Debit Market Competition: Interactions between Profit Maximizing and Welfare Maximizing Networks

by

Alexa Gendron-O'Donnell

An essay submitted to the Department of Economics
in partial fulfillment of the requirements for
the degree of Masters of Arts

Queen's University
Kingston, Ontario, Canada
August 2010

Abstract

The increased use of payment cards has prompted the development of economic literature to determine the appropriate interchange fee to effectively balance both sides of the market. In this paper, we focus on the debit card market and extend the model in Manenti and Somma (2010) to include competition between two profit maximizing networks as well as competition between a profit maximizing and a welfare maximizing network. We find that a single profit maximizing network will set interchange fees higher than is socially optimal while two such competing networks will set identical fees. A welfare maximizing network will set the lowest interchange fee possible to induce the profit maximizing network to reduce their fee and increase social welfare.

Acknowledgements

I would first like to thank Professor Roger Ware for his insight, expertise, and guidance as my professor and advisor. I would also like to thank my family and the Economics MA '10 class for their support.

Contents

List of Figures

1 Introduction

The use of debit cards as a means of payment has grown significantly worldwide. Canadians, in particular, use this payment card[1] for nearly 30 percent of their retail transactions, making them the second heaviest debit card users, after Sweden.[2] Payment cards exist as two-sided markets; the network must convince both the buyer and the seller to use its services before a transaction can take place. This is done by influencing the prices charged to both of these end-users through an interchange fee which is exchanged from one side of the market to the other.

The economic literature about payment card networks has also developed in the past few decades, as theorists attempt to determine the optimal interchange fee. However, the results tend to be highly dependent on market assumptions (like level of competition, ownership of one or more payment cards, ability to set multiple prices, etc.), making it difficult to apply the theory to an individual country's market. This increases the importance of creating and testing new models with varying assumptions to heighten our knowledge of debit systems.

This paper will adapt the model laid out in Manenti and Somma (2010) to observe the effects of inter and intra-network competition on profit maximizing and welfare maximizing networks. Section 2 of this paper will examine the unique characteristics of debit markets and the intuition behind interchange fee changes. Section 3 will explore the existing payment card network literature and the various assumptions applied in different models. Section 4 introduces the model we will use while Section 5 studies equilibrium interchange fee conditions with a single profit maximizing network, two competing profit

[1] Payment card is a generic term that includes credit, debit, and charge cards.
[2] Bergevin (2009)

maximizing networks, and finally competition between a profit maximizing and a welfare maximizing network. Lastly, section 6 offers concluding remarks.

2 The Debit Market

The market for debit cards differs substantially from ordinary markets and thus it is important to take these distinctions into account when formulating any sort of analysis. The mechanics of this market involve many parties who participate in a series of interrelated transactions, facilitated by an organizational payment network. The four main parties include consumers, issuers, merchants, and acquirers.

The consumer is the cardholder who receives debit services from the issuing bank. This is a depository institution which issues all payment cards to consumers. The merchant is the retailer from whom consumers will make purchases and their depository banks are known as acquirers. These parties interact within a payment network which coordinates the transmission of money and information between the issuing and acquiring sides of the market.

The two main types of organizational payment networks are known as three-party or four-party. Four-party systems are used by Interac, MasterCard, and Visa debit and involve cardholders, merchants, issuers, and acquirers as described above. Within the literature they are also referred to as payment associations or open networks. Alternatively, in the three-party system (also known as a proprietary network), the network acts as both the issuer and the acquirer. This is not prevalent for debit cards but is used by companies like American Express and Discover in the United States for their credit card networks.

