Interchange Fees and Payment Card Networks: Economics, Industry Developments, and Policy Issues

Robin A. Prager, Mark D. Manuszak, Elizabeth K. Kiser, and Ron Borzekowski

2009-23

Interchange Fees and Payment Card Networks: Economics, Industry Developments, and Policy Issues

Robin A. Prager
Mark D. Manuszak
Elizabeth K. Kiser
Ron Borzekowski

May 13, 2009

Executive Summary

In many countries around the world, electronic card-based payments have been replacing older types of payments at a rapid rate. In the United States, use of both debit cards and credit cards has been rising rapidly, while check volumes have been declining. The increased use of electronic payment methods has generated a number of public policy debates. One prominent debate concerns *interchange fees*. This paper is intended to provide background for understanding the interchange fee debate. The paper describes the operation of a typical payment card system, presents a summary of the economic theory underlying interchange fees, and discusses various developments in the U.S. payment cards industry, as well as legal and regulatory developments abroad. The paper concludes with a discussion and critical evaluation of a number of potential policy interventions.

Interchange fees typically involve a payment from a merchant's bank to a card user's bank for each debit card or credit card transaction, are determined at the network level, and are the same for all banks participating in a network. These fees are generally passed through to merchants by their banks and comprise a large fraction of the fees that merchants pay to their banks for processing card transactions. Card-issuing banks often use a portion of their interchange fee revenue to encourage card use by offering their cardholders rewards, such as cash rebates or airline miles, that increase with card use. In recent years, increases in interchange fee rates, together with growth in the volume of card transactions, have led to a dramatic rise in the total value of interchange fee payments and, consequently, in merchants' cost of accepting payment cards. These cost increases have given rise to significant concerns among merchants.

Merchant concerns focus on the level and sometimes the very existence of interchange fees. Merchant groups and their supporters contend that a network-determined interchange fee serves as a means for otherwise competing banks to collude on their fees for card transactions and to avoid negotiation of fees with merchants. These parties argue further that the primary alternative available to merchants – rejection of a network's cards – is not a viable option in the case of the major card networks because customers have come to expect acceptance of those cards. As a result, merchants contend

1

that the major card networks are able to set excessively high interchange fees, which lead to correspondingly high merchant fees, inflated incentives (i.e., rewards programs) for consumers to use the cards, and supranormal profits for the networks and their card-issuing banks. Merchants also raise concerns about restrictive network rules, such as *honor-all-cards rules* and *no-surcharge rules*.[1]

In contrast, the card networks and their supporters contend that payment cards provide substantial value to both consumers and merchants, and that interchange fees are essential for the proper operation of the card networks. They note that a common interchange fee standardizes the terms of exchange between the merchant's bank and the card user's bank and plays an economically important role in influencing the incentives of merchants to accept cards and consumers to hold and use them. These groups argue that, when setting its fees, a card network must recognize the need to attract consumers and merchants, both of whom are necessary for the card network to exist. In their view, the resulting fees attempt to balance the two sides of the payment card market to maximize the value of the network, including the value of card services for both consumers and merchants.

The core economic and policy issues concern the effects of interchange fees on the extent of card use and the welfare of different parties in the economy. In terms of efficiency, the key issue is whether current patterns of retail payments appropriately reflect the social costs and benefits of the various payment methods. If the transaction fees faced by either merchants or card users are too high (or too low), then some payment methods will be overused and others will be underused relative to the socially optimal outcome. However, determining whether observed patterns of card fees (including interchange fees) and card usage are socially optimal is an extremely difficult task.

In terms of equity, interchange fees and associated transaction fees can generate transfers among non-card users, card users, merchants, and banks that increase the welfare of some parties while reducing the welfare of others. High interchange fees

---

[1] Honor-all-cards rules require a merchant that accepts a network's cards of a given type (e.g., its credit cards) to accept all of the network's cards of that type regardless of the identity of the card issuer or the specific card program (e.g., premium or basic credit cards). No-surcharge rules prohibit a merchant from imposing a surcharge on a customer for using a network's payment card.

could support low transaction fees (or high rewards) for cardholders, thereby potentially making card users better off. At the same time, if the net effect of card acceptance is to increase merchant costs, and if merchants do not set prices that vary by payment method, then merchant acceptance of cards could lead to higher retail prices for all consumers, including those who pay with alternative methods and receive none of the direct benefits associated with card use.

To consider these efficiency and equity issues, economists have recently developed theoretical models of the pricing of card services in payment card markets. These models generally suggest that a transfer payment between the merchant's bank and the consumer's bank (i.e., an interchange fee) may be necessary to induce efficient use of card-based payment methods. This transfer payment affects the costs of card transactions for those banks and, in turn, the transaction fees that they charge merchants and cardholders (i.e., merchant discounts and fees or rewards for card use). These fees ultimately influence the volume of transactions on a card system through their effect on the willingness of merchants to accept cards and consumers to use them for purchases. With an appropriately chosen interchange fee, a payment card will be used in a transaction whenever doing so yields a higher level of overall social welfare than would be obtained by using an alternative payment method. With such an interchange fee, the number of card transactions will be economically efficient.[2] A few characteristics of an efficient interchange fee are worth noting:

- In general, an efficient interchange fee is not solely dependent on the cost of producing a card-based transaction nor is it equal to zero.
- An efficient interchange fee may yield prices for card services to each side of the market that are "unbalanced" in the sense that one side pays a higher price than the other.
- The efficient interchange fee for a particular card network is difficult to determine empirically.

---

[2] Economic efficiency occurs when the marginal social cost of the last card transaction is just equal to its marginal social benefit.

Whether the private marketplace and competition in that marketplace will yield interchange fees that are efficient is an open question. The conclusions of the theoretical literature vary substantially, depending on the assumptions underlying the models. Although no findings are completely robust, most models suggest that, when merchant prices do not vary by payment method,

- Profit maximization does not, in general, lead a network to set an interchange fee at the efficient level.
- In theory, privately-set interchange fees can be either too high or too low relative to the efficient interchange fee, depending on a number of factors, including the cost and demand considerations underlying the merchant decision to accept cards and the extent of competition among issuing and acquiring banks.
- In most markets, an increase in the level of competition among firms generates downward pressure on prices; however, this is not necessarily true for interchange fees. In general, competition among payment networks is unlikely to exert downward pressure on interchange fees because the networks tend to focus their competitive efforts on getting their card to be the favored card of a consumer. This objective is facilitated by having a higher interchange fee that can be used to fund more attractive terms (e.g., lower fees and higher rewards) for the consumer. In the special case where all consumers hold and use the cards of multiple networks and are indifferent with respect to which network's card they use for a given transaction, competition among card networks can lead to lower interchange fees.

Extensive legal and regulatory activity worldwide reflects the intensity of the debate over interchange fees and payment cards. In several countries, central banks or competition authorities have taken action aimed at reducing interchange fees. In the United States, a consolidated class action antitrust lawsuit pitting numerous merchants and merchant groups against Visa, MasterCard, and their member banks is pending in the U.S. District Court for the Eastern District of New York. In addition, Congress has held

hearings on interchange fees, and members of Congress have introduced several bills concerning them.

Motivated by this increased attention to pricing in payment card markets, this paper discusses and evaluates a number of potential policy interventions for the payment cards industry, which reflect both actions that authorities in other countries have taken and proposals that have been advanced by various parties in the United States. These interventions include

- Clarifying or eliminating restrictions on differential retail pricing across payment methods;
- Prohibiting network determination of interchange fees;
- Regulating the level of interchange fees;
- Relaxing card acceptance requirements;
- Mandating "multi-bugged" cards that can perform transactions on multiple networks with merchant control of network routing rules; and
- Doing nothing.

A number of common concerns arise in connection with all of these policy options. In particular, the effects of any intervention are uncertain and may involve unintended consequences. Among these possible consequences is the unintended redistribution of costs and benefits of card transactions across merchants, banks, card users, and non-card users. In addition, although much of the debate over payment cards has focused on interchange fees in credit card systems, similar issues arise for debit card systems as well as other card systems that do not have explicit interchange fees. A narrow intervention that targets interchange fees for credit cards could have effects on competition and pricing throughout the retail payments market.

In addition to these general concerns, each of these options has specific benefits and costs that should be carefully considered by policymakers before taking any legislative or regulatory action. The various options differ in their transparency and ease of implementation, as well as in the extent to which they may be able to redress any potential inefficiencies in the payment card market.

5

I. Introduction

In many countries around the world, electronic card-based payments have been replacing older types of payments at a rapid rate. In the United States, use of both debit cards and credit cards has been rising, while check volumes have been declining. In addition, the amount of cash in circulation has been growing more slowly in recent years.[3] This transition from paper-based payment methods to electronic payment methods has certainly modernized the payment system; however, it is not clear whether the incentives inherent in the current payment system infrastructure will lead participants to make socially optimal choices among alternative payment methods. In addition, increased use of electronic payment methods has generated a number of public policy debates.

One prominent policy debate concerns *interchange fees*.[4] These fees, which typically involve a payment from a merchant's bank to a card user's bank for each debit card or credit card transaction, are determined at the network level and are generally the same for all banks participating in a network.[5] Merchants' banks generally pass the costs associated with interchange fees through to merchants. In recent years, increases in interchange fee rates, together with growth in the volume of card transactions, have led to a dramatic rise in interchange fee payments, and consequently in merchants' cost of accepting payment cards.[6] As a result of these developments, merchants have increasingly expressed concern about their costs associated with card transactions.

---

[3] Recent financial and macroeconomic developments may have affected some of these trends. For example, distribution of cash by Federal Reserve Banks increased sharply in the latter part of 2008. In addition, recent evidence suggests that the growth of credit card transactions slowed in 2008, while the growth of debit card use may have accelerated (Sidel 2009).

[4] Appendix I provides definitions for a number of key terms and concepts used throughout this document, including interchange fees.

[5] As explained in section II, explicit interchange fees exist in payment card systems in which depository institutions serve as both card issuers and merchant acquirers (e.g., Visa and MasterCard). Payment card systems in which the network serves one or both of these roles (e.g., American Express and Discover) do not have explicit interchange fees, but they do charge fees to merchants for card transactions.

[6] Neither the total value of fees paid by merchants for card transactions, nor the total value of interchange fee payments, is publicly available. Calculations by the authors suggest that the total value of interchange fee payments on the Visa and MasterCard credit, signature debit, and PIN debit card systems was approximately $35 billion to $45 billion in 2007. These estimates suggest a substantial increase as compared with an analogous calculation of around $20 billion in 2002.

Merchant concerns focus on the level and sometimes the very existence of interchange fees. Merchant groups and their supporters contend that a common interchange fee serves as a means for otherwise competing banks to collude on their fees for card transactions and to avoid negotiation of fees with merchants. These parties argue further that the primary alternative available to merchants – rejection of a network's cards – is not a viable option in the case of the major card networks because customers have come to expect acceptance of those cards. As a result, merchants contend that the major card networks are able to set excessively high interchange fees, which lead to correspondingly high merchant fees, inflated incentives (i.e., rewards programs) for consumers to use the cards, and supranormal profits for the networks and their card-issuing banks.[7]

In contrast, the card networks and their supporters contend that payment cards provide substantial value to both consumers and merchants, and that interchange fees are essential for the proper operation of the card networks.[8] They note that a common interchange fee standardizes the terms of exchange between the merchant's bank and the card user's bank and plays an economically important role in influencing the incentives of merchants to accept cards and consumers to hold and use them. These groups argue that, when setting its fees, a card network must recognize the need to attract consumers and merchants, both of whom are necessary for the card network to exist. In their view, the resulting fees attempt to balance the two sides of the payment card market to maximize the value of the network, including the value of card services for both consumers and merchants.

Extensive legal and regulatory activity worldwide reflects the intensity of the debate over interchange fees and payment cards. In several countries, central banks or competition authorities have taken action aimed at reducing interchange fees. In the United States, a consolidated class action antitrust lawsuit pitting numerous merchants and merchant groups against Visa, MasterCard, and their member banks is pending in the

---

[7] Throughout this document, the term "consumers" should be interpreted to include all purchasers of goods and services, regardless of whether they are households, businesses, nonprofits, or government entities.

[8] Given that interchange fees are, by definition, payments between the merchant's bank and the card user's bank, this argument applies only to systems (like Visa and MasterCard) in which banks (rather than the network itself) are involved in issuing cards and acquiring transactions.

U.S. District Court for the Eastern District of New York. In addition, Congress has held hearings on interchange fees, and members of Congress have introduced several bills concerning them.

The core economic and policy issues concern the effects of interchange fees on the extent of card use and the welfare of different parties in the economy. In terms of efficiency, the key issue is whether current patterns of retail payments appropriately reflect the costs and benefits of the various payment methods. If the transaction fees faced by either merchants or card users are too high (or too low), then some payment methods will be overused and others will be underused relative to the socially optimal outcome. In particular, high fees for merchants and low or negative fees for card users may cause overuse of payment cards, or certain types of payment cards, relative to alternative payment methods. However, determining whether observed patterns of card fees (including interchange fees) and card usage are socially optimal is an extremely difficult task.

In terms of equity, interchange fees and associated transaction fees can generate transfers among non-card users, card users, merchants, and banks that increase the welfare of some parties while reducing the welfare of others. For example, high interchange fees could support low transaction fees (or high rewards) for cardholders, thereby potentially making card users better off. However, if the net effect of card acceptance is to increase merchant costs (that is, if card acceptance fees exceed any reduction in merchants' transaction costs due to card acceptance), and if merchants do not set prices that vary by payment method, then merchant acceptance of cards could lead to higher retail prices for all consumers, including those who pay with alternative methods and receive none of the direct benefits associated with card use.[9,10]

This paper is intended to provide background for understanding the current debate surrounding interchange fees. It presents a comprehensive, yet accessible, description of the economic theory underlying interchange fees, as well as a discussion of industry

---

[9] This result holds even if card acceptance leads to increased demand and increased profits for merchants.

[10] To the extent that card users, on average, have higher incomes than non-card users (as noted for credit cards in, e.g., Stavins 2001, Klee 2006, and Zinman 2008), such transfers would be regressive.

trends and the regulatory environment relevant to the debate. More specifically, section II describes the operation of a typical payment card system and the various parties and fees, including interchange fees, associated with each transaction. Section III then examines the economic theory that has been developed to analyze payment card systems. It examines the role that interchange fees play in this market, and in particular, the role that interchange fees play when externalities are present. The determination of the socially optimal level of interchange fees and the question of whether the private market will attain that level are also discussed.

Section IV turns from the theoretical to the institutional and empirical, describing the recent history of payment card use and pricing in the United States. It discusses the structure and nature of competition within the industry in order to highlight the various dynamics affecting interchange fees. The next two sections discuss recent developments affecting the payments industry: Section V describes recent innovations by participants in the payment card industry, and section VI focuses on regulatory and legal developments in the United States and abroad. Finally, section VII discusses and evaluates a series of options that are representative of the types of policy measures that have been or could be proposed to address concerns associated with interchange fees.

II. Basic Features of U.S. Payment Card Systems

There are three main types of payment card transactions: credit, signature debit, and PIN debit.[11] Two main organizational forms for general purpose payment card systems, often referred to as *four-party* and *three-party* systems, currently operate in the United States.[12] The so-called four-party system is the model used for most card transactions and is employed by Visa and MasterCard for their *credit card* and *signature*

---

[11] Charge cards, which require a cardholder to pay his or her balance in full at the end of a billing cycle, operate in a manner similar to credit cards. In recent years, prepaid cards have emerged as an additional important type of payment card. Although prepaid cards have raised various policy concerns, the debate over interchange fees has generally focused on the three more established types of payment cards mentioned in the text.

[12] Industry participants sometimes refer to four-party systems as "open loop" systems and three-party systems as "closed loop" systems.

debit card operations, as well as by all the *PIN debit card* networks.[13] The four parties in the system are the consumer, the depository institution that issued the payment card to the consumer (the *issuer*), the retail merchant, and the merchant's depository institution (the merchant *acquirer*). The network coordinates monetary transfers and the transmission of information between the issuing and acquiring sides of the market.[14] In a three-party system, the network itself acts as both issuer and acquirer. Thus, the three parties involved in a transaction are the consumer, the merchant, and the network. American Express and Discover have traditionally used this form, although both have expanded their card programs in recent years to include depository institution issuers.[15]

As illustrated in figure 1, a typical transaction over a four-party system proceeds as follows. The consumer initiates a purchase by presenting his or her card or card information to a merchant.[16] An electronic authorization request with a specific dollar amount and the cardholder's identity is sent from the merchant to the acquirer to the network, which forwards the request to the card-issuing institution.[17] The transaction is checked against a file of active card accounts that resides with either the card issuer or its processor. A message authorizing (or declining) the transaction is returned to the merchant via the reverse path.

Subsequently, the issuer posts a charge for the transaction to the cardholder's account, and the acquirer posts a credit for the transaction to the merchant's account. The timing of these charges and credits varies depending on the arrangements that cardholders and merchants have with their respective banks. For PIN debit transactions, the cardholder charges are posted immediately following the receipt of the authorization

---

[13] All general purpose credit card transactions are carried by one of the major credit card networks (Visa, MasterCard, American Express or Discover); signature debit card transactions are carried by MasterCard, Visa or, very recently, Discover; and PIN debit transactions are carried by either a regional EFT network (such as Star or Pulse) or by Visa's or MasterCard's PIN debit network (Interlink or Maestro, respectively).

[14] Note that the term "four-party system" is something of a misnomer because the network is, in fact, a fifth party involved in a transaction.

[15] This "hybrid" model still differs from the so-called four-party system in that the network itself acts as the sole acquirer. Even more recently, American Express and Discover have started using outside acquirers to provide card services to merchants (Costanzo 2008). This development has further reduced the differences between those networks and Visa and MasterCard.

[16] In the case of PIN debit, the consumer also enters a security code.

[17] Specialized payment processors may carry out some functions between the merchant and the network or between the network and the issuer.

message. For credit and signature debit transactions, the issuer usually posts the transaction to the consumer's account within one day of the transaction.[18] The acquirer usually credits the merchant's account within four days of the transaction (though the acquirer often credits the merchant's account much sooner, especially in the case of PIN debit).[19]

At the end of a business day, the merchant submits records of all card transactions for that day to its acquirer.[20] The acquirer reconciles these data against the earlier authorization information and transfers the reconciled data to the network. The network then clears the transactions; that is, it determines the net financial positions, including interchange fees and other network fees, as discussed below, of all issuers and acquirers. Usually within two days for credit and signature debit card transactions and one day for PIN debit card transactions, banks settle their accounts; that is, they receive and send payments based on their net financial positions through their accounts at settlement banks associated with the network.

Various fees are involved in every payment card transaction.[21] In a four-party system, an *interchange fee* is paid by the merchant acquirer to the card issuer.[22] The

---

[18] In signature debit transactions, the issuer generally places a hold on funds in a cardholder's account after authorization. Once a debit card transaction has been posted to a cardholder's account, the funds are withdrawn from the available account balance. For credit cards, the posting of a transaction decreases the cardholder's available credit line by the purchase amount. At the end of a billing cycle, if the cardholder had no prior balance, he or she can pay the entire balance, thereby receiving an interest-free loan on transactions performed during the billing cycle. Alternatively, the cardholder can pay at least the minimum required payment and will incur finance charges on the unpaid balance.

[19] As long as the merchant has followed the system's rules for submitting and receiving authorization for a transaction, it is guaranteed payment regardless of subsequent cardholder behavior.

[20] The merchant can also submit transaction records at multiple pre-determined times during a business day. In PIN debit networks, the initial authorization message also serves as a full transaction record for the acquirer, so the merchant does not send a second message of transaction records.

[21] Users of payment cards may face various other fees and charges, such as finance charges on revolving credit card balances, late payment fees for credit cards, and overdraft fees for debit cards. In addition, because debit cards are typically one component of a broader deposit account relationship, debit card transaction fees may be linked to other account features (e.g., maintenance fees or interest rates). While these considerations are worth noting, the discussion in this paper and in much of the literature regarding interchange fees focuses on those fees that are directly related to the transaction function of payment cards.

network collects a switch fee (or assessment) from the acquirer and the issuer. The acquirer charges the merchant a *merchant discount*, which is the difference between the face value of the transaction and the amount the acquirer transfers to the merchant. In some cases, consumers pay fees to or receive rewards from their card issuers for each card transaction.

An interchange fee typically comprises a large fraction of the merchant discount for a particular card transaction.[23,24] An interchange fee may take the form of a flat fee per transaction, a percentage of the purchase price, or a combination of the two. As discussed in more detail in section IV, interchange fees for credit and signature debit transactions generally exceed those for PIN debit transactions. In the United States, interchange fees for PIN debit typically average $0.35 to $0.50 per transaction; interchange fees for a typical signature debit transaction are about 1.2 percent of the transaction value; and interchange fees for a typical credit card transaction for Visa and MasterCard are in the range of 1.5 to 2 percent of the transaction value. Interchange fees also vary by merchant type (e.g., grocery store, department store, fast food restaurant), merchant sales volume, and credit card program within a network (e.g., "premium" cards such as gold or platinum cards may carry higher interchange fees than "basic" cards).

A noteworthy feature of an interchange fee is that it is set by the network that carries the transaction, and not by the individual card issuers that receive the fee from

---

[22] Theoretically, interchange fees can flow in either direction, but in most cases they flow from the merchant's bank to the card user's bank. The Australian debit card system (EFTPOS) is an example of a rare case in which the interchange fee is paid by the card user's bank to the merchant's bank. Automated teller machine (ATM) networks are another example where interchange fees flow in the opposite direction (i.e., card issuing banks pay interchange fees to ATM operators for each transaction). Historically, this arrangement arose because ATM networks prohibited ATM operators from charging transaction fees (i.e., surcharges) to ATM users whose cards were issued by other banks. Under this restriction, interchange fees encouraged deployment of network-connected ATMs by providing a means for ATM operators to recover their costs. Since restrictions on ATM surcharges were lifted in the 1990s, ATM operators can recover their costs through direct transaction fees to ATM users, rather than through interchange fees (McAndrews 2003).

