Market Structure and Credit Card Pricing: What Drives the Interchange?

Zhu Wang*

October 30, 2008

Abstract

This paper provides a new theory to explain empirical puzzles regarding credit card interchange fees. Our model departs from the existing two-sided market theories by arguing card adoption externalities are less important in a mature card market. Instead, we focus on card issuer entry, elastic consumer demand and the role of card transaction value. Our analysis suggests that card networks demand higher interchange fees to maximize member issuers' profits as card payments become more efficient and convenient. At equilibrium, consumer rewards and card transaction values increase with interchange fees, while consumer surplus and merchant profits may not. Based on the theoretical framework, we discuss pros and cons of policy interventions.

JEL classification: D4; L1; G2

Keywords: Credit cards; Oligopolistic markets; Interchange fees

*Economic Research Department, Federal Reserve Bank of Kansas City, 1 Memorial Drive, Kansas City, MO 64198. Email: zhu.wang@kc.frb.org. I thank Sujit Chakravorti, Michael Gort, Fumiko Hayashi, Robert Hunt, Thomas Jeitschko, Boyan Jovanovic, Robert Lucas, James McAndrews, Jean-Charles Rochet, Marc Rysman, Randy Wright and seminar participants at Fed Kansas City, NYU-Stern, SUNY Buffalo, 2008 Fed Atlanta "Economics of Payments III" conference, 2007 International Industrial Organization Conference and 2007 Fed New York "Money and Payments" conference for comments, and Nathan Halmrast for research assistance. The views expressed herein are solely those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

1 Introduction

1.1 Motivation

As credit cards become increasingly prominent forms of payments, the structure and performance of this industry have attracted intensive scrutiny.[1] At the heart of the controversy are interchange fees - the fees paid to card issuers when merchants accept their cards for purchase.

Interchange fees are set by credit card networks. Two major card networks, Visa and MasterCard, each set their interchange fees collectively for tens of thousand member financial institutions that issue and market their cards.[2] For one simple example of how interchange functions, imagine a consumer making a \$100 purchase with a credit card. For that \$100 item, the retailer would get approximately \$98. The remaining \$2, known as the merchant discount fees, gets divided up. About \$1.75 would go to the card issuing bank as interchange fees, and \$0.25 would go to the merchant acquiring bank (the retailer's account provider). Interchange fees serve as a key element of the credit card business model and generate significant revenues for card issuers.[3] In 2007, the US card issuers made \$42 billion revenue in interchange fees.

In recent years, merchants have become increasingly critical on interchange fees, claiming the fees are excessively high. They pointed out that, despite of falling costs in the card industry, interchange rates in the US have been rising over the last ten years and are among the largest and fast-growing costs of doing business for many retailers (see

[1] There are four types of general purpose payment cards in the US: (1) credit cards; (2) charge cards; (3) signature debit cards; and (4) PIN debit cards. The analysis of this paper applies to the first three types of cards, which are routed over credit card networks and account for 90% of total card purchase volume.

[2] Visa and MasterCard provide card services through member financial institutions (card-issuing banks and merchant-acquiring banks). They are called "four-party" systems and account for approximately 80% of the US credit card market. Amex and Discover primarily handle all card issuing and acquiring by themselves. They are called "three-party" systems and account for the remaining 20% of the market. In a "three-party" system, interchange fees are internal transfers and hence not directly observable. This paper provides a model for four-party systems, but the analysis can also be applied to three-party systems.

[3] Note that credit cards may serve two functions: payment and credit. The payment function allows cardholders to make transaction with cards and generate interchange revenues to card issuers. The credit function allows cardholders to borrow funds and generate finance revenues. While this paper focuses on card payment function and interchange revenues, we need to note that interchange fees may help increase card transaction values, so they also contribute to finance revenues for card issuers.

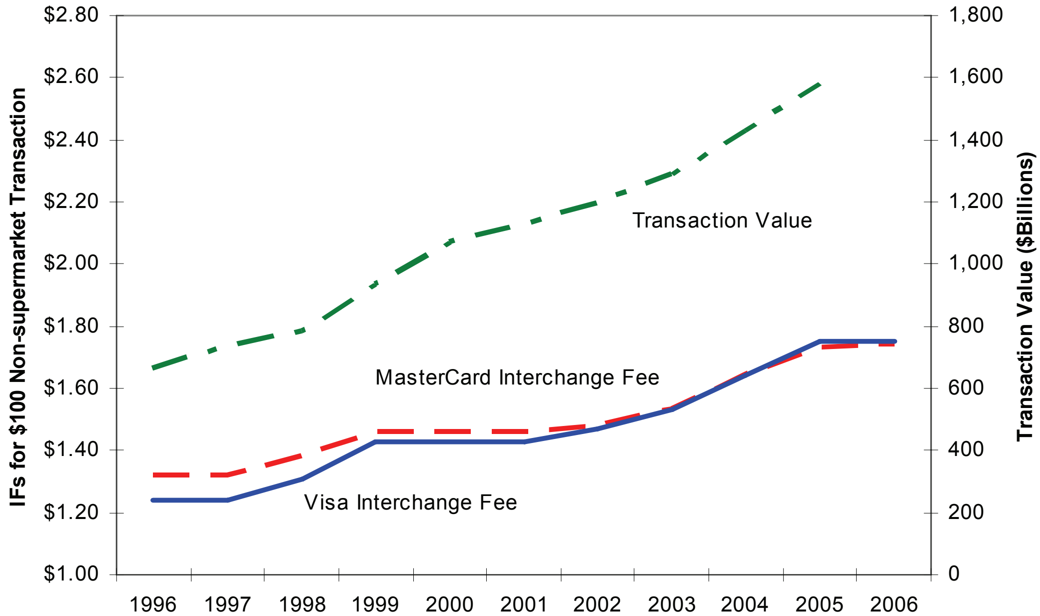

Figure 1: Credit Card Interchange Fees (IFs) and Transaction Values in the US

Figure 1).[4] However, card networks disagree, arguing interchange fees serve the needs of all parties in the card system, including funding better consumer reward programs that could also benefit merchants.

In the meantime, many competition authorities and central banks around the world have taken action on interchange fees. In Australia, the Reserve Bank of Australia mandated a sizeable reduction in credit card interchange fees in 2003, and is currently reevaluating the regulation. EU, UK, Belgium, Israel, Poland, Portugal, Mexico, New Zealand, Netherlands, Spain and Switzerland have made similar decisions and moves. In the US, interchange fees have been mainly challenged by private litigation. Since 2005, more than 50 antitrust cases have been filed by merchants contesting interchange fees.

The performance of the credit card industry raises following challenging questions:

- Why have interchange fees been increasing given falling costs and increased competition in the card industry (card processing, borrowing and fraud costs have all declined, while the number of issuers and card solicitations have been rising over

[4] Data sources: *Nilson Report* and *American Banker,* various issues.

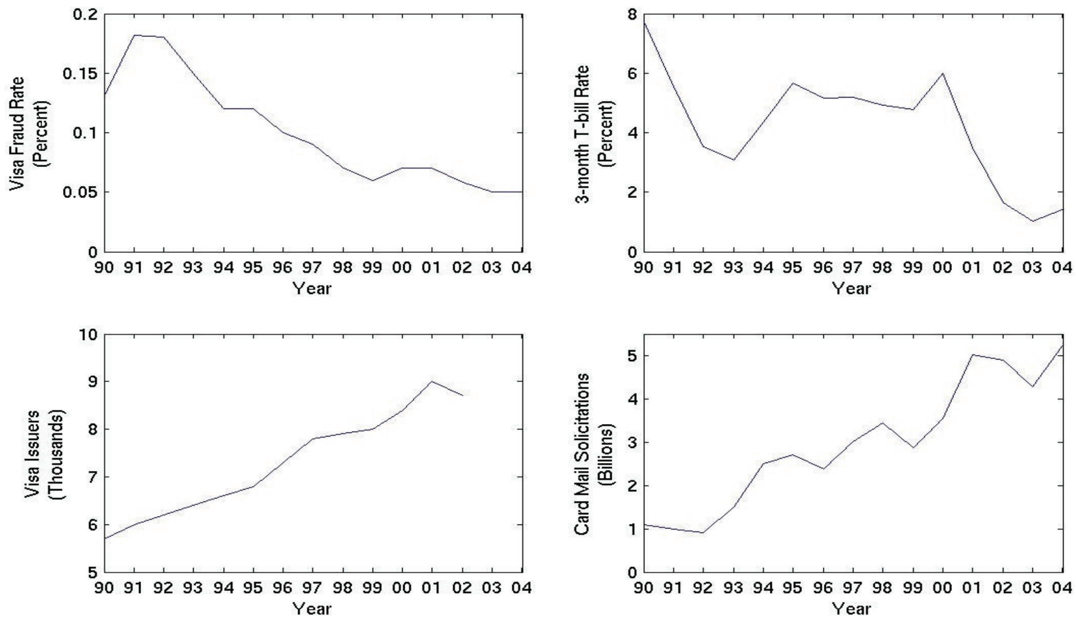

Figure 2: Credit Card Industry Trends: Costs and Competition

recent years, as shown in Figure 2)?[5]

- Given the rising interchange fees, why can't merchants refuse to accept cards? Why have total card transaction values been growing rapidly?

- What are the causes and consequences of the increasing consumer card rewards?

- What are the choices and consequences of policy interventions?

In order to answer these questions, a growing literature on payment card markets has been developed recently.[6] These models, following the pioneering work of Baxter (1983), emphasize the two-sided market externalities in card payment systems emanating from the fact that every card transaction must necessarily involve two users — a consumer and a merchant. According to this literature, it is reasonable for card networks to impose different fees for consumers and merchants in order to balance the demand on the two

[5] Data sources: Visa USA, Federal Reserve Board, Evans and Schmalense (2005) and Frankel (2006).

[6] For example, Schmalensee (2002), Rochet and Tirole (2002), Wright (2003), (2004), Schwartz and Vincent (2006), Hayashi (2006), McAndrews and Wang (2008).

sides of the market.[7] However, it remains an unsettled issue to systematically explain and evaluate the performance of the card industry.[8]

1.2 A New Approach

The present paper provides a new theory that addresses the aforementioned issues. The theory assumes a realistic framework of credit card markets, which consists of competing payment instruments, e.g., credit cards vs. alternative payment methods;[9] rational consumers (merchants) that always use (accept) the lowest-cost payment instruments; oligopolistic card networks that set profit-maximizing interchange fees; and competitive card issuers that join the most profitable network and compete with one another via consumer rewards.

The model yields equilibrium industry outcome that explains the puzzles regarding card interchange fees. It is shown that card networks, given their market power, demand higher interchange fees to maximize member issuers' profits as card payments become more efficient and convenient. Therefore, falling costs in the card industry could have indeed driven up interchange fees. At equilibrium, consumer rewards and card transaction values increase with interchange fees, but consumer surplus and merchant profits may not.