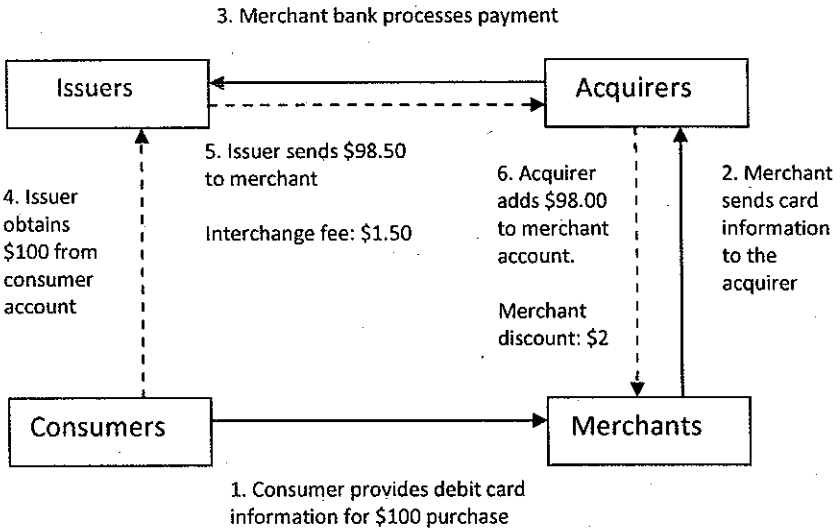

Figure 1: Typical Transaction in a Four-Party Network

Figure 1 provides an example of the informational and monetary flows within a four-party system that occur after a $100 debit purchase[3]. The consumer first provides their debit information to the merchant which is then relayed to the merchant's acquiring bank. The acquiring bank passes this information along to the consumer's issuing bank which will obtain the funds from the cardholder's debit account to settle the purchase along with a processing fee. The fee in this example is $0.50, making the total debit from the cardholder's account $100.50.[4] The payment network (Interac, MasterCard, or Visa) provides the link between the acquiring and issuing banks to coordinate the information and money flows.

[3] Example and chart adapted from Pancheco and Sullivan (2006)
[4] In some cases, the fee charged by issuers to consumers may actually be negative, indicating that consumers are receiving rewards or rebates.

There are additional fees which exist within the network, the first of which is known as an interchange fee which is charged by the issuing bank to the acquiring bank[5]. Instead of passing the $100 to the acquirer, the issuing bank retains a portion of this money (in this case $1.50). The acquiring bank therefore receives only $98.50.

The acquirer will also charge the merchant a processing fee (for the purposes of our example, it is $0.50) and the merchant will therefore receive a total of $98.00 in their account from the $100 sale. Overall, the acquiring bank charged the merchant a total of $2, known as the merchant discount fee, which is the sum of the interchange fee and the processing fee.[6]

It is important to note that this is a stylized and simplified example. Within real payment systems, the network (Interac, MasterCard, or Visa) is sometimes seen as a kind of fifth party. They may also charge switch fees to the acquiring and/or issuing banks for coordinating information and monetary exchanges. Additionally, a switch fee could allow a private network to extract profit. There may also exist another intermediary between merchants and their acquirers known as "merchant acquirers" which act only as informational processors for merchants. Finally, consumer card fees are often charged on a yearly basis, though they are often modeled as per-transaction fees. Within our analysis, we will only deal with a standard four-party system which charges per-transaction interchange fees and merchant discount fees.

---

[5] In most countries the interchange fee is paid by the acquirer to the issuer. However this is not always the case; for Australia's EFTPOS (debit) transactions the fee flows the other way (from the issuer to the acquirer) implying a negative interchange fee. It is also possible for the interchange fee to be zero, as is the case for Interac debit in Canada.

[6] The fee amounts in this example are larger than in most debit markets. For example, in Canada, interchange fees range from zero for Interac and MasterCard Maestro, to a maximum of 1.15% of the purchase price + $0.05 for Visa debit. (Bergevin, 2009)