[23] The merchant discount is generally equal to the sum of the interchange fee, the acquirer switch fee, other acquirer costs, and an acquirer markup.

[24] In a three-party system, because the network serves as both the issuer and the acquirer, there is no explicit interchange fee. Instead, the merchant directly pays a merchant discount or service charge to the network. In the United States, the merchant service charge set by American Express typically exceeds 2 percent of the transaction value (Quittner 2006); Discover's merchant fees are not publicly disclosed.

acquirers. Until recently, both the Visa and MasterCard networks were organized as joint ventures of their member banks. Some observers have argued that, under this ownership structure, the setting of a common interchange fee by the network for all member banks could be interpreted as collective price determination by the member banks. This approach towards setting interchange fees, and the possibility that it could constitute illegal collusion under the antitrust laws, is one of the points of controversy surrounding payment card networks. As discussed in section V, however, both Visa and MasterCard have recently converted to publicly traded companies. The implications of these reorganizations for the disposition of antitrust complaints regarding collective price determination are unclear.

In addition to setting the structure and level of interchange fees, each card network specifies operating rules that govern how network participants interact. Although contracts are written only between the network and its issuers and acquirers, merchants and processors must also comply with the network rules or risk losing access to that network.[25] Network operating rules cover a broad range of activities, including merchant card acceptance practices, technological specifications for cards and terminals, risk management, and determination of transaction routing when multiple networks are available for a given transaction.

The operating rules include some key provisions that have important ramifications for the operation of the networks. First, each network requires a merchant that agrees to accept a given type of card to accept that type of card regardless of the identity of the card issuer or specific card program. For example, a merchant that accepts Visa credit cards must accept all Visa credit cards and may not reject a particular bank's Visa credit card or a Visa credit card associated with a particular rewards program. This type of universal acceptance requirement is commonly referred to as an *honor-all-cards rule*.[26]

---

[25] Merchant groups have frequently expressed concern that merchants themselves were not permitted to view the network operating rules by which they were bound because they were not network members (though their acquiring banks were members). In response to these concerns, Visa and MasterCard have recently begun to make their rules available to merchants. As of February 2009, online versions of the rules for the two networks were available at http://usa.visa.com/merchants/operations/op_regulations.html for Visa and http://www.mastercard.com/us/merchant/support/rules.html for MasterCard.

[26] Prior to 2003, both Visa and MasterCard imposed broader honor-all-cards rules that required that any merchant that accepted Visa (MasterCard) credit cards must also accept Visa

Such rules are significant because a network's interchange fees may vary across card programs. Specifically, premium cards, typically associated with generous rewards programs for card users, tend to carry higher interchange fees than basic cards.

Another important operating rule is the *no-surcharge rule*, which prohibits a merchant from adding a surcharge to a customer's bill if the customer pays with one of the network's payment cards.[27,28] (Exceptions to this rule include certain payments to government entities and tuition payments to educational institutions.) As required by the Cash Discount Act, the networks generally do allow merchants to offer discounts for cash payments.[29] Merchant groups allege that the card networks place restrictions on the form that such discounts may take, but other industry observers deny that such restrictions exist.[30] In addition, although discounts for cash and surcharges for cards may appear at first to be equivalent, allowing only a cash discount effectively prohibits differential pricing across types of cards.

III. The Economics of Payment Cards

Payment card markets are often described by economists as being *two-sided*. A two-sided market is a market for the provision of a product whose value is realized only

---

(MasterCard) signature debit cards. The scope of these rules was narrowed to allow separate merchant acceptance decisions for credit and signature debit cards as part of the settlement agreement in a lawsuit discussed in section VI.

[27] From 1976 to 1984, federal law prohibited surcharges for credit card payments. Currently, a small number of states have laws that prohibit surcharges for various types of card payments.

[28] Merchants have also raised concerns about restrictions that they claim the networks impose on merchants' ability to use non-price incentives, commonly known as "steering," to influence the payment choices of their customers.

[29] The relevant part of the U.S. Code (15 USC 1666f(a)) says "the card issuer may not. . . prohibit any. . . seller from offering a discount to a cardholder to induce the cardholder to pay by cash, check or similar means rather than use a credit card." Note that this provision refers to card issuers, not networks. Also, it is unclear what constitutes a payment method that is "similar" to cash or check. Finally, although the publicly available versions of the networks' rules do not explicitly mention discounts for check payments, the networks presumably allow such discounts as required by the Cash Discount Act.

[30] Mallory Duncan (2007), testimony delivered at the Hearing on Credit Card Interchange Fees before the Judiciary Committee Antitrust Task Force, U.S. House, 110 Cong., July 19. (Transcript available at http://judiciary.house.gov/hearings/July2007/hear_071907.html, p. 63.) Timothy Muris (2007), testimony delivered at the Hearing on Credit Card Interchange Fees before the Judiciary Committee Antitrust Task Force, U.S. House, 110 Cong., July 19. (Transcript available at http://judiciary.house.gov/hearings/July2007/hear_071907.html, p. 60.)

if a member of each of two distinct and complementary sets of users simultaneously agrees to its use (Rochet and Tirole 2006a). A payment card has value only if a merchant and the merchant's customer agree on its use to carry out a transaction. The two-sided nature of demand for payment cards has important implications for pricing that are absent in standard markets. In particular, the prices faced by the two sides of the market must be set at levels that "balance demand," because a payment card that appeals to one side of the market will not be used if it does not also attract the other side of the market.

In a four-party card system, a payment card transaction can be characterized as a product jointly supplied by the card issuer and the merchant acquirer, and jointly demanded by the consumer and the merchant. Each of these parties experiences benefits and costs from engaging in a payment card transaction; the difference between a party's benefits and costs constitutes that party's economic *surplus* from a card transaction. A card transaction is economically efficient if the sum of the surplus accruing to all of the parties involved is non-negative – that is, if the sum of the benefits that all of the parties derive from using a card is greater than or equal to the sum of the costs incurred by all of the parties. However, the parties will agree to conduct a card transaction only if the surplus realized by each of the individual parties is non-negative. The fact that efficiency depends on aggregate costs and benefits while actual use of a card depends on the separate decisions of several distinct parties implies that, in the absence of transfer payments, some efficient card transactions may not in fact take place (Baxter 1983).

The example in table 1 illustrates how transfer payments can facilitate efficient card transactions that would not otherwise occur. Suppose that the issuer and acquirer each incur a production cost of $.50 for a card transaction. Assuming that they do not derive any direct benefit from the transaction, the issuer and acquirer each obtain a surplus of -$.50. In addition, suppose the merchant derives benefits of $2.25 while incurring costs of $.25, yielding a surplus of $2 for the merchant. Similarly, suppose the consumer experiences a benefit of $.50 and a cost of $.75, yielding a surplus of -$.25.[31]

---

[31] The costs and benefits in this example are intended to represent purely hypothetical costs and benefits associated with use of a card for a particular transaction, instead of an alternative, such as cash. Examples of production costs for issuers and acquirers are processing, billing, and fraud mitigation costs. For merchants, benefits could include faster checkout times and decreased cash handling. For consumers, benefits could include security, convenience, and ease of budgeting.

As shown in the first column of table 1, use of the card to perform the transaction between the merchant and consumer is efficient, producing an aggregate surplus of $.75. However, in the absence of a transfer payment, none of the parties, except for the merchant, would be willing to participate in a card transaction.

Now, assume that the merchant is required to pay a transfer equal to $1.50, divided equally among the other three parties, as illustrated in the second column of table 1. With this transfer, each party would derive a non-negative surplus from the card transaction and therefore all parties would be willing to participate. Thus, transfer payments can provide a mechanism for improving efficiency in a two-sided market by redistributing private benefits and costs in a way that leads private decision makers to choose a more socially efficient outcome.

More formally, the two-sided nature of the market for payment cards introduces the possibility of *externalities*. An externality arises when one agent's action affects the welfare of another agent, without any compensation for the effect. Because an agent's private incentives in the presence of an externality do not reflect the true social cost or benefit of his or her actions, socially inefficient outcomes can result.

The economics literature has emphasized the importance of two potential externalities in the context of payment cards (Rochet 2003). The first, described in the numerical example above, has been termed the usage externality and arises because each party in a given transaction evaluates his or her own costs and benefits associated with a particular payment method, but does not consider the costs and benefits of the other party. A second type of externality, referred to as a network (or adoption) externality, reflects the fact that the value of a payment card network increases for both merchants and consumers as the card becomes more ubiquitous. In other words, a larger cardholder base makes card acceptance more valuable for merchants, while broader merchant acceptance makes cardholding more attractive for consumers. When deciding whether to accept cards, the merchant may not account for the increased network value associated with his or her acceptance. Similarly, a consumer's decision about whether or not to hold

---

As in the example, consumers (or merchants) may dislike certain features of a payment card (e.g., the lack of anonymity for card users) and, as a result, derive negative surplus from a card transaction.

a card may not account for the effect of his or her decision on the value of the overall network.[32]

A standard economic prescription to deal with externalities involves adjusting prices (through taxes, subsidies, or transfers) so that agents' private incentives reflect the true social costs and benefits of their decisions. In the context of payment cards, this involves appropriate pricing of card services for all of the parties involved in a card transaction. How this could be accomplished depends on the organizational form of the card system. In the three-party card system (in which the network itself is both issuer and acquirer), the network could, theoretically, set transaction prices for consumers and merchants that would align private and social costs and benefits. In a four-party card system, the network cannot directly set transaction fees for merchants and consumers. It can, however, establish an interchange fee – a transfer payment between the merchant's bank (acquirer) and the consumer's bank (issuer) – that influences the costs of card transactions for those banks and the transaction fees that they subsequently charge merchants and consumers (i.e., merchant discounts and card use fees or rewards). As a result, an interchange fee can, at least theoretically, improve efficiency by internalizing externalities.

It is important to recognize that an interchange fee is not the only tool that can serve this purpose. The four-party systems also charge switch fees (payments from acquirers and issuers directly to the network, sometimes called assessments), which could be used to influence prices of card transactions for the two sides of the market, even if the interchange fee were zero (Katz 2005). Particularly in light of the recent corporate restructurings of Visa and MasterCard, the four-party card systems could substitute these types of direct fees (and possibly rebates) to acquirers and issuers for explicit interchange fees. However, the economic considerations underlying the pricing of payment card

---

[32] Some have argued that as a network matures and the "chicken and egg" problem that faces a nascent network is overcome, the marginal network externality (the marginal benefit from expanding the network) declines (Katz 2001, Wang 2008). Even for a mature network, however, the externality associated with the use decision (the misalignment between private and social costs and benefits) still remains (Rochet 2003).

17

services, from both a social and a private perspective, do not depend on whether a card system uses an interchange fee between banks or direct fees to participating banks.[33]

A. EFFICIENT INTERCHANGE FEES

An efficient interchange fee would lead to a socially optimal number of payment card transactions. That is, with an efficient interchange fee, a payment card would be used in a transaction if and only if using the card would result in a non-negative change in *total surplus*. Technically, the number of card transactions would be such that the marginal social benefit of the last card transaction would be equal to its marginal social cost. A few characteristics of an efficient interchange fee are worth noting:[34]

- An efficient interchange fee transfers surplus (i.e., costs and benefits) from one side of the market to the other in order to internalize the external effect that one party has on the other. In general, an efficient interchange fee is not solely related to the cost of producing a card-based transaction nor is it equal to zero.

- In order to provide appropriate incentives to parties in a transaction, an efficient interchange fee may yield prices for card services to each side of the market that are "unbalanced" in the sense that one side pays a higher price than the other. Indeed, efficiency may, in some cases, require a negative price for one side of the market.

- Even in the simplest case, the efficient interchange fee can be difficult to determine. At a minimum, calculation of the efficient interchange fee requires estimation of the demand curves for card services for heterogeneous consumers and merchants, in addition to precise cost data for acquirers, issuers, merchants, and consumers.

In summary, the basic economic role of an interchange fee is to affect the prices of card services for the merchant and the consumer in a transaction. Both prices are

---

[33] In fact, the same economic considerations are relevant for a three-party card system that directly sets fees to merchants and cardholders.
[34] Appendix II explains some of these results in more detail.

important because, unlike in a standard market where a single agent makes an independent purchase decision, two parties must jointly agree to use a particular payment service. The card transaction will not take place unless both parties agree to participate given the prices they face (Schmalensee 2002, Rochet and Tirole 2002). Efficient prices induce each party in a transaction to account for the external costs and benefits resulting from his or her behavior, factors that a self-interested agent might otherwise not consider. An appropriate interchange fee is one way to achieve such prices for card services.

B. A COMMON INTERCHANGE FEE IN FOUR-PARTY CARD SYSTEMS

As noted in section II, the fact that interchange fees (for a given combination of transaction type, merchant category, and card program) are the same for all issuing and acquiring banks is a highly controversial aspect of the four-party card networks. Because the previous discussion about efficient interchange fees does not rely on a common interchange fee, it is natural to ask whether a common, pre-determined interchange fee between banks in a four-party card network is necessary at all.

In the absence of a common interchange fee, banks would need to negotiate bilateral fee agreements with each other. In a system with thousands of banks, the number of such bilateral agreements would be quite large. Even if bilateral negotiation of fees between banks were feasible, the existence of a common interchange fee plays a key role in establishing the value of a four-party card network. In particular, a common interchange fee facilitates the maintenance of an honor-all-cards rule, and some sort of honor-all-cards rule is a core feature of a four-party card network. The logic is as follows: In the absence of common terms of exchange between banks, each issuer could set its own interchange fee. However, an honor-all-cards rule would require that every merchant that accepts a particular network brand and type of card (e.g., Visa credit card) accept every card with that brand and type, regardless of the level of the interchange fee set by the issuer. As a result, an honor-all-cards rule with bilaterally negotiated interchange fees would introduce the possibility of a "holdup problem" (Small and Wright 2002, Klein et al. 2006). That is, an individual issuing bank could demand very high interchange fees from acquiring banks. The acquirers would then factor the high

fees charged by this issuer into the merchant discounts that they charge their merchants.[35] A merchant, in turn, could avoid those fees only by rejecting all of a network's cards of that type, in which case card acceptance could be inefficiently low. A common interchange fee avoids this holdup problem.

As an alternative solution to this holdup problem, a card network could drop its honor-all-cards rule, allowing merchants to reject an individual issuer's cards if they carried inordinately high fees. However, without the acceptance guarantee provided by some sort of honor-all-cards rule, a consumer would need to ascertain whether each individual merchant's acceptance policy applied to the consumer's specific card. The complexity involved in obtaining or providing this information would substantially diminish the value of the network. Consequently, some observers have argued that a common interchange fee does not serve as a purely collusive mechanism for banks to set excessive and anticompetitive fees; rather, it allows a four-party card network to maintain an honor-all-cards rule, some form of which is an important factor underlying the value of the network (Klein et al. 2006).[36]

C. PRIVATE DETERMINATION OF INTERCHANGE FEES

In light of these potentially important roles of a common interchange fee, a crucial question remains: Will the private marketplace, and competition in that marketplace, yield interchange fees that are efficient? Competition is generally viewed as an effective mechanism for providing incentives that lead private agents to make choices that yield efficient outcomes. Hence, the central question is whether a card network, in competition with other card networks and payment methods, has appropriate incentives to choose an interchange fee that yields efficient prices to both sides of the market.

---

[35] Depending on the contract between the acquirer and the merchant, the acquirer could either set a different merchant discount for each issuer's cards or could adjust a uniform merchant discount to reflect the possibility of high interchange fees for some transactions.

[36] This discussion does not address the appropriate breadth of an honor-all-cards rule across different types (e.g., debit or credit) or classes (e.g., business or personal, premium or basic) of cards within a network.

The conclusions of the theoretical literature vary substantially depending on the assumptions underlying the models.[37] Assumptions about the degree of market power for acquiring banks, issuing banks, merchants, or networks, and the elasticities of demand for card services and final goods all influence the results. While no findings are completely robust, a number of common themes do emerge regarding privately-set interchange fees, when merchant prices do not vary by payment method.

- Although a card network's objective to maximize profits (for itself or its member banks) requires the network to recognize the two-sided nature of demand, profit maximization does not, in general, lead the network to set the interchange fee at the level that maximizes social welfare.

- Privately optimal interchange fees will typically yield an unbalanced price structure (i.e., a price structure in which the two sides of the market pay different prices), a common feature of two-sided markets, both in theory and in practice. Many models of payment card systems find that unbalanced prices are necessary to achieve efficiency; however, it does not follow that *any* privately determined pattern of unbalanced prices is efficient.

- In theory, privately-set interchange fees can be too high or too low relative to the efficient interchange fee, depending on a number of factors.[38] However, the incentives underlying merchants' card acceptance decisions in the theoretical models tend, all else equal, to support interchange fees that are higher than the

---

[37] See, for example, Rochet and Tirole (2002), Wright (2004), Rochet and Tirole (2006b), Guthrie and Wright (2007). Hayashi and Weiner (2006) survey many of the models and highlight differences in their assumptions and conclusions.

[38] See Rochet and Tirole (2006b). The presence of market power among issuing or acquiring banks is one factor that can influence the relationship between privately and socially optimal interchange fees. For example, when issuing banks have market power, a higher interchange fee can lower the costs of those banks, thereby offsetting issuer mark-ups that could otherwise inefficiently curtail card use. However, this role of an interchange fee is not related to the fundamental role of internalizing externalities and has been criticized as a means to subsidize firms (i.e., banks) with market power to induce lower prices (Farrell 2006). Heterogeneity across merchants in the transactional benefits from cards can also affect the relationship between privately and socially optimal interchange fees.

social optimum. In such a situation, merchant fees will be inefficiently high and card use fees will be inefficiently low (or card rewards will be inefficiently high), leading to excessive card use.[39]

The economic theory literature has emphasized two closely related reasons why merchants may be willing to accept cards with inefficiently high merchant fees. First, by increasing a merchant's "quality of service," card acceptance makes a merchant more attractive to consumers, leading to an increase in sales volume.[40] The merchant will take into account this private benefit when he or she evaluates the costs and benefits of card acceptance. To the extent that this increase in sales represents a diversion of transactions away from other merchants that do not accept cards, without any increase in aggregate sales, the private benefit to the merchant from its decision to accept cards will exceed the social benefit. As a result, the merchant will be willing to pay an inefficiently high merchant discount.[41]

Second, as long as some merchants are willing to accept cards despite an inefficiently high merchant discount, others will feel compelled to do so in order to avoid losing business. Thus, even if merchant discounts are high enough that merchants as a whole would be better off rejecting cards, they may nonetheless all choose to accept cards because no single merchant would find it profitable to unilaterally reject them.[42]

Although most of the theoretical literature examines the price-setting behavior of a monopoly payment card network, some recent models consider the effects of competition among networks for issuers and their associated cardholders (Guthrie and Wright 2007, Rochet and Tirole 2002, 2006b, Chakravorti and Roson 2006). In these models, network competition has varying effects on interchange fees, depending crucially on the cardholding behavior of consumers. Specifically, if consumers tend to hold and use the cards of multiple networks and are indifferent with respect to which network's

---

[39] Similar distortions can also arise with regard to the transaction fees established by a profit-maximizing three-party card system.

[40] The "quality of service" enhancement associated with merchant card acceptance encompasses a number of features associated with card payments, including convenience and, in the case of credit cards, the extension of credit itself.

[41] See Farrell (2006), Hayashi (2006), and Rochet and Tirole (2006b, 2006c).

[42] See Katz (2001) and Hayashi (2006).

card they use for a given transaction, then competition between card networks should tend to lower interchange fees because merchants can choose to accept only the cards that provide the most favorable terms for them, without losing customers. The empirical evidence, however, suggests that even though most consumers do hold multiple cards, they tend to favor a single card when making purchases (e.g., the card that provides the greatest rewards).[43] In this case, competition is unlikely to exert downward pressure on interchange fees because competition between networks is focused on becoming the favored card of a consumer, an objective that is facilitated by a higher interchange fee, which can be used to fund more attractive terms for the consumer. Moreover, heterogeneity among consumers with respect to their preferred card makes it difficult for merchants to reject cards from networks that have high interchange fees. Regardless of the direction in which it pushes interchange fees, competition between card systems is not a sufficient condition to yield an efficient interchange fee.[44]

Finally, the actual distribution of total surplus associated with card transactions (i.e., the difference between benefits and costs for each of the parties involved) under a privately determined interchange fee depends on how merchant fees affect the end prices of goods and services. Merchants that accept cards may adjust their product prices to reflect the effect of card transactions on their costs. Thus, if the net effect of card acceptance is to increase merchant costs, merchants may increase their product prices, leading to a possible decline in both sales volume and merchant surplus. At the same time, if all consumers pay the same price regardless of the payment method used, consumers who use other payment methods (e.g., cash) would bear some of the costs associated with card transactions without receiving the benefits that accrue to card users.[45] The possibility of a transfer of surplus to card users from users of other payment

---

[43] See Rysman (2007).
[44] In theory, an important exception to this result can arise when all cardholders have cards for multiple networks, sometimes referred to as "multi-homing" (Guthrie and Wright 2007, Rochet and Tirole 2006b). This scenario is discussed in section VII where the implications of "multi-bugged" cards, one mechanism for inducing multi-homing by cardholders, are considered.
[45] See Rochet and Tirole (2002, 2006c). In contrast, McAndrews and Wang (2008) find that when merchants operate in contestable markets and price at average cost, merchants accept cards only if doing so lowers their costs. In this model, cash users benefit from lower prices than would prevail without card acceptance.

methods is often viewed as a regressive redistribution due to the fact that card users tend to have higher incomes than non-card users.

