An Introduction to the Economics of Payment Card Networks

ROBERT M. HUNT *

Research Department, Federal Reserve Bank of Philadelphia

Abstract

Open payment card networks typically coordinate the activities of thousands of financial institutions that issue cards, millions of retail locations that accept them, and several hundred million consumers that use them. This coordination can include the collective setting of certain prices and other controversial network rules. Such practices have recently come under the scrutiny of antitrust authorities in the U.S. and abroad. This article provides a brief overview of the economics of the payment card industry, explaining some of the differences from the textbook model of competitive markets. Such differences are important factors for the antitrust analysis of payment card networks.

1   Introduction

In the United States, general-purpose payment cards – Visa or MasterCard, ATM cards, or debit cards – are ubiquitous and easy to use. In 2000, there were about 900 million general-purpose payment cards in the U.S., or about four for every adult.[1] These cards were used in 28 billion transactions worth \$1.9 trillion (BIS, 2002). Indeed, payment cards are displacing the paper check at the point of sale – the number of consumer checks written peaked during the 1990s and is now in decline (Gerdes and Walton, 2002).

This article describes how payment cards work and explains how the market for consumer payment methods differs from most other markets economists study. These differences have implications for when, why, and how the rules of antitrust law – which regulate how firms may exercise market power – should be applied to this industry. This is not a purely academic question: in the U.S. there are currently two important antitrust cases involving payment cards. Australia recently introduced new regulations for the payment card industry in that country while the U.K. and the European Union have contemplated similar measures.

[1] In this article, we focus on general-purpose credit cards, such as Visa or MasterCard, and debit cards. We do not discuss department store cards, oil company cards, or bankcards when they are used at ATMs.

Payment cards have two important distinguishing features. First, they exhibit network externalities – payment cards are more valuable to consumers when more merchants accept them. Second, in the U.S. at least, thousands of banks and other firms provide payment card services to millions of cardholders and millions of merchants who accept cards. In an environment with network externalities and so many participants, economic theory suggests that some form of coordination is beneficial, possibly essential. In the U.S., this is usually done by forming a payment card network.

Payment networks coordinate the behavior of banks, merchants, and consumers by setting certain prices and rules. In many other contexts, such practices might be considered anti-competitive. And it is also possible they can have anti-competitive effects in the market for consumer payments. Yet a careful examination of economic theory tells us this is not always the case.

The challenge to policymakers is to decide, based on the available information, whether a network's pricing strategy and rules are likely to advance or retard economic efficiency. Such conclusions are complicated by dynamic considerations – a network that exercises market power may stimulate the development of superior payment technologies, hastening its own replacement.

2   The organization and economics of payment card networks

The U.S. payment card industry involves thousands of banks participating in a number of networks, millions of consumers who find it valuable to use a payment card, and millions of merchants who find it valuable to accept those cards. Successful general-purpose payment cards were first developed in the late 1950s and their use expanded rapidly after the late 1960s (Evans and Schmalensee, 1999; Hunt, 2002; and Mandell, 1990).

2.1   Pricing and rulemaking in payment card networks

Banks engage in two types of activities within a payment card network. Card issuers – banks that offer cards to consumers – determine the level of any fees or finance charges their customers see on their regular statements.
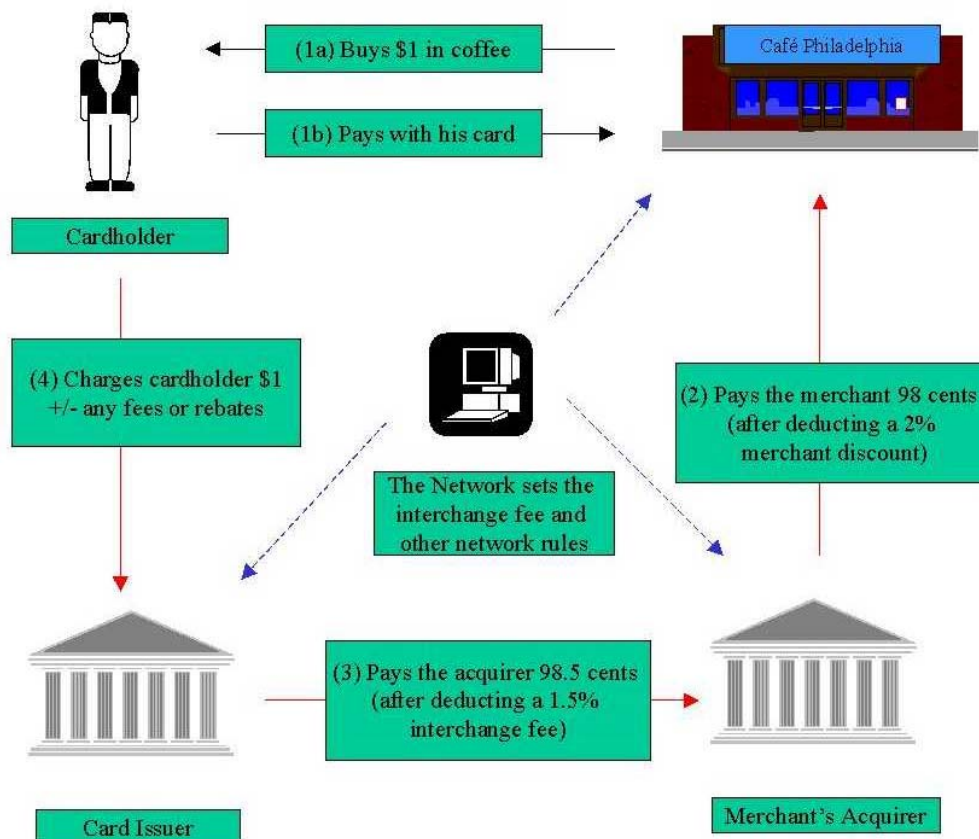
Merchants also have banks – called acquirers – that process card payments on their behalf. Merchants pay their acquirer for these services by accepting a merchant discount – when a consumer makes a $1 purchase using a payment card, the acquiring bank pays the merchant slightly less than $1 for that transaction (see Figure 1).