A significant concept that arises when analyzing the transaction process is that of interchange fee neutrality. Under a special set of circumstances, the neutrality result implies that the level of the interchange fee has no net effect on any of the four parties. First, issuers and acquirers must be perfectly competitive and simply pass through any and all costs to the consumers and merchants through consumer card fees and merchant discount fees. In the case of negative consumer fees (rewards), issuers must costlessly rebate higher interchange fee revenue to consumers. Second, retailers must have the freedom to charge different prices based on a consumer's method of payment. For example, the price of a good paid with cash would likely be the cheapest, followed by a slightly higher price if paid by debit (to cover the merchant discount fee incurred). A higher interchange fee would then simply result in a higher price charged to the consumer which they would receive back as higher rebates from the issuer, leaving the net price unchanged. The level of the interchange fee then becomes irrelevant as all parties remain unaffected.[7]

Using our previous numerical example let the interchange fee increase from $1.50 to $2.00. The higher interchange fee also drives up the merchant discount fee by $0.50 and if the merchant had to retain the original $100 price, they would only receive $97.50 from the acquirer, thereby decreasing their profit. Now let us assume that the conditions for the neutrality result explained above hold and the merchant is free to increase the price to $100.50 to compensate for the higher fee. Consumers get the higher interchange fee rebated to them so their card fee goes down by $0.50 to zero and a $100.50 debit is made from their account by the issuer. This latter party now retains $2.00 from the

[7] For more literature pertaining to interchange fee neutrality see Gans and King (2003) and Carlton and Frankel (1995).

interchange fee and passes $98.50 to the acquirer who in turn charges an additional $0.50 processing fee and passes $98.00 to the merchant. In the end, all parties receive exactly the same net amount of money and are indifferent between the two interchange fees.

Payment networks may, subject to regulatory authority, set certain rules for their participating members, including one which could affect the neutrality result. The two most common include the no surcharge rule (NSR) and the honour all cards (HAC) rule. To a certain extent, the interchange fee can also be used by the networks to dictate the structure of relationships and fees between different parties.

The NSR ensures that affiliated merchants may only set one price for each good, regardless of the payment method used by the consumer to buy it. Since cash, debit, and credit purchases all cost the merchant different amounts to process, they will end up setting a price which covers their costs for even the most expensive payment option (usually credit). In effect, consumers may be paying a distorted price which doesn't reflect the true cost to the merchant of the payment method they choose. This is especially true for consumers who pay with cash; the price they face tends to be higher than it would be if the NSR wasn't in place. Note that this directly affects the neutrality result as being able to set different prices is crucial in that case.

Payment card companies (especially credit networks) prefer to impose a NSR so as to ensure that consumers will continue to demand and use their cards. With only one price in place, consumers may find it most beneficial to use their payment card, particularly if it provides them with rewards. When multiple prices exist, the cheapest

"cash price" may be more attractive, especially if consumers are paying card fees. However, even when surcharging is permitted, it is not always observed to be prevalent.[8]

The HAC rule relates to the different payment products available from payment networks. Companies like Visa and MasterCard offer many types of debit and credit cards, each of which tends to cost a different amount for the merchant when used by consumers. To ensure that affiliated merchants accept all of these products, payment networks prefer to contractually require an HAC promise from them.

Finally, since the network does not directly provide services to end users (consumers and merchants) it may not directly determine the fees which they are charged. Instead, it can set an interchange fee which structures the incentives of the affiliated parties to maximize profit. To illustrate this point, it is useful to consider how acquirers and issuers respond to changes in the interchange fee.

Acquirers view the interchange fee as a cost incurred for the services they provide to merchants. In other words, if the interchange fee rises, so do their costs for every transaction processed. Acquirers will want to pass this cost through to merchants by raising the merchant discount fee. If pass through is perfect (as was described in our neutrality example above with perfect competition) then acquirer profits will remain unchanged, but if this is not the case acquirer profits may decrease, making them less likely to promote merchant acceptance of that particular payment card.

Conversely, issuers view the interchange fee as a payment received for the services they provide to cardholders. As the interchange fee rises, so does the payment from every transaction. To encourage more consumer transactions (and receive more payments), issuers will respond by decreasing consumer fees and/or increasing rewards

[8] Prager et al. (2009) and Hayashi and Weiner (2006)

or rebates. If pass through is perfect then their profits remain the same, but if it is less than perfect, issuers' profits will increase and they will be more likely to promote consumer use of that particular payment card.