IV.   Payment card use and pricing in the United States

Two major trends underlie the recent increase in attention to interchange fees among both merchants and policymakers: the shift away from checks and cash toward card payments, and changes in the level and structure of interchange fees. As shown in figure 2 and table 2, both the number and the total value of debit card payments grew dramatically between 1990 and 2006.[46] The Federal Reserve estimates that, by either measure, debit card payments approximately tripled between 2000 and 2006, with average annual growth rates close to 20 percent. At the same time, the number and value of credit card payments increased at annual rates of roughly six percent and nine percent, respectively. Given the relatively rapid growth in debit card use as compared with credit card use, the Federal Reserve estimates that the number of debit card transactions exceeded the number of credit card transactions by 2006.[47] However, because the average value of a credit card transaction substantially exceeds the average value of a debit card transaction, the total value of credit card transactions is still significantly larger than that of debit card transactions.[48]

In contrast, as seen in table 2, the Federal Reserve estimates that the number of checks paid by depository institutions declined at a rate of approximately 5 percent per year between 2000 and 2006. Although direct evidence regarding the number of cash

---

[46] The data in table 2 are based on a triennial survey of depository institutions, electronic payment networks, card issuers, and card processors conducted by the Federal Reserve System. The methodology and findings of these surveys are described in Gerdes et al. (2005) and Gerdes (2008). Table 2 also contains values and volumes for some other noncash payment methods to allow comparison between card systems and those systems. Notably, table 2 does not include large-value transfer systems, such as the Fedwire Funds Service or CHIPS, which would represent a small share of the number of transactions in the payment system, but a very high share of the transaction value.

[47] Debit's share of total card transactions rose from around 35 percent in 2000 to around 54 percent in 2006.

[48] Nonetheless, debit's share of the value of total card transactions grew from around 21 percent in 2000 to 32 percent in 2006. Some recent evidence suggests that debit card dollar volumes may have begun to exceed credit card dollar volumes for certain card networks, due to slower growth (and perhaps declines) in credit card volumes along with continued (and perhaps accelerating) growth in debit card volumes (Sidel 2009).

payments is not available, indirect evidence suggests that cash use may be declining (or at least growing more slowly) as well.[49] As of 2005, card transactions are estimated to have comprised more than half of total retail sector transactions, and their share has continued to rise.[50]

A. INTERCHANGE FEES

Contemporaneous with the increase in the share of purchases conducted with cards, network-determined fees for card payments have risen, and the pricing structures have become more complex. For many years, the interchange fee for an individual card transaction at a given merchant has varied with transaction type (credit, signature debit, or PIN debit) and purchase value. Interchange fees also vary across merchants according to merchant type (e.g., supermarket or gas station) and merchant sales volume, and the number of distinct merchant types listed in interchange fee schedules continues to proliferate. Recently, Visa and MasterCard have introduced interchange fees that vary across card programs, even for a given merchant type and sales volume.

Figure 3 shows the path of interchange rates over time (monthly), by network and transaction type.[51] Visa and MasterCard carry all three types of transactions, with separate interchange schedules for each. PIN debit interchange is shown for Visa's Interlink network, MasterCard's Maestro network, and the four leading regional PIN debit networks (Star, NYCE, Pulse, and Accel/Exchange). The three-party credit card systems (American Express and Discover) are excluded from this diagram because they

---

[49] For example, the amount of cash paid into circulation by Federal Reserve Banks has been growing more slowly in recent years (although that growth rate increased sharply in the latter part of 2008, likely in response to financial and macroeconomic developments). In addition, the number of small-value payments made with payment cards has increased, suggesting that cards may have replaced cash for such transactions. Similarly, the number of traditionally "cash-only" businesses (such as fast food restaurants) that now accept cards has increased. Finally, the number of ATM withdrawals has declined (although the number of debit transactions involving cash back has increased). See Gerdes (2008) and Bauer and Littman (2007).

[50] The share of card transactions was estimated at 56 percent in 2005, up from 43 percent in 1999 (American Bankers Association/Dove Consulting 2006).

[51] Data for the figures that report interchange fees are from the EFT Data Book (published by ATM & Debit News) and individual networks' published fee schedules. MasterCard and Visa have recently begun to make their current interchange fee schedules available online. As of February 2009, the networks' interchange fee schedules were available at http://usa.visa.com/merchants/operations/interchange_rates.html for Visa and http://www.mastercard.com/us/merchant/support/interchange_rates.html for MasterCard.

do not have explicit interchange fees.[52] Instead, they negotiate merchant service charges directly with merchants and do not report the rates publicly.

As figure 3 shows, interchange rates for basic credit cards rose during the early 2000s, a trend that began in the 1990s. These rates have plateaued over the past 3 years, possibly due, at least in part, to increased regulatory scrutiny abroad and civil litigation in the United States. It is worth noting, however, that the interchange fees associated with Visa's and MasterCard's premium rewards cards are markedly higher than the interchange fees associated with their basic credit cards.

Credit card interchange fees have been above signature debit card interchange fees since 2003, and both credit and signature debit interchange fees have consistently exceeded PIN debit interchange fees. Several possible explanations account for these differences. A possible demand- or market-power-based reason for higher credit card interchange fees is the broad penetration of Visa and MasterCard and the associated consumer expectation that merchants accept these credit cards (that is, merchants agree to accept cards with high interchange fees to avoid losing business), as described in section III.C. Another reason why credit interchange fees may be higher than those of debit is to compensate for the credit risk taken on by the issuing bank in extending credit (over the billing cycle or longer) for the full amount of the transaction.

The difference between signature and PIN debit interchange fees, which has diminished considerably in recent years, may be due to several factors. First, prior to 2003, Visa's (MasterCard's) honor-all-cards rule required merchants who accepted Visa's (MasterCard's) credit cards to also accept its signature debit cards. This rule placed pressure on merchants to accept signature debit. During that period, Visa and MasterCard maintained signature debit interchange rates at essentially the same levels as their credit card interchange rates. In 2003, a settlement agreement reached in a lawsuit that had been filed against Visa and MasterCard eliminated the tying of signature debit

---

[52] For cards issued through a bank, American Express and Discover likely have some exchange fees (or rebates) that they charge (or pay) to their issuers. Similarly, those networks charge wholesale service fees to acquirers associated with their recent moves to use outside acquirers (Costanzo 2008). However, neither issuer fees (or rebates) nor acquirer fees are publicly available for those networks. When a transaction involves both an outside issuer and an outside acquirer, the fees charged by these networks to issuers and acquirers serve a similar economic role as an interchange fee, even if any direct payments between acquirers and issuers (i.e., interchange fees) are set to zero.

card acceptance to credit card acceptance and required a reduction in signature debit interchange fees.[53] As shown in figure 3, implementation of this settlement agreement led to a decline in signature debit interchange fees, but did not completely eliminate the differential between signature debit and PIN debit fees.

Second, the differential between PIN and signature debit interchange fees may reflect more intense competition among PIN debit networks due to the presence of multiple brands on a given card. In the early days of PIN debit, each network covered a fairly small geographic area, and many banks sought to offer their customers wider merchant acceptance of their PIN debit cards by "multi-bugging" the cards (i.e., issuing cards that bore the logos of and could be used on multiple PIN networks).[54] Over time, individual networks expanded their geographic coverage through a combination of mergers and organic growth, and in some cases, networks' geographic regions began to overlap one another. In this environment, merchant acquirers or their processors could often choose which one of the networks whose brands appeared on a card would carry the transaction. Merchants generally prefer that their acquirers route PIN debit transactions over the network with the lowest interchange fee, resulting in direct price competition among PIN debit networks. More recently, this price competition appears to have diminished (and, as shown in figure 3, PIN debit interchange fees have risen) as the largest national PIN debit networks have increasingly required issuers to sign exclusive agreements under which they become the sole PIN network whose logo appears on an issuer's cards (Breitkopf 2007).

The need to encourage merchant investment in PIN terminals may also have contributed to relatively low PIN debit interchange fees, particularly in the early years of PIN debit. In contrast, signature debit, which utilizes the same infrastructure as MasterCard and Visa credit cards, does not require merchant investment in new equipment. In addition, higher fraud costs (due to the lack of a PIN requirement) for

---

[53] This lawsuit is discussed further in Section VI.
[54] PIN debit networks were originally quite limited in their geographic scope, resulting in a large number of small, regional networks. At the peak (in the late 1980s), over 150 such networks operated in the United States. Rapid consolidation reduced the number of PIN debit networks to fewer than 20 as of 2008.

27

signature debit may have led to higher signature debit interchange fees compared to PIN debit.

Turning to variation in the structure of interchange fees at a given point in time, figure 4 illustrates interchange fee schedules over a range of possible purchase values at a moderate-sized general-purpose retailer, by transaction type and network, at year-end 2008. Each interchange fee has a minimum value and then rises with the purchase amount.[55] Until very recently, all of the PIN debit networks capped their interchange fees; however, both Interlink and Pulse removed their caps in 2008 (Musante 2008), while NYCE removed its cap in early 2009.[56] Fees for credit and signature debit are not capped.[57]

Although interchange fees have historically varied by merchant type, the number of merchant types has recently proliferated. In addition, over the past three to five years, Visa and MasterCard have established rate schedules that vary across credit card programs. Figure 5 shows a few examples of credit card interchange fees for different types of merchants and card programs.[58] These schedules show wide variation – the basic card used at a large-volume supermarket generates a fee just over 1 percent of the purchase value, while the fee for the highest-reward card used at any merchant exceeds 2 percent. Merchants frequently express concern about the variation in fee schedules across card programs for two reasons. First, at the time of a transaction, they cannot determine the interchange fee associated with the particular card used by the customer.[59]

---

[55] Several of the PIN networks have flat fees per transaction for particular merchants (e.g., a flat rate of 22 cents for all supermarket transactions). However, such a fee structure, while once the norm for PIN debit transactions, is becoming less common.

[56] Star and a number of other PIN debit networks still had caps as of January 2009, with the highest possible PIN debit interchange fee being 66 cents per transaction for Star.

[57] The fact that interchange fees increase with the purchase value is one of the sources of concern for merchants, which argue that, with the exception of credit or default risk, the cost of processing a transaction does not rise proportionally with the purchase value. Gas station owners were particularly vocal about this point during the recent period of rising gasoline prices. In 2007, MasterCard capped the interchange fees charged for gasoline purchases of more than $50. In June 2008, Visa reduced the interchange rate applied to gasoline credit card purchases and capped the interchange fee for gasoline debit card purchases (Aspan 2008).

[58] The full schedule of fees is far too complex to reproduce here.

[59] Merchants have noted that their inability to distinguish between premium and basic cards prevents them from verifying the accuracy of their merchant fees. Although merchant acquirers and processors generally have access to this information, they do not necessarily share the

Second, and perhaps more important, each network's honor-all-cards rule requires merchants to accept all cards of a given type (i.e., credit or signature debit) within that network and not to exclude cards associated with specific rewards programs.

B. INDUSTRY STRUCTURE

The pricing developments described above have occurred during a period of consolidation and shifting market shares among payment networks, card issuers, and merchant acquirers. The underlying causes of consolidation vary across segments of the payment card industry; nonetheless, the resulting market structure places a large share of consumer payment system operations in the hands of a relatively small number of service providers. The discussion below highlights some of the trends in concentration and market shares for each set of providers and explores some competitive implications of these trends.

1. Network Market Structure

Concentration among card networks is important from a policy perspective; recent litigation and calls for regulatory intervention focus on the market power of the largest networks. Specifically, the largest networks are setting the terms of trade among merchants, cardholders, acquirers, and issuers for a large share of all purchase transactions.

Figure 6 shows network shares of transactions over time by type of transaction.[60,61] As shown in the top panel, Visa's and MasterCard's individual shares of credit card transactions have remained relatively constant over the past several years, with Visa and MasterCard holding respective shares of 44 percent and 31 percent as of 2007. American Express and Discover followed with shares of 18 percent and 7 percent, respectively.

---

information with merchants, nor do merchants have the ability to verify any information provided by their servicers.

[60] Data for this figure, and various other tables and figures in this paper that report transaction volumes, are from The Nilson Report. Although The Nilson Report does not publicly report the sources of its data, the data are thought to incorporate direct reports from the major networks and issuers. For some historical statistics, the Nilson Report is the only source of nationwide information for the United States.

[61] The market share data throughout this section exclude transactions associated with proprietary credit cards issued by retailers and oil companies.

Signature debit transactions run almost exclusively over the Visa and MasterCard credit card infrastructure and were estimated by the Federal Reserve to comprise about 63 percent of total debit card transactions in 2006, as shown in table 2.[62] This large share is likely due to a number of factors that influence merchant acceptance and customer card use, including (1) Visa's and MasterCard's earlier honor-all-cards rules (overturned in the 2003 court settlement discussed in section VI), which required merchants that accepted a network's credit cards also to accept like-branded signature debit cards; (2) aggressive marketing programs by Visa and MasterCard, aimed at both issuers and consumers, including higher interchange fees for issuers and the associated rewards programs for card users; and (3) the ease of adoption of signature debit relative to PIN debit due to the fact that no new equipment beyond credit card readers need be installed.[63] Within signature debit, Visa has held a large and fairly stable share, with about 75 percent of transactions in 2007, as shown in the second panel of figure 6.

PIN debit transactions are less concentrated than either credit card transactions or signature debit transactions, and the PIN debit transaction shares of the networks have exhibited greater variation over time. The third panel of figure 6 shows the transaction shares of what are now the top four PIN debit networks (Interlink, Star, NYCE, and Pulse), over time.[64] There is a marked contrast between the positions of Visa and MasterCard in this market segment: Interlink, Visa's PIN debit network, experienced rapid growth in recent years and was the largest PIN debit network in 2007, with a transaction share of 37 percent, while Maestro, MasterCard's PIN debit network, had only a negligible transaction share.

The various developments in the debit card arena over the past decade have resulted in an increase in Visa's and MasterCard's shares of all debit card transactions. In particular, the continued growth of signature debit and the dominant position of Visa and MasterCard in that segment have yielded a combined share near 77 percent of all debit card transactions in 2007, as shown in the fourth panel of figure 6. Due to the

---

[62] In 2006, Discover announced plans to offer a signature debit card, but deployment of that card has been limited.

[63] Another technological factor that historically has made signature debit acceptance more available to merchants was the ability to conduct Internet and other types of card-not-present transactions with signature debit cards but not with PIN debit cards.

[64] Since 2005, Pulse has been owned by Discover.

additional expansion of its PIN debit operations, Visa alone has seen its share grow to almost 61 percent of all debit card transactions as of 2007.

Across all card transactions, Visa and MasterCard's combined share has increased slightly since 2001, mainly due to Visa's growing share of debit card transactions, as shown in the last panel of figure 6. As of 2007, the combined share of Visa and MasterCard stood at about 76 percent of all card transactions. The remaining card transactions were divided among American Express (7.4 percent), Star (6.3 percent), Discover (including the Pulse PIN debit network) (5 percent), and a number of small PIN debit networks.

2. Issuer Market Structure

Issuer concentration is of interest for several reasons. First, credit card issuers serve as important points of access for consumer credit, in addition to serving as providers of a means of conducting purchase transactions. Second, debit cards now represent an important tool for accessing deposits, through either purchases or withdrawals. Finally, card issuer concentration is important for network competition because (1) an individual issuer may serve on the board of a four-party payment card network, and therefore may be able to influence its pricing behavior, and (2) large issuers could potentially serve as competitors to either credit or debit card networks by abandoning the existing networks and forming new networks, either on their own or in conjunction with other issuers. The potential effect of this type of entry on interchange fees is unclear.

Figure 7 shows that the 3-firm issuer concentration ratio of total credit card dollar volume (black line) rose from 48 percent in 2003 to 57 percent in 2007, if the three-party networks (American Express and Discover) are included as issuers.[65] Excluding the three-party networks, the increase in the 3-firm concentration ratio was more rapid, from 49 to 63 percent over the same period (red line). The two identifiable increases in

---

[65] The analysis here focuses on shares of dollar volumes, rather than shares of the number of transactions, because the only complete data available are dollar volumes.

concentration are the direct result of the Bank One – JP Morgan Chase merger in 2004 and the purchase of MBNA by Bank of America in 2005.[66]

The top panel of table 3 shows market shares of the major credit-card issuers in 2007.[67] American Express had the largest market share, with 25 percent of volume. Excluding the three-party issuers, JP Morgan Chase was the top bank card issuer in 2007 with a market share of 25 percent of volume, followed by Bank of America and Citigroup with shares of 21 percent and 18 percent, respectively.[68] The distribution of market shares drops off quickly after the top three bank card issuers. If the three-party issuers are included, then the top ten issuers accounted for approximately 90 percent of all credit card volume in 2007.

As seen in middle and bottom panels of table 3, the top issuers for signature and PIN debit differ somewhat from those for credit cards because, unlike credit cards, debit shares are closely linked to relationships with deposit account customers. Debit shares are also affected by the marketing efforts – including rewards programs – of both the individual banks and the national networks to induce customers to use their debit cards (particularly for signature transactions). Some regional variation exists in overall debit card use due to merchant acceptance as well as marketing by the regional PIN debit networks.[69] The 3-firm concentration ratios in figure 7 show both signature and PIN debit issuing, as measured by dollar volume, to be historically less concentrated than credit card issuing. As of the end of 2007, the signature debit 3-firm concentration ratio had risen slightly since 2000 to 37 percent, and the analogous PIN ratio stood at 48 percent. However, the recent acquisitions of Wachovia by Wells Fargo, Washington Mutual by JP Morgan Chase, and Merrill Lynch by Bank of America have substantially increased the concentration of debit card issuance. Adjusting the 2007 year-end numbers

---

[66] The Bank of America / MBNA merger closed Jan 1, 2006; however, the 2005 data already reflect the deal.

[67] Because the available data do not separately report credit and charge card volumes, volumes for both types of cards are included in the credit card numbers.

[68] Accounting for its 2008 acquisition of Washington Mutual, JP Morgan Chase's adjusted 2007 share of four-party credit-card issuer volume would be 26.6 percent; other recent acquisitions had smaller effects on shares of credit-card issuer volume.

[69] As of 2006, rates of debit card use per capita were higher in the midwestern and western United States than in the South or the Northeast (Gerdes 2008).

to account for these acquisitions yields 3-firm concentration ratios of 53 percent and 62 percent for signature debit and PIN debit, respectively.

3. Acquirer Market Structure

Since 2000, concentration levels for acquiring payment transactions have also increased. As shown in figure 8, the 3-firm acquiring concentration ratio for all types of card transactions rose from 48 percent in 2000 to 62 percent in 2007.[70] By transaction type, the 3-firm concentration ratio for credit and signature debit was 49 percent in 2000, rising to 62 percent by 2007.[71] The increase for PIN debit is more substantial, with a rise in the 3-firm concentration ratio from 52 percent to 78 percent over the same period. As shown in table 4, First Data held the largest share of all acquired transactions (credit, signature debit and PIN debit combined) as of 2007, with its share conservatively estimated at around 32 percent.[72] Bank of America and Chase ranked second and third, with shares of about 18 and 12 percent, respectively.

The market for acquiring differs substantially from the market for payment networks or for card issuing.[73] Acquiring services are not strongly "brand-based;" that is, acquiring services obtained from different providers are very close substitutes for each other since no provider's service is particularly unique. Although there are substantial (indeed, overwhelming) scale economies, there is no obvious network effect. In other words, a merchant's benefit from using a given acquirer will depend on the prices and services of that acquirer, but will not depend on the number of other merchants who also use that acquirer. Furthermore, a merchant's customers generally have no preference for or knowledge of the identity of its acquirer. A merchant may therefore choose relatively freely among existing acquirers. (The limits on this free choice derive from multi-year contracts, bundled services, and other switching costs.) In addition, although fixed costs

---

[70] These concentration measures are based on card transactions (as opposed to dollar volume).
[71] Data that separate signature debit from credit transactions are unavailable, as acquirers do not report these totals separately. In the calculation of market shares for each firm in a joint venture, each party is credited with a share of transactions equal to its ownership share in the joint venture.
[72] First Data is not technically a merchant acquirer. The payment card networks require that a merchant acquirer be a member of the respective network. To qualify for membership, a firm must be an insured depository institution. First Data is granted access to the respective networks by sponsoring banks.
[73] Kjos (2007) provides a more detailed discussion of many of these issues associated with acquirers.

may serve as a barrier to entry, a new entrant could likely provide the data processing functions of merchant acquirers after an initial investment and scaling-up period. Nonetheless, some evidence suggests that the share of the total merchant discount received by acquirers (as opposed to flowing through to issuers in the form of interchange) has increased recently. One recent estimate is that for smaller merchants, roughly one third of their fees is paid to the acquirers (including independent sales organizations, which serve as resellers of acquirer services), while two thirds is passed along as interchange.[74]

4. Issuers, Acquirers, and Processors as Potential Competitors to Networks

Individual issuer and acquirer market shares – particularly in combination – are important in the context of network competition. A bank with a substantial share of card issuance, large transaction volume, and well-established merchant acquiring relationships has considerably greater potential to enter the market successfully as a new network than would a *de novo* firm. Entry for a new network is difficult in part because of the two-sided nature of the market: In order to attract issuers to join the network, enough merchants must have joined on the acquiring side, while in order to attract merchants to join, a sufficient number of individuals must hold the network's cards. A firm that already serves as both a large issuer and a large acquirer would begin with a substantial built-in base of users on each side of the market, and could therefore enter more easily than could a network starting from scratch to attract either set of users.