An open payment network, like the bankcard associations Visa and MasterCard and most electronic funds transfer (EFT) networks, allows many banks to participate. The association builds and maintains much of the infrastructure: the lines and switches required to route transaction information between different acquiring and issuing banks. The associations specify that, for each transaction, an interchange fee be paid to the bank issuing a card by the bank acting as the acquirer for the merchant.[2] In a *closed* network, the card issuer also acts as the acquirer. Such networks carry a merchant discount but not an

---

[2] In the U.S., about 1.5 percent of the value of all general-purpose-card transactions flows to issuers in the form of interchange fees – about $23 billion a year (see the Nilson Report, February 2002, No. 758).

interchange fee. Examples include American Express and Discover, which together account for about one-quarter of credit card transaction volume in the U.S.[3]

(1a) Buys $1 in coffee

(1b) Pays with his card

Cardholder

Café Philadelphia

(4) Charges cardholder $1 +/- any fees or rebates

The Network sets the interchange fee and other network rules

(2) Pays the merchant 98 cents (after deducting a 2% merchant discount)

(3) Pays the acquirer 98.5 cents (after deducting a 1.5% interchange fee)

Card Issuer

Merchant's Acquirer

Figure 1: Flow of payments in a card network

An interchange fee is one way to ensure that network participants are able to recover their costs. But, as we will see, it can also be used to coordinate the activities of banks that issue payment cards and acquiring banks that process transactions on behalf of merchants. Even though consumers do not directly pay the interchange fee, it often affects the cost and benefits of using a payment card.

The bankcard association also acts as the rule-making body for the network. In recent years, a number of these rules have received a great deal of attention. The honor-all-cards rule says that merchants wishing to accept a card brand must accept all cards issued under that brand. For example, a merchant who accepts a Visa card issued by ABC Bank must also accept Visa cards issued by XYZ Bank. In addition, merchants must accept all types of a particular brand – from platinum to plain vanilla cards.

---

[3] Schmalensee (2002) compares the fee structure in open and closed payment networks. See Guthrie and Wright (2003) for a model of competition between open and closed payment networks.

The no-surcharge rule says that merchants may not charge customers more for a transaction using one brand of card than for a transaction involving any other card brand.[4] Taken together, these rules require that merchants treat all cards issued under a given brand equally and must not favor another card brand by offering its users better prices.

Some card associations enforce an exclusivity rule, prohibiting issuers from offering certain other brands of cards. For example, in the U.S., issuers of the Visa or MasterCard brands were not permitted to issue American Express or Discover cards. This rule was struck down in the recent Department of Justice antitrust suit against the associations (see the Appendix).[5]

2.2   The new kid on the block: The debit card

Debit cards allow customers to pay for goods and services at the point of sale by authorizing a withdrawal from their checking or savings account. In the U.S., debit transactions at the point of sale only became common in the 1990s, but they have increased extremely rapidly. In the 20 years ending in 2000, consumer purchases made via debit cards rose from essentially zero to account for nearly 12 percent of all noncash consumer transactions. During this same period, the share of these transactions paid via credit card increased from 14 to 21 percent; the share paid by check fell from 86 to 59 percent (Gerdes and Walton, 2002).[6]

Most ATM cards can be used at the point of sale as debit cards. Such transactions are called PIN debit transactions because the cardholder must enter a four-digit personal identification number (PIN) to authorize the transaction. Funds are then immediately withdrawn from the associated bank account. The transaction itself is routed through an electronic funds transfer (EFT) network, for example Star, NYCE, and Pulse.[7] But in order to accept PIN debit transactions, the merchant must first install a PIN pad and have a contract with one or more EFT networks. Today, about 1.3 million different store locations in the U.S. can accept a PIN debit transaction.

Visa and MasterCard offer their own brand of debit card, but they function differently. These cards can typically be used for PIN debit transactions, but they can also authorize transactions with just the signature of the cardholder. A purchase paid for in this manner is often called a signature debit transaction. Unlike a PIN debit transaction, a signature debit transaction does not immediately remove funds from the cardholder's account; it typically takes a day or two for the transaction to clear. This delay creates some credit risk for the issuing bank because the cardholder may have insufficient funds in her account at the time the transaction clears. So, unlike with ATM cards, banks offer signature debit cards only to account holders that meet minimum credit standards.

Another important difference between the two types of debit transactions is that a signature debit transaction can be carried out with the same equipment used to authorize credit card transactions. In fact, under the honor-all-cards rule, stores that accept Visa or MasterCard must also accept the comparable brand of debit card. Currently, 4.9 million store locations in the U.S. accept one or both of these cards, offering a huge merchant base for signature debit cards (see the Nilson Report, June 2002, No. 768). Signature debit

---

[4] It also says more generally that merchants may not set a higher price to customers because they use a particular card.
[5] Balto (1998, 1999) reviews a number of arguments that enter into an antitrust analysis of such rules.
[6] These statistics are shares of the *number* of transactions, not the *value* of transactions.
[7] These networks are also the backbone of the 350,000 ATM machines around the country.

transactions are routed over the card associations' network, and the card issuer receives an interchange fee comparable to the interchange fee on a credit card transaction.[8]

2.3   Network effects and fixed costs in payment card networks

Payment networks function differently from most markets, in part because of network effects. A payment card is more valuable to consumers when more merchants accept the card. At the same time, merchants are more willing to accept a card if they know many consumers use it. Every consumer who obtains a card and every retailer who accepts a card increase the value of the network to all other cardholders and all other merchants who accept it.