Our model departs quite bit from the existing two-sided market theories. First, because we are studying a mature card market, we do not consider card adoption externalities between merchants and consumers. Instead, we assume the set of card users is exogenously given. Second, we relax many restrictive assumptions in previous studies including consumers have fixed demand for goods; merchants engage in imperfect competition (e.g., Hotelling); and there is no entry/exit of card issuers. Instead, we assume

[7] Payment card systems are not the only case of such two-sided markets. Rochet and Tirole (2003) provide a detailed analysis of other examples, such as the software industry, video games, internet portals, medias, and shopping malls. In all these industries as well, the platforms may price differently to each side of the markets in order to balance the demand, while making a profit overall.

[8] These theories show that, although the socially optimal and privately optimal levels of card fees both depend on the same factors (e.g., issuing costs, acquiring costs, cardholders' and merchants' demand elasticities, market structure, and bargaining power of the parties), they are not equal in general. However, given various complications of the models, there is generally no way to tell that the card fees are systematically too high or too low, as compared with socially optimal levels (Katz 2001, Hunt 2003, Rochet 2003, Rochet and Tirole 2006).

[9] Alternative payment methods may include cash, check, PIN debit cards, stored value cards, automated clearing houses (ACH) and etc.

competitive merchants, free entry/exit of card issuers, oligopolistic card networks, and allows for elastic consumer demand.[10]

As a result, our model views the credit card industry as a vertical control system with monopolistic networks on top of price taking intermediates and end users. Card networks, in order to pursue their profits, set card fees to exploit the intensive margin of card usage by inflating the card transaction value. This is in contrast with the typical analyses of two-sided market theories that focus on the extensive margin of card usage (e.g., inducing more merchants and consumers to use cards). We found that card networks, by charging high interchange fees, can inflate retail prices to create more demand for their payment services. As card payments become more efficient and convenient than alternative payment instruments, card networks are able to further raise interchange fees and extract more profits (efficiency rents) out of the system. Meanwhile, due to higher retail prices, consumer surplus and merchant profits may not improve.

Our theory provides a new perspective that complements the two-sided market literature. In reality, card networks' incentive of inflating retail prices may coexist with two-sided market externalities. This is shown in the work of McAndrews and Wang (2008), which extends the analysis to an emerging card market where card networks set card fees to balance the "two-sided market effect" and the "inflation effect." Their results show that a monopoly card network charges a higher interchange fee than the social optimum, and exploits both intensive and extensive margins of card usage.

1.3 Road Map

Section 2 sets up a model of a "four-party" card system with merchants, consumers, acquirers, issuers and card networks. The model shows that a monopoly card network demands higher interchange fees to maximize member issuers' profits as card payments become more efficient or convenient. At equilibrium, consumer rewards and card transaction values increase with interchange fees, while consumer surplus and merchant profits

[10] Our model is also different from the existing theories by assuming that payment cards charge proportional fees instead of fixed per-transaction fees. This is motivated by the fact that only cards charging proportional fees that have pricing controversies in reality. However, assuming either proportional fees or fixed per-transaction fees does not change the main findings of our analysis (see Shy and Wang 2008).

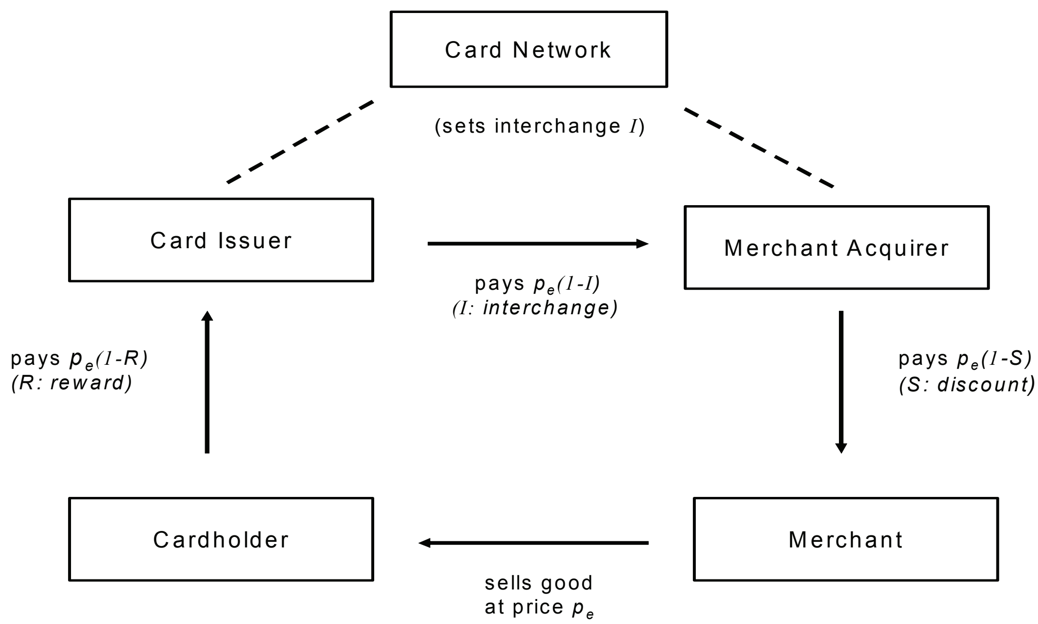

Figure 3: A Four-Party Credit Card System

do not. We show these findings could also hold under oligopolistic card networks. Section 3 extends the model to evaluate various policy interventions. Section 4 concludes.

2 The Model

2.1 Basic Setup

A four-party card system is composed of five players: merchants, consumers, acquires, issuers, and card networks, as illustrated in Figure 3. They are modeled as follows.

Merchants: A continuum of identical merchants sell a homogenous good in the market.[11] The competition leads to zero profit. Let p and k be price and non-payment cost for the good respectively. Merchants have two options to receive payments. Accepting non-card payments, such as cash, costs merchants $\tau_{m,a}$ per dollar, which includes the handling, storage, and safekeeping expenses that merchants have to bear. Accepting card

[11] Assuming identical merchants implies merchants always make zero profit regardless of interchange fees. In Appendix B, we relax this assumption and instead assume that merchants are heterogenous in costs, in which case merchants' profits are shown to be negatively affected by interchange fees.

payments costs merchants $\tau_{m,e}$ per dollar plus a merchant discount rate S per dollar paid to merchant acquirers. Therefore, a merchant who does not accept cards (i.e., cash store) charges p_a, while a merchant who accepts cards (i.e., card store) charges p_e:

$$p_a = \frac{k}{1-\tau_{m,a}}; \quad p_e = \max(\frac{k}{1-\tau_{m,e}-S},\ p_a).$$

We require $p_e \geq p_a$ so that $(1-\tau_{m,a})p_e \geq k$, which ensures card stores do not incur losses in case someone uses cash for purchase. This condition implies $S \geq \tau_{m,a} - \tau_{m,e}$. Moreover, we require $1 - \tau_{m,e} > S$ so that p_e is positive.

Consumers: There are two types of consumers. One is cash users, who do not own cards and have to pay with cash. The other is card users, who have option to pay either with card or cash. To use each payment instrument, consumers incur costs on handling, storage and safekeeping. Using cash costs consumers $\tau_{c,a}$ per dollar while using card costs $\tau_{c,e}$. In addition, card users receive a reward R from card issuers for each dollar spent on cards.[12] Therefore, card users do not shop at cash stores if and only if

$$(1+\tau_{c,a})p_a \geqslant (1+\tau_{c,e}-R)\ p_e \iff \frac{1+\tau_{c,a}}{1-\tau_{m,a}} \geqslant \frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-S}\ .$$

Meanwhile, given $p_a \leq\ p_e$, cash users prefer shopping at cash stores, and card users have no incentive to ever use cash in card stores.[13]

When making a purchase decision, card users face the after-reward price

$$p_r = (1+\tau_{c,e}-R)\frac{k}{1-\tau_{m,e}-S},$$

and have the total demand for card transaction values TD:

$$TD = p_e D(p_r) = \frac{k}{1-\tau_{m,e}-S} D[\frac{k}{1-\tau_{m,e}-S}(1+\tau_{c,e}-R)],$$

[12] Although our analysis focuses on the payment but not the credit function of credit cards, the reward R could be interpreted to include some benefits that consumers receive from the credit function of cards. See Chakravorti and To (2007) for related discussions.

[13] In reality, some consumers may use cash in stores that accept credit cards. In theory, this can happen if cash stores have a higher unit cost k than card stores. However, to keep our analysis focused, we do not explicitly explore this issue in the paper.

where D is the demand function for goods.

Acquirers: The acquiring market is competitive, where each acquirer receives a merchant discount rate S from merchants and pays an interchange rate I to card issuers. Acquiring incurs a constant cost C for each dollar of transaction. For simplicity, we normalize $C = 0$ so acquirers play no role in our analysis but pass through the merchant discount as interchange fee to the issuers, i.e., $S = I$ (see Rochet and Tirole 2002 for a similar treatment).[14]

Issuers: The issuing market is competitive, where each issuer receives an interchange rate I from acquirers and pays a reward rate R to consumers for each dollar spent on card. An issuer α incurs a fixed cost K each period, and an issuing cost V_α^β/α to handle its card transaction value V_α, where $\beta > 1$.[15] Issuers are heterogenous in their operational efficiency α, which is distributed with pdf $g(\alpha)$ over the population. They also pay the card network a processing fee T per dollar transaction and share their profits with the network.[16]

Issuer α's profit π_α (before sharing with the network) is determined as follows:

$$\pi_\alpha = \underset{V_\alpha}{Max}(I - R - T)V_\alpha - \frac{V_\alpha^\beta}{\alpha} - K$$

$$\Longrightarrow V_\alpha = (\frac{\alpha}{\beta}(I - R - T))^{\frac{1}{\beta-1}}; \quad \pi_\alpha = \frac{\beta-1}{\beta}(\frac{\alpha}{\beta})^{\frac{1}{\beta-1}}(I - R - T)^{\frac{\beta}{\beta-1}} - K.$$

Free entry condition requires that the marginal issuer α^* breaks even, so we have

[14] Note $C = 0$ is an innocuous assumption because C is mathematically equivalent to the network processing cost T in the following analysis. Moreover, we could instead model acquirers with heterogenous costs, but that would just duplicate our analysis of issuers.

[15] Assuming convex issuing costs is critical for our analysis. This is in contrast with some early studies (e.g., Wright 2003, Gans and King 2003) who found interchange fees are neutral (undetermined) under constant (zero) issuing costs. By assuming convex costs and issuer competition, we are able to pin down a unique interchange fee and characterize its properties. In reality, the issuing costs include the costs for providing credit float, fraud protection and other customer services. These costs depend on the gross value of card transactions but not the net (after-reward) transaction value. Also note that our model does not pin down the aggregate price level of the economy, but only the price levels for those sub-markets using cards. Therefore, as nominal card transaction values increase, card issuers incur increasing real costs.