Overall, a rise in the interchange fee will tend to increase the merchant discount fee and decrease cardholder fees (or increase benefits). If pass through is less than perfect, it can also lead to a decrease in acquirer profits and an increase in issuer profits which affects the extent to which the payment card is promoted on each side of the market. Therefore, interchange fee changes can alter the structure of fees within the entire payment network.[9]

This highlights the use of the interchange fee as a balancing instrument used by networks on both sides of the market. An increase or decrease allows them to alter the fee structure and expand or contract one side relative to the other. For example, a higher interchange fee encourages more card usage by consumers while a lower one expands merchant acceptance of cards. Furthermore, networks can raise or lower the overall level of fees in the system if they are aware of less than perfect pass through from issuers and/or acquirers. This could allow some members of the network to increase their profits at the expense of consumers and/or merchants. It is important to note that though networks have some control over the structure of their payment system, they will not necessarily set the socially optimal level of interchange fees.[10]

Along with its two-sided nature, there are also externalities which make the debit market unique. The first is that these markets are subject to network effects; in other

[9] It is possible for the level of interchange fees to remain the same if the increase in the merchant discount fee exactly offsets the decrease in the consumer fee.

[10] The Reserve Bank of Australia is among the first in the world to regulate credit card interchange fees in their domestic market and compel them to be lower. As a result, consumer card fees are increasing (rewards are declining) while merchants have experienced lower merchant discount fees. Weiner and Wright (2005).

words, they become more valuable the more they are used. Consumers tend to place a higher value on a card and be more inclined to use it if it is accepted by many merchants (and equivalently for merchant acceptance with higher consumer use). However, consumers and merchants only evaluate their own costs and benefits when joining the network and do not take into consideration the fact that the network as a whole will gain in value. Therefore, consumers and merchants may under-use a card or accept it less than is socially optimal. As a result, there tend to be only a few large, established networks operating in each market. In the Canadian case, though Visa and MasterCard are relatively new entrants joining Interac in the debit market, they were able to capitalize on their marketing power and financial strength in the credit card market to do so.

The second externality revolves around competition; in a regular market, competition drives prices down to cost and induces efficient resource allocation, but this is not the case with the debit market. Here, there is no guarantee that increased competition between networks will lead to a socially optimal outcome. Competition could decrease interchange fees if consumers hold multiple cards (i.e. multi-home) and merchants feel confident accepting only the lowest priced cards. Payment networks may therefore be more likely to cater to merchants by decreasing merchant discount fees. Alternatively, networks may cater to future consumers and raise interchange rates to try to induce them to hold their card exclusively. In the United States, competition between Visa and MasterCard increased interchange rates as each network tried to create more favourable conditions to attract potential cardholders.[11]

The reason that increasing interchange fees can even be sustained within this two sided market relates to a third externality – low merchant resistance. Despite the fact that

[11] Simon (2005)

high interchange fees should induce lower merchant acceptance of payment cards, merchants continue allowing consumers to use them in their stores. One reason may be that the underlying value of the transaction exceeds the cost of the payment method. Therefore, the merchant perceives that it is not worth the sales loss to decline the card. This is exacerbated when high interchange fees lead to high consumer rewards. The consumer then has extra incentive to use the payment card and will be more likely to go to another store if their card is declined, increasing the chance that the merchant will lose the sale.

Furthermore, merchants may want to accept payment cards for strategic reasons. As mentioned above, accepting a wider range of cards may lead to a higher sales volume. An externality is created if some of these sales have been diverted from other merchants (in essence, "business-stealing"), as opposed to increases in aggregate sales.[12] In this case the merchant's private benefit is exceeding his social one and he is willing to pay a higher merchant discount fee (resulting from a high interchange fee) than is socially optimal. Yet this strategy is prevalent since no merchant wants to be the only one declining major cards and it may lead to merchants accepting certain cards when they would be better off collectively rejecting them.