This interdependence is sometimes called an *adoption externality*, and it has a number of important implications for payment networks. If a network is introduced on a small scale, there is inertia – consumers and merchants have little incentive to join. Since such a network would clearly be unprofitable, it would never be launched. But if a network is launched on a large scale, it is possible that many consumers will carry the card and many stores will accept it. If that happens, even more stores are likely to accept the card, which may induce even more people to carry the card, and so on. In this environment, the expectations of both consumers and merchants are clearly important.

A second implication is that there may be a significant barrier to entry in payment cards, which might permit established networks to maintain prices above costs for some time. The reason is that in order to successfully enter the market, a rival network must do so on a large scale. But the market may not be large enough to support more than a few networks at such a scale.

Payment cards may also exhibit *usage externalities*.[9] For example, after consumers decide whether to get a payment card, they must also decide how often they will use it. But when a consumer uses a card, this may provide benefits to the merchants they purchase from. Similarly, when a merchant decides to accept a card for a transaction, this may create a benefit to the user of the card. While each party faces its own private fees to use the card for a transaction, each card transaction provides benefits to both types of users. An externality exists whenever these private fees do not reflect the benefits or costs conferred to the other party.

Another factor that distinguishes payment cards from many, but certainly not all, industries is the large fixed cost associated with establishing a viable payment network. Modern payment card networks require large investments in communications and computing facilities in order to make card transactions convenient to customers and merchants and to minimize fraud. The latter is especially important to card issuers because network rules typically promise to pay merchants for fraudulent transactions as long as the network's procedures are followed.[10] If networks did not make this promise to merchants, fewer merchants would be willing to accept payment cards. But card issuers will accept the

---

[8] A $40 signature debit transaction generates an interchange fee of about 60 cents, or 1.5 percent, while a comparable PIN debit transaction generates an interchange fee of about 18 cents, or 0.5 percent (ATM and Debit News 2002).

[9] Rochet (2003) describes this, rather than the adoption externality, as the fundamental externality in payment systems. It is also the basis for the analysis in Baxter (1983), one of the first theoretical models of a payment card network.

[10] An exception to this rule for credit cards in the U.S. is card-not-present transactions, such as Internet or telephone orders, where network rules stipulate that fraud losses are borne by the merchant.

cost of fraudulent transactions only if the network's antifraud technology is sufficiently effective.

Network effects and fixed costs at the network level may explain why bankcard associations and EFT networks use many of the strategies described earlier (Chang and Evans, 2000). Setting an appropriate interchange fee is one way to ensure that network participants take into account network effects, presumably increasing card usage (see section 3). This in turn reduces the unit cost of card transactions, making payment cards more attractive to use or accept.

Setting the fee at the network level eliminates costs associated with bargaining between individual card issuers and acquirers and uncertainty about the actual costs of a card transaction (Baxter, 1983; Small and Wright, 2001).[11] Consumers are more likely to use a payment card if they know where it will be accepted and on what terms. The honor-all-cards rule and the no-surcharge rule reduce the uncertainty consumers would otherwise face. This was especially important in the late 1960s and 1970s, when the card associations were trying to build nationwide acceptance of credit cards issued primarily by small banks.

But do these mechanisms remain essential after a payment card network becomes well established? Katz (2001) questions whether network effects are relevant for mature payment networks, especially if consumers and merchants would incur a fixed cost to switch to an alternative means of payment. The important question is really whether there remains a large usage externality and it is not internalized by some other means. This question has yet to be thoroughly addressed in empirical research.

There could also be a dark side to all this coordination – these rules might be used to enhance a dominant network's market power. Such allegations form the basis of an important antitrust case in the U.S., regulation of the payment card industry in Australia, and far-reaching inquiries in Europe (see the Appendix for a summary of recent legal and regulatory challenges around the world). An ingredient common to most regulatory proposals is the setting of a cap on interchange fees based on the marginal or average cost.[12]

The following sections examine the role of interchange fees, the no-surcharge rule, and the-honor-all-cards rule in consumer payment networks.

3   The pros and cons of interchange fees

An interchange fee can play two roles in solving what might first appear to be an intractable problem: how to maximize the value of an open payment network, while also ensuring that participating members are at least able to cover their costs. We will illustrate each of these roles separately. The first step is to understand how a profit-maximizing network would set prices for its services.

---

[11] Why doesn't the collective setting of the interchange fee – an obvious example of price fixing among competitors – violate U.S. antitrust law? Because federal courts recognize there are situations where such arrangements may promote competition. See *Broadcast Music, Inc. v. Columbia Broadcasting Co.*, 441 U.S. 1 (1979) and *National Bancard Corp. (NaBANCO) v. Visa U.S.A., Inc.* 596 F.Supp. 1231 (S.D. FL 1984), 779 F2d 592 (11th Cir 1986).

[12] Gans and King (2003) compare a number of the proposed regulatory approaches to the (socially) optimal interchange fees derived in the theoretical literature.