Similarly, processors that provide services to merchants or acquirers on one side of the market and issuers on the other side could potentially serve as competitors to the existing networks. A processor serving both sides of the market could, at least in theory, provide on-us processing in which transactions are authorized, cleared, and settled by the processor itself without using any of the existing networks.[75] Alternatively, the existence

---

[74] Cohen (2008) reports that the interchange fee comprises about 65 percent of the merchant discount. The remaining 34 percent is divided between the network (3 percent) and the acquirer (31 percent).

[75] The possibility of bypassing the network when processing on-us transactions has been the subject of litigation between card networks and processors. For example, in 2006, Visa and First Data settled a series of lawsuits that concerned on-us transactions performed by First Data. In the settlement, First Data agreed to process transactions using Visa's systems, even when First Data

of such processors could facilitate the entry of new card networks affiliated with issuers that have large existing cardholder bases or acquirers with large numbers of merchant contracts.

V. Recent industry developments

A number of recent industry developments may affect the competitive environment in which payment card systems operate. These include changes in the organizational forms of Visa and MasterCard, adaptations of the business model of American Express and Discover, and the development of new credit and debit card products, some of which rely on the automated clearinghouse (ACH) system.

One notable development over the past few years has been the conversion of Visa and MasterCard from not-for-profit, membership-based organizations to for-profit, publicly traded firms. MasterCard converted from a membership organization to a private share corporation in 2002, and its initial public offering (IPO) on the New York Stock Exchange occurred in 2006. In March 2008, Visa completed its IPO. Commentators have speculated that these conversions were motivated, at least in part, by a desire to limit possible antitrust liability, although the implications of the conversions for such liability are unclear (Shinder 2009). Even if these organizational changes alleviate antitrust concerns, it is unclear how they may affect the behavior of the two companies. The networks state that their objective is to maximize the number of transactions that they carry, an objective that is furthered by establishing an interchange fee that rewards issuers and consumers for using the network.

In recent years, both American Express and Discover have begun to allow bank issuance of their respective brands' credit cards. This change followed a 2001 court decision (discussed in the next section) that Visa and MasterCard could not prohibit banks that issue their cards from also issuing American Express or Discover cards. This change is likely to intensify competition among the networks for card-issuing banks. However, as discussed earlier, competition among networks for issuers does not

---

has existing relationships with both the issuer and the acquirer. See *Stipulation of Dismissal with Prejudice, Visa U.S.A. Inc. v. First Data Corporation, et al.,* No. 02-01786 (N.D.Cal. entered July 6, 2006).

necessarily lead to downward pressure on interchange rates, because networks can attract issuers and card users by setting higher interchange fees.

Several companies have tried, with varying degrees of success, to introduce new payment card products. While some of these products aim to compete directly with the existing networks by reducing merchants' card acceptance costs, others build on the existing card networks. The types of entry that have occurred largely represent either competition to undercut high merchant fees or attempts to capture interchange fee revenue. Although the former may be expected to place downward pressure on interchange fees, the latter is unlikely to do so. Appendix III describes some of the specific products that have been introduced.

VI. Regulatory and Legal Developments in the United States and Abroad

In a number of countries throughout the world, payment card systems have been targeted for investigation and, in some cases, for regulatory or legal action, by the courts, the competition authority, or the central bank. This section reviews regulatory and legal developments in the United States, Australia, and other countries.

A. UNITED STATES

In the United States, concerns about interchange fees and rules imposed on merchants by the card networks have been addressed mainly through private litigation, with occasional government intervention through antitrust lawsuits. Thus far, there have been no direct regulatory interventions.[76]

The first and most prominent civil case that addressed interchange fees was filed in 1979. In this case, National Bancard Corp. (NaBanco) asserted that Visa's interchange fees constituted illegal price fixing. The courts decided in favor of Visa in 1984, stating that interchange fees were a legitimate mechanism to transfer costs from the side of the market that the court determined to have lower costs (merchant acquirers) to the side with

---

[76] Wildfang and Marth (2006) provide more detail about major litigation involving the credit card networks.

higher costs (card issuers), thereby inducing the side bearing the greater cost to participate.[77]

In another influential case initiated in 1998, the U.S. Department of Justice filed a Sherman Act complaint against Visa and MasterCard that addressed two practices.[78] First, both networks contractually prohibited card issuers from issuing cards on other competing networks, such as Discover and American Express (exclusivity).[79] Second, Visa and MasterCard had overlapping members on their corporate boards (duality), resulting in a mechanism by which the two networks could potentially collude. The court's decision, reached in 2001, stated that Visa and MasterCard's dual board memberships were not a Sherman Act violation, but that exclusivity was. As a result of this decision, no credit card issuer may be prohibited by any payment network from issuing cards on a competing network. As previously noted, following the decision, a number of banks that issue cards on the Visa and MasterCard networks have begun to issue cards on the American Express or Discover networks as well.

In 1996, a number of merchants and retail trade associations filed lawsuits against Visa and MasterCard challenging the networks' rules that required merchants that accepted their credit cards to also accept their signature debit cards (the honor-all-cards rules). The various lawsuits were combined into a single, consolidated action, which became known as the *Wal-Mart* case.[80] In 2003, the *Wal-Mart* case was settled, with Visa and MasterCard agreeing to pay over $3 billion in damages and to rescind partially the honor-all-cards rules.[81]

---

[77] *National Bancard Corp. (NaBanco) v. Visa U.S.A., Inc.*, 596 F. Supp. 1231, 1265 (S.D. Fla. 1984), *aff'd*, 779 F.2d 592 (11th Cir. 1986).

[78] *United States v. Visa U.S.A. et al.*, 163 F. Supp. 2d 322 (S.D.N.Y. 2001), *aff'd*, 344 F.3d 229 (2d Cir. 2003), *cert. denied*, 543 U.S. 811 (2004).

[79] Since the 1970s, banks have been able to be members of both Visa and MasterCard following a private legal challenge in 1970 to Visa's (then called National BankAmericard, Inc.) exclusivity rules and a Department of Justice business review letter in 1975 that declined to approve those rules. *Worthen Bank & Trust Co. v. National BankAmericard, Inc.*, 345 F. Supp. 1309 (E.D.Ark. 1972), *rev'd*, 485 F.2d 119 (8th Cir. 1973). U.S. Department of Justice Antitrust Division, Business Review Letter No. 75-15 from Thomas E. Kauper, Assistant Attorney General, to Francis R. Kirkham and Allan N. Littman, Counsel for NBI (Oct. 7, 1975).

[80] *In re: Visa Check/MasterMoney Antitrust Litigation*, 192 F.R.D. 68 (E.D.N.Y. 2000).

[81] Following the settlement, merchants were allowed to make separate acceptance decisions for credit cards and signature debit cards.

In 2005 and 2006, a large number of retail merchants and trade associations filed approximately fifty civil lawsuits in U.S. courts against Visa, MasterCard, and several card-issuing banks alleging, among other charges, that interchange fees are too high and that the collective setting of interchange fees by members of the payment card associations constitutes illegal price fixing under U.S. antitrust laws. A large number of these lawsuits were consolidated into a single class action lawsuit that is pending in the U.S. District Court for the Eastern District of New York.[82] In recent testimony before Congress, a witness on behalf of MasterCard stated that the two sides have agreed to mediation, in an effort to resolve the lawsuit.[83]

In the past few years, merchant organizations have expressed their concerns to Congress about interchange fees and network operating rules. Both the Senate and the House have held multiple hearings on those topics.[84] In 2008, members of both chambers introduced legislation aimed at addressing a number of concerns associated with interchange fees and various rules and practices of payment card networks.[85]

---

[82] *In re Payment Card Interchange Fee and Merchant Discount Antitrust Litigation*, 2006 WL 2038650 (E.D.N.Y. 2006).

[83] Joshua Peirez (2008), statement delivered at the Hearing on H.R. 5546: The Credit Card Fair Fee Act of 2008 before the Judiciary Committee Antitrust Task Force, U.S. House, 111 Cong., May 15. (Transcript available at http://judiciary.house.gov/hearings/hear_051508.html, p. 73.)

[84] Energy and Commerce Committee Subcommittee on Commerce, Trade and Consumer Protection (2006), The Law and Economics of Interchange Fees, U.S. House, 109 Cong., February 15. Judiciary Committee (2006), Credit Card Interchange Fees: Antitrust Concerns, U.S. Senate, 109 Cong., June 12. Judiciary Committee Antitrust Task Force (2007), Hearing on Credit Card Interchange Fees, U.S. House, 110 Cong., July 19. Judiciary Committee Antitrust Task Force (2008), Hearing on H.R. 5546: The Credit Card Fair Fee Act of 2008, U.S. House, 110 Cong., May 15.

[85] Representative John Conyers introduced H.R. 5546, the Credit Card Fair Fee Act of 2008, which would have authorized merchants and card networks that met a market share threshold to negotiate network fees and other terms associated with merchant access to a card network. (At the time the bill was introduced, only Visa and MasterCard would have met the threshold.) The bill would have granted a limited antitrust exemption to allow merchants to negotiate collectively. Senator Richard Durbin introduced a similar companion bill (S. 3086) in the Senate. Representative Peter Welch introduced H.R. 6248, the Credit Card Interchange Fees Act of 2008, which would have prohibited certain rules and practices of card networks and would have mandated greater disclosure of system rules and transaction fees. Representative Stephen Lynch introduced H.R. 6620, the Gas Pump Fair Payment Act of 2008, which would have set a maximum merchant charge of one percent on sales of motor vehicle fuel.

B. AUSTRALIA

The Australian government established a Payments System Board (PSB) at the Reserve Bank of Australia (RBA) in 1998, which was given responsibility for promoting competition, efficiency, and stability in the payment system. The PSB undertook extensive data collection and analysis, leading to the publication in October 2000 (together with the Australian Competition and Consumer Commission) of *Debit and Credit Card Schemes in Australia: A Study of Interchange Fees and Access* (RBA Study).

The RBA Study concluded that credit and debit card interchange fees were well above the levels that could be justified on the basis of costs, and that these fees were not subject to the normal forces of competition.[86] To address these issues, the RBA imposed a number of reforms on the Australian payment system. Key elements of these reforms included (i) requiring the Visa and MasterCard networks to remove their no-surcharge rules, and (ii) imposing standards that required substantial reductions in credit and debit card interchange fees.

Since the reforms were implemented, the number of merchants levying surcharges on credit card transactions has increased steadily, and interchange fees have fallen. As of year-end 2007, about 10 percent of small merchants, 15 percent of large merchants, and 23 percent of very large merchants imposed surcharges.[87] The average interchange fee in the MasterCard and Visa systems fell from about 0.95 percent prior to the reforms to about 0.50 percent after the reforms, and the decline in interchange fees led to a significant reduction in merchant service fees. At the same time, the average value of cardholder rewards has declined, and average annual cardholder fees have risen.[88]

In September 2006, the RBA launched a comprehensive review of its reforms to Australia's payment card systems. Preliminary conclusions from this review were

---

[86] As discussed in Section III, theoretical models of interchange fees in two-sided markets demonstrate that welfare-maximizing fees are not necessarily based purely on costs.

[87] See Reserve Bank of Australia (2008a), pp. 17-18. "Small merchants" refers to those with annual revenue between AU$1 million and AU$20 million; "large merchants" to those with revenue between AU$20 million and AU$340 million; and "very large merchants" to those with revenue greater than AU$340 million. As of year-end 2006, the share of merchants in these three categories that levied surcharges were 5 percent, 9 percent, and 14 percent, respectively (Reserve Bank of Australia 2007, pp. 18-19).

[88] Reserve Bank of Australia (2007), pp. 19-23.

released in April 2008, and final conclusions were published in September 2008.[89] The

RBA concluded that

> (T)he reforms have met their key objectives. They have: increased
> transparency; improved competition by removing restrictions on
> merchants and liberalising access; and promoted more appropriate price
> signals to consumers. Nevertheless … the Board remains of the view that
> the competitive forces acting on interchange fees are still relatively
> weak.[90]

In light of the improvements in the competitive environment, the RBA is prepared to

remove the existing interchange fee regulations if industry participants take appropriate

steps to ensure that interchange fees do not rise above their current levels. In August

2009, the RBA will assess the extent to which industry participants have addressed its

concerns. If it is satisfied with the measures they have taken, it will deregulate

interchange fees; otherwise, regulation will be continued and strengthened.

C. OTHER COUNTRIES

In a large and growing number of countries, authorities have taken steps aimed at

reducing or eliminating common interchange fees.[91] In several countries, including New

Zealand, Poland, and the United Kingdom, competition authorities have determined that

multilateral setting of interchange fees is illegal and must be discontinued. In Spain, the

competition authority denied the legally-required authorization of the payment card

schemes' interchange fee agreements and instructed the schemes to submit new plans for

authorization. In other cases, either the competition authority (e.g., Switzerland) or the

central bank (e.g., Mexico) has reached an agreement with the card issuers to reduce

interchange fees. No-surcharge rules have been examined in a number of countries and

have been eliminated in the Netherlands, Sweden, and the United Kingdom.[92]

The European Commission (EC) has also intervened with regard to cross-border

interchange fees in the European Economic Area (EEA). In 2002, the EC issued a

decision exempting Visa's cross-border interchange fees from antitrust prosecution for a

period of five years, ending December 31, 2007, provided that Visa met certain

---

[89] Reserve Bank of Australia (2008a, 2008b).
[90] Reserve Bank of Australia (2008b), p. 14.
[91] Bradford and Hayashi (2008) catalog many of these interventions.
[92] See Organization for Economic Cooperation and Development (2006), p. 15.

conditions.[93] These conditions required Visa to (i) develop a cost-based method for setting cross-border interchange fees; (ii) cap interchange fees at a specified level until December 31, 2007; and (iii) provide merchants with more information about interchange fees. Following the expiration of the exemption, the EC announced on March 26, 2008 that it was initiating formal antitrust proceedings related to Visa's cross-border interchange fees, as well as the network's honor-all-cards rule.[94]

On December 19, 2007, the EC announced that MasterCard's multilateral interchange fees for cross-border transactions were in violation of EC Treaty rules on restrictive business practices.[95] As required by the decision, MasterCard temporarily repealed its cross-border interchange fees in the EEA on June 21, 2008.[96] However, the network increased various other system fees in October 2008.[97] In April 2009, the EC announced that MasterCard had agreed to withdraw the recent fee increases by July 2009 and would begin using a new methodology to determine its cross-border interchange fees. According to the EC, this new methodology will substantially decrease cross-border interchange fees.[98] In its announcement, the EC stated that it would closely monitor the implementation of the proposed changes.

VII. Policy Interventions

This section discusses a number of possible policy interventions for the payment cards industry. The options considered reflect actions that authorities in other countries have taken, as well as proposals that have been advanced by various parties in the United States. We emphasize that the discussion in this section does not necessarily reflect a list of interventions that the U.S. government should undertake. Instead, this discussion is intended to present and critically evaluate some of the policy prescriptions that have been proposed to address perceived problems in the payment cards industry.

---

[93] European Commission (2002).
[94] European Commission (2008a).
[95] European Commission (2007).
[96] European Commission (2008b).
[97] European Commission (2009a).
[98] European Commission (2009a). MasterCard's methodology to determine its cross-border interchange fees in the EEA is based on a criterion known as the "tourist test" (European Commission 2009b). We discuss this approach to setting interchange fees in the next section.

From the viewpoint of economic efficiency, intervention in a market may be warranted to correct a misallocation of resources, in this case the over- or underutilization of certain payment methods. Although many proponents of intervention in payment card markets assert that the policy objective should be to lower merchant fees, the primary policy objective would presumably be to produce an overall structure of card fees that is more efficient, and not simply to alter certain fees.

Some common concerns arise in connection with all of the policy options discussed in this section. First, the possible effects of any intervention are highly uncertain. Although economic models can provide some insights regarding the qualitative effects of a policy intervention, they typically have little to say about the quantitative magnitudes of these effects. Furthermore, the theoretical models tend to be highly stylized, and therefore may fail to capture important real-world features that influence the effect of an intervention. For example, in response to a policy that required a reduction of interchange fees in four-party card systems, a large issuer could conceivably leverage its base of existing cardholders and spin off a proprietary (three-party) network that would not be subject to such regulation. The current generation of economic models does not provide much insight into whether such an outcome would be likely or desirable.

A second concern common to all of the interventions considered here is the possible redistribution of surplus (i.e., the difference between benefits and costs) across parties. Most of the interventions seek a structure of card fees that is more balanced between merchants and card users. Any rebalancing of card fees will necessarily involve a transfer of surplus from some parties to others. A lower interchange fee, for example, might involve lower surplus for card users and issuing banks and higher surplus for non-card users and merchants. Such concerns do not relate to economic efficiency, but rather involve judgments about equity in the distribution of surplus among different groups in the economy.

Finally, much of the debate surrounding the payment card industry has focused on interchange fees in four-party credit card systems. However, issues regarding the structure of fees in a transaction also apply to debit and three-party card systems. Moreover, because these various payment methods and systems operate in the same

42

broadly construed retail payments market, a narrow intervention that targets interchange fees for credit cards (or even interchange fees for both credit and debit cards) could have effects on competition and pricing throughout the entire retail payments market.[99]

Thus, policymakers should proceed with caution, carefully assessing the potential effects of any contemplated intervention upon the payments system as a whole. This assessment should include consideration of the possible responses of all parties involved in the payments system (e.g., merchants, targeted networks, other networks, issuing banks, acquiring banks, and consumers) to legal or regulatory changes.

A. CLARIFY OR ELIMINATE RESTRICTIONS ON DIFFERENTIAL RETAIL PRICING ACROSS PAYMENT METHODS

As noted in section II, most of the card systems have rules, generally referred to as no-surcharge rules, that restrict the ability of merchants to set different prices across payment methods. Currently, the rules of the card systems appear to allow discounts for cash (and presumably checks, as required by the Cash Discount Act), but prohibit surcharges for card payments as well as differential pricing across different brands or types of cards. As noted in section II, however, recent Congressional testimony indicates that merchant associations and the card systems disagree regarding the extent to which current rules allow for cash discounts and the form that any discounts may take.

One possible policy intervention would address the existing restrictions on differential pricing.[100] The most limited such intervention could amend the Cash Discount Act in order to clarify the exact payment methods (e.g., cash, check, debit card)

---

[99] A further consideration reflects possible differences in the roles of various payment methods in the broader economy. In particular, while both debit and credit cards serve the same general payment purpose, debit cards have arguably become an integral part of the banking system as a means to access deposit accounts. Whether differences stemming from the link to deposit accounts (i.e., debit cards) versus provision of consumer credit (i.e., credit cards) may warrant different regulatory treatment is an open question that is beyond the scope of this paper.

[100] In the early 1980s, the Federal Reserve Board commented on differential pricing across payment methods in regard to the Cash Discount Act. In testimony, the Board encouraged Congress to allow both cash discounts and credit card surcharges with minimal restrictions. Ultimately, only the provision allowing cash discounts was enacted. See Nancy H. Teeters (1981), statement delivered before the Banking, Housing and Urban Affairs Committee Subcommittee on Consumer Affairs, U.S. Senate, 97th Cong., February 18; and Nancy H. Teeters (1984), statement delivered before the Banking, Housing and Urban Affairs Committee Subcommittee on Consumer Affairs, U.S. Senate, 98th Cong., February 7.

for which discounts must be allowed by the card networks and the forms that any such discounts may take. A more significant intervention could prohibit any restrictions by card networks on differential pricing across payment methods.

If unrestricted differential pricing were allowed, it is not clear to what extent merchants would choose to set different prices for various payment methods. Indeed, in those countries where surcharging is currently permitted, a relatively small proportion of merchants actually impose surcharges, although the experience of some countries suggests that surcharging becomes more prevalent over time after the abolition of no-surcharge rules.[101] However, even the threat of surcharging may improve the bargaining position of merchants as a whole, which could induce the networks to lower interchange fees.

If merchants were actually to engage in differential pricing, retail prices faced by consumers would adjust to reflect the benefits and costs that merchants experience from different payment methods. Relative to a single uniform price, prices for those methods with high net costs for merchants would rise, while prices for those with lower net costs would fall.[102] When choosing a payment method, a consumer would face a menu of prices and would take into account both the price differentials and his or her own private benefits and costs associated with different payment methods. As a result of the price differences, overall use of different payment methods would change, with use falling for high-price methods and rising for low-price methods.

Eliminating restrictions on differential pricing may be an attractive policy option for a number of reasons. First, it would remove an existing restraint of trade, the economic justification for which is generally questionable. Second, it is a transparent policy that is relatively easy to implement. Third, it does not require a great deal of information about costs and benefits accruing to the various parties in a transaction; rather, it relies upon market forces to yield retail prices that reflect those costs and benefits. Fourth, it can potentially eliminate the effect of the level of the interchange fee

---

[101] See, for example, Reserve Bank of Australia (2008a), pp. 17-18.
[102] Heterogeneity across merchants in the relative costs of accepting different types of payments would likely lead to considerable variation in payment pricing across merchants.

on card use and on the welfare of the parties involved in a transaction.[103] Fifth, its applicability and effect would not depend on the network type (i.e., PIN debit, signature debit, or credit), organizational form (i.e., three-party or four-party system), or pricing strategy (i.e., interchange fees or switch fees) of different card networks. Finally, and perhaps most importantly, it would allow for the possibility that prices could adjust to reflect costs and benefits of card transactions to merchants, potentially leading consumers to internalize the externalities associated with their use of payment cards that might otherwise prevent efficient payment card use.[104]

Permitting unrestricted differential pricing may also have some drawbacks. Most importantly, in order for price differences to accurately internalize externalities, those price differences must reflect true social costs and benefits. If merchants have market power, the retail price differences across payment methods that result from removing restrictions on differential pricing may not reflect social costs and benefits.[105] Similarly, if issuing or acquiring banks have market power, card user fees or merchant discounts need not reflect the social costs of card services. In general, the presence of market power can prevent differential pricing from inducing efficient use of payment methods by distorting prices and fees away from the underlying social costs and benefits.[106]

Another consideration regarding differential pricing is that it may create confusion on the part of consumers, particularly if it involves a different price for each of numerous alternative payment methods. Merchants would clearly have an incentive to minimize such confusion, but they may not be able to eliminate the problem completely. Nonetheless, removing restrictions on differential pricing would simply leave this pricing decision to individual merchants.