Part of the reason lies in the way in which strategic merchants evaluate their card acceptance decisions. While consumers take into account only their transactional benefit from holding a card, merchants internalize both their own and their consumers' benefits. This is due to the fact that when consumers make their decision about which sellers to

[12] Authors like Vickers (2005) and Katz (2001) have concluded that it is unlikely that more purchasing with payment cards leads to an expansion in aggregate consumption.

purchase from, they take into account which ones will accept their payment card. Thus, the merchant actually internalizes both transactional benefits.

3 Review of the Literature

The majority of papers which model the payment card industry focus on credit cards though many of them can represent the payment card industry as a whole. The four-party system of credit cards which interact over a payment network is the same as the debit market. The intuition behind these models, as well as any conclusions about interchange fees and parties' behaviour can be considered to be identical. The only credit card models which are not appropriate to use for the debit market are those which explicitly consider the credit function and have the parties behaving accordingly.[13]

One of the first influential papers about payment card systems comes from Baxter (1983) who performed a normative analysis for finding the optimal interchange fee. He assumes a competitive market operating within one time period and concludes that total demand for credit cards is determined by the joint demand from both consumers and merchants. Analogously, total cost must include the joint costs of both issuers and acquirers. Ultimately, the equilibrium occurs where joint demand equals joint cost, thus incorporating both sides of the market and using the interchange fee to balance demands and costs for payment services. Baxter also states that since costs and demands are unlikely to be symmetric, the interchange fee will rarely be zero.

A major criticism of Baxter's paper is that the participants do not act as strategic players and are therefore not fully rational.[14] In Baxter's analysis, merchants do not

[13] See Rochet and Tirole (2009) for an example of a model incorporating the function of credit.
[14] Rochet and Tirole (2002)

believe that accepting cards will help attract more consumers and this causes him to overstate merchant resistance. This would only be a legitimate outcome if consumers were unaware of which stores accepted their payment card and would still be willing to make the purchase if it turned out the store did not.

Many theorists have expanded on Baxter's work and created models representing the payment industry including strategic parties as well as varying market assumptions. One such major paper comes from Rochet and Tirole (2002) who construct a mathematical model representing the interaction between the parties of the payment network. Their two main focuses are to compare the socially efficient and privately efficient interchange fees under different circumstances and to determine the change in welfare when the NSR is applied or lifted.

With consumer knowledge of which stores accept their payment card or a costless search to find the product at a store which does, homogenous merchants treat card acceptance as a strategic tool to attract customers. They are therefore prepared to pay a higher merchant discount fee to accept these cards, giving the network the ability to raise the interchange fee.

Rochet and Tirole's model presents the acquiring side of the market as perfectly competitive while the issuing side retains some market power. Since costs are fully passed through on the acquiring side but revenues on the issuing side are not, the higher interchange fee may lead to higher overall profits for members.

Under these conditions with full consumer knowledge, the model also predicts an overprovision of card services since the interchange fee is high. More specifically, Rochet and Tirole determine that the privately optimal interchange fee could be equal to

or higher than the socially optimal fee depending on total issuer and acquirer cost as well as cardholders' surplus. If the private interchange fee is too high, the market will be flooded with more cards than is socially optimal. This also implies that there could exist a lower level of consumer knowledge where the profit maximizing interchange fee is equal to the socially optimal one despite the overprovision of cards.

When focusing on the effects of the NSR on welfare, Rochet and Tirole predict that for a given interchange rate, allowing merchants to set multiple prices will increase the price for cardholders and decrease it for consumers who use cash to make their purchases. Consumers will therefore reduce their demand for payment cards due to the higher markups and issuers will focus on a high end clientele which makes large enough purchases that they don't wish to use cash.

If the total cost of issuers and acquirers minus merchant benefits is less than the cardholder's fee then changing to a NSR would reduce welfare. This is exacerbated if there was already a high cardholder fee and under-provision of cards. If, however, the total cost of issuers and acquirers minus merchant benefits is greater than the cardholder's fee, then the change in welfare from a NSR is ambiguous; it could increase or decrease. This result also holds if there was originally an overprovision of cards from low cardholder fees. The degree of downward pressure caused by the multiple prices could cause welfare to rise or fall depending on its intensity.

