has increased at  $m = \bar{m}^1$ .

$$\begin{aligned} \{(1 - \alpha) D(p^1 + t_c) + \alpha D(p^1 + f)\} - D(p^0 + t_c) &= \frac{\alpha b (t_c - f)^2 \{D(p^0 + t_c) + tb\}}{(AB + t^2 b^2) (t + t_c - f) D(p^0 + f)} \\ \{D(p^0 + f) D(p^1 + t_c) + tb D(p^1 + f) + b (t + t_c - f) D(p^0 + f) + t^2 b^2\} &> 0 \end{aligned}$$

The difference of aggregate consumers' surplus with and without cards is:

$$\begin{aligned} &(1 - \alpha) v(p^1 + t_c) + \alpha v(p^1 + f) - v(p^0 + t_c) \\ &= \frac{1}{2b} \{(1 - \alpha) D(p^1 + t_c)^2 + \alpha D(p^1 + f)^2 - D(p^0 + t_c)^2\} \\ &= \frac{1}{2b} [\{D(p^1 + f)^2 - D(p^0 + t_c)^2\} - (1 - \alpha) \{D(p^1 + f)^2 - D(p^1 + t_c)^2\}] \\ &> \frac{1}{2b} \{D(p^1 + f) + D(p^0 + t_c)\} \{(1 - \alpha) D(p^1 + t_c) + \alpha D(p^1 + f) - D(p^0 + t_c)\} > 0 \end{aligned}$$

since  $D(p^1 + t_c) < D(p^0 + t_c) < D(p^1 + f)$ .

(d) From the results of 3.1.2 and 3.2.1, this is likely. ■

## Appendix B: Inflexible Price Setting by Competing Merchants

### Footnote 31

From equations 16:

$$\pi_l(\text{accept; accept}) = \frac{t + \bar{p}^n - p^0 - \alpha(m - t_m)}{2}. \quad (\text{B1})$$

To obtain  $\pi_l(\text{reject; accept})$ , suppose Merchant A accepts cards and Merchant B does not. In this case, Merchant A cannot set profit maximizing price but sets  $p_A^n = \bar{p}^n$ . Given  $p_A^n = \bar{p}^n$ ,

Merchant B's profit maximizing price is  $p_B^n = p^0 + \frac{1}{2}\{\bar{p}^n - p^0 - \alpha(t_c - f)\}$ . From equation 20,

Merchant B's profit is  $\pi_B(\text{reject; accept}) = \frac{1}{2t}\left\{t + \frac{\bar{p}^n - p^0 - \alpha(t_c - f)}{2}\right\}^2$ .

#### Footnote 32

By plugging equation 44 into equation B1,  $\pi_l(\text{accept; accept}) = \frac{t - \{p(\bar{m}^N) - \bar{p}^n\}}{2}$ .

### **Appendix C: Effects of a Lower Cardholder Fee on the Merchant Fee**

#### Footnote 35

First, we show that equation 45 is smaller than zero. When  $\alpha = 1$ , equation 45 is equal to  $-1$ . For  $0 < \alpha < 1$  it is equivalent to show that:

$$2(1 - \alpha)(t_c - f + A) < (2 - \alpha)\sqrt{(1 - \alpha)(t_c - f)^2 + 2(1 - \alpha)(t_c - f)A + A^2}.$$

Since both sides of the inequality are positive, it is equivalent to show:

$$4(1 - \alpha)^2(t_c - f + A)^2 < (2 - \alpha)^2\{(1 - \alpha)(t_c - f)^2 + 2(1 - \alpha)(t_c - f)A + A^2\}. \quad (\text{C1})$$

Since for any  $0 < \alpha < 1$ ,  $4(1 - \alpha)^2 < (2 - \alpha)^2(1 - \alpha) < (2 - \alpha)^2$ , inequality C1 always holds. Now we show that equation 45 is greater than  $-1$  for any  $0 < \alpha < 1$ . It is equivalent to show that:

$$\sqrt{(1 - \alpha)(t_c - f)^2 + 2(1 - \alpha)(t_c - f)A + A^2} < t_c - f + A. \quad (\text{C2})$$

Obviously, inequality C2 holds for any  $0 < \alpha < 1$ .