---

[103] With differential pricing, a high interchange fee could generate a high retail price for card transactions, in addition to rewards for card use. The high card price and rewards could cancel, yielding no net effect of the interchange fee on consumers' incentives to use cards. See appendix II.D or Rochet and Tirole (2002), Wright (2003), and Gans and King (2003).

[104] Appendix II.D presents some of these results associated with differential pricing in more detail.

[105] This possible outcome reflects the standard result that firms with market power do not set prices that correspond to social costs.

[106] See Rochet and Tirole (2002), Wright (2003), and Schwartz and Vincent (2006).

B. PROHIBIT NETWORK DETERMINATION OF INTERCHANGE FEES

Some observers have argued that the use of a common interchange fee among banks in a network should be disallowed as a violation of antitrust laws.[107] One possible alternative could involve bilateral negotiation of interchange fees between merchants and their banks on one side of the market and cardholders and their banks on the other. Supporters of this policy option often point to bilateral negotiations between merchants and the three-party systems as examples of the feasibility of such an approach. However, in a country with thousands of banks, the number of bilateral arrangements required in a four-party system could be prohibitively high.

More importantly, as discussed in section III, abolition of a common network-determined interchange fee would make it difficult for the network to maintain an honor-all-cards rule. Some type of honor-all-cards rule plays a central role in creating network value in a four-party system by providing assurance to the consumer that a card bearing the network logo will be accepted by any merchant that accepts the network's cards. Thus, a policy prohibiting network determination of interchange fees, which might undermine the networks' honor-all-cards rules and the associated card acceptance guarantees, could eliminate the four-party system as an organizational form.

C. REGULATE THE LEVEL OF INTERCHANGE FEES

Another possible policy involves intervening to alter the level of interchange fees. Such an intervention could take any of several forms, ranging from a negotiated reduction in interchange fees to a cap on the level of interchange fees. The most extreme form of this regulatory intervention would involve a regulator directly setting interchange fees, perhaps to zero.

Because the typical concern is that interchange fees are too high, this type of policy would likely involve a reduction of those fees. Economic theory provides guidance about some of the effects of such a reduction, although the empirical magnitudes of these effects are currently unknown. On the merchant side of the market,

---

[107] Critics of common interchange fees sometimes point to the check and automated clearinghouse (ACH) systems as examples of four-party retail payment systems that operate "without an interchange fee" (e.g., Frankel and Shampine 2006). In fact, these systems do have a common interchange fee; it is simply set to zero.

merchant discounts, which typically equal the interchange fee plus other acquirer costs and a markup, would be expected to fall. Correspondingly, merchant acceptance of the card would likely increase; however, the magnitude of this effect would be small for a card that already enjoys widespread merchant acceptance. On the card user side of the market, card user fees (rewards) would be expected to rise (fall), thereby making the card less attractive to consumers, compared with other payment methods. Overall, the effect of consumer substitution away from the regulated card system would likely dominate that of increased merchant acceptance, leading card use for the system to decline. If merchant costs were to decline, final goods prices could decline, if merchants pass on some or all of their cost savings to their customers. The extent to which merchant costs decline would depend on the merchant costs associated with the payment methods to which consumers switch (and could theoretically increase, if consumers switch to a payment method that is more costly to merchants). The extent to which any cost savings are passed through to customers would depend on the merchants' market power in the final product market.

This policy approach has a number of advantages that likely explain its use by central banks, such as the RBA, and competition authorities elsewhere. First, once the regulated values for the interchange fee have been chosen, implementation of the policy is straightforward and transparent. Moreover, this intervention directly addresses the primary concern of merchants and some policymakers, namely that interchange fees are too high.

However, determining an appropriate regulated value for the interchange fee can be quite challenging. Ideally, the regulator would want to set the interchange fee equal to the efficient level (i.e., the level that internalizes externalities between the parties to a transaction). As discussed in Section III, calculation of that fee requires knowledge of social costs and benefits that are difficult, if not impossible, to measure accurately.

This difficulty has led authorities in other countries, such as the RBA, to establish "cost-based" fees in which the regulated value of or cap on interchange fees is computed so as to reimburse card issuers for costs that they cannot recover from card users without a substantial increase in card user fees.[108] Others have called for setting the interchange

---

[108] See, for example, Reserve Bank of Australia (2005).

fee to zero. Proponents of these two approaches argue that they provide a transparent methodology to determine the regulated levels of fees. However, the determination of which costs should be included in a cost-based fee is necessarily arbitrary, and measuring those costs is nontrivial, particularly if frequent re-estimation of costs is necessary. More importantly, the economic theory underlying the efficient interchange fee provides no rationale for either a strictly cost-based interchange fee or an interchange fee of zero.

As noted in the previous section, the EC has recently announced that MasterCard will use a criterion known as the "tourist test" when setting its cross-border interchange fees in the EEA.[109] Introduced by Rochet and Tirole (2006b), the tourist test requires that the transactional costs of cards, including the merchant discount, to merchants not exceed the transactional benefits of cards, relative to some baseline payment method, such as cash. In other words, the interchange fee would be set such that card transactions do not raise merchant costs, relative to cash, at the point of sale.[110]

From a theoretical point of view, this criterion may be attractive because, under certain conditions discussed in Rochet and Tirole (2006b), an interchange fee that satisfies the tourist test corresponds to the efficient interchange fee. Despite this theoretical appeal, however, use of the tourist test as a means to set interchange fees involves substantial problems.[111] Most importantly, the tourist test requires detailed information on transactional costs and benefits of cards and cash to merchants. These benefits and costs are likely complex and may vary substantially across merchants as well as across types of transactions, even for a given merchant. Moreover, it is unclear whether or how those benefits and costs should be aggregated across different types of merchants or different types of transactions.

---

[109] This change reflects an action by MasterCard, rather than a regulatory intervention. However, it was made in the face of substantial antitrust scrutiny and will presumably be subject to approval by the EC, as discussed in section VI.

[110] Rochet and Tirole (2006b, p. 2) describe the tourist test as answering the question "would the merchant want to refuse a card payment when a non-repeat customer with enough cash in her pocket is about to pay at the cash register?" They explain that "(t)he merchant discount passes the tourist test if and only if accepting the card does not increase the merchant's operating costs."

[111] One practical problem is that the tourist test technically concerns the merchant discount, not the interchange fee. Knowledge of acquirer costs and mark-ups would be required to translate the tourist test's restrictions on the merchant discount into restrictions on interchange fees.

Regulation of interchange fees is also problematic because the key issue is the overall structure of fees across parties to a transaction, not the interchange fee itself. When faced with a constraint on the level of its interchange fee, a card network might adjust other prices, such as switch fees, in order to achieve the network's desired prices for the two sides of the market. Furthermore, some systems, notably the three-party systems, operate without an explicit interchange fee. Narrow regulation of interchange fees would not apply to the prices or fees of these three-party systems (as was the case in Australia), even though concerns about the efficiency of their prices or fees are presumably the same as for systems with explicit interchange fees. These concerns suggest that regulation of fees charged to merchants might appear more appropriate than narrow regulation of interchange fees. As with interchange fee regulation, however, regulation of merchant fees involves the fundamental difficulty of choosing the appropriate level for those fees.

D. RELAX CARD ACCEPTANCE REQUIREMENTS

Another policy option would involve limiting the scope of the honor-all-cards rules. Although an honor-all-cards rule is often stated as the requirement that merchants honor cards of different issuers with the same network brand, it also applies to different classes or categories of credit cards Thus, a merchant that accepts Visa or MasterCard credit cards must accept all credit cards of those systems regardless of the card program. As noted earlier, this requirement has become increasingly controversial with the recent introduction by the networks of higher interchange fees for premium cards. Critics argue that the card networks are exploiting their honor-all-cards rules to tie acceptance of premium cards to acceptance of basic cards. The critics further argue that, although interchange fees for basic cards have not increased in the past few years, the introduction of higher interchange fees for premium cards has effectively served to increase interchange fees.[112]

This issue is similar to the dispute that surrounded signature debit cards leading to the class action lawsuit (the *Wal-Mart* case) discussed in section VI, the settlement of

---

[112] Merchants have also expressed concern that their inability to distinguish between premium and standard cards prevents them from verifying the accuracy of the merchant fees assessed by their acquirer.

which separated acceptance of Visa/MasterCard signature debit and credit cards. Reflecting the similarity between the premium-card example and this precedent, the pending class action lawsuit discussed in section VI includes complaints about the application of honor-all-cards rules to different classes of credit cards.

On the one hand, policy action to alter honor-all-cards rules might be advisable to avoid the cost and uncertainty associated with such litigation. On the other hand, whereas differences between debit and credit cards are clear enough to allow their acceptance to be separated, distinctions between credit card programs are arguably more subtle. Both Visa and MasterCard have a number of different credit card programs, and any intervention targeting honor-all-cards rules would need to classify the cards associated with these programs in some way. The most straightforward classification could involve a threshold based on interchange fees (e.g., any card with an interchange fee above some pre-specified level is designated a premium card), but even this approach would necessarily involve somewhat arbitrary criteria. Moreover, defining card types based on interchange fees would leave open the possibility that networks could alter other fees, such as switch fees (assessments) and rebates to issuers, in such a way as to maintain their desired pricing structure while keeping interchange fees below the premium threshold, thereby circumventing the intent of the policy. In addition, to avoid widespread confusion among consumers regarding card acceptance, any such classification would require a re-labeling of cards to clarify which cards are or are not accepted by a particular merchant. Even with such a re-labeling, cardholders could experience significant confusion about acceptance policies. Re-labeling would also be necessary for merchants to distinguish different classes of cards in order to apply their specific acceptance policies. Finally, technological standards would have to be changed and implemented to allow merchants to distinguish electronically among these card types.

E. MANDATE MULTI-BUGGED CARDS AND ALLOW MERCHANT CONTROL OF ROUTING RULES

As noted in section III, an important factor influencing the effect of competition among card networks on interchange fees is the cardholding behavior of consumers. Some of the more intriguing results in the economics literature arise when *all* consumers hold cards of multiple networks. This behavior, known as "multi-homing" by consumers,

tends to place downward pressure on interchange fees and can, in certain circumstances, yield interchange fees that are economically efficient.

The result that multi-homing consumers place downward pressure on interchange fees is straightforward. When all consumers hold and use cards of multiple networks, merchants can choose to accept only a subset of those cards without losing business. In particular, merchants can choose to accept cards that carry lower interchange fees.[113] In general, this will lead to lower interchange fees than would be observed either with a monopoly card network or with multiple networks and single-homing consumers.[114]

This result suggests that one method to encourage efficient competition among card networks would be to induce multi-homing on the part of consumers. One obvious way to induce multi-homing by consumers would be to mandate that each card be able to operate on two or more networks (Frankel and Shampine 2006). Such a situation would involve what have been called "multi-bugged" cards. Contingent on the consumer choosing to use the card, the merchant could then choose the "routing rule;" that is, the rule dictating which network carries the transaction.[115,116]

---

[113] A merchant's tradeoff between costs and benefits of card acceptance implies that a merchant will not necessarily choose to accept cards from the network with the lowest interchange fee. The benefits that a merchant derives from card acceptance include both the transactional benefits (e.g., reduced handling costs or increased certainty of payment, relative to other payment methods) and increased demand for its products. The magnitude of the latter effect depends, in part, on the surplus that consumers derive from the use of a given network's card – the greater the consumer's surplus from a card, the more likely he or she will be to choose a merchant based on whether the merchant accepts that card. In general, consumer surplus from using a network's card will be higher (e.g., rewards will be greater or fees will be lower) when the interchange fee is higher. Thus, when making its card acceptance decision, a merchant will weigh the cost of the interchange fee (or, more precisely, the merchant discount) against the benefit of increased final product demand generated by card acceptance (Farrell 2006, Rochet and Tirole 2006b).

[114] If all consumers multi-home, it is theoretically possible that the resulting interchange fees will be efficient (Guthrie and Wright 2007, Rochet and Tirole 2006b).

[115] As discussed in section IV, PIN debit is an example of an existing system in which some cards carry multiple network brands. The networks typically have rules to determine how transactions involving multi-bugged cards should be routed. In some cases, the rules applied by different networks contradict each other, and the merchant processor makes the routing decision. Conditional on choosing to use a multi-bugged PIN debit card, a consumer typically does not know which network carries his or her transaction. Hermalin and Katz (2006) analyze routing rules, network membership, and card use for situations, like PIN debit, in which multiple networks may be available to perform transactions.

[116] Voluntary multi-bugging would not be likely to work because card issuers (which generally prefer higher interchange fees) would have no incentive to issue cards with multiple network brands.

Despite the apparent simplicity of this intervention, implementation could require some difficult or arbitrary decisions. For example, which networks' logos, or "bugs," must a card carry and how would such a determination be made? Is it reasonable to require banks, both issuers and acquirers, to enter contracts with multiple networks as would be necessary under this policy?[117] What information, if any, would the merchant and card issuer need to provide to the cardholder regarding the routing of transactions? What technological standards would need to be introduced to provide a common platform for multi-bugging? How would three-party card networks fit into an environment of multi-bugged cards? Finally, by reducing the distinctions among card networks, mandating multi-bugged cards would be likely to degrade the values of the network brands and to reduce incentives to innovate.

In summary, a mandate that all cards be capable of operating on multiple networks, while attractive theoretically, would raise a number of concerns. In general, these concerns reflect the likely difficulties associated with implementing this option. However, despite these concerns, this policy option is interesting because it illustrates conditions under which, in theory, competition among networks has the potential to yield efficient outcomes. In particular, the theoretical literature suggests that, in certain circumstances, competition among networks can yield efficient interchange fees and card use in a setting where all cardholding consumers multi-home. Mandating multi-bugged cards could be one way to achieve universal multi-homing.

F. DO NOTHING

A final alternative for policymakers regarding concerns about the payment card industry would be to do nothing. A policy of non-intervention in card markets has a number of advantages. To begin with, it is not entirely clear that, under the status quo, the market is in fact generating outcomes that are inefficient (or at least less efficient than would result from intervention). While the theoretical economics literature lends some support to merchant concerns that interchange fees, and hence merchant discounts, are too high, no rigorous empirical evidence supports these claims. This dearth of empirical evidence results, in part, from the difficult nature of the underlying empirical question

---

[117] Although most acquiring banks provide services for multiple networks and most issuing banks issue cards for multiple networks, some banks deal with only one network.

that must be answered. In order to demonstrate inefficiency in the market, one must document a misallocation of resources. In this case, the misallocation would involve excessive use of certain payment methods (presumably credit cards and perhaps signature debit cards) at the expense of others relative to the "efficient" level. Absent such evidence of an inefficient outcome, some would argue that government intervention is unwarranted.[118]

In addition, as noted in the introduction to this section, considerable uncertainty exists about the possible effects of any intervention (though clearly some interventions are more substantial than others). The extent to which various interventions could alleviate concerns in payment card markets is unknown and any intervention could have unintended adverse consequences for the payment card market and the overall retail payments system. Whether such unintended consequences might yield distortions that outweigh any efficiency gains from intervention is an important consideration.

To the extent that problems do exist in the payment card market, a policy of non-intervention relies purely on market responses to deal with those problems. Market responses could involve either actions taken by the existing card networks to address merchants' concerns or the entry of new payment networks. For example, Visa and MasterCard have responded to merchant concerns regarding their lack of access to network operating rules by making versions of network rules available online. Also, as discussed in section V and Appendix III, entry and innovation have occurred in retail payment markets, and some of this activity reflects an attempt to address merchant concerns about high merchant fees for credit card transactions. However, due to network effects, the establishment of new payment networks is difficult; therefore, the extent to which these entrants will serve as effective competitors for the established networks is unclear. Nevertheless, actions taken by Visa and MasterCard and by new entrants suggest that concerns in the payment card market have led to market responses that could arguably mitigate various problems to some degree, even in the absence of government intervention.

---

[118] Documentation of inefficient card use would require information about both societal costs *and* benefits associated with different payment methods.

However, a policy of non-intervention also has some potential drawbacks. Currently, most activity to address concerns associated with interchange fees and other card network rules occurs in the courts as a result of antitrust litigation. On the one hand, private or government litigation based on the antitrust statutes often provides an effective means to deal with the type of competition issues that surround payment cards. On the other hand, litigation involves substantial costs and a high degree of uncertainty with regard to the outcome. For example, if the current class action lawsuit against Visa and MasterCard were to yield a decision that the joint setting of interchange fees is a violation of antitrust laws, the economic ramifications could be far-reaching.[119] Legislative or regulatory intervention in the market might not forestall litigation; nevertheless, it could mitigate some of the principal concerns that underlie much of the legal activity.

Finally, while it is not clear that current market outcomes are inefficient, it is also not clear that they are efficient. As discussed in the section III, the theoretical economics literature broadly suggests that market-determined prices and outcomes will be inefficient due to the nature of the economic problem that interchange fees are intended to correct (i.e., that the incentives of the card networks, banks, merchants, and consumers do not reflect externalities). Moreover, although some entry has occurred in the market, entry into network industries is difficult, particularly when entrants are faced with incumbent firms as well-established as the major card networks. A policy of non-intervention clearly does not alleviate any inefficiencies that are present.

---

[119] Some observers have speculated that a decision in favor of the plaintiffs could yield an award as high as $50 billion (e.g., Berry 2006). However, the judge in the pending class action lawsuit recently ruled that the settlement in the *Wal-Mart* case nullified the plantiffs' claims for damages incurred prior to January 1, 2004. *In re Payment Card Interchange Fee and Merchant Discount Antitrust Litigation*, 2008 WL 115104 (E.D.N.Y. 2008). Some observers have argued that this decision, along with the recent corporate reorganizations of Visa and MasterCard, greatly limits the potential claims (e.g., MacDonald 2006). The plaintiffs in the lawsuit argue that the networks still facilitate collusive pricing by banks and have challenged the IPOs of Visa and MasterCard (Shinder 2009).

APPENDIX I: Key terms and concepts

Acquirer – The depository institution or other party that provides payment card services to merchants.

Credit card – A payment card that draws on a (typically uncollateralized) line of credit when the cardholder makes a purchase.

Debit card – A payment card linked to a deposit account that withdraws funds from that deposit account when the cardholder makes a purchase. Many debit cards can function as either a PIN debit card or a signature debit card.

- **PIN debit card** – A debit card that requires a cardholder to enter a security code (typically a four digit number) when making a purchase.

- **Signature debit card** – A debit card that does not require a cardholder to enter a security code (and often requires the cardholder to sign a receipt instead) when making a purchase.

Externality – An effect (either negative or positive) that the actions of one party have on the welfare of another party, without any compensation being paid to or by the affected party. Externalities are relevant in two-sided markets when a party on one side of the market (e.g., a consumer or a merchant) does not consider the effect of his or her actions on the welfare of a party on the other side of the market.

Four-party payment card system – A payment card system in which four parties are involved in a card transaction: the consumer, the bank that issued the card to the consumer, the merchant, and the bank that provides card services to the merchant. (Technically, the card network is itself a fifth party that also participates in such a system.) Examples include Visa, MasterCard, and all of the PIN debit networks.

Honor-all-cards rule – A network rule under which a merchant that accepts a network's card (of a given type) must accept all of the network's cards (of that type) regardless of the card issuer or specific card program.

Interchange fee – A fee in a four-party payment card network that is (typically) paid by the acquirer to the issuer for each card transaction. The fee is set at the network level and is typically common to all issuers and acquirers associated with a network.

Issuer – The depository institution or other party that provides a payment card to a cardholder and maintains the cardholder's account.

Merchant discount – The difference between the dollar value of payment card transaction and the amount that the merchant receives from its acquirer.

No-surcharge rule – A network rule that prohibits merchants from imposing surcharges on purchases made with a network's cards. (Network rules typically allow discounts for cash purchases.)

Surplus – The difference between benefits and costs experienced by a party. Total surplus is the sum of the surplus of each party in the economy.

Three-party payment card system – A payment card system in which the network itself both issues cards to consumers and provides card services to merchants. As a result, only three parties are involved in a card transaction: the consumer, the merchant, and the card network. Examples include the traditional structure of American Express and Discover.

Two-sided market – A market in which two distinct and complementary sets of users must simultaneously agree to use a product provided by a third party. Payment card markets are two-sided markets in that a merchant and a merchant's customer must both agree to use a payment card.

APPENDIX II: The economics of interchange fees illustrated in a model

This appendix provides more detail about some major results in the economics literature regarding payment card networks and interchange fees. It presents a simple model of consumers, merchants, banks, and a card network to evaluate the effects of an interchange fee in that model. The model is a simplification of the seminal model of a card network developed by Rochet and Tirole (2002).[120]

Consider an economy with a single good sold by multiple merchants with identical costs. Each consumer buys one unit of the good, but must decide where to buy it and how to pay for it.[121] For the latter decision, consumers have two options: a payment card or cash.[122] All merchants accept cash payments, but each merchant must decide whether also to accept cards. Each payment method provides different *transactional benefits* to consumers and merchants.[123] Consumers (i.e., buyers) are heterogeneous in their transactional benefits, b_B, associated with card transactions, while merchants are assumed to receive homogeneous transactional benefits, b_M, from card transactions. The transactional benefits that both consumers and merchants receive from cash transactions are normalized to zero.[124] Thus, the overall (social) benefit associated

---

[120] Many of the assumptions and numerical values in this appendix are intended to be purely illustrative of key issues surrounding interchange fees.