[16] In reality, T refers to the Transaction Processing Fees that card networks collect from their members to process each card transaction through its central system, which is typically cost-based. In addition, card networks charge their members Service Fees based on each member's contribution to the network including the number of card issued, total transaction and sales volume. (Source: Visa USA By-Laws).

$$\pi_{\alpha^*} = 0 \Longrightarrow \alpha^* = \beta K^{\beta-1}(\frac{\beta}{\beta-1})^{\beta-1}(I-R-T)^{-\beta}.$$

As a result, the total number of issuers is

$$N = \int_{\alpha^*}^{\infty} g(\alpha)d\alpha,$$

and the total supply of card transaction value is

$$TV = \int_{\alpha^*}^{\infty} V_\alpha g(\alpha)d\alpha = \int_{\alpha^*}^{\infty} [(\frac{I-R-T}{\beta})\alpha]^{\frac{1}{\beta-1}} g(\alpha)d\alpha.$$

Networks: Each period, a card network incurs a variable cost T per dollar for processing card transactions. In return, it charges its member issuers a processing fee T to cover the variable costs and receives a share of their profits (the share is determined by bargaining between the card network and member issuers). As a result, the card network sets the interchange fee I to maximize the total profits for its member issuers, which also maximizes its own profit.

2.2 Monopoly Outcome

A monopoly network maximizing its member issuers' profits Ω^m solves the following problem each period:

$$\underset{I}{Max} \;\; \Omega^m = \int_{\alpha^*}^{\infty} \pi_\alpha g(\alpha)d\alpha \qquad \text{(Card Network Profit)}$$

$$s.t. \qquad \pi_\alpha = (\frac{\beta-1}{\beta})(\frac{\alpha}{\beta})^{\frac{1}{\beta-1}}(I-R-T)^{\frac{\beta}{\beta-1}} - K, \qquad \text{(Profit of Issuer } \alpha\text{)}$$

$$\alpha^* = \beta K^{\beta-1}(\frac{\beta}{\beta-1})^{\beta-1}(I-R-T)^{-\beta}, \qquad \text{(Marginal Issuer } \alpha^*\text{)}$$

$$N = \int_{\alpha^*}^{\infty} g(\alpha)d\alpha, \qquad \text{(Number of Issuers)}$$

$$\frac{1+\tau_{c,a}}{1-\tau_{m,a}} \geqslant \frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-I}, \qquad \text{(Pricing Constraint I)}$$

$$1-\tau_{m,e} > I \geq \tau_{m,a}-\tau_{m,e}, \qquad \text{(Pricing Constraint II)}$$

$$TV = \int_{\alpha^*}^{\infty} V_\alpha g(\alpha) d\alpha = \int_{\alpha^*}^{\infty} [(\frac{I-R-T}{\beta})\alpha]^{\frac{1}{\beta-1}} g(\alpha) d\alpha, \qquad \text{(Total Card Supply)}$$

$$TD = \frac{k}{1-\tau_{m,e}-I} D(\frac{k}{1-\tau_{m,e}-I}(1+\tau_{c,e}-R)), \qquad \text{(Total Card Demand)}$$

$$TV = TD. \qquad \text{(Card Market Clearing)}$$

To simplify the analysis, we assume that α follows a Pareto distribution so that $g(\alpha) = \gamma L^\gamma/(\alpha^{\gamma+1})$, where $\gamma > 1$ and $\beta\gamma > 1+\gamma$;[17] the consumer demand function takes the isoelastic form $D = \eta p_r^{-\varepsilon}$; and the pricing constraint $1-\tau_{m,e} > I \geq \tau_{m,a}-\tau_{m,e}$ is not binding.[18] Therefore, the above maximization problem can be rewritten as

$$\underset{I}{Max} \;\; \Omega^m = A(I-R-T)^{\beta\gamma} \qquad \text{(Card Network Profit)}$$

$$s.t. \qquad B(I-R-T)^{\beta\gamma-1} = (1-\tau_{m,e}-I)^{\varepsilon-1}(1+\tau_{c,e}-R)^{-\varepsilon}, \qquad \text{(Card Market Clearing)}$$

$$\frac{1+\tau_{c,a}}{1-\tau_{m,a}} \geqslant \frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-I}, \qquad \text{(Pricing Constraint I)}$$

where

$$A = KL^\gamma \beta^{-\gamma}(\frac{K\beta}{\beta-1})^{(1-\beta)\gamma}(\frac{\gamma}{\gamma-\frac{1}{\beta-1}}-1), \quad B = \frac{L^\gamma \beta^{-\gamma} k^{\varepsilon-1}}{\eta}(\frac{\gamma}{\gamma-\frac{1}{\beta-1}})(\frac{K\beta}{\beta-1})^{1+\gamma-\beta\gamma}.$$

[17] The size distribution of card issuers, like firm size distribution in many other industries, is highly positively skewed. Although possible candidates for this group of distributions are far from unique, Pareto distribution has typically been used as a reasonable and tractable example in the empirical IO literature.

[18] For simplicity, we assume the consumer demand D to be a fixed function of price p_r. Allowing the demand function to shift, e.g., by an exogenous increase of η due to income growth, would not affect our theoretical analysis, though empirically it may help explain the increase of card transaction values.

To simplify notation, we hereafter refer to the "Card Market Clearing Equation" as the "CMC Equation"; and refer to "Pricing Constraint I" as the "API Constraint", where API stands for "Alternative Payment Instruments". We denote the card markup $Z = I - R$, and further rewrite the above maximization problem as:

$$\underset{I}{Max} \;\; \Omega^m = A(Z-T)^{\beta\gamma} \qquad \text{(Card Network Profit)}$$

$$s.t. \qquad B(Z-T)^{\beta\gamma-1} = (1-\tau_{m,e}-I)^{\varepsilon-1}(1+\tau_{c,e}+Z-I)^{-\varepsilon}, \qquad \text{(CMC Equation)}$$

$$\frac{1+\tau_{c,a}}{1-\tau_{m,a}} \geqslant \frac{1+\tau_{c,e}+Z-I}{1-\tau_{m,e}-I}, \qquad \text{(API Constraint)}$$

where A, B are defined as before. Now it has become clear that a monopoly network would like to choose an interchange fee I^m to maximize the card markup Z. To fully characterize the monopoly outcome, we need to discuss two scenarios: elastic demand ($\varepsilon > 1$) and inelastic demand ($\varepsilon \leq 1$).

2.2.1 Elastic Demand: $\varepsilon > 1$

When demand is elastic ($\varepsilon > 1$), the CMC equation implies there is an interior maximum Z^m where

$$\partial Z^m/\partial I^m = 0 \Longrightarrow \frac{1+\tau_{c,e}+Z^m-I}{1-\tau_{m,e}-I^m} = \frac{\varepsilon}{\varepsilon-1} \quad \text{and} \quad \partial^2(Z^m)/\partial(I^m)^2 < 0.$$

Therefore, if the API constraint is not binding, the maximum is determined by the following conditions:

$$\frac{1+\tau_{c,e}+Z-I}{1-\tau_{m,e}-I} = \frac{\varepsilon}{\varepsilon-1}, \tag{1}$$

$$B(Z-T)^{\beta\gamma-1} = (1-\tau_{m,e}-I)^{\varepsilon-1}(1+\tau_{c,e}+Z-I)^{-\varepsilon}, \tag{2}$$

$$\frac{1+\tau_{c,a}}{1-\tau_{m,a}} \geqslant \frac{\varepsilon}{\varepsilon-1} \Longrightarrow \varepsilon \geqslant \frac{1+\tau_{c,a}}{\tau_{c,a}+\tau_{m,a}} > 1. \tag{3}$$

Proposition 1 then characterizes the monopoly interchange fee I^m as follows.

Proposition 1 *Given consumer demand is elastic and the API constraint is not binding (i.e., $\varepsilon \geqslant \frac{1+\tau_{c,a}}{\tau_{c,a}+\tau_{m,a}} > 1$), the monopoly interchange fee I^m increases if card service becomes less costly (i.e., T, $\tau_{m,e}$ or $\tau_{c,e}$ is lower) or the issuers' fixed cost K becomes lower, but I^m is not affected by the costs of using non-card payments $\tau_{m,a}$ or $\tau_{c,a}$.*

Proof. Equations (1)-(3) suggest that $\partial I^m/\partial T < 0$, $\partial I^m/\partial \tau_{m,e} < 0$, $\partial I^m/\partial \tau_{c,e} < 0$, $\partial I^m/\partial K < 0$, but $\partial I^m/\partial \tau_{m,a} = 0$, $\partial I^m/\partial \tau_{c,a} = 0$. ■

Similarly, we can derive comparative statics for the other endogenous variables at the monopoly maximum, including the card markup Z^m, the consumer reward $R^m = I^m - Z^m$, the issuer α's profit π_α and transaction value V_α, the number of issuers N, the card network's profit Ω^m and transaction value TV, before-reward retail price p_e, after-reward retail price p_r, and card users' consumption D. All the analytical results are reported in Table 1 (See Appendix A for proofs).

Table 1. Comparative Statics: $\varepsilon \geqslant \frac{1+\tau_{c,a}}{\tau_{c,a}+\tau_{m,a}} > 1$

(Signs of Partial Derivatives)

	I^m	R^m	Z^m	π_α	V_α	N	Ω^m	TV	p_e	p_r	D
$\tau_{m,e}$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	0	0
$\tau_{c,e}$	$-$	$\pm$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	0	0
T	$-$	$-$	$+$	$-$	$-$	$-$	$-$	$-$	$-$	0	0
K	$-$	$-$	$+$	$\pm$	$+$	$-$	$+$	$-$	$-$	0	0
$\tau_{m,a}$	0	0	0	0	0	0	0	0	0	0	0
$\tau_{c,a}$	0	0	0	0	0	0	0	0	0	0	0

Table 1 suggests that everything else being equal, we have the following findings:

- As it becomes less costly for merchants to accept cards (a lower $\tau_{m,e}$), both interchange fee and consumer reward increase, but interchange fee increases more and leads to a higher card markup. Meanwhile, the profit and transaction value of individual issuers increase, the number of issuers increases, total profit and transaction value of the card network increase, and before-reward retail price increases. However, after-reward retail price and card users' consumption stay the same.

- The above effects also hold if it becomes less costly for consumers to use card (a lower $\tau_{c,e}$) or it costs less for the network to provide card services (a lower T). Note for a lower $\tau_{c,e}$, consumer reward can either increase or decrease; and for a lower T, card markup decreases.

- As the fixed cost K for card issuers falls, both interchange fee and consumer reward increase, but consumer reward increases more and leads to a decrease of card markup. As a result, all incumbent issuers suffer a decline in transaction value, while large issuers' profits decrease, but small issuers' profits increase. Meanwhile, the number of issuers increases, card network profit decreases while transaction value increases, and before-reward retail price increases. However, after-reward retail price stays the same and there is no change in card users' consumption.

- Merchants or consumers' costs of using non-card payment instruments, $\tau_{m,a}$ and $\tau_{c,a}$, have no effect on any of the endogenous variables.

Alternatively, if the API constraint is binding, the monopoly maximum satisfies the following conditions:

$$B(Z-T)^{\beta\gamma-1} = (1-\tau_{m,e}-I)^{\varepsilon-1}(1+\tau_{c,e}+Z-I)^{-\varepsilon}, \tag{4}$$

$$\frac{1+\tau_{c,a}}{1-\tau_{m,a}} = \frac{1+\tau_{c,e}+Z-I}{1-\tau_{m,e}-I}, \tag{5}$$

$$\frac{1+\tau_{c,a}}{1-\tau_{m,a}} < \frac{\varepsilon}{\varepsilon-1} \Longrightarrow \frac{1+\tau_{c,a}}{\tau_{c,a}+\tau_{m,a}} > \varepsilon > 1. \tag{6}$$

Proposition 2 then characterizes the monopoly interchange fee I^m as follows.