3.1   Some intuition on the pricing of payment card transactions

Suppose that a payment network can directly set the prices that merchants and consumers pay for its services. We have already noted that consumers' and merchants' demand for these services are interdependent.[13] In order to maximize profits, the network must set prices taking into account how consumers and merchants respond to price. In a number of theoretical models of this problem, the prices chosen will typically have the following properties:

(1) Consider a small reduction in the price chosen by the network: The profits earned on the resulting increase in transaction volume are exactly offset by the decline in profits earned on all existing transactions.

(2) A small reduction in the price charged to consumers will increase transaction volume, in percentage terms, by as much as a comparable reduction in the price charged to the merchant.

But (2) has an additional implication:

(3) A higher share of the total price paid (the sum of the prices charged to the consumer and merchant) is paid by that side of the transaction whose demand curve is less responsive (elastic) to changes in price.

The first property occurs in models where a network can increase profits by limiting output (Rochet and Tirole, 2002b). The second and third results are common in models that capture the two-sided nature of card usage and acceptance (e.g. Chang and Evans, 2000; Schmalensee, 2002; and Wright, 2001)

Somewhat surprisingly, (3) does not imply the prices consumers' and merchants' pay are set according to an inverse elasticity rule (Rochet and Tirole, 2002b). It turns out that the elasticity of the less elastic demand curve, calculated at the relevant equilibrium price, can be higher than the elasticity for the more elastic demand curve, calculated at the relevant equilibrium price.[14]

3.2   What interchange fees can do

Now it is possible that when prices are set in the manner described above, the consumer's price (or the merchant's price) is below the cost incurred by the consumer's bank (or the merchant's bank). In that case, an interchange fee can be used to transfer resources from one part of the network to another to ensure that all network participants are able to recover the costs of enabling the transaction (Baxter, 1983).[15]

---

[13] In a number of papers, this interdependence is modeled by defining network transaction volume as the product of consumers' and merchants' demands for payment services.

[14] To see this, suppose the consumer's demand curve is $Q_c(p_c) = A - 2b \cdot p_c$ and the merchant's demand curve is $Q_m(p_m) = A - b \cdot p_m$. Condition (2) is satisfied where $p_m - p_c = A/2b > 0$. Conditions (1) and (2) imply that $p_c = c/3$ and $p_m = A/2b + c/3$, where c represents the total cost of providing a transaction. At those prices, $\eta_m > \eta_c$.

[15] Continuing the preceding example, if costs at the consumer's bank account for more than a third of total costs (in the absence of any transfers from acquiring banks), the consumer's bank will lose money on every transaction.

It turns out that in most open payment networks, interchange fees play another important role. The reason is that such networks typically do not exercise direct control over the prices that consumers and merchants pay for a transaction. Instead, they influence those prices by setting the interchange fee.

For example, suppose the network raises an interchange fee paid to card issuers so that each card transaction is more profitable for those banks. Card issuers will seek out more cardholders either by offering them more benefits or by reducing cardholder fees. Merchants will observe more cardholders using the card. The trade-off is that raising the interchange fee raises costs for acquiring banks, and at least some of that cost is passed on to merchants. The higher cost of card transactions may cause some merchants to stop accepting the card. By selecting the appropriate interchange fee, a network can ensure the conditions described in the previous section are satisfied.

Some models take into account differences in market power on different sides of a network (e.g. issuers vs. acquirers). In that case, an interchange fee can also be used to shift profits to the side of the network where less profit is dissipated by competition between network participants (Schmalensee, 2002; Wright, 2001).

To summarize, economic theory suggests two factors that are likely to influence the size of a privately optimal interchange fee, that is, one that maximizes the value of a payment network. The first is the relative size of costs borne by issuing and acquiring banks – banks will not willingly participate if they cannot recover their costs. The second is the degree of price sensitivity exhibited by cardholders on the one hand and merchants on the other.

3.3   Networks may not choose the interchange fee best for society

But the best interchange fee for a particular payment network need not be the best from the standpoint of consumers. The economic literature has explored a variety of reasons a privately determined interchange fee may not correspond to the fee that maximizes social welfare (Carlton and Frankel, 1995; Chakravorti and To, 2002; Gans and King, 2002; Rochet and Tirole, 2002a; and Wright, 2001).[16] But for this article, we will focus on just one.

Suppose that a merchant charges the same prices regardless of the manner in which consumers pay. For example, customers who pay with a credit card pay the same price as customers who pay with cash. While different payment instruments mean different costs for the merchant, those costs are not reflected in the prices paid by every customer. Users of the cheaper form of payment bear some of the costs created by customers who use a more expensive form of payment. Consumers tend to overuse the more expensive form of payment because they enjoy all the benefits but do not bear all the costs. At the economywide level, this could mean a payment network would grow too large because purchases outside the network are subsidizing purchases made within the network.[17]

It is possible a network can exploit this implicit subsidy by raising interchange fees while reducing cardholder fees. Merchants would pass on these costs to all their customers, increasing the subsidy noncard users pay to card users. If some of the increased interchange revenues are passed on to cardholders (through lower fees or more perks), this would, in turn, increase the number of cardholders (Carlton and Frankel, 1995; Frankel,

---

[16] See the references for a more complete list of the relevant literature.
[17] It is possible the subsidy can work across different dimensions. In Chakravorti and Emmons (2001), card users that carry balances from period to period subsidize transactors.

1998; Schwartz and Vincent, 2002).[18] Merchants may not like this outcome, but they may be reluctant to stop accepting the card if they think cardholders will take their business elsewhere.

The actual outcome depends crucially on how consumers react to changes in prices and the nature of competition among merchants. If customers who do not use the card respond to small price increases by switching to merchants that accept only cheaper cards, or just cash, this implicit subsidy must be small. Thus, an important insight gleaned from theoretical models of payment networks is that the effects of interchange fees depend on the degree of market power enjoyed by merchants (Gans and King, 2001).