[121] In other words, aggregate demand for the good is assumed to be completely inelastic. This assumption ignores the fact that consumers may decrease their overall consumption of the good to the extent that card acceptance yields higher costs and subsequently higher prices. At the same time, however, the assumption also precludes the possibility that card acceptance may provide benefits that increase overall consumption at any given price. Because these two effects move in opposite directions with an ambiguous net effect on consumption, the literature has typically assumed inelastic demand to focus on the choice of payment method.

[122] As is universal in the theoretical literature, this discussion abstracts from the role of credit cards as a source of financing and focuses instead on the choice of payment method.

[123] These transactional benefits are intrinsic benefits that merchants and consumers receive from using the card technology. Notably, these intrinsic benefits are fixed in the short run and do not depend on any transaction fees including any interchange fee. For consumers, these benefits could include security, convenience, and ease of budgeting. For merchants, these benefits could include faster check-out times and decreased cash handling.

[124] Technically, the benefits b_B and b_M measure card benefits net of cash benefits. Even more technically, because a consumer may like certain aspects of card transactions, but not others, the consumer benefits b_B reflect consumer transactional benefits less transactional costs of card use,

with a particular card transaction is $b_B + b_M$.

The payment card technology is supplied by a card network that links merchant and consumer banks. Issuing banks provide card services to consumers and incur a cost c_I per card transaction. Similarly, acquiring banks provide card services to merchants and incur a cost c_A per card transaction. Hence, the overall social cost of a card transaction is $c_I + c_A$.[125]

A. EFFICIENT CARD USE AND PRIVATE USE DECISIONS

Efficient card use describes which transactions should be performed with the card in order to maximize social welfare. Use of a card to complete a particular transaction is efficient if the social benefits of a card transaction exceed the social costs:

$$b_B + b_M \geq c_I + c_A. \qquad \text{(efficient card use criterion)}$$

If this condition is not satisfied in an interaction between a consumer and merchant, then the transaction should be performed with cash.

Although efficient card use considers social benefits and costs, each party will consider only his or her own private benefits and costs when deciding how to perform a transaction. Suppose that a consumer faces a card use fee (or, if negative, a reward) f set by his or her issuing bank. If merchants do not set prices for the products they sell that vary across payment methods, then a consumer will compare his or her transactional benefits to the card fee when deciding how to pay. In other words, given that a merchant accepts cards, the consumer will use the card whenever

$$b_B \geq f \qquad \text{(consumer use criterion)}.$$

Similarly, an acquirer may charge a merchant discount m for each card transaction. The surplus or net benefit of a card transaction for a merchant is then any transactional benefits that the merchant experiences less the merchant discount: $b_M - m$. If the merchant sets the same price for cash and card purchases, then the merchant's cost per transaction of card acceptance relative to cash is this difference, weighted by the share of payments that are made with cards. Assuming that a merchant accepts cards when acceptance yields a positive net benefit, the merchant will accept cards whenever $b_M - m$

relative to cash.

[125] For simplicity, we assume that the costs of card transactions to the network are negligible. As discussed in the main text, the networks typically levy small fees on banks to cover those costs.

≥ 0 or

$$b_M \geq m. ^{126} \qquad \text{(merchant acceptance criterion)}$$

As noted in section III, use of the card requires the joint participation of both merchants and consumers. Hence, both the consumer use criterion and the merchant acceptance criterion must be satisfied simultaneously for the card to be used in a transaction.

B. PRIVATE CARD USE DECISIONS WITH AN INTERCHANGE FEE OF ZERO

We begin by considering card use when the interchange fee is fixed at zero. Issuing and acquiring banks will set fees to their respective customers that depend on the respective costs of card transactions for those banks. Assuming that issuing and acquiring banks are perfectly competitive, they will set fees exactly equal to their marginal costs.[127] With a zero interchange fee, consumers will face a card use fee equal to issuer costs (i.e., $f = c_I$) while the merchant discount will equal acquirer costs (i.e., $m = c_A$.) In light of these fees, the criteria for merchant card acceptance and consumer card use given above imply that a card will be used in a transaction whenever

$$b_B \geq c_I \qquad \text{and} \qquad b_M \geq c_A.$$

However, these conditions do not coincide with the condition for socially efficient card use, $b_B + b_M \geq c_I + c_A$. In particular, otherwise efficient card transactions may not take place because either the consumer or the merchant may be unwilling to participate in a card transaction that yields a net social gain.[128]

To see the inefficiencies that could arise without an interchange fee, suppose that

---

[126] The merchant's acceptance criterion could equivalently be expressed in terms of the net cost of card transactions for merchants, $m - b_M$. Using this symmetric criterion, a merchant would accept cards as long as card transactions do not raise its costs (i.e., $m - b_M \leq 0$). As discussed in section III, a merchant's acceptance decision likely involves considerations other than just the impact of card transactions on costs, including improved quality of service for consumers and competitive considerations.

[127] The assumption of perfectly competitive banks provides a useful benchmark case in which to evaluate some basic economics of payment card use. Various authors (e.g., Rochet and Tirole 2002, 2006b) have considered the implications of market power among issuing or acquiring banks. Those studies find that an interchange fee can enhance efficiency by lowering costs for banks with market power to offset excessive mark-ups. This role of an interchange fee is not related to the fundamental role of internalizing externalities and has been criticized as a means to subsidize firms (i.e., banks) with market power to induce lower prices (Farrell 2006).

[128] In this model, with an interchange fee equal to zero, socially inefficient card transactions will not occur. The key insight, however, is that some otherwise socially efficient card transactions will not take place.

acquirers and issuers both experience costs of 0.5 from a card transaction, so that $c_I = c_A$ = 0.5 with total social costs $c_I + c_A = 1$. Suppose further that merchants experience net transactional benefits of 2 from a card transaction. [129] In this case, setting the marginal social cost equal to the marginal social benefit implies that the marginal consumer for whom card use would be efficient has a transactional benefit from the card equal to -1 (= 0.5 + 0.5 – 2). With a zero interchange fee, consumers would face a transaction fee of f = 0.5 (i.e., the issuer cost of a card transaction). As a result, consumers with card benefits below 0.5 would not use the card. Because efficient card use requires consumers with valuations above -1 to use the card, but only those with valuations greater than 0.5 would do so with a zero interchange fee, some otherwise efficient card transactions will not occur. [130]

The inefficiency in this example stems from a simple usage externality. By considering only his or her own transaction benefits and card usage fees (the latter being equal to issuer costs), the consumer does not account for the social surplus that accrues on the merchant side of the market (i.e., merchant transactional benefits less acquirer costs). Although the social surplus on the other side of the market is high enough to justify having a consumer with a low, indeed negative, benefit use the card, such a consumer would not do so because the card use fee that he or she faces is too high. [131]

C. PRIVATE CARD USE DECISIONS WITH THE EFFICIENT INTERCHANGE FEE

Although somewhat abstract and difficult to implement, theoretical analysis reveals the potential efficiency enhancing role of an interchange fee. Moreover, the analysis reveals certain features that an efficient interchange fee might exhibit. Finally, consideration of the efficient interchange fee establishes a benchmark for comparison with market-determined interchange fees.

Practically speaking, an interchange fee can alter the transaction fees facing consumers and merchants through its effect on the costs of their respective banks. In turn, this effect on transaction fees alters the incentives of consumers and merchants to

---

[129] These hypothetical costs and benefits correspond to the values presented in table 1.

[130] By assumption, the transactions would still take place, but would involve use of cash rather than the more efficient card.

[131] This example also illustrates that, due to externalities, perfect competition among banks is not sufficient to guarantee efficient card use.

use cards, thereby affecting the overall volume of card transactions. As illustrated below, an appropriate interchange fee can yield transaction fees such that the individual decisions of consumers and merchants will yield efficient card use.

Assume acquirers pay issuers an interchange fee equal to a. Then issuer costs equal $c_I - a$ and acquirer costs equal $c_A + a$.[132] Continuing to assume that issuers and acquirers are perfectly competitive, consumers then face card fees equal to $f = c_I - a$, while the merchant discount equals $m = c_A + a$. The criteria for consumer card use and merchant acceptance imply that the card will be used to perform a transaction whenever

$$b_B \geq c_I - a \qquad \text{and} \qquad b_M \geq c_A + a.$$

Suppose that the interchange fee is set equal to merchant transactional benefits less acquirer costs so that

$$a^* = b_M - c_A.\text{[133]}$$

At this interchange fee, the merchant discount would equal

$$m = c_A + (b_M - c_A) = b_M$$

while the card user fee would equal

$$f = c_I - (b_M - c_A) = c_I + c_A - b_M.$$

Faced with these fees, merchants would continue to accept cards because the merchant discount would exactly equal the merchant's transactional benefit from cards.[134] Substituting the card user fee into the consumer use criterion, a consumer would now choose to use the card whenever

$$b_B \geq c_I + c_A - b_M$$

or

$$b_B + b_M \geq c_I + c_A.$$

In other words, when faced with the card use fee associated with the interchange fee a^*, a *consumer will use the card whenever it is socially efficient to do so.*

Continuing the earlier numerical example, because merchant transactional

---

[132] In theory, a could be positive (involving a payment from acquirers to issuers) or negative (involving a payment from issuers to acquirers). In the previous subsection, a implicitly was present, but was set equal to zero.

[133] This interchange fee is sometimes (e.g., Rochet 2003) called the "Baxter fee" reflecting Baxter's (1983) early analysis of interchange fees. In this simple model, a^* would also be the interchange fee that exactly satisfies the "tourist test" discussed in section VII.

[134] Formally, the merchant would be indifferent between accepting and rejecting cards.

benefits are 2 while acquirer costs are 0.5, the efficient interchange fee is $a* = 1.5 (= 2 - 0.5)$. The card use fee for consumers, including both issuer costs and the interchange fee, equals -1 $(= 0.5 - 1.5)$. Faced with this fee, all consumers with transactional benefits greater than -1 would use the card as required for efficient card use.

As noted above, the general difficulty with obtaining efficient card use occurs because each party does not account for the social surplus (i.e., social benefits less social costs) on the other side of the market. The efficient interchange fee $a* = b_M - c_A$ in this simple model exactly transmits the social surplus (i.e., social benefits less costs) on the merchant side of the market to the consumer by lowering the card use fee. As a result, when deciding whether to use the card, the consumer internalizes the social surplus that card use generates on the merchant side of the market.

While simple, this model illustrates a number of results common to more complex models of interchange fees. First, the interchange fee alters the transaction fees facing the merchant and the consumer, thereby influencing their respective incentives to accept or use the card and the subsequent volume of card transactions. Second, the efficient interchange fee transfers social benefits and costs generated on one side of the market (e.g., the merchant) to the other side of the market (e.g., the consumer). In doing so, the efficient interchange fee induces the latter group to base decisions on social costs and benefits rather than private costs and benefits. Third, the efficient interchange fee is not generally cost-based, nor is it generally zero. In particular, the efficient interchange fee is not generally related to issuer costs.[135] Fourth, even if the costs associated with card transactions are small, a positive interchange fee may still be necessary to yield transaction fees that promote efficient use of the card.[136] Finally, the efficient

---

[135] Interchange fees are often justified as a way to reimburse issuers for the costs they incur in supplying card transactions. In general, this view of interchange fees is incorrect, because an issuer could always recover its costs through an appropriate fee to card users. A correct statement of this view of interchange fees is that the efficient interchange fee allows issuers to cover the costs of the efficient transaction fee for consumers (i.e., the fee that induces efficient card use by consumers). For example, in the numerical example, the efficient transaction fee for consumers is $f = -1$. An issuer facing production costs of $c_I = 0.5$ would make negative profits at the efficient transaction fee without payment of the interchange fee greater than or equal to $a* = 1.5$.

[136] Indeed, if issuer and acquirer costs are zero, a higher interchange fee would be necessary to induce appropriate consumer participation in this simple model. This result reflects the fact that, when costs are extremely low, card transactions are so (socially) attractive relative to cash that it is necessary to provide even stronger incentives for consumers to use the card.

interchange fee depends on underlying social costs and benefits that are difficult, if not impossible, to measure.

A few features of the socially efficient transaction fees (i.e., the card user fee and the merchant discount) that result from the efficient interchange fee are also worth noting. First, the socially efficient transaction fee faced by a party does not generally equal the costs incurred by that party's bank. As the previous section illustrated, efficient merchant fees should not generally equal costs of acquiring banks, nor should card user fees equal costs of issuers. Second, the socially efficient transaction fees may be unbalanced, with one group paying larger fees than the other. In fact, as in the numerical example, the socially efficient structure of transaction fees may involve one side of the market paying a negative fee (i.e., receiving rewards) for card transactions. Third, inframarginal card users (i.e., those with transactional benefits greater than the transaction fee) may derive substantial surplus from card transactions, but efficiency is concerned with providing appropriate incentives for the marginal card user.

D. CARD USE WITH DIFFERENTIAL PRICING ACROSS PAYMENT METHODS

The previous discussion of the efficient interchange fee was based on the assumption that merchants do not set different prices across payment methods. If this assumption is relaxed, differential pricing can, in certain circumstances, yield efficient card use, regardless of the level of the interchange fee.[137] This result broadly reflects the standard economic prescription for dealing with externalities: Adjust prices so that agents' incentives reflect the external effects of their actions. In the case of a payment card, differential pricing can achieve this objective through a retail price for card transactions (distinct from that of cash) that reflects costs and benefits of cards for the merchant.

To illustrate how differential pricing can yield efficient card use, suppose that merchants can costlessly set two prices, p^{cash} and p^{card}. Suppose further that the

---

[137] Carlton and Frankel (1995) were the first to make this point. Rochet and Tirole (2002) and Wright (2003) consider the effects of differential pricing in more detail and construct various situations in which differential pricing will not yield efficient card use. They find that differential pricing when merchants or banks have market power can lead to inefficient card use because price differences will not reflect differences in underlying social costs and benefits in the presence of market power.

interchange fee is initially set to zero, but that the acquirer may still set a merchant

discount. Because the merchant discount alters the merchant's cost of card transactions,

a merchant's card price will depend on the merchant discount. Moreover, competition

among merchants should lead a merchant to incorporate any transactional savings (i.e.,

benefits) from card transactions in the card price. Thus, in the simplest case, the

merchant will set $p^{card} = p^{cash} + m - b_M$. As before, perfectly competitive acquirers will

set a merchant discount equal to $m = c_A$ implying a card price of $p^{card} = p^{cash} + c_A - b_M$

and a price differential of $p^{card} - p^{cash} = c_A - b_M$.[138]

    A consumer deciding how to pay will now consider any price difference across

payment methods in addition to his or her transactional benefits and fees from card use.

In particular, the consumer will use the card whenever the total card price, including

consumer fees and transactional benefits, is less than the cash price:

$$p^{card} + f - b_B \le p^{cash},$$

or

$$b_B \ge p^{card} - p^{cash} + f.$$

Assuming, as before, that issuing banks are perfectly competitive so that $f = c_I$ and

recalling that the price differential for merchants is $p^{card} - p^{cash} = c_A - b_M$, the condition

for consumer card use becomes

$$b_B \ge c_A - b_M + c_I$$

or, rearranging terms,

$$b_B + b_M \ge c_A + c_I$$

which is the condition for efficient card use.

    This simple example illustrates that differential pricing by merchants can lead to

efficient card use, even when the interchange fee is set to zero. The key insight is that a

merchant can set a price differential that reflects costs and benefits on the merchant side

of the market. This price differential can then transmit information about those costs and

benefits to the incentives facing a consumer. Given those incentives, the consumer's

payment choice would then depend on his or her own costs and benefits as well as those

---

[138] In theory, this difference could be positive or negative. Indeed, if merchant benefits exceed acquirer costs, as would likely be the case when acquirer costs are small, then the card price in the case of a zero interchange fee would be less than the cash price .

of the merchant.

A further interesting implication of differential pricing in this example is that introduction of a non-zero interchange fee would have no effect on card use. To illustrate this result, suppose that a non-zero interchange fee a is introduced and, as before, issuers and acquirers set fees equal to their costs, including costs introduced by the interchange fee. A merchant would still set the card price to reflect the merchant discount including the non-zero value of the interchange fee, so that

$$p^{card} = p^{cash} + c_A + a - b_M.$$

When choosing how to pay, a consumer would continue to use the card whenever the card price, adjusted for any fees or consumer benefits, is less than the cash price,

$$p^{card} + f - b_B \leq p^{cash}$$

or

$$b_B \geq p^{card} - p^{cash} + f.$$

Now, however, the consumer's transaction fee would reflect the non-zero interchange fee paid to the issuer. Thus, the consumer would choose to use the card whenever

$$b_B \geq c_A + a - b_M + c_I - a$$

or, rearranging terms,

$$b_B + b_M \geq c_A + c_I$$

which is again the condition for efficient card use.

In this example, the interchange fee does not enter the consumer's criterion for card use and, as a result, has no real impact on card use (that is, the interchange fee is "neutral" to the card use decision). Instead, any impact of the interchange fee on the card user fee f is exactly offset by its effect on the merchant discount and the resulting price for card transactions. Thus, in this case, the interchange fee is ineffective as a means to influence card use. While this model provides a particularly stark illustration of this result, it reflects a broader result that the effect of an interchange fee on card use diminishes under differential pricing.[139]

---

[139] Rochet and Tirole (2002) and Wright (2003) illustrate this result under various assumptions. When each card transaction involves fixed fees, Gans and King (2003) obtain results about the neutrality of interchange fees in a much more general setting.

Appendix III: Entry in the Payment Card Market

As described in section V, some firms have attempted to introduce new payment card products. This appendix describes several of these innovations and explains how each relates to existing payment card systems.

The first of these payment card innovations aimed at reducing merchant cost was the deployment of "decoupled debit" by Tempo, a merchant-owned and -operated payment card network.[140,141] Under this model, merchants serve as both issuers and acquirers. To obtain a card, a customer provides his or her bank routing number and account number along with other information to a participating Tempo merchant issuer. This information is then encoded onto a Tempo card. When this card is presented at the point of sale, merchants that accept Tempo cards originate ACH debits against the customer's deposit account. Each member merchant agrees to accept all Tempo-branded cards, and may choose to issue the card as well.

Tempo sets its own interchange fees for transactions between its issuing and acquiring merchants, currently on the order of five to fifteen cents per transaction. The reason for this low fee is that the network is essentially piggybacking on the ACH system, which charges less than one cent per transaction.

Revolution Card, a three-party credit card system, is also trying to challenge the existing networks by offering low merchant fees.[142] The concept behind this credit card system is to encourage merchant acceptance by offering very low interchange fees (0.5 percent, far below the prevailing average of nearly 2.0 percent) and to generate cardholder use by offering enhanced security and rewards.[143]

To date, both Tempo and Revolution have signed up merchants, but have limited card deployment. If successful, the Tempo model has the potential to put downward pressure on interchange fees, because the ultimate recipients of Tempo interchange fees

---

[140] The card is "decoupled" from a customer's deposit account because the customer has a debit card issued by an entity other than the institution at which he or she has a deposit account.

[141] This network was formerly called Debitman. Earlier implementations of this model have also been attempted by individual merchants – primarily supermarket chains – but due to the presence of network effects, the current network model is likely to be more viable than single-merchant initiatives.

[142] This card, created by Steve Case of AOL, was originally launched as the Gratis Card.

[143] The card may also be used as a prepaid card, for which the customer pre-loads a balance in advance of any purchases.

are the merchants themselves. It is not obvious, however, that the Revolution model would have the same potential, given that its business model is not clearly different from that of the existing three-party networks.

Other entrants into the payment card system have adopted a strategy of building on the existing network structures. These include the decoupled debit product from Capital One Financial and many of the online and mobile payment initiatives. Capital One's decoupled debit product, launched in a pilot program that ran from mid-2007 to mid-2008, was accepted like any other MasterCard debit card, but used the ACH system to withdraw funds from a cardholder's deposit account at a bank other than Capital One. Capital One received the MasterCard-set interchange fee from the merchant's acquirer and paid the ACH transaction fee (as mentioned above, typically less than one cent per transaction) itself. As such, the Capital One model was designed to compete with other debit card issuers, but, unlike Tempo, not with the major credit or debit card networks, and therefore would be unlikely to place downward pressure on interchange fees.[144]

---

[144] Many other attempts to enter the payments market for online transactions and mobile payments also tie into the existing card networks and the ACH system. The more established players are PayPal and Google Checkout, systems that tie into existing debit and credit card networks in order to process online purchase payments. Revolution also offers an Internet payment product for online purchases. Similarly, several companies now offer mobile payment systems in which consumers can make payments for purchases using their mobile phones. Like the online products, most of the mobile products rely on connections to the established credit and debit card networks in order to conduct transactions.

References

American Bankers Association/Dove Consulting (2006) *Study of Consumer Payment Preferences 2005/2006*.

Aspan, Maria (2008) "Visa to Cut Interchange for Gas Purchases," *American Banker*, June 27.

ATM and Debit News (various years) "EFT Data Book."

Australian Competition and Consumer Commission and Reserve Bank of Australia (2000) *Debit and Credit Card Schemes in Australia: A Study of Interchange Fees and Access*.

Bauer, Paul W. and Daniel A. Littman (2007) "Are Consumers Cashing Out?," Federal Reserve Bank of Cleveland, *Economic Commentary* (October 1).

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

Berry, Attilla (2006) "As IPO Looms, Visa's Outside Counsel See Litigation Bonanza," *Legal Times*, March 7.

Bradford, Terry and Fumiko Hayashi (2008) "Developments in Interchange Fees in the United States and Abroad," Federal Reserve Bank of Kansas City, Payment System Research Briefing.