Proposition 2 *Given consumer demand is elastic and the API constraint is binding (i.e., $\frac{1+\tau_{c,a}}{\tau_{c,a}+\tau_{m,a}} > \varepsilon > 1$), the monopoly interchange fee I^m increases if card service becomes less costly (i.e., T, $\tau_{m,e}$ or $\tau_{c,e}$ is lower) or issuers' fixed cost K becomes lower, but decreases if the cost of using non-card payments ($\tau_{m,a}$ or $\tau_{c,a}$) becomes lower.*

Proof. Equations (4)-(6) suggest that $\partial I^m/\partial T < 0$, $\partial I^m/\partial \tau_{m,e} < 0$, $\partial I^m/\partial \tau_{c,e} < 0$, $\partial I^m/\partial K < 0$, but $\partial I^m/\partial \tau_{m,a} > 0$, $\partial I^m/\partial \tau_{c,a} > 0$. ■

Table 2. Comparative Statics: $\frac{1+\tau_{c,a}}{\tau_{c,a}+\tau_{m,a}} > \varepsilon > 1$

(Signs of Partial Derivatives)

	I^m	R^m	Z^m	π_α	V_α	N	Ω^m	TV	p_e	p_r	D
$\tau_{m,a}$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$-$
$\tau_{c,a}$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$-$
$\tau_{m,e}$, $\tau_{c,e}$, T, K	Same signs as Table 1										

Similarly, we can derive comparative statics for the other endogenous variables at the maximum. As shown in Table 2, we have the following findings:

- As it becomes less costly for merchants or consumers to use non-card payment instruments (a lower $\tau_{m,a}$ or $\tau_{c,a}$), interchange fee decreases more than consumer reward, which leads to a decrease in card markup. Meanwhile, the profit and transaction value of individual issuers decrease, the number of issuers decreases, and the card network profit and transaction value decrease. In addition, before-and-after reward retail prices decrease and card users' consumption increases.

- The effects of other variables are the same as Table 1.

Figure 4 provides an intuitive illustration for the analysis. In the two graphs, the CMC equation describes a concave relationship between the card markup Z (Note the network profit Ω^m increases with Z) and the interchange fee $I \in [\tau_{m,a} - \tau_{m,e}, 1 - \tau_{m,e})$. In Case (1), the API constraint is not binding so the monopoly card network can price at the interior maximum, on which $\tau_{m,a}$ or $\tau_{c,a}$ has no effect. Alternatively, in Case (2), the

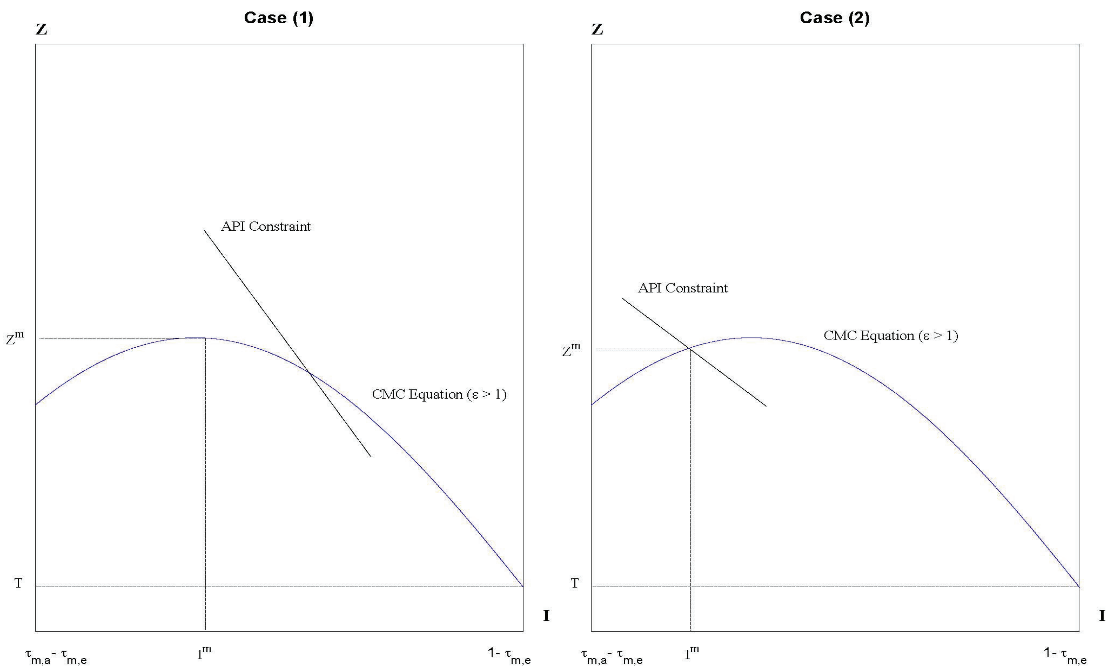

Figure 4: Monopoly Interchange Fee under Elastic Demand

API constraint is binding so $\tau_{m,a}$ or $\tau_{c,a}$ does affect the interchange pricing. Particularly, at the constrained maximum $(I^m,\ Z^m)$, the curve of the CMC equation has a slope less than 1. As a result, a local change of $\tau_{m,a}$ or $\tau_{c,a}$ shifts the line of the API constraint, but Z^m changes less than I^m so that $\partial R^m/\tau_{m,a} > 0$ and $\partial R^m/\tau_{c,a} > 0$. Furthermore, in Cases (1) and (2), changes of other parameters, such as $\tau_{m,e}$, $\tau_{c,e}$, T, K, shift the curve of CMC equation and affect the interchange pricing as described in Tables 1 and 2.

2.2.2 Inelastic Demand: $\varepsilon \leqslant 1$

When demand is inelastic ($\varepsilon \leqslant 1$), the CMC equation suggests that Z is an increasing function of I (i.e., $\partial Z/\partial I > 0$) and there is no interior maximum. Therefore, the API constraint is binding. The maximum satisfies the following conditions:

$$B(Z-T)^{\beta\gamma-1} = (1-\tau_{m,e}-I)^{\varepsilon-1}(1+\tau_{c,e}+Z-I)^{-\varepsilon}, \tag{7}$$

$$\frac{1+\tau_{c,a}}{1-\tau_{m,a}} = \frac{1+\tau_{c,e}+Z-I}{1-\tau_{m,e}-I}. \tag{8}$$

Proposition 3 then characterizes the monopoly interchange fee I^m as follows.

Proposition 3 *Given consumer demand is inelastic (i.e., $\varepsilon \leqslant 1$), the API constraint is binding and the monopoly interchange fee I^m increases if card service becomes less costly (i.e., T, $\tau_{m,e}$ or $\tau_{c,e}$ is lower) or issuers' fixed cost K becomes lower, but decreases if the cost of using non-card payments ($\tau_{m,a}$ or $\tau_{c,a}$) becomes lower.*

Proof. Equations (7)-(8) suggest that $\partial I^m/\partial T < 0$, $\partial I^m/\partial \tau_{m,e} < 0$, $\partial I^m/\partial \tau_{c,e} < 0$, $\partial I^m/\partial K < 0$, but $\partial I^m/\partial \tau_{m,a} > 0$, $\partial I^m/\partial \tau_{c,a} > 0$. ■

Table 3. Comparative Statics: $\varepsilon \leq 1$[19]

(Signs of Partial Derivatives)

	I^m	R^m	Z^m	π_α	V_α	N	Ω	TV	p_e	p_r	D
$\tau_{m,a}$	$+$	$\pm$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$-$
$\tau_{c,a}$	$+$	$\pm$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$-$
$\tau_{m,e}$, $\tau_{c,e}$, T, K	Same signs as Tables 1 and 2										

Similarly, we can derive comparative statics for the other endogenous variables at the maximum. As shown in Table 3, we have the following findings:

- The effects of $\tau_{m,a}$ and $\tau_{c,a}$ are the same as Table 2 except that consumer reward may either increase or decrease.
- The effects of other variables are the same as Tables 1 and 2.

Figure 5 provides an intuitive illustration of the analysis. In the two graphs, the CMC equation describes an increasing and convex relationship between the card markup Z (Note the network profit Ω^m increases with Z) and the interchange fee $I \in [\tau_{m,a} - \tau_{m,e}, 1 - \tau_{m,e})$. Therefore, the API constraint is binding so $\tau_{m,a}$ and $\tau_{c,a}$ affect the interchange pricing. In Case (3), at the constrained maximum (I^m, Z^m), the curve of the CMC

[19] Notice that for $\varepsilon = 0$, we have $\partial D/\partial \tau_{m,a} = \partial D/\partial \tau_{c,a} = 0$.

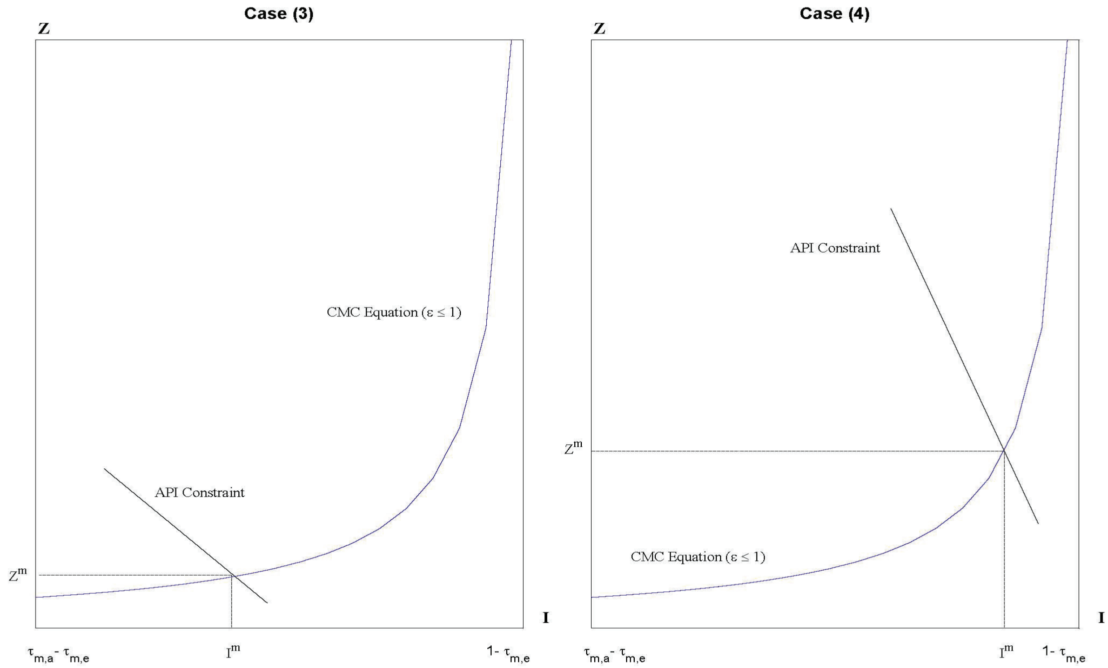

Figure 5: Monopoly Interchange Fee under Inelastic Demand

equation has a slope less than 1. As a result, a local change of $\tau_{m,a}$ or $\tau_{c,a}$ shifts the line of the API constraint, but Z^m changes less than I^m so that $\partial R^m/\tau_{m,a} > 0$ and $\partial R^m/\tau_{c,a} > 0$. Alternatively, in Case (4), at the constrained maximum (I^m, Z^m), the curve of the CMC equation has a slope greater than 1 so that $\partial R^m/\tau_{m,a} < 0$ and $\partial R^m/\tau_{c,a} < 0$. Furthermore, changes of other parameters, such as $\tau_{m,e}$, $\tau_{c,e}$, T, K, shift the curve of CMC equation and affect the interchange pricing as described in Table 3.