From the standpoint of policy, the literature suggests three important conclusions. First, a variety of factors may result in a privately determined interchange fee that diverges from the socially optimal fee, resulting in payment networks that are either too large or too small.[19] Second, a zero interchange fee is generally not socially optimal. And third, an interchange fee based purely on costs, ignoring the effect of changes in prices on consumers' and merchants' demand for payment services, is generally not socially optimal.

4   The pros and cons of surcharges

So far, we have assumed that merchants do not set different prices for different card transactions. What happens if payment networks permit merchants to add a fee to transactions when consumers use a more expensive payment method?

In principle, merchants could pass on any difference in their cost of using different payment cards to the customers using those cards. This would eliminate any subsidy implicitly paid by customers using cheaper payment methods to those who use more expensive means, and it would encourage consumers to use the most efficient forms of payment. So if we think that a costly payment instrument is being used too much, allowing merchants to surcharge may be a useful remedy.

But permitting surcharges may have a second effect. If merchants are willing to pass on costs in this way, an open payment network cannot use an interchange fee to influence the prices paid by merchants and consumers (Gans and King, 2001). To see this, imagine what would happen if the network raised the interchange fee and card issuers passed on the additional revenue to cardholders via lower fees or additional perks. Suppose also that card acquirers pass on the higher interchange fee to retailers by raising the merchant discount.

Merchants could respond by increasing the surcharge on card transactions, essentially negating the increased benefits provided by card issuers. So with surcharging, raising interchange fees in order to stimulate card use will not be very effective. If network effects are important and cannot be taken into account by some other means, the result could be underutilization of payment cards.[20]

---

[18] Wright (2001) specifies conditions where the subsidy can work in the opposite direction, with benefits flowing from card users to cash users.

[19] Schmalensee (2002) presents a model where the interchange fee set by a profit maximizing network, assuming linear demand functions, is also welfare maximizing. But this result does not hold for more general assumptions about demand or the objective function of the network.

[20] Even in the absence of network effects, card usage may be too low in the presence of surcharging if merchants equate card usage with a high willingness to pay and price discriminate accordingly (Gans and King, 2002). This is another example of why the nature of merchant competition is relevant to any welfare analysis.

4.1   But will merchants surcharge?

So far, we have assumed that given the opportunity, merchants actually would impose surcharges.[21] But there is some reason to doubt that much surcharging would occur. In the U.S., since 1975, federal law has permitted merchants to offer a discount for cash purchases. Yet in the early 1980s, less than 10 percent of retailers offered cash discounts (Board of Governors, 1983). The practice was observed more frequently in certain segments, such as gas stations, but even there it became less common once those stores began to accept bank-issued credit and debit cards (Evans and Schmalensee, 1999). In Sweden and the Netherlands, two countries that banned no-surcharge rules during the 1990s, less than 10 percent of retailers report surcharging their customers and there has been little change in the merchant discount (IMA Market Development, 2000).[22]

There are different opinions about why so little surcharging occurs. Kitch (1990) argues that, in the U.S., other regulatory barriers made it relatively costly for merchants to offer cash discounts. Frankel (1998) argues that merchants believe any benefit of charging different prices is not worth risking a negative reaction from customers. He also wonders whether consumers react more strongly to fees charged at the point of sale than to fees that appear later on their bank statement. This raises an important empirical issue: the theoretical models assume that, other than float, consumers respond to transaction fees in the same way, regardless of where they appear.

5   The pros and cons of an honor-all-cards rule

Few question whether payment networks should be permitted to have an honor-all-cards rule; the question is how broadly it should be applied. Suppose a network issues two types of cards – a credit card and a debit card – under the same brand. Should a merchant be required to accept both types of cards even it prefers to accept only one type?

This is the central argument in the Wal-Mart antitrust case: The plaintiffs argued that the bankcard associations used the honor-all-cards rule to impose an illegal tying arrangement, forcing merchants that accept a brand of credit card to accept the same brand of signature debit card (Balto, 2000).[23] Because card issuers receive more interchange revenue from signature debit transactions, they have an incentive to subsidize consumers' use of these cards. This, in turn, influences consumers' choices about signature vs. PIN debit transactions. Merchants pay these higher fees and pass at least some of the cost on to consumers via higher prices. During the 1990s, the use of both types of debit cards grew immensely (see Figure 2). But the share of all debit transactions using a PIN fell from about 60 percent in 1993 to about 38 percent in 2002.

---

[21] At least in theory, we might worry there could be too much surcharging. This might occur if a no-surcharge rule plays the role of a commitment device, which prevents opportunistic merchant behavior (Wright, 2003).

[22] The data for the Netherlands are found in the 2001 European Competition Commission decision (COMP/29.373). Katz (2001) points out that among retailers in the Netherlands who were aware of the legal change, 20 percent did surcharge.