Breitkopf, David (2007) "Exclusive Network Pacts for PIN Debit Gain Favor," *American Banker*, September 28.

Carlton, Dennis W. and Alan S. Frankel (1995) "The Antitrust Economics of Credit Card Networks," *Antitrust Law Journal*, 63:2, pp. 643-668.

Chakravorti, Sujit and Roberto Roson (2006) "Platform Competition in Two-Sided Markets: The Case of Payment Networks," *Review of Network Economics*, 5:1, pp. 118-142.

Cohen, Greg (2008) "Where Does the Money Go? Analyzing Merchant Discount Rates and Fees," *Transaction World* (February), pp. 15-16.

Costanzo, Chris (2008) "Acquirers Open Doors in Card Acceptance," *American Banker*, May 6.

European Commission (2002) "Commission exempts multilateral interchange fees for

cross-border Visa card payments," IP/02/1138.

European Commission (2007) "Antitrust: Commission prohibits MasterCard's intra-EEA Multilateral Interchange Fees," IP/07/1959.

European Commission (2008a) "Antitrust: Commission initiates formal proceedings against Visa Europe Limited," MEMO/08/170.

European Commission (2008b) "Antitrust: Commission notes MasterCard's decision to temporarily repeal its cross-border Multilateral Interchange Fees within the EEA," MEMO/08/397.

European Commission (2009a) "Antitrust: Commissioner Kroes takes note of MasterCard's decision to cut cross-border Multilateral Interchange Fees (MIFs) and to repeal recent scheme fee increases," IP/09/515.

European Commission (2009b) "Antitrust: Commissioner Kroes takes note of MasterCard's decision to cut cross-border Multilateral Interchange Fees (MIFs) and to repeal recent scheme fee increases – frequently asked questions," MEMO/09/143.

Evans, David and Richard Schmalensee (2005) "The Economics of Interchange Fees and Their Regulation: An Overview," MIT Sloan Working Paper No. 4548-05.

Farrell, Joseph (2006) "Efficiency and Competition Between Payment Instruments," *Review of Network Economics*, 5:1, pp. 26-44.

Frankel, Alan S. and Allan L. Shampine (2006) "The Economic Effects of Interchange Fees," *Antitrust Law Journal*, 73:3 , pp. 627-673.

Gans, Joshua and Stephen P. King (2003) "The Neutrality of Interchange Fees in Payment Systems," *BEP Press, Topics in Economic Analysis and Policy*, 3, Article 1.

Gerdes, Geoffrey R., Jack K. Walton, May. X Liu, and Darrel W. Parke (2005) "Trends in the Use of Payment Instruments in the United States," *Federal Reserve Bulletin*, 91, pp. 180-201.

Gerdes, Geoffrey R. (2008) "Recent Payment Trends in the United States," *Federal Reserve Bulletin*, 94, pp. A75-A106.

Guthrie, Graeme and Julian Wright (2007) "Competing Payment Schemes," *Journal of Industrial Economics*, 55:1, pp. 37-67.

Hayashi, Fumiko (2006) "A Puzzle of Card Payment Pricing: Why Are Merchants Still Accepting Card Payments?" *Review of Network Economics*, 5:1, pp. 144-174.

Hayashi, Fumiko and Stuart E. Weiner (2006) "Interchange Fees in Australia, the UK, and the United States: Matching Theory and Practice," *Federal Reserve Bank of Kansas City Economic Review*, 3rd Quarter, pp. 75-112.

Hermalin, Benjamin E. and Michael L. Katz (2006) "Your Network or Mine? The Economics of Routing Rules," *Rand Journal of Economics*, 37:3, pp. 692-719.

HSN Consultants, Inc. (various months) *The Nilson Report*.

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

Katz, Michael L. (2005) "Commentary: What Do We Know About Interchange Fees and What Does It Mean for Public Policy," in *Interchange Fees in Credit and Debit Card Industries: What Role for Public Authorities*, Federal Reserve Bank of Kansas City.

Klee, Elizabeth (2006) "Families' Use of Payment Instruments During a Decade of Change in the U.S. Payment System," Federal Reserve Board Finance and Economics Discussion Series, 2006-01.

Klein, Benjamin, Andres V. Lerner, Kevin Murphy, and Lacey L. Plache (2006) "Competition in Two-Sided Markets: The Antitrust Economics of Payment Card Interchange Fees," *Antitrust Law Journal*, 73:3, pp. 571-626.

Kjos, Ann (2007) "The Merchant-Acquiring Side of the Payment Card Industry: Structure, Operations, and Challenges," Payment Cards Center Discussion Paper, Federal Reserve Bank of Philadelphia.

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

McAndrews, James (2003) "Automated Teller Machine Pricing – A Review of the Literature," *Review of Network Economics*, 2:2, pp. 146-158.

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

Musante, Ken (2008) "The Pinch of PIN Debit," *The Green Sheet*, April 14, 2008.

Organization for Economic Cooperation and Development (2006) "Competition and Increased Efficiency for Retail Payment Cards," DAF/COMP/WP2(2006)3

Quittner, Jeremy (2006) "Rewards Programs Feeling Interchange Rate Pressure,"

American Banker, May 3.

Reserve Bank of Australia (2005) "Common Benchmark for the Setting of Credit Card Interchange Fees."

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

Reserve Bank of Australia (2008a) "Reform of Australia's Payments System: Preliminary Conclusions of the 2007/08 Review."

Reserve Bank of Australia (2008b) "Reform of Australia's Payments System: Conclusions of the 2007/08 Review."

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

Rochet, Jean-Charles and Jean Tirole (2002) "Cooperation Among Competitors: Some Economics of Payment Card Associations," *RAND Journal of Economics*, 33:4, pp. 549-570.

Rochet, Jean-Charles and Jean Tirole (2006a) "Two-Sided Markets: A Progress Report," *RAND Journal of Economics*, 37:3, pp. 645-667.

Rochet, Jean-Charles and Jean Tirole (2006b) "Must Take Cards and the Tourist Test," IDEI, University of Toulouse, mimeo.

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

Rysman, Marc (2007) "An Empirical Analysis of Payment Card Usage," *Journal of Industrial Economics*, 55:1, pp. 1-36.

Schwartz, Marius and Daniel R. Vincent (2006) "The No Surcharge Rule and Card User Rebates," *Review of Network Economics*, 5:1, pp. 72-102.

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

Shinder, Jeffrey I. (2009) "Viewpoint: Interchange Facing an Unsteady Future," *American Banker*, April 8.

Sidel, Robin (2009) "Debit-Card Use Overtakes Credit," *Wall Street Journal*, May 1.

Small, John and Julian Wright (2002) "The Bilateral Negotiation of Interchange Fees in Payment Schemes," working paper, University of Auckland.

Stavins, Joanna (2001) "Effect of Consumer Characteristics on the Use of Payment Instruments," *New England Economic Review*, Issue 3.

Wang, Zhu (2008) "Market Structure and Credit Card Pricing: What Drives the Interchange?," Payments System Research Working Paper, Federal Reserve Bank of Kansas City.

Wildfang, K. Craig and Ryan W. Marth (2006) "The Persistence of Antitrust Controversy and Litigation in Credit Card Networks," *Antitrust Law Journal*, 73:3, pp. 675-707.

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

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

Zinman, Jonathan (2008) "Debit or Credit?" *Journal of Banking and Finance*, forthcoming.

**Figure 1: A Typical Transaction in a Four Party Card System**
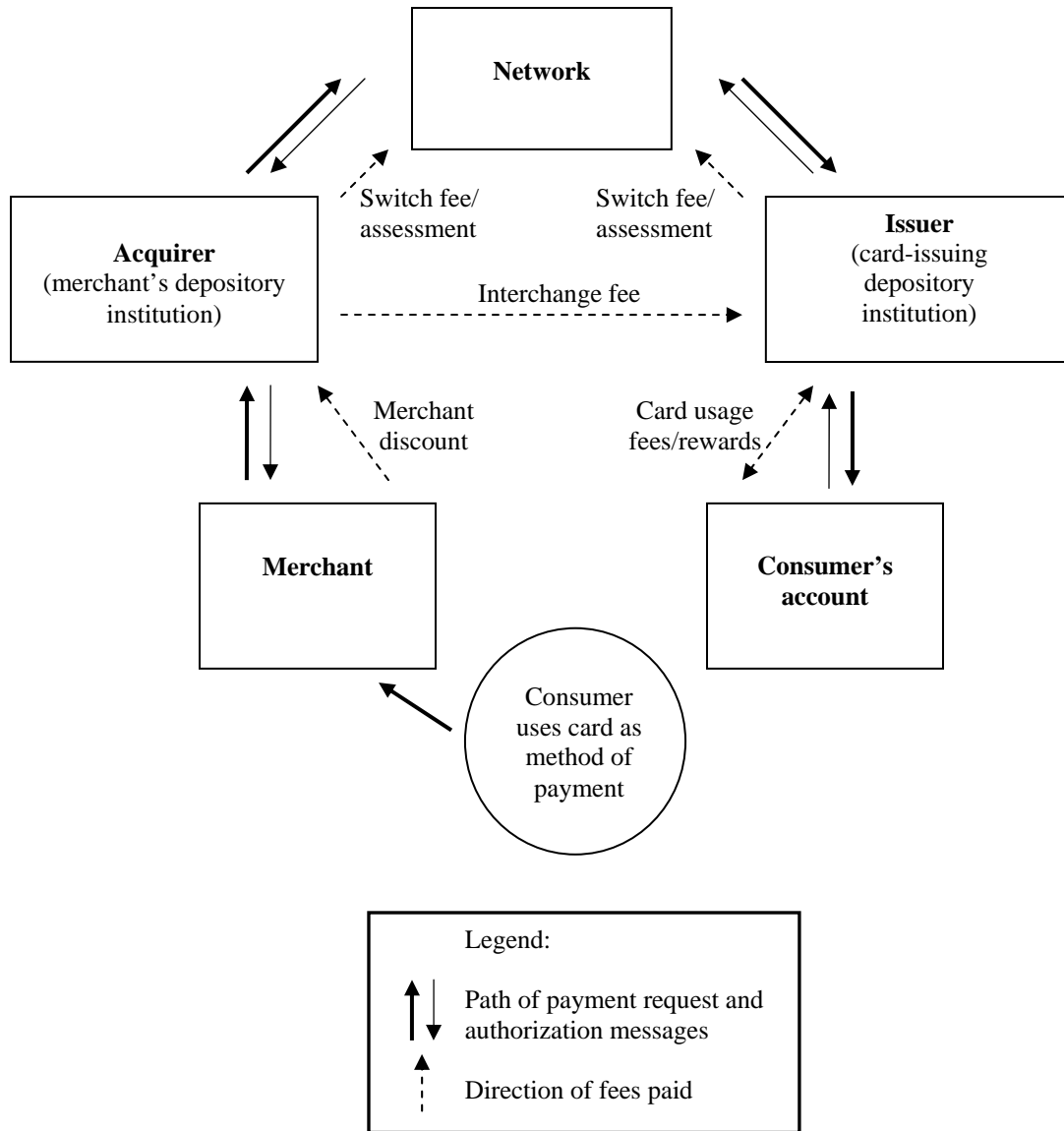
Figure 2: Total Number of U.S. Purchase Transactions by Transaction Type
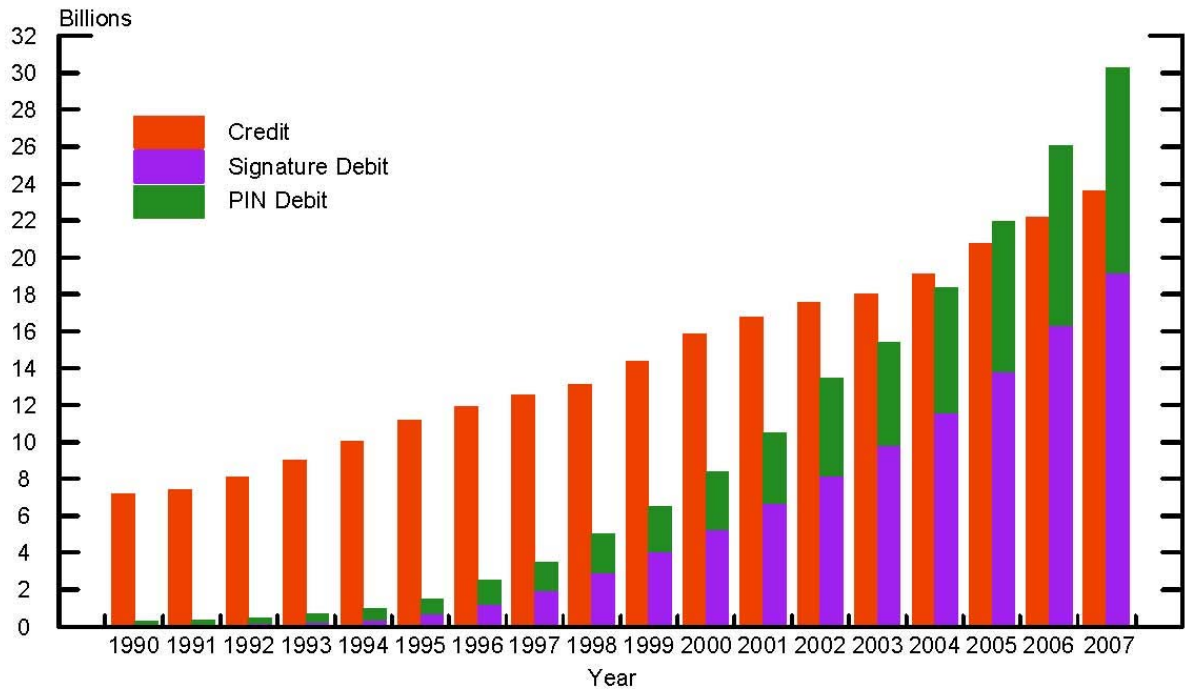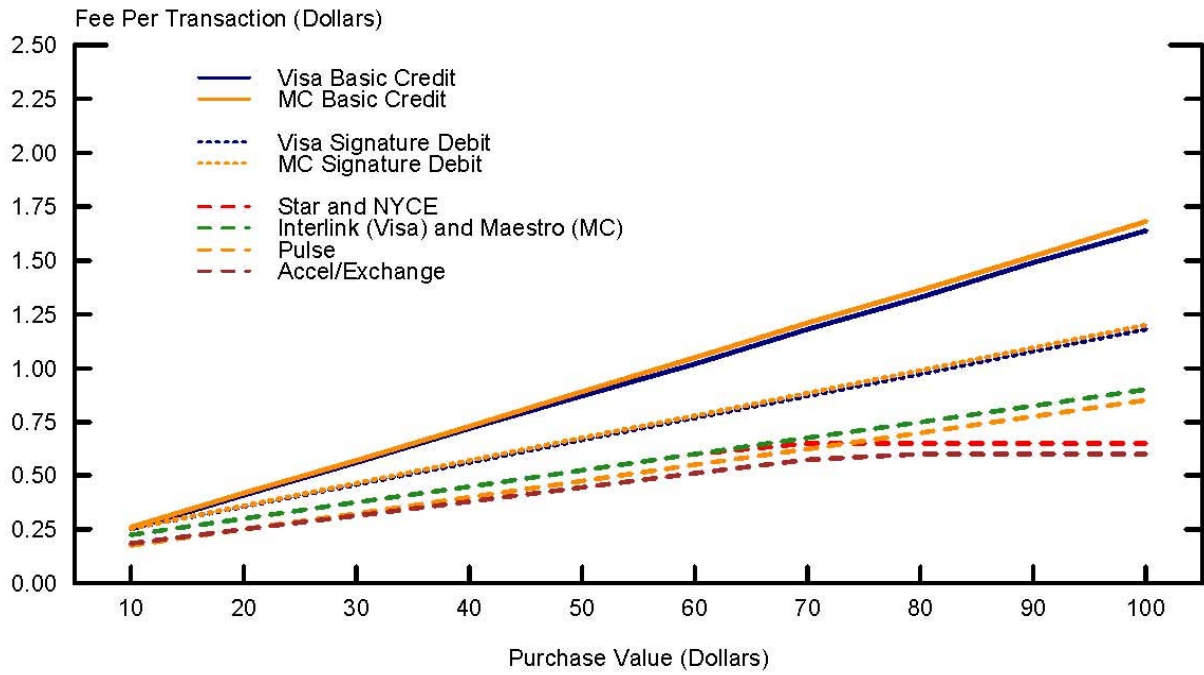

Source: Authors' calculations based on the ATM and Debit News EFT Data Book, The Nilson Report, and other industry sources.
Note: Credit includes charge cards and private-label credit cards issued by retailers.

Figure 3: Interchange Fees by Network and Transaction Type
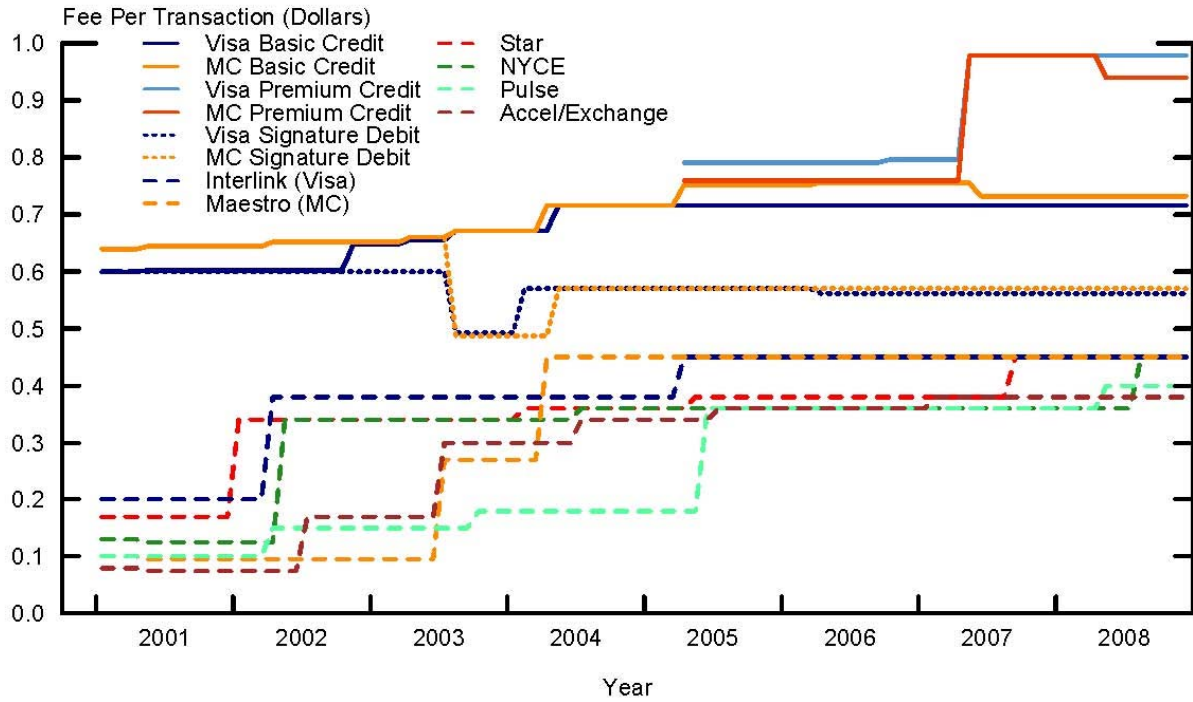
Source: Networks' published rate schedules, ATM and Debit News EFT Data Book, and industry sources.
Note: Fees displayed assume a $40 purchase transaction at a typical merchant. Fees are reported for each month in each year. Solid lines denote credit cards. Dotted lines denote signature debit cards. Dashed lines denote PIN debit cards

Figure 4: Interchange Fees for Different Purchase Values by Network and Transaction Type (December 2008)

Fee Per Transaction (Dollars)

Legend:
- Visa Basic Credit
- MC Basic Credit
- Visa Signature Debit
- MC Signature Debit
- Star and NYCE
- Interlink (Visa) and Maestro (MC)
- Pulse
- Accel/Exchange

Purchase Value (Dollars)

Source: Networks' published rate schedules, ATM and Debit News EFT Data Book, and industry sources.
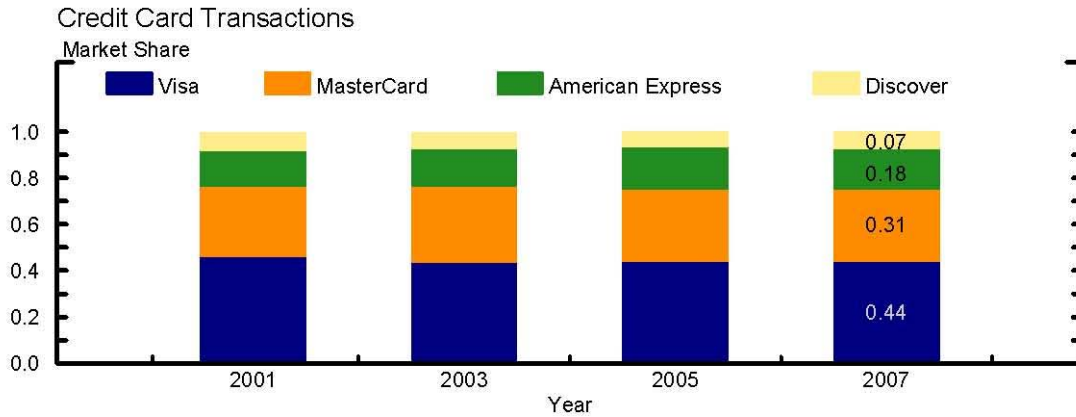Note: Fees displayed assume a card-present transaction at a moderate-sized general purpose retailer.
Solid line denote credit cards. Dotted lines denote signature debit cards. Dashed lines denote PIN debit cards.