2.2.3 Recap and Remarks

As shown in the above analysis, under a monopoly card network, equilibrium interchange fees tend to increase as card payments become more efficient/convenient (a lower $\tau_{m,e}$, $\tau_{c,e}$ or T) or as the issuers' market becomes more competitive (a lower K).[20] These findings offer a consistent explanation for the puzzle of rising interchange fees. Meanwhile, we show that consumer rewards and card transaction values increase with interchange fees, but consumer welfare may not improve.

[20] As mentioned, the network processing cost T is mathematically equivalent to the acquiring cost C. Hence, a decrease of acquiring costs may also contribute to the increase of interchange fees.

The theory also explains other puzzles in the credit card market. For example, why can't merchants refuse to accept cards given the rising interchange fees? The answer is simple: As card payments become increasingly more efficient and convenient than alternative payment instruments, card networks can afford charging higher interchange fees but still keep cards as a competitive payment service to merchants and consumers. Another puzzle is why card networks, from a cross-section point of view, charge lower interchange fees on transaction categories with lower fraud costs, e.g., face-to-face purchases with card present are generally charged a lower interchange rate than online purchases without card present. This might seem to contradict the time-series evidence that interchange fees increase as fraud costs decrease. Our analysis suggests that the answer lies on the different API constraints that card networks face in different payment environments. In an environment with higher fraud costs for cards, such as online shopping, the costs of using a non-card payment instrument are also likely to be higher, which allows card networks to demand higher interchange fees.

The card networks underwent structural changes recently. Allegedly not-for-profit, MasterCard and Visa changed their status to for-profit and went public in 2005 and 2008 respectively. However, many industry observers feel that both networks have always been for-profit. In fact, no evidence shows that the networks changed their behavior after the IPO, which provides further support for our assumption of profit-maximizing networks.[21]

2.3 Duopoly Outcome

So far, we have discussed the monopoly outcome in the credit card market. To extend our analysis to a more realistic setting, we may consider a duopoly card market where two card networks (e.g., Visa and MasterCard) that produce homogenous card services have the same cost structure as specified in Section 2.1. Let $\Omega^i(I_{it}, I_{jt})$ denote network i's profit at period t when it charges interchange fee I_{it} and its rival charges I_{jt}. Network i maximizes the present discounted value of its profits, $U_i = \sum_{t=0}^{\infty} \delta^t \Omega^i(I_{it}, I_{jt})$, where δ is

[21] Note that because monopoly pricing is not by itself considered an antitrust issue, being an independent public firm may help card networks to get round the antitrust charges. According to many industry observers, this is the main reason for the networks to undertake the organizational changes (MacDonald 2006).

the discount factor.

First, consider the case that the two networks engage in a simple Bertrand competition. At each period t, the networks choose their interchange fees (I_{it}, I_{jt}) simultaneously. If the two networks charge the same interchange fee $I_{it} = I_{jt} = I_t$, they share the market, that is $\Omega^i = \Omega^j = \frac{1}{2}\Omega^m(I_t)$, where $\Omega^m(I_t)$ is the monopoly network profit at the interchange level I_t. Otherwise, the lower-interchange network may capture the whole market. This is suggested by the following proposition.

Proposition 4 *Everything else being equal, the after-reward retail price p_r increases with the interchange fee I.*

Proof. The CMC equation suggests that $\partial p_r / \partial I > 0$. ■

Proposition 4 says that a lower interchange fee results in a lower after-reward retail price, so a lower-interchange network is able to attract all the merchants and card consumers.[22] This implies that two card networks, if engaging in a Bertrand competition, should both set interchange fee at the minimum level $I = \tau_{m,a} - \tau_{m,e}$, given by the Eq. (Pricing Constraint II).[23] This is the competitive equilibrium outcome.

However, the outcome could be different in a repeated game. Particularly, if each network's interchange strategy at period t is allowed to depend on the history of previous interchanges $H_t \equiv (I_{i0}, I_{j0}; ...; I_{it-1}, I_{jt-1})$, it is possible for the monopoly outcome to be supported at equilibrium.

Consider a simple Forgiving Trigger (FT) strategy, which prescribes collusion in the first period, and then n periods of defection for every defection of any player, followed by reverting to cooperation no matter what has occurred during the punishment phase. Therefore, if a network undercuts the monopoly interchange fee I^m, it may earn a maximum profit $\Omega^m(I^m)$ during the period of deviation (indeed it earns approximately $\Omega^m(I^m)$ by slightly undercutting) but then it receives much lower profit for n periods. As shown

[22] Rysman (2007) found that consumers tend to concentrate their spending on a single payment network (single-homing), but many of them maintain unused cards that allow the ability to use multiple networks (multihoming). Therefore, consumers and merchants can easily switch between networks.

[23] We reasonably assume $\tau_{m,a} > \tau_{m,e}$, so the minimum interchange fee is positive. Otherwise, consumers have to pay for the card use (i.e., the reward is negative).

by the Folk Theorem, for a given n, if I^p is low and δ is large, (FT, FT) is a subgame perfect Nash equilibrium, and I^m can be supported at equilibrium.

This result suggests that a collusion could happen in the card market for two reasons. First, the model shows that the collusion can be supported at equilibrium if δ is large. This is because tacit collusion is enforced by punishment, which can occur only when deviation is detected. In the card industry, most card issuers are members of both Visa and MasterCard, so interchange fees tend to be public knowledge between the two networks. Second, an infinitely repeated game may have multiple equilibria. But as a natural method, the networks may coordinate on an equilibrium that yields a Pareto-optimal point for the two networks, that is the monopoly outcome. Meanwhile, a symmetric equilibrium appears to be consistent with the empirical observation that Visa and MasterCard have almost identical network structures and market shares.

Although it remains to be an empirical question to tell which equilibrium (competitive equilibrium or collusive equilibrium) the card market is at, our theory suggests that a possible collusive equilibrium could cause concerns about the card industry performance.

3 Policy and Welfare Analysis

The above analysis suggests that oligopolistic card networks could set interchange fees at the monopoly level at a collusive equilibrium. Although this remains an empirically testable hypothesis, in case it is true, competition authorities then have reason to consider intervening the card markets.

3.1 Policy Interventions

In many countries, public authorities have chosen to regulate down interchange fees. Our theory provides a formal framework to study the implications of these policy interventions.

3.1.1 Price Cut

As shown in Proposition 4, $\partial p_r/\partial I > 0$, which says a lower interchange fee results in a lower after-reward retail price and hence higher card users' consumption. Therefore, in order to increase consumer surplus, public authorities may want to cut interchange fees. Recall the CMC equation

$$B(Z-T)^{\beta\gamma-1}=(1-\tau_{m,e}-I)^{\varepsilon-1}(1+\tau_{c,e}+Z-I)^{-\varepsilon}.$$

The following proposition predicts the likely effects:

Proposition 5 *Everything else being equal, reducing the interchange rate results in a lower card markup, lower profits and transaction value for card issuers, fewer issuers, lower before-and-after-reward retail prices, and higher card users' consumption. For elastic demand, consumer reward decreases; and for inelastic demand, consumer reward may either decrease or increase.*

Proof. The CMC equation suggests that for $I<I^m$, we have $\partial Z/\partial I>0$, $\partial\pi_\alpha/\partial I>0$, $\partial V_\alpha/\partial I>0$, $\partial N/\partial I>0$, $\partial\Omega/\partial I>0$, $\partial p_e/\partial I>0$, $\partial p_r/\partial I>0$, $\partial D/\partial I<0$, and $\partial R/\partial I>0$ for $\varepsilon>1$, $\partial R/\partial I\gtrless 0$ for $\varepsilon\leq 1$. ■

3.1.2 Price Ceiling

A one-time price cut, however, may only have temporary effects because the interchange fees can easily come back. Alternatively, public authorities may set an interchange ceiling $I^c<I^m$.[24] Given a binding interchange ceiling I^c, the market outcome is determined by the modified CMC equation:

$$B(Z-T)^{\beta\gamma-1}=(1-\tau_{m,e}-I^c)^{\varepsilon-1}(1+\tau_{c,e}+Z-I^c)^{-\varepsilon},$$

where I^c is a constant. As a result, any changes of environmental parameters will then affect the industry differently from the non-intervention scenario.

Table 4 reports comparative statics of endogenous variables for an elastic demand ($\varepsilon>1$), which suggests that a binding interchange ceiling yields the following results:

[24] In 2003, Reserve Bank of Australia introduced a price ceiling for credit card interchange fees. At the time, the interchange fees averaged around 0.95% of the card transaction value. The regulation required that the weighted-average interchange fee for both Visa and MasterCard systems could not exceed 0.5% of the transaction value. The regulation is currently due for review, and one notable finding is that card rewards have been effectively reduced.

Table 4. Comparative Statics: $\varepsilon > 1$ and I^c is binding

(Signs of Partial Derivatives)

	I^c	R^c	Z^c	π_α	V_α	N	Ω^c	TV	p_e	p_r	D
$\tau_{m,e}$	0	+	−	−	−	−	−	−	+	+	−
$\tau_{c,e}$	0	+	−	−	−	−	−	−	0	+	−
T	0	−	+	−	−	−	−	−	0	+	−
K	0	−	+	±	+	−	+	−	0	+	−
$\tau_{m,a}$	0	0	0	0	0	0	0	0	0	0	0
$\tau_{c,a}$	0	0	0	0	0	0	0	0	0	0	0

- As it becomes less costly for merchants or consumers to use card (a lower $\tau_{m,e}$ or $\tau_{c,e}$), consumer reward decreases, which leads to an increase in card markup. As a result, the profit and transaction value of individual issuers increase, the number of issuers increases, card network profits and transaction values increase, after-reward retail price decreases, and card users' consumption increases. Meanwhile, a lower $\tau_{m,e}$ results in a lower before-reward price, but a lower $\tau_{c,e}$ does not affect the before-reward price.

- The above effects also hold if it costs less for card networks to process card transactions (a lower T). Note for a lower T, consumer reward increases and card markup decreases.

- As the fixed cost K for card issuers falls, consumer reward increases and leads to a decrease of card markup. As a result, all incumbent issuers suffer a decline in transaction value, while large issuers' profits decrease, but small issuers' profits increase. Meanwhile, the number of issuers increases, card network profits decrease but transaction values increase, after-reward retail price decreases, and card users' consumption increases. However, before-reward retail price stays the same.

- Merchants or consumers' costs of using non-card payment instruments ($\tau_{m,a}$ and $\tau_{c,a}$) have no effect on any of the endogenous variables.