[23] In a tying case, the plaintiff tries to prove that the defendant is using the market power it enjoys in one market to extract profits from another, typically more competitive, market.
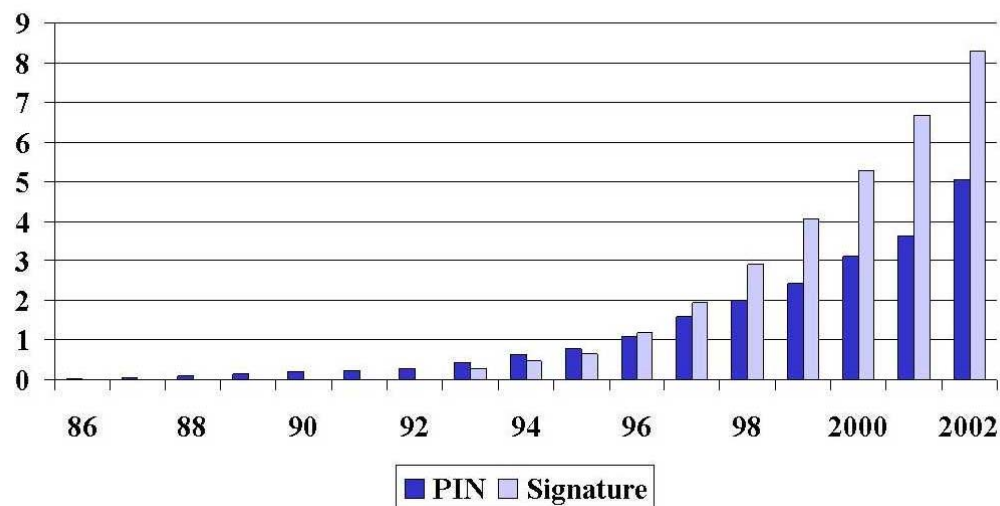
Figure 2: PIN debit vs. signature debit transaction volume (billions per year)

Sources: EFT Network Data Book, Debit Card and POS Market Data Book, and Card Industry Directory.

If we view credit and debit cards as distinct products, it seems reasonable that an honor-all-cards rule should be enforced separately for each type of card. In other words, merchants could be allowed to refuse the debit card but still be required to accept all credit cards issued under that brand. Similarly, merchants could separately decide whether to accept a debit card brand but would then be required to accept all debit cards issued under that brand.

The likely result would be that merchants would pay different fees for credit card and signature debit transactions. If signature and PIN debit transactions offer merchants the same benefits, the interchange fee for a signature debit would have to fall in order for these cards to remain competitive with a PIN debit, since the cost of processing a PIN debit is less than the cost of processing a signature debit. Users of signature debit cards would presumably pay higher fees or enjoy fewer benefits, since the use of the cards could not be subsidized as heavily.[24]

5.1   Do credit and debit cards compete in the same market?

Applying the honor-all-cards rule to both credit and debit cards of the same brand seems more reasonable if these products actually compete in the same market. But do they? Economists typically define the boundaries of a market based on how consumers respond to price changes: If a change in the price of one good induces consumers to switch to another good, these goods are substitutes. If only a small change in price is sufficient to cause consumers to switch to the substitute product, they compete in the same market.

---

[24] The Wal-Mart case was settled in April 2003, on terms favorable to the merchants (see the Appendix for details). So we will soon see if this forecast is accurate.

What can we say about consumer substitution between credit cards and debit cards? There is at least some evidence that, in the U.S., consumers do not use credit and debit cards in the same way. For example, consumers are more likely to use debit cards in drug and grocery stores than they are at department stores (see Table 1). In addition, consumers often initiate a PIN debit transaction larger than their purchase and receive the difference in cash, a feature not available in a credit card transaction. While credit cards provide an explicit line of credit, debit cards do not. Such differences suggest that credit and debit cards are not pure substitutes as a means of payment.

	Stores with PIN Pads (percent)	Percentage of Store Sales Paid via			
		Cash	Check	Credit Card	Debit Card
All Stores	50	35	21	25	8
Discount	43	47	17	27	3
Drug	73	41	17	26	14
Supermarket	100	44	32	11	12
Department Store	20	29	15	26	2
Home Center	7	21	27	26	6
Apparel	38	28	19	32	10

Table 1: Data on the use of debit and other forms of payment (in 1999)

Note: These statistics are derived from a survey of large retail chains. It is not a representative sample of the retail sector.   Source: "Survey of Retail Payment Systems," Chain Store Age (December 1999).

Two federal courts recently reached this conclusion in separate antitrust suits brought against the bankcard associations by the Department of Justice and a class of merchants.[25] But in a number of previous antitrust decisions, the courts defined the relevant market more broadly to include cash, checks, department store cards, and ATM cards (e.g. the *NaBanco* cases). Even in the recent Department of Justice decision, the judge recognized that the emergence of all-in-one cards – a single card that can be used for credit, signature debit, or PIN debit transactions – may increase consumers' willingness to substitute between these different forms of payment.

6   Yesterday, today, and tomorrow

At the end of the day, policymakers need to know the answer to the following question: Does the conduct of a payment network benefit or harm consumers? In antitrust cases, judges are often forced to weigh the static costs of certain conduct against any dynamic benefits it may offer. This is not easy to do when it is not clear how a market would have developed in the absence of the conduct under scrutiny.

Suppose we return to the 1980s, before a thriving debit card market developed. How might such a market be developed? The method chosen by Visa and MasterCard was to

---

[25] See *U.S. v. Visa U.S.A., Inc.* 163 F. Supp. 2d. 322 (S.D.N.Y. 2002) and *In re Visa Check/MasterMoney Antitrust Litigation*, 96-CV-5238 (E.D.N.Y. 2003).

graft debit cards on to the existing credit card networks. Using their honor-all-cards rule, the associations ensured that millions of merchants would accept signature debit cards. Using their no-surcharge rule, the associations ensured that these cards would be accepted on terms equal to those of any other debit card (Evans and Schmalensee, 1999). Even with these advantages, signature debit cards were not immediately successful. Their success occurred only after credit cards were commonly accepted in more price-sensitive retail segments, and virtually all merchants were using modern electronic terminals to authorize transactions. The networks encouraged these developments through targeted reductions in interchange fees (Felgran, 1985; Felgran and Ferguson, 1986).