Figure 5: Credit Card Interchange Fees for Different Purchase Values by Merchant Type and Card Program (December 2008)
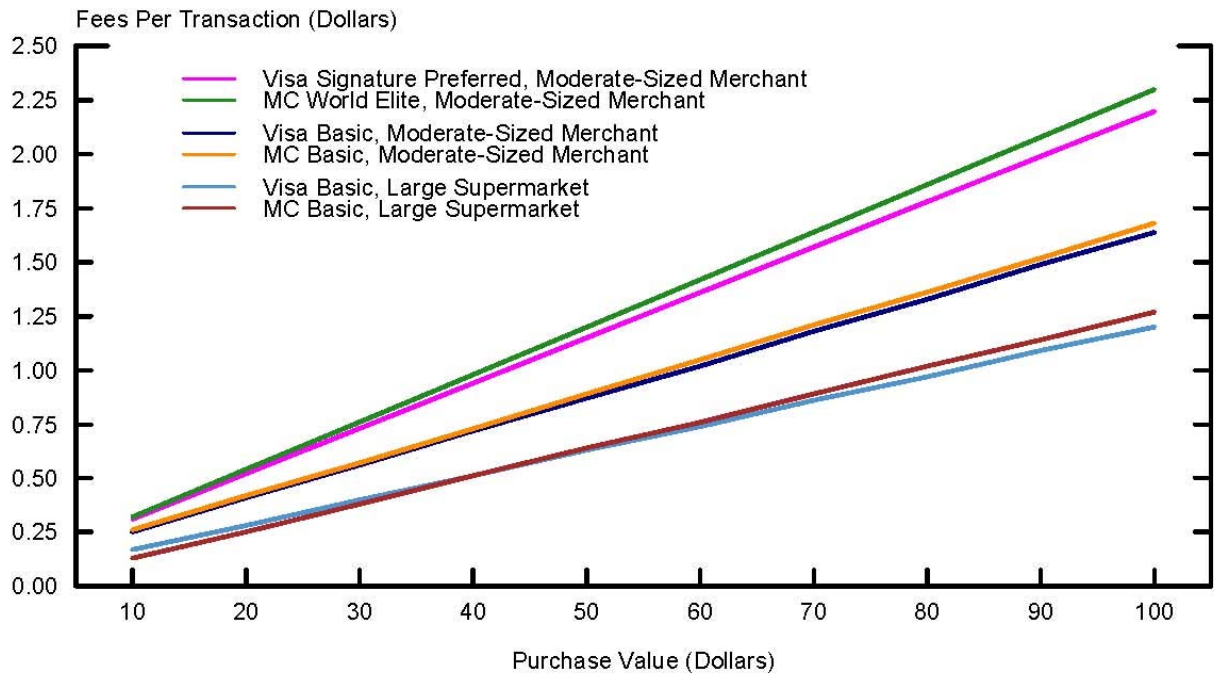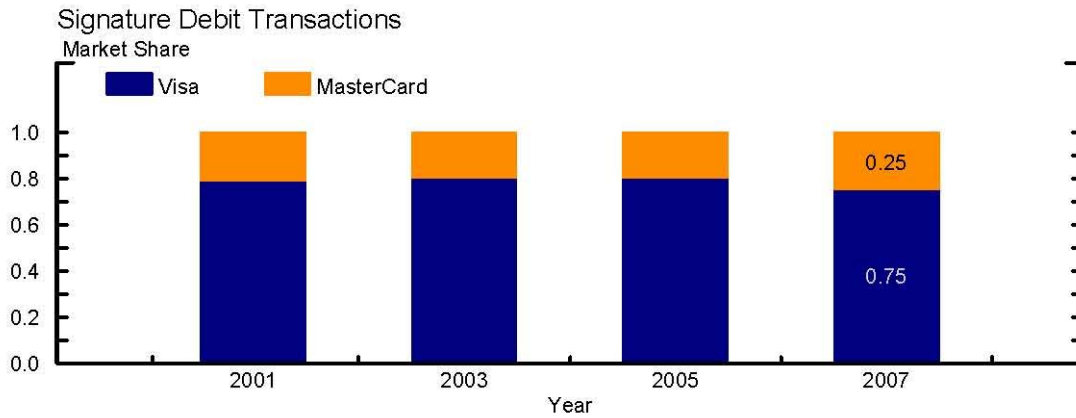

Source: Networks' published rate schedules.

Figure 6: Network Market Shares (Transactions)

Credit Card Transactions

Market Share

| Visa | MasterCard | American Express | Discover |

1.0				
0.8				
0.6				
0.4				
0.2				
0.0				

2001   2003   2005   2007
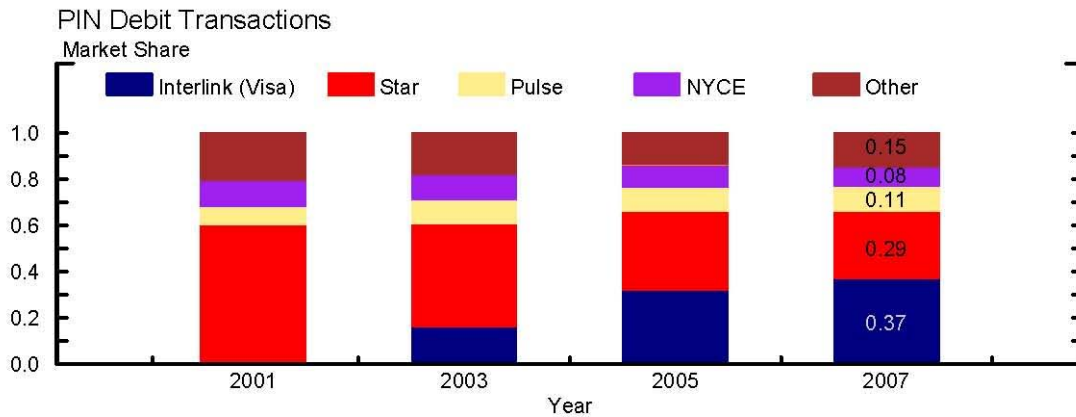
Year

0.07
0.18
0.31
0.44

Source: Authors' calculations from the Nilson Report.
Note: Excludes private-label credit cards issued by retailers. Includes charge cards. Shares for American Express and Discover include transactions performed with cards issued by third-party issuers.

Signature Debit Transactions

Market Share

| Visa | MasterCard |

1.0	
0.8	
0.6	
0.4	
0.2	
0.0	

2001   2003   2005   2007

Year

0.25
0.75

Source: Authors' calculations from the Nilson Report.

PIN Debit Transactions

Market Share

| Interlink (Visa) | Star | Pulse | NYCE | Other |

1.0				
0.8				
0.6				
0.4				
0.2				
0.0				

2001   2003   2005   2007

Year

0.15
0.08
0.11
0.29
0.37

Source: Authors' calculations from the Nilson Report and other industry sources.

78

All Debit Transactions

Market Share

Legend: Visa ■ | MasterCard ■ | Star ■ | Pulse ■ | Other ■

Values for 2007: 0.08, 0.04, 0.11, 0.16, 0.61

Source: Authors' calculations from the Nilson Report and other industry sources.
Note: Shares for Visa and MasterCard include both signature and PIN debit systems.
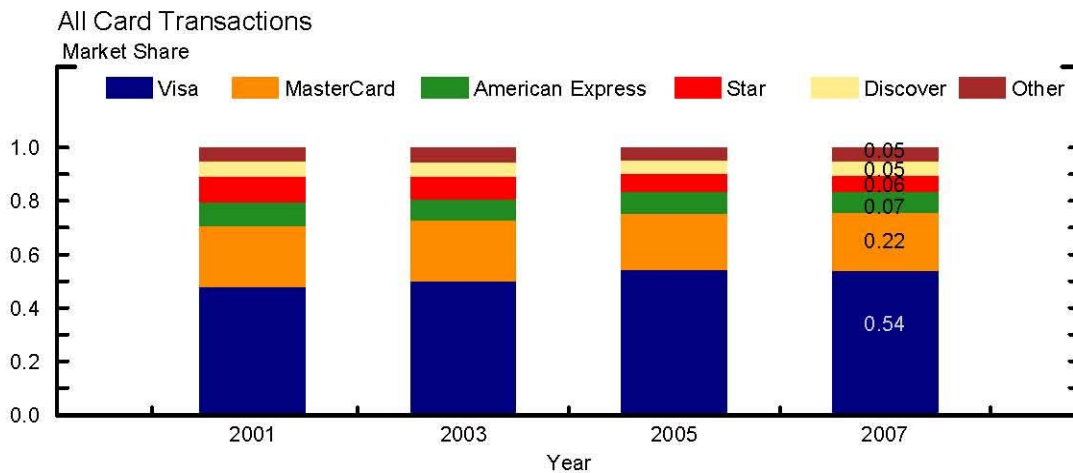
All Card Transactions

Market Share

Legend: Visa ■ | MasterCard ■ | American Express ■ | Star ■ | Discover ■ | Other ■

Values for 2007: 0.05, 0.05, 0.06, 0.07, 0.22, 0.54

Source: Authors' calculations from the Nilson Report and other industry sources.
Note: Shares for Visa and MasterCard include credit, signature debit, and PIN debit systems. Shares for Discover include credit and the PIN debit network Pulse, which Discover has owned since 2005. Includes charge cards. Excludes private-label credit cards issued by retailers. Shares for American Express and Discover include transactions performed wih cards issued by third-party issuers.

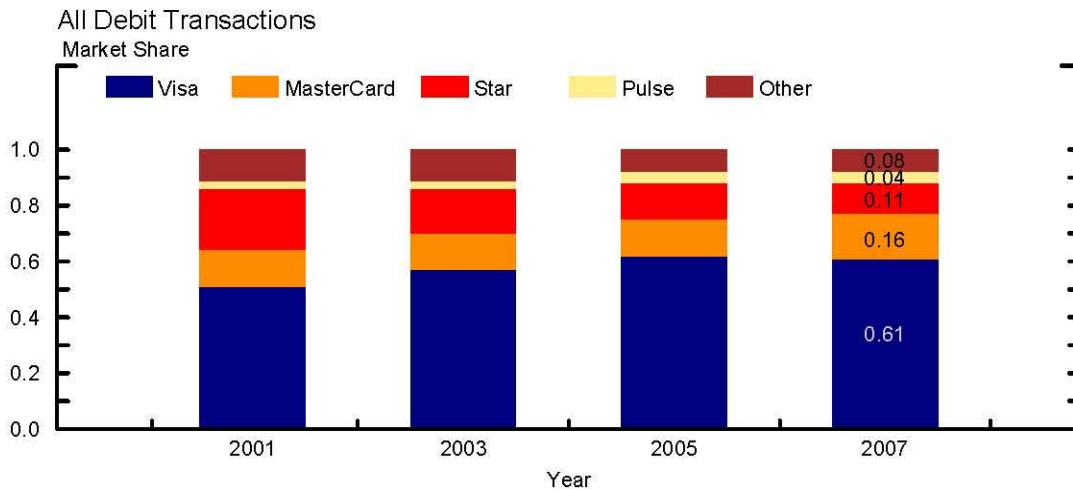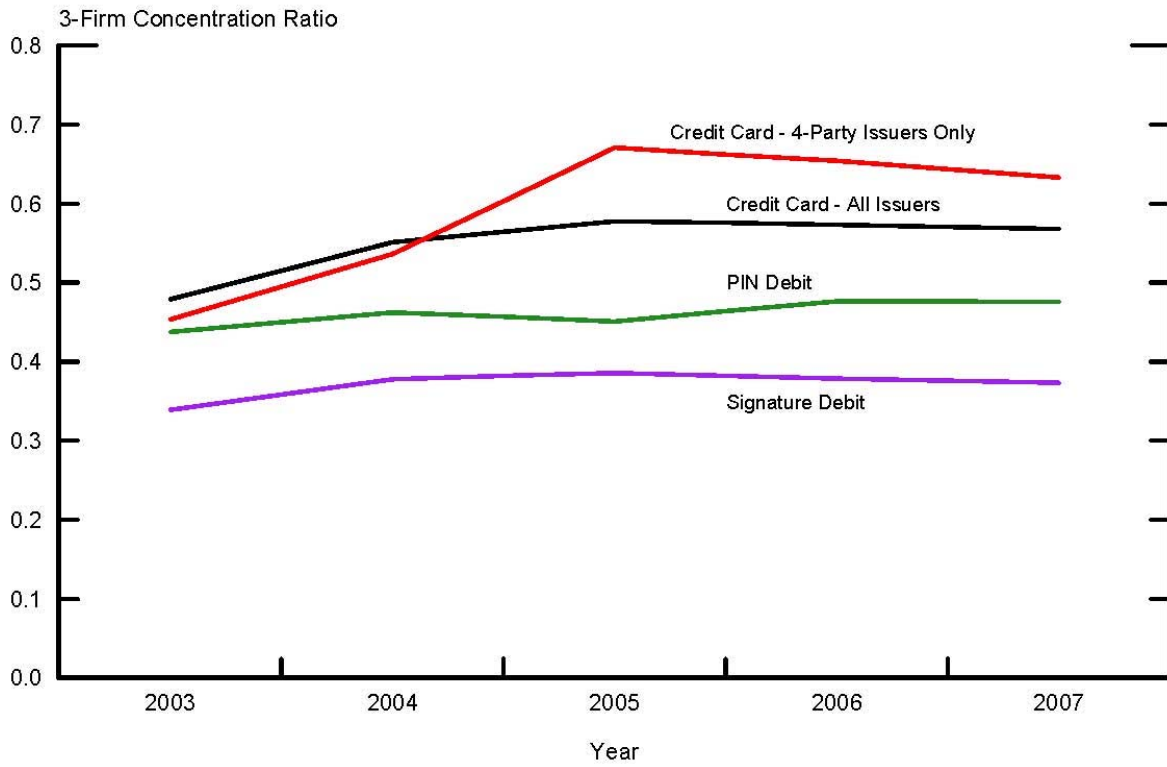**Figure 7: Issuer Concentration (Dollar Value)**

3-Firm Concentration Ratio

Credit Card - 4-Party Issuers Only

Credit Card - All Issuers

PIN Debit

Signature Debit

Year

Source: Authors' calculations from data in the Nilson Report.

Figure 8: Acquirer Concentration (Transactions)
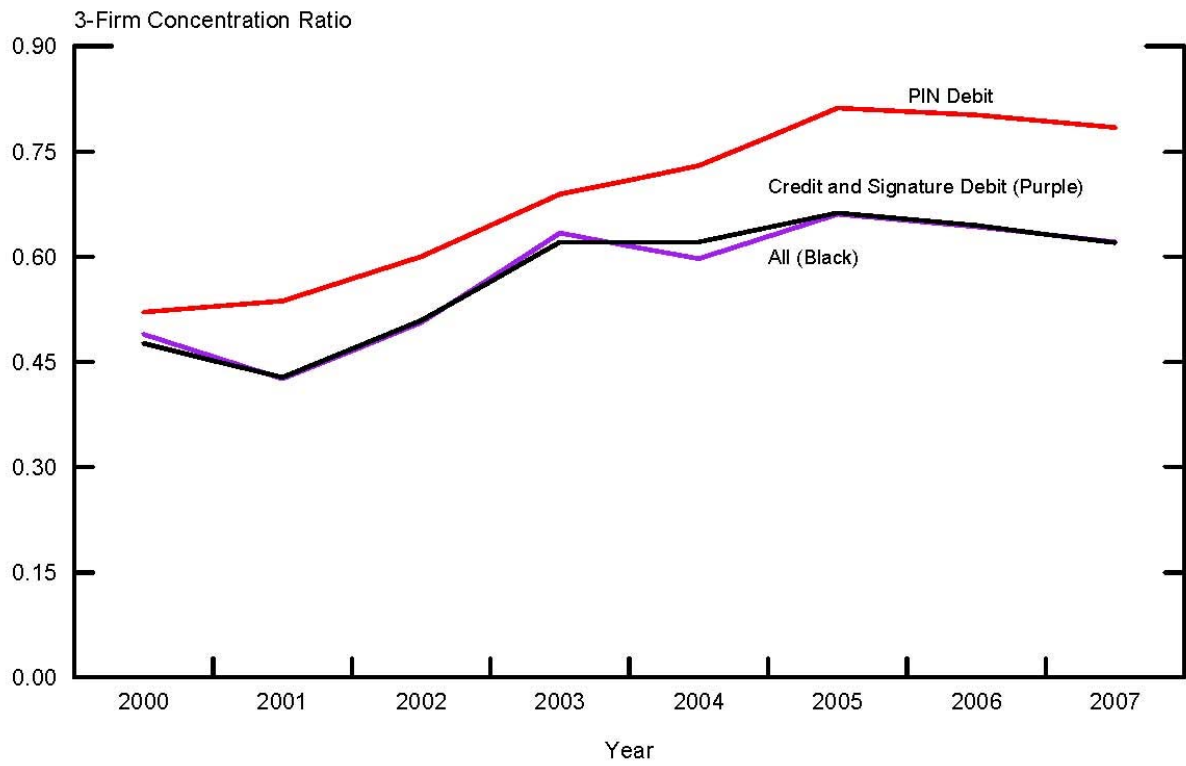
Source: Authors' calculations from the Nilson Report.
Note: For joint ventures, calculations assume transactions are shared in proportion to ownership share.

Table 1: An example of a transfer payment that induces an efficient transaction to take place

Costs and benefits from a hypothetical transaction	No transfer payments	With transfer payments of… $1.50 from merchant $0.50 to issuer $0.50 to acquirer $0.50 to consumer
Acquirer benefit	$0.00	$0.00
Less acquirer cost	-$0.50	-$0.50
Plus transfer	$0.00	$0.50
Acquirer surplus	**-$0.50**	**$0.00**
Issuer benefit	$0.00	$0.00
Less issuer cost	-$0.50	-$0.50
Plus transfer	$0.00	$0.50
Issuer surplus	**-$0.50**	**$0.00**
Merchant benefit	$2.25	$2.25
Less merchant cost	-$0.25	-$0.25
Less transfer	$0.00	-$1.50
Merchant surplus	**-$2.00**	**$0.50**
Consumer benefit	$0.50	$0.50
Less consumer cost	-$0.75	-$0.75
Plus transfer	$0.00	$0.50
Consumer surplus	**-$0.25**	**$0.25**
Social surplus	**$0.75**	**$0.75**
Private decision outcomes	**No transaction occurs; issuer, acquirer and consumer do not participate**	**Transaction occurs; all parties agree to participate**

Table 2: Number and value of transactions for select noncash payment methods in various years

	Number of transactions (billions)				Value of transactions ($ billions)[1]			
	2000	2003	2006	CAGR[2]	2000	2003	2006	CAGR[2]
Payment cards								
Credit	15.6	19.0	21.7	5.7	1,277	1,694	2,124	8.8
General-purpose[3]	12.3	15.2	19.0	7.5	1,073	1,410	1,870	9.7
Private-label[4]	3.3	3.8	2.8	-2.9	205	284	254	3.6
Debit	8.3	15.6	25.3	20.5	348	631	986	18.9
Signature	5.3	10.3	16.0	20.3	210	427	637	20.3
PIN	3.0	5.3	9.4	20.8	138	204	349	16.7
All	23.9	34.6	47.1	12.0	1,625	2,324	3,110	11.4
Select other noncash payment methods								
Check[5]	41.9	37.3	30.5	-5.1	39,796	41,147	41,600	0.7
Automated clearinghouse (ACH)[6]	6.1	8.8	14.6	15.7	18,237	24,105	31,017	9.3

Source: Gerdes, Walton, Liu and Parke (2005), Gerdes (2008).

Note: [1] Nominal value. [2] Compound annual growth rate (percent), 2000 - 2006. [3] Includes four widely accepted credit and charge card networks. [4] Includes private-label credit cards issued by retailers. [5] Checks paid, that is, checks that were on-us (involving only one institution) and checks processed through the interbank check-clearing system. See Gerdes (2008) for additional information. [6] Includes retail and CCD ACH payments. See Gerdes (2008) for additional information.

Table 3: Top Issuers by Transaction Type (Dollar Value, 2007)

Credit Card Issuers	All Issuers Share of Dollar Value (%)	Excluding 3-party Issuers Share of Dollar Value (%)
American Express	24.7	-
JP Morgan Chase	17.6	24.9
Bank of America	14.6	20.8
Citigroup	12.3	17.6
Capital One	5.9	8.4
Discover	5.0	-
U.S. Bank	3.6	5.2
HSBC	2.3	3.3
Wells Fargo	2.1	2.9
GE Money	1.5	2.1

Signature Debit Card Issuers	Share of Dollar Value (%)
Bank of America	20.0
Wells Fargo	10.3
JP Morgan Chase	7.0
Wachovia	7.0
Washington Mutual	6.6
U.S. Bank	4.3
Regions Bank	2.9
Fifth Third	2.3
USAA Savings	2.3
RBS Citizens Bank	2.2

PIN Debit Card Issuers	Share of Dollar Value (%)
Bank of America	26.8
Wells Fargo	11.4
Washington Mutual	9.3
Wachovia	8.1
JP Morgan Chase	6.7
Regions Bank	3.8
U.S. Bank	2.4
SunTrust	2.4
Citigroup	2.3
BB&T	1.8

Source: Authors' calculations from the Nilson Report.

Table 4: Top Acquirers (Transactions, 2007)

Acquirer	Share of Transactions (%)
First Data Corp.	32.4
Bank of America	17.8
Chase Paymentech Solutions	11.8
Fifth Third	11.7
Nova Information	3.8
Global Payments	2.8
Heartland Payment Systems	2.5
Alliance Data Systems	2.4
RBS Lynk	2.0
First National	1.5

Source: Authors' calculations from the Nilson Report.
Note: Includes all transaction types. For joint ventures, calculations assume transactions are shared in proportion to ownership share.