Table 5. Comparative Statics: $0 < \varepsilon \leq 1$ and I^c is binding

(Signs of Partial Derivatives)

		I^c	R^c	Z^c	π_α	V_α	N	Ω^c	TV	p_e	p_r	D
$\tau_{m,e}$	$(\varepsilon < 1)$	0	$-$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$-$
	$(\varepsilon = 1)$	0	0	0	0	0	0	0	0	$+$	$+$	$-$
$\tau_{c,e}$, T, K, $\tau_{m,a}$, $\tau_{c,a}$		Same signs as Table 4										

Table 5 reports comparative statics of endogenous variables for an inelastic demand ($0 < \varepsilon \leq 1$), which suggests that a binding interchange ceiling yields the following results:[25]

- For a unit elastic demand ($\varepsilon = 1$), a lower $\tau_{m,e}$ has no effect on card pricing, output and profits. For an inelastic demand ($\varepsilon < 1$), a lower $\tau_{m,e}$ will have opposite effects on card pricing, output and profits as the elastic demand. However, regardless of demand elasticity, a lower $\tau_{m,e}$ always lowers the before-and-after-reward retail prices and raises card users' consumption (except for a perfectly inelastic demand).

- The effects of other variables are the same as Table 4.

The findings in Tables 4 and 5 suggest that a binding interchange fee ceiling allows card users to benefit from technology progress or enhanced competition in the card industry. These results are in sharp contrast with what we have seen in Tables 1, 2 and 3 for the non-intervention scenarios.

Figure 6 illustrates the effects of the interchange ceiling. In the two graphs for Cases (5) and (6), the API constraint is not binding so $\tau_{m,a}$ and $\tau_{c,a}$ have no effects. Furthermore, changes in the other parameters, such as $\tau_{m,e}$, $\tau_{c,e}$, T, K, shift the curve of the CMC equation. However, given a binding interchange ceiling, these changes can not raise the level of the interchange fee, but may affect other industry variables as described in Tables 4 and 5.

[25] For a perfectly inelastic demand ($\varepsilon = 0$), the results are reported in Table 6 in Appendix A.

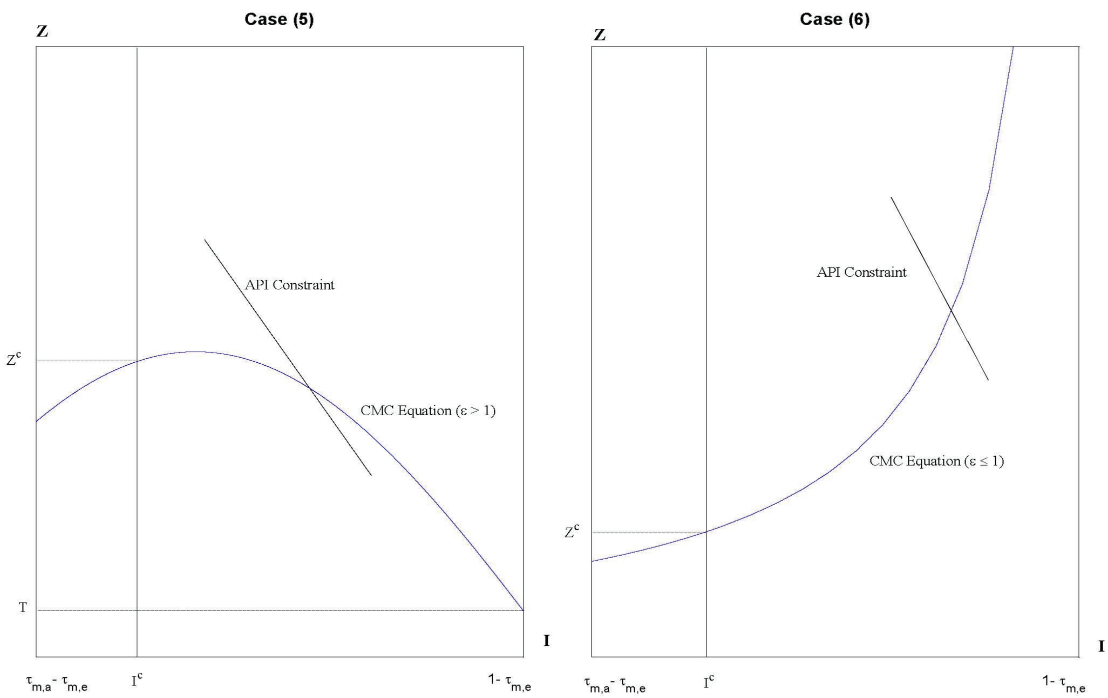

Figure 6: Interchange Fee Ceiling under Elastic/Inelastic Demand

3.2 Socially Optimal Pricing

Given the structure of credit card industry, our analysis shows consumer surplus decreases with interchange fees. However, it may not be socially optimal to set the interchange fee at its minimum level. In fact, the social planner aims to maximize the social surplus, which is the sum of issuers' profits and consumer surplus. Accordingly, the social planner's problem is:

$$\underset{I}{Max}\ \Omega^s = \int_0^{Q^*} D^{-1}(Q)dQ - \frac{k(1+\tau_{c,e}-R)}{1-\tau_{m,e}-I}Q^* + \int_{\alpha^*}^{\infty} \pi_\alpha g(\alpha)d\alpha \qquad \text{(Social Surplus)}$$

$$s.t. \qquad Q^* = D(\frac{k}{1-\tau_{m,e}-I}(1+\tau_{c,e}-R)), \qquad \text{(Demand of Goods)}$$

$$\pi_\alpha = (\frac{\beta-1}{\beta})(\frac{\alpha}{\beta})^{\frac{1}{\beta-1}}(I-R-T)^{\frac{\beta}{\beta-1}} - K, \qquad \text{(Profit of Issuer } \alpha\text{)}$$

$$\alpha^* = \beta K^{\beta-1}(\frac{\beta}{\beta-1})^{\beta-1}(I-R-T)^{-\beta}, \tag{Marginal Issuer α^*}$$

$$N = \int_{\alpha^*}^{\infty} g(\alpha)d\alpha, \tag{Number of Issuers}$$

$$\frac{1+\tau_{c,a}}{1-\tau_{m,a}} \geqslant \frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-I}, \tag{Pricing Constraint I}$$

$$1-\tau_{m,e} > I \geq \tau_{m,a}-\tau_{m,e}, \tag{Pricing Constraint II}$$

$$TV = \int_{\alpha^*}^{\infty} V_\alpha g(\alpha)d\alpha = \int_{\alpha^*}^{\infty} [(\frac{I-R-T}{\beta})\alpha]^{\frac{1}{\beta-1}} g(\alpha)d\alpha, \tag{Total Card Supply}$$

$$TD = \frac{k}{1-\tau_{m,e}-I} D(\frac{k}{1-\tau_{m,e}-I}(1+\tau_{c,e}-R)), \tag{Total Card Demand}$$

$$TV = TD. \tag{Card Market Clearing}$$

As before, we assume that α follows a Pareto distribution $g(\alpha) = \gamma L^\gamma/(\alpha^{\gamma+1})$, the consumer demand function takes the isoelastic form $D(p_r) = \eta p_r^{-\varepsilon}$, and the pricing constraint $1-\tau_{m,e} > I \geq \tau_{m,a}-\tau_{m,e}$ is not binding.

For $\varepsilon > 1$, the above maximization problem can then be rewritten as

$$\underset{I}{Max} \;\; \Omega^s = A(Z-T)^{\beta\gamma} + \frac{\eta}{\varepsilon-1} p_r^{1-\varepsilon} \tag{Social Surplus}$$

$$s.t. \qquad B(Z-T)^{\beta\gamma-1} = (1-\tau_{m,e}-I)^{\varepsilon-1}(1+\tau_{c,e}+Z-I)^{-\varepsilon}, \tag{CMC Equation}$$

$$\frac{1+\tau_{c,a}}{1-\tau_{m,a}} \geqslant \frac{1+\tau_{c,e}+Z-I}{1-\tau_{m,e}-I}, \tag{API Constraint}$$

where $Z = I - R$, $p_r = \frac{k(1+\tau_{c,e}+Z-I)}{(1-\tau_{m,e}-I)}$, and A, B are defined as before. Similarly, we can derive the social planner's problem for $\varepsilon \leq 1$ (see Appendix A).

Let I^s denote the socially optimal interchange fee. Note that the social surplus consists of two parts. One is card issuers' profits, which increase with the interchange fee. The other is consumer surplus, which decreases with the interchange fee. Therefore, we expect that this yields an interchange fee I^s lower than the monopoly level I^m, as shown in the following proposition.

Proposition 6 *The socially optimal interchange fee I^s is generally lower than the monopoly interchange fee I^m, i.e., $I^s \leq I^m$.*

Proof. This result holds for both elastic and inelastic demand. See Appendix A for the proof. ■

3.3 Further Issues

Our policy and welfare analysis offers some justification for the concerns and actions that public authorities worldwide have on the credit card interchange fees. Meanwhile, our analysis also provides a framework to discuss additional issues with various policy interventions.

First, we treated technological progress in payments (both cards and non-card payments) as exogenous in the model. Based on this, regulating down interchange fees appears to be desirable. However, it is likely in reality that advances in card technology are driven by intended R&D efforts by the card networks, and network profits provide important incentives and resources for these efforts. With endogenous technology progress, the social surplus calculation becomes more complicated. On one hand, regulating down interchange fees may raise consumer surplus, but on the other hand, it could hurt technology progress in the card industry and cause efficiency losses in the long run. Moreover, the extra profits in the card industry may also provide incentives for inventing and developing alternative payment products/technologies. All these endogenous and dynamic factors may make the welfare results of interchange regulation less clear and obvious.

Second, our analysis assumed that the market costs of payment instruments reflect their social costs. In reality, this may not be true. In some cases, when market costs of alternative payment instruments are lower than their social costs, the binding API constraint of card pricing may already lower interchange fees from where they otherwise would be. Therefore, learning about total social costs of various payment instruments is a prerequisite for designing and implementing good policy in payment markets.

Third, we abstracted from some potentially important issues in our analysis. For example, we assume that merchants are perfectly competitive, so we do not consider their strategic motives of accepting cards. And the no-surcharge rule does not play a role in our model because competitive merchants specialize on serving either card users or cash users. Assuming competitive merchants might be a reasonable assumption for many markets, but certainly not for all. It would be interesting to relax this assumption and investigate the implications.

Fourth, direct price regulation is not the only option or necessarily the best option for public authorities to improve market outcomes. There are other policy mixes worthy of exploring. In the case of credit card industry, regulating interchange fee is a quick solution but might be arbitrary and less adaptable. Policy makers may consider alternative approaches that target the market structure (e.g., enforcing competition between card networks)[26] or competing products (e.g., encouraging technology progress in non-card payments). In addition, increasing public scrutiny and rising regulatory threat may also be effective policy measures (see Stango 2003).

Last but not least, policy interventions may render unintended consequences. This is more likely to happen in a complex environment like the credit card industry. Therefore, a thorough study of the market structure can not be over emphasized. This paper is one of the beginning steps toward this direction, and many issues need further research, including the market definition of various payment instruments, the competition between four-party systems and three-party systems, and the causes and consequents of credit card rules, just to name a few.