When network effects and dynamic issues are both important, as they appear to be in this industry, policymakers face a difficult problem in deciding what remedies, if any, will benefit consumers in the long run. On the one hand, network rules and pricing strategies may be essential elements in the successful launch of a payment card network and its subsequent expansion. On the other hand, once a payment network is well established, it is possible the same rules can lead to their overutilization and to pricing well in excess of costs.

A further complication is that the pricing strategy of existing payment networks affects how and when newer, and presumably better, forms of payment are adopted. New forms of payment must overcome any subsidies consumers receive when using today's payment instruments, and this may require offering subsidies of their own. For example, it is possible that the relatively slow adoption of debit cards in the U.S. was due in part to subsidies implicit in the pricing and processing of consumer checks (Caskey and Sellon Jr., 1994). Today, higher interchange fees paid on debit transactions coincide with debit cards' replacement of checks in many consumer transactions. Policymakers should take these factors into account when examining competition among consumer payment networks or between competing payment instruments.

7   References

ATM and Debit News (2002) "Letting Cardholder Pick Debit Function May Bolster Profits," August 8.

Bank for International Settlements (2002) *Statistics on Payment and Settlement Systems in Selected Countries*. Basel: Bank for International Settlements.

Balto, David (1998) "Access Demands and Network Joint Ventures," in David Gabel and David F. Weiman, eds., *Opening Networks to Competition: The Regulation and Pricing of Access*. Boston: Kluwer Academic Publishers.

Balto, David (1999) "Networks and Exclusivity: Antitrust Analysis to Promote Network Competition," *George Mason Law Review*, 7: 523-76.

Balto, David (2000) "Creating a Payment System Network: The Tie That Binds or an Honorable Peace?" *The Business Lawyer*, 55: 1391-1408.

Baxter, William F. (1983) "Bank Interchange of Transactional Paper: Legal and Economic Perspectives," *Journal of Law and Economics*, 26: 541-88.

Betz, Kip (2003) "MasterCard, Visa to Pay $3 Billion to Resolve Card Suit; Will Modify Debit Card Policy, Fees," *BNA Banking Report*, 80: 739-40.

Board of Governors of the Federal Reserve System (1983). *Credit Cards in the U.S. Economy: Their Impact on Costs, Prices, and Retail Sales*. Washington: Board of Governors of the Federal Reserve System.

Carlton, Dennis W., and Alan S. Frankel (1995) "The Antitrust Economics of Credit Card Networks," *Antitrust Law Journal*, 63: 643-68.

Caskey, John P. and Gordon H. Sellon Jr. (1994) "Is the Debit Card Revolution Finally Here?" *Federal Reserve Bank of Kansas City Economic Review*, Fourth Quarter, 79-95.

Chakravorti, Sujit and William R. Emmons (2001) "Who Pays for Credit Cards?" Federal Reserve Bank of Chicago Public Policy Series (EPS-2001-1).

Chakravorti, Sujit and Ted To (2002) "A Theory of Credit Cards," mimeo, Federal Reserve Bank of Chicago.

Chang, Howard and David S. Evans (2000) "The Competitive Effects of the Collective Setting of Interchange Fees by Payment Card Systems," *Antitrust Bulletin*, 45: 641-77.

Evans, David S. and Richard Schmalensee (1995) "Economic Aspects of Payment Card Systems and Antitrust Policy Toward Joint Ventures," *Antitrust Law Journal*, 63: 861-901.

Evans, David S. and Richard Schmalensee (1999) *Paying with Plastic: The Digital Revolution in Buying and Borrowing*. Cambridge, MIT Press.

Evans, David S. and Richard Schmalensee (2001) "Some Economic Aspects of Antitrust Analysis in Dynamically Competitive Industries," NBER Working Paper No. 8268.

Felgran, Steven D. (1985) "From ATM to POS Networks: Branching, Access, and Pricing," *New England Economic Review*, May/June: 44-61.

Felgran, Steven D. and R. Edward Ferguson (1986) "The Evolution of Retail EFT Networks," *New England Economic Review*, July/August: 42-56.

Frankel, Alan S. (1998) "Monopoly and Competition in the Supply and Exchange of Money," *Antitrust Law Journal*, 66: 313-62.

Gans, Joshua S. and Steven P. King (2001) "The Neutrality of Interchange Fees in Payment Systems," mimeo, University of Melbourne.

Gans, Joshua S. and Steven P. King (2002) "A Theoretical Analysis of Credit Card Regulation," mimeo, University of Melbourne.

Gans, Joshua S. and Steven P. King (2003) "Approaches to Regulating Interchange Fees in Payment Systems," *Review of Network Economics*, this issue.

Gerdes, Geoffrey R. and Jack K. Walton II (2002) "The Use of Checks and Other Noncash Payment Instruments in the United States," *Federal Reserve Bulletin*, 360-374.

Guthrie, Graeme and Julian Wright (2003) "Competing Payment Systems," Department of Economics, University of Auckland Working Paper No. 245.

Hunt, Robert M. (2002) "The Development and Regulation of Consumer Credit Reporting in America," Federal Reserve Bank of Philadelphia Working Paper No. 02-21.

IMA Market Development (2000) "Study Regarding the Effects of the Abolition of the Non-Discrimination Rule in Sweden," February.

Katz, Michael L. (2001) "Reform of Credit Card Schemes in Australia II: Commissioned Report," Reserve Bank of Australia.

Katz, Michael L. and Carl Shapiro (1994) "Systems Competition and Network Effects," *Journal of Economic Perspectives*, 8: 93-115.

Kitch, Edmund W. (1990) "The Framing Hypothesis: Is It Supported by Credit Card Issuer Opposition to a Surcharge on a Cash Price?" *Journal of Law, Economics and Organization*, 6: 217-33.

Liebowitz, S.J. and Stephen E Margolis (1994) "Network Externality: An Uncommon Tragedy," *Journal of Economic Perspectives*, 8: 133-50.

Mandell, Lewis (1990). *The Credit Card Industry: A History*. Boston: Twayne Publishers.

Reserve Bank of Australia (2000) "Debit and Credit Card Schemes in Australia: A Study of Interchange Fees and Access," Reserve Bank of Australia.

Reserve Bank of Australia (2002) "Reform of Credit Card Schemes in Australia IV: Final Reforms and Regulation Impact Statement," Reserve Bank of Australia.

Rochet, Jean-Charles (2003) "The Theory of Interchange Fees: A Synthesis of Recent Contributions," *Review of Network Economics*, this issue.

Rochet, Jean-Charles and Jean Tirole (2002a) "Cooperation Among Competitors: Some Economics of Payment Card Associations," *RAND Journal of Economics*, 33: 549-70.

Rochet, Jean-Charles and Jean Tirole (2002b) "Platform Competition in Two-Sided Markets," IDEI working paper.

Schmalensee, Richard (2002) "Payment Systems and Interchange Fees," *Journal of Industrial Economics*, 50: 103-22.

Schwartz, Marius and Daniel R. Vincent (2002) "Same Price, Cash or Card: Vertical Control by Payment Networks," Working Paper No. 02-01, Department of Economics, Georgetown University.

Small, John and Julian Wright (2001) "The Bilateral Negotiation of Interchange Fees in Payment Schemes," mimeo, NECG and University of Auckland.

Wright, Julian (2001) "The Determinants of Optimal Interchange Fees in Payment Systems," Department of Economics, University of Auckland Working Paper No. 220.

Wright, Julian (2003) "Optimal Card Payment Systems," *European Economic Review*, forthcoming.

8   Appendix: Legal and regulatory challenges to payment card systems

8.1   United States

In October 1996, Wal-Mart and other retailers filed an antitrust suit against Visa and MasterCard.[26] This suit later became a class action, and the plaintiffs now represent several million retail locations. In April 2003, the district court ruled on the pre-trial motions, reaching a number of conclusions favorable to the plaintiffs' tying claim.[27] The case was settled shortly thereafter. The bankcard associations agreed to revise their honor-all-cards rules so that merchants can separately decide whether to accept their brands of credit and debit cards, to reduce interchange fees charged on signature debit transactions, and to pay $3 billion in damages over a ten year period. The reduction in interchange fees in 2003 alone is expected to save merchants $1 billion (Betz, 2003).

In 1998 the U.S. Department of Justice (DOJ) filed a separate antitrust suit against Visa and MasterCard. Among other things, DOJ objected to the associations' exclusivity rules, which prevent banks that issue Visa or MasterCard credit cards from simultaneously issuing a Discover or American Express card. In October 2001, the trial court invalidated these rules.[28] The case is currently under appeal.

8.2   Europe

In July 2002, the Competition Directorate of the European Commission announced a settlement with Visa that addresses multilateral interchange fees levied on certain credit and debit transactions that involve banks in more than one member state.[29] Under the terms of the agreement, Visa pledges to reduce those fees gradually over the next five years and to keep them below a cap that will be calculated each year on the basis of card issuers' costs.

Allowable costs include transaction processing, financing the interest free period enjoyed by cardholders, and certain payment guarantees provided to merchants. Visa also agreed to amend its rules so that its interchange fee can be disclosed to merchants.

---

[26] *In re Visa Check/MasterMoney Antitrust Litigation*, No. 00-7699 (2d Cir 2001).
[27] *In re Visa Check/MasterMoney Antitrust Litigation*, 96-CV-5238 (E.D.N.Y. 2003).
[28] *U.S. v. Visa U.S.A., Inc*. 163 F. Supp. 2d. 322 (S.D. NY 2001).
[29] Case No. COMP/29.373—Visa International-Multilateral Interchange Fees, *Official Journal of the European Community*, July 24, 2002.

In a separate decision published in November 2001, the Commission concluded that Visa's honor-all-cards rule did not restrict competition even when applied to different types of cards (e.g. credit and debit) within the same brand (for example, Visa).[30]

8.3   Australia

In August 2002, the Reserve Bank of Australia (RBA) announced regulations that apply to domestic credit card transactions using Visa, MasterCard, or Bankcard credit cards.[31] As of January 2003, merchants were permitted to surcharge transactions using these cards. In October 2003, credit card interchange fees will be capped according to a cost-based formula that will be revised every three years.[32]

Allowable costs include authorizing and processing transactions, financing the interest-free grace period enjoyed by cardholders, and costs resulting from card fraud and its prevention. The card associations must provide the RBA with audited data on these costs each year. The RBA also invalidated certain card association rules that it concluded were inhibiting entry by monoline credit card banks and merchant acquirers.

---

[30] Case No. COMP/29.373—Visa International, *Official Journal of the European Community*, November 10, 2001.

[31] Reserve Bank of Australia (2002). Visa and MasterCard have challenged the decision on procedural grounds.

[32] The RBA is imposing the cap against the volume weighted average interchange fee levied on card transactions rather than specifying specific caps for different kinds of transactions. The RBA expects that once implemented, the caps will reduce average interchange fees about 40 percent.