[26] There may be many ways to re-design the card market to enforce competition, for example, introducing multi-network cards, requiring bilateral interchange fees between issuers and merchants, or reforming the network ownership/governance structure.

4 Conclusion

As credit cards become an increasingly prominent form of payments, the structure and performance of this industry have attracted intensive scrutiny. This paper presents an industry equilibrium model to better understand this market.

Our model takes a different approach from the existing literature. First, we model a mature card market without network effects given the fact that card adoption externality has become less important at this stage. Second, we relax many restrictive assumptions used in the previous studies: consumers have a fixed demand for goods; merchants engage in imperfect competition; and there is no entry/exit of card issuers. Instead, we assume competitive merchants, free entry/exit of card issuers, oligopolistic networks, and allows for elastic consumer demand.

The new model offers a more realistic and arguably better analytical framework. It is shown that card networks tend to charge high interchange fees to inflate retail prices in order to create more demand for their payment services. As card payments become more efficient and convenient, card networks are able to further raise the interchange fees and extract more profits (efficiency rents) out of the system. Meanwhile, due to higher retail prices, consumer surplus and merchant profits may not improve. Based on the theoretical framework, the pros and cons of policy interventions are discussed.

Appendix A.

Proof. (Table 1): Results in the first column are given by Proposition 1. Note Eqs. FOC and CMC imply

$$B(Z-T)^{\beta\gamma-1} = (\varepsilon-1)^{\varepsilon-1}(\varepsilon)^{-\varepsilon}(\tau_{c,e}+Z+\tau_{m,e})^{-1}.$$

The results in column 3 then are derived by implicit differentiation. Recall that all other endogenous variables are functions of Z, I and parameters:

$$
\begin{array}{ll}
R = I - Z, & \pi_\alpha = (\frac{\beta-1}{\beta})(\frac{\alpha}{\beta})^{\frac{1}{\beta-1}}(Z-T)^{\frac{\beta}{\beta-1}} - K, \\
V_\alpha = (\frac{\alpha}{\beta}(Z-T))^{\frac{1}{\beta-1}}, & \alpha^* = \beta K^{\beta-1}(\frac{\beta}{\beta-1})^{\beta-1}(Z-T)^{-\beta}, \\
N = \int_{\alpha^*}^{\infty} g(\alpha)d\alpha = (L/\alpha^*)^\gamma, & \Omega^m = A(Z-T)^{\beta\gamma}, \\
TV = \ B(Z-T)^{\beta\gamma-1}k^{1-\varepsilon}, & p_e = \frac{k}{1-\tau_{m,e}-I}, \\
p_r = \frac{(1+\tau_{c,e}+Z-I)}{(1-\tau_{m,e}-I)}k, & D = \eta p_r^{-\varepsilon}, \\
A = KL^\gamma\beta^{-\gamma}(\frac{K\beta}{\beta-1})^{(1-\beta)\gamma}(\frac{\gamma}{\gamma-\frac{1}{\beta-1}} - 1), & B = \frac{L^\gamma\beta^{-\gamma}k^{\varepsilon-1}}{\eta}(\frac{\gamma}{\gamma-\frac{1}{\beta-1}})(\frac{K\beta}{\beta-1})^{1+\gamma-\beta\gamma}.
\end{array}
$$

The other results in the table then are derived by differentiation. ■

Proof. Table 6 below reports comparative statics for the case of perfectly inelastic demand ($\varepsilon = 0$). ■

Table 6. Comparative Statics: $\varepsilon = 0$ and I^c is binding

(Signs of Partial Derivatives)

	I^c	R^c	Z^c	π_α	V_α	N	Ω^c	TV	p_e	p_r	D
$\tau_{m,e}$	0	$-$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	0
$\tau_{c,e}$	0	0	0	0	0	0	0	0	0	$+$	0
T	0	$-$	$+$	0	0	0	0	0	0	$+$	0
K	0	$-$	$+$	$\pm$	$+$	$-$	$+$	0	0	$+$	0
$\tau_{m,a}$	0	0	0	0	0	0	0	0	0	0	0
$\tau_{c,a}$	0	0	0	0	0	0	0	0	0	0	0

Proof. (Proposition 6): For $\varepsilon > 1$, the social planner's problem is

$$\underset{I}{Max} \ \ \Omega^s = A(Z-T)^{\beta\gamma} + \frac{\eta}{\varepsilon - 1} p_r^{1-\varepsilon}.$$

Consider the following two cases. First, if the API constraint is not binding, the monopoly's problem requires $\partial Z^m/\partial I^m = 0$ for the CMC equation. Accordingly, the social planner's problem implies

$$\frac{\partial \Omega^s}{\partial I^m} = \frac{\partial \Omega^s}{\partial Z}\frac{\partial Z}{\partial I^m} + \frac{\partial \Omega^s}{\partial p_r}\frac{\partial p_r}{\partial I^m} = -\eta p_r^{-\varepsilon}\frac{\partial p_r}{\partial I^m} < 0$$

since Proposition 4 shows $\partial p_r/\partial I > 0$. Therefore, $I^s < I^m$. Alternatively, if the API constraint is binding, $(Z^m,\ I^m)$ have to satisfy both the CMC equation and the API constraint, and $\partial Z^m/\partial I^m > 0$ for the CMC equation. Accordingly, the social planner's problem implies

$$\frac{\partial \Omega^s}{\partial I^m} = \frac{\partial \Omega^s}{\partial Z}\frac{\partial Z}{\partial I^m} + \frac{\partial \Omega^s}{\partial p_r}\frac{\partial p_r}{\partial I^m} = A\beta\gamma(Z-T)^{\beta\gamma-1}\frac{\partial Z}{\partial I^m} - \eta p_r^{-\varepsilon}\frac{\partial p_r}{\partial I^m}.$$

Then, if $\partial\Omega^s/\partial I^m < 0$, we have $I^s < I^m$; otherwise, if $\partial\Omega^s/\partial I^m \geqslant 0$, $I^s = I^m$.

For $\varepsilon \leq 1$, the analysis would be very similar. However, we then need a technical assumption to ensure that consumer surplus is bounded, e.g., $D(p_r) = \eta p_r^{-\varepsilon}$ for $D(p_r) \geq Q_0 > 0$, and $\int_0^{Q_0} D^{-1}(Q)dQ = H < \infty$. If $\varepsilon = 1$, the social planner's problem can be written as

$$\underset{I}{Max}\ \ \Omega^s = A(Z-T)^{\beta\gamma} + H - \eta\ln Q_0 - \eta + \eta\ln\eta - \eta\ln p_r.$$

Alternatively if $\varepsilon < 1$, the social planner's problem can be written as

$$\underset{I}{Max}\ \ \Omega^s = A(Z-T)^{\beta\gamma} + H + \frac{\varepsilon}{1-\varepsilon}\eta^{1/\varepsilon}Q_0^{1-1/\varepsilon} + \frac{\eta}{\varepsilon-1}p_r^{1-\varepsilon};$$

or if $\varepsilon = 0$, we have

$$\underset{I}{Max}\ \ \Omega^s = A(Z-T)^{\beta\gamma} + H - p_0Q_0 + (p_0 - p_r)\eta,$$

where p_0 is consumers' highest willingness to pay for $Q \in (Q_0, \eta)$. In each case, a similar proof as the elastic demand case then shows that $I^s \leq I^m$. ■

Appendix B.

In the paper, merchants are assumed to be identical. As a result, they always break even regardless of interchange fees. Although this assumption help simplify our analysis, it does not explicitly explain merchants' motivation for lowering interchange fees. In this appendix, we show that under a more realistic assumption that merchants are heterogenous in costs, their profits are indeed negatively affected by interchange fees in the same way as the consumer surplus of card users.

As before, we assume a continuum of merchants sell a homogenous good in a competitive market. A merchant θ incurs a fixed cost W each period and faces an operational cost q_θ^φ/θ for its sale q_θ, where $\varphi > 1$. Merchants are heterogenous in their operational efficiency θ, which follows a Pareto distribution over the population with pdf $f(\theta) = \phi J^\phi/(\theta^{\phi+1})$, $\phi > 1$ and $\phi\varphi > 1 + \phi$. Merchants have two options to receive payments. Accepting non-card payments, such as cash, costs merchants $\tau_{m,a}$ per dollar. Accepting card payments costs merchants $\tau_{m,e} + I$ per dollar. Therefore, a merchant who does not accept cards (i.e., cash store) charges p_a, while a merchant who accepts cards (i.e., card store) charges p_e. The share of card merchants is λ and the share of cash merchants is $1 - \lambda$. The values of p_a, p_e, and λ are endogenously determined as follows.

A merchant θ may earn profit $\pi_{\theta,e}$ for serving the card consumers:

$$\pi_{\theta,e} = \underset{q_\theta}{Max}(1 - \tau_{m,e} - I)p_e q_\theta - \frac{q_\theta^\varphi}{\theta} - W.$$

Alternatively, it may earn profit $\pi_{\theta,a}$ for serving the cash consumers:

$$\pi_{\theta,a} = \underset{q_\theta}{Max}(1 - \tau_{m,a})p_a q_\theta - \frac{q_\theta^\varphi}{\theta} - W.$$

At equilibrium, firms of the same efficiency must earn the same for serving either card or cash consumers. Therefore, it is required that

$$(1 - \tau_{m,e} - I)p_e = (1 - \tau_{m,a})p_a. \tag{9}$$

Note that the pricing of p_e requires $p_a \leq p_e$ so that card stores do not attract cash users. Eq. (9) then implies

$$I \geq \tau_{m,a} - \tau_{m,e}.$$

Meanwhile, card users do not shop cash stores if and only if

$$(1 + \tau_{c,a})p_a \geqslant (1 + \tau_{c,e} - R)\ p_e.$$

Eq. (9) then implies

$$\frac{1+\tau_{c,a}}{1-\tau_{m,a}} \geqslant \frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-I}.$$

In addition, $1-\tau_{m,e} > I$ so that p_e is positive. Note all these interchange pricing constraints are the same as what we derived for identical merchants.

Solving the profit-maximizing problem, a merchant θ has sale q_θ and profit π_θ for serving card users,

$$q_\theta = [\frac{\theta}{\varphi}(1-\tau_{m,e}-I)p_e]^{\frac{1}{\varphi-1}}, \quad \pi_\theta = \frac{\varphi-1}{\varphi}(\frac{\theta}{\varphi})^{\frac{1}{\varphi-1}}[(1-\tau_{m,e}-I)p_e]^{\frac{\varphi}{\varphi-1}} - W,$$

which would be the same at the equilibrium if it serves cash users.

Free entry condition requires that the marginal card merchant θ^* breaks even, so we have

$$\pi_{\theta^*,e} = 0 \Longrightarrow \theta^* = \varphi(\frac{\varphi W}{\varphi-1})^{\varphi-1}[(1-\tau_{m,e}-I)p_e]^{-\varphi}.$$

Then, the total supply of goods by card stores is

$$Q_{s,e} = \lambda\int_{\theta^*}^{\infty} q_{\theta,e} f(\theta)d\theta = \Psi\lambda[(1-\tau_{m,e}-I)p_e]^{\phi\varphi-1},$$

where $\Psi = \varphi^{-\phi}(\frac{W\varphi}{\varphi-1})^{1+\phi-\phi\varphi}\phi J^{\phi}(\frac{1}{\phi-\frac{1}{\varphi-1}})$. At the same time, the total demand of goods by card users is

$$Q_{d,e} = \eta_e[(1+\tau_{c,e}-R)p_e]^{-\varepsilon},$$

where η_e is related to the measure of card users. Therefore, the good market equilibrium achieved via card payments requires

$$Q_{s,e} = Q_{d,e} \Longrightarrow \Psi\lambda[(1-\tau_{m,e}-I)p_e]^{\phi\varphi-1} = \eta_e[(1+\tau_{c,e}-R)p_e]^{-\varepsilon},$$

which implies the price charged in a card store is

$$p_e = [\frac{\Psi\lambda}{\eta_e}(1-\tau_{m,e}-I)^{\phi\varphi-1}(1+\tau_{c,e}-R)^{\varepsilon}]^{\frac{1}{1-\phi\varphi-\varepsilon}}.$$

Similarly, the price charged in a cash store is

$$p_a = [\frac{\Psi(1-\lambda)}{\eta_a}(1-\tau_{m,a})^{\phi\varphi-1}(1+\tau_{c,a})^{\varepsilon}]^{\frac{1}{1-\phi\varphi-\varepsilon}},$$

where η_a is related to the measure of cash users.

At equilibrium, Eq. (9) can then pin down the share of merchants accepting cards versus cash:

$$\frac{\lambda}{1-\lambda} = \frac{\eta_e}{\eta_a}(\frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-I})^{-\varepsilon}(\frac{1+\tau_{c,a}}{1-\tau_{m,a}})^{\varepsilon}.$$

In the market, the total demand of card transaction value now becomes

$$\begin{aligned} TD &= p_e\eta_e[(1+\tau_{c,e}-R)p_e]^{-\varepsilon} \\ &= \Psi^{\frac{1-\varepsilon}{1-\phi\varphi-\varepsilon}}\eta_e[\eta_a(\frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-I})^{\varepsilon}(\frac{1+\tau_{c,a}}{1-\tau_{m,a}})^{-\varepsilon}+\eta_e]^{\frac{\varepsilon-1}{1-\phi\varphi-\varepsilon}} \\ &\quad (1-\tau_{m,e}-I)^{\frac{(1-\varepsilon)(\phi\varphi-1)}{1-\phi\varphi-\varepsilon}}(1+\tau_{c,e}-R)^{\frac{\varepsilon\phi\varphi}{1-\phi\varphi-\varepsilon}}. \end{aligned}$$

Recall the total supply of card transaction value derived in Section 2.2:

$$\begin{aligned} TV &= \int_{\alpha^*}^{\infty}[(\frac{I-R-T}{\beta})\alpha]^{\frac{1}{\beta-1}}g(\alpha)d\alpha \\ &= \gamma L^{\gamma}\beta^{-\gamma}(\frac{1}{\gamma-\frac{1}{\beta-1}})(\frac{K\beta}{\beta-1})^{1+\gamma-\beta\gamma}(I-R-T)^{\beta\gamma-1}. \end{aligned}$$

Therefore, the card market equilibrium $TD = TV$ implies

$$\begin{aligned} \Theta(I-R-T)^{\beta\gamma-1} &= [\eta_a(\frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-I})^{\varepsilon}(\frac{1+\tau_{c,a}}{1-\tau_{m,a}})^{-\varepsilon}+\eta_e]^{\frac{\varepsilon-1}{1-\phi\varphi-\varepsilon}} \\ &\quad (1-\tau_{m,e}-I)^{\frac{(1-\varepsilon)(\phi\varphi-1)}{1-\phi\varphi-\varepsilon}}(1+\tau_{c,e}-R)^{\frac{\varepsilon\phi\varphi}{1-\phi\varphi-\varepsilon}}. \end{aligned}$$

where $\Theta = \frac{\gamma}{\eta_e}L^{\gamma}\beta^{-\gamma}(\frac{1}{\gamma-\frac{1}{\beta-1}})(\frac{K\beta}{\beta-1})^{1+\gamma-\beta\gamma}\Psi^{\frac{\varepsilon-1}{1-\phi\varphi-\varepsilon}}$.

As before, assuming the pricing constraint $1-\tau_{m,e} > I \geq \tau_{m,a}-\tau_{m,e}$ is not binding, the monopoly card network then solves the following problem:

$$\underset{I}{Max} \;\; \Omega^m = A(I-R-T)^{\beta\gamma} \qquad \text{(Card Network Profit)}$$

$$s.t. \quad \frac{1+\tau_{c,a}}{1-\tau_{m,a}} \geqslant \frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-I}, \qquad \text{(API Constraint)}$$

$$\begin{aligned} \Theta(I-R-T)^{\beta\gamma-1} &= [\eta_a(\frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-I})^{\varepsilon}(\frac{1+\tau_{c,a}}{1-\tau_{m,a}})^{-\varepsilon}+\eta_e]^{\frac{\varepsilon-1}{1-\phi\varphi-\varepsilon}} \\ & (1-\tau_{m,e}-I)^{\frac{(1-\varepsilon)(\phi\varphi-1)}{1-\phi\varphi-\varepsilon}}(1+\tau_{c,e}-R)^{\frac{\varepsilon\phi\varphi}{1-\phi\varphi-\varepsilon}}, \end{aligned}$$

$$\text{(CMC Equation)}$$

where

$$A = KL^{\gamma}\beta^{-\gamma}(\frac{K\beta}{\beta-1})^{(1-\beta)\gamma}(\frac{\gamma}{\gamma-\frac{1}{\beta-1}}-1), \quad \Theta = \frac{\gamma}{\eta_e}(\frac{L^{\gamma}\beta^{-\gamma}}{\gamma-\frac{1}{\beta-1}})(\frac{K\beta}{\beta-1})^{1+\gamma-\beta\gamma}\Psi^{\frac{\varepsilon-1}{1-\phi\varphi-\varepsilon}}.$$

Following a similar analysis as for identical merchants, we then can show merchants' profits are affected by interchange fees in the same way as the card consumer surplus. Particularly, when the API constraint is binding, the monopoly maximum satisfies the following conditions:

$$\Theta(I-R-T)^{\beta\gamma-1} = (\eta_a+\eta_e)^{\frac{\varepsilon-1}{1-\phi\varphi-\varepsilon}}(1-\tau_{m,e}-I)^{\frac{(1-\varepsilon)(\phi\varphi-1)}{1-\phi\varphi-\varepsilon}}(1+\tau_{c,e}-R)^{\frac{\varepsilon\phi\varphi}{1-\phi\varphi-\varepsilon}},$$

$$\frac{1+\tau_{c,e}-R}{1-\tau_{m,e}-I} = \frac{1+\tau_{c,a}}{1-\tau_{m,a}}.$$

Define $Z = I - R$ and $\nu = \frac{-\varepsilon\phi\varphi}{1-\phi\varphi-\varepsilon}$. The above condition then can be rewritten as

$$\Theta(\eta_a+\eta_e)^{\frac{1-\varepsilon}{1-\phi\varphi-\varepsilon}}(Z-T)^{\beta\gamma-1} = (1-\tau_{m,e}-I)^{\nu-1}(1+\tau_{c,e}-R)^{-\nu},$$

$$\frac{1+\tau_{c,e}+Z-I}{1-\tau_{m,e}-I} = \frac{1+\tau_{c,a}}{1-\tau_{m,a}}.$$

Note that $\nu \gtreqless 1$ if and only if $\varepsilon \gtreqless 1$, so the equilibrium conditions are indeed equivalent to what we derived for identical merchants.

Now merchants' motivation for lowering interchange fees becomes clear. Credit card networks, given their market power, may charge higher interchange fees to maximize card issuers' profits as card payments become more efficient. Consequently, technology progress or enhanced competition in the card industry drives up consumer rewards and card transaction values, but may not increase consumer surplus or merchant profits. Our analysis suggests that by forcing down the interchange fee, after-reward retail prices may decrease and card users' consumption may increase. This could subsequently raise market demand for merchant sales, and hence increase merchant profits.

References

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

[2] Chakravorti, Sujit and Ted To (2007). "A Theory of Credit Cards," *International Journal of Industrial Organization*, 25, (3), 583-595.

[3] Evans, David and Richard Schmalensee (2005), *Paying with Plastic: The Digital Revolution in Buying and Borrowing,* 2nd Ed., MIT Press, Cambridge, MA.

[4] Frankle, Alan (2006), "Interchange Fees in Various Countries: Commentary on Weiner and Wright," Fed Kansas City conference proceedings: *Interchange Fees in Credit and Debit Card Industries: What Role for Public Authorities*? 51-61.

[5] Hayashi, Fumiko (2006), "A Puzzle of Card Payment Pricing: Why Are Merchants Still Accepting Card Payments?" *Review of Network Economics* 5: 144-174.

[6] Hunt, Robert (2003), "An Introduction to the Economics of Payment Card Networks," *Review of Network Economics*, 2: 80-96.

[7] MacDonald, Duncan (2006), " What MasterCard's IPO Means for Merchant Suit," *American Banker*, June 9, 80.

[8] McAndrews, James and Zhu Wang (2008), "The Economics of Two-Sided Payment Card Markets: Pricing, Adoption and Usage," Working Paper, Federal Reserve Bank of Kansas City.

[9] Katz, Michael (2001), "Reform of Credit Card Schemes in Australia II: Commissioned Report," RBA Public Document, August.

[10] Rochet, Jean-Charles (2003), "The Theory of Interchange Fees: A Synthesis of Recent Contributions," *Review of Network Economics* 2: 97-124.

[11] Rochet, Jean-Charles and Jean Tirole (2002), "Cooperation among Competitors: Some Economics of Payment Card Associations," *RAND Journal of Economics* 33: 549-570.

[12] Rochet, Jean-Charles and Jean Tirole (2003), "An Economic Analysis of the Determination of Interchange Fees in Payment Card Systems," *Review of Network Economics* 2: 69-79

[13] Rochet, Jean-Charles and Jean Tirole (2006), "Externalities and Regulation in Card Payment Systems," *Review of Network Economics* 5: 1-14.

[14] Rysman, Marc (2007), "An Empirical Analysis of Payment Card Usage," *Journal of Industrial Economics* 60: 1-36.

[15] Schmalensee, Richard (2002), "Payment Systems and Interchange Fees," *Journal of Industrial Economics* 50: 103-122.

[16] Schwartz, Marius and Daniel Vincent (2006), "The No Surcharge Rule and Card User Rebates: Vertical Control by a Payment Network," *Review of Network Economics* 5: 72-102.

[17] Shy, Oz and Zhu Wang (2008), "Why Do Card Issuers Charge Proportional Fees?" Working Paper, Federal Reserve Bank of Kansas City.

[18] Stango, Victor (2003), "Strategic Responses to Regulatory Threat in the Credit Card Market," *Journal of Law and Economics*, October, 427-451.

[19] Wright, Julian. (2003), "Optimal Card Payment Systems," *European Economic Review* 47: 587-612.

[20] Wright, Julian. (2004), "Determinants of Optimal Interchange Fees in Payment Systems," *Journal of Industrial Economics* 52: 1-26.