Wright (2003) also looks at the effects of a NSR by comparing markets in which there are monopolistic merchants or perfectly competitive merchants. Since the focus is on the merchants, both issuing and acquiring are perfectly competitive with full pass through. He finds that with monopolistic merchants, allowing multiple prices causes the

merchants to extract surplus from their card users which leads to lower issuer revenue. In a situation where annual consumer card fees are introduced, all customers revert to cash since the merchant has extracted any and all surplus needed to induce consumers to hold cards. As a result, networks prefer to put a NSR in place since it will increase card demand and improve welfare.

When merchants are perfectly competitive, the NSR will induce them to cater to only one kind of consumer clientele, either cardholders or cash users, setting their prices accordingly. When the NSR is removed, merchants will choose to set a price equal to the cost of the payment instrument used (cash or debit) minus any transactional benefit they obtain from accepting the card, causing them to divide into cash only and card only merchants again. In effect, allowing multiple prices has no effect on overall welfare with perfectly competitive merchants.

Schmalensee (2002) has also based his paper on Baxter's seminal work and extends it by introducing issuers and acquirers who have market power though he retains the assumption that merchants are not strategic entities. He supports Baxter's conclusions that the interchange fee serves as a powerful tool to balance the joint demands and costs for payment card services as well as the unlikelihood that this fee will be zero.

With the introduction of market power amongst issuers and acquirers, Schmalensee finds that it may be necessary to shift revenue from one to the other. This can arise if one party is more powerful within the payment card association or if one side of the market has more market power than the other. Taking the case where issuers are less competitive, the fee revenue would be shifted to them. In this case, costs would pass through more easily on the acquiring side but revenues would not on the issuing side. A

higher interchange fee might generate a higher collective profit for the network's members. With more voting power, issuers would receive more revenue from any increase in the interchange fee and therefore would push the network to set a higher fee than is optimal.

Schmalensee further clarifies the role the interchange fee can play in balancing costs across network members and creating the appropriate incentives to structure fees. He states that high issuing costs and/or high demand by merchants will lead to higher interchange fees while high acquiring costs and/or demand by cardholders will lead to lower interchange fees. The network can therefore steer interchange revenue to flow to the high cost side of the market and the side which has lower demand to create a profit maximizing strategy.

Wright (2004) also examines the balancing considerations of the interchange fee when both issuers and acquirers have some market power. When the payment card is used, the interchange fee in his model is decreasing in the transactional benefits for consumers and increasing in the transactional benefits for merchants. These benefits can be thought of as cardholder and merchant demand respectively. The interchange fee also increases in the price-cost margin of issuers but decreases in the price-cost margin of acquirers.

Again using the example of an issuer with more market power (and therefore a higher margin), Wright agrees with Schmalensee that a higher interchange fee is required to balance both sides of the market in order to maximize the volume of card transactions and profit for the network. However, Wright stresses that the degree of pass through is critically important. If raising the interchange fee results in a greater or equal decrease in

consumer card fees (or increase in rewards) then he concludes that issuers and acquirers do not share a joint incentive to set a higher interchange fee and shift revenues to issuers. His theory instead predicts that the interchange fee should increase as the degree of acquirer pass through increases and decrease as issuer pass through increases. When both pass through costs at the same rate, Wright predicts that the profit maximizing interchange fee will be higher than the socially optimal fee if the average transactional benefit for card accepting merchants is lower than their merchant discount fee at the profit maximizing interchange rate.

The majority of the theoretical literature about payment card systems focuses on a single network and conducts analyses by altering model assumptions and observing the resulting interchange fee or welfare changes. Recently, the study of competing networks has become more common as it pertains more prevalently to existing real world conditions and social policy.

Rochet and Tirole (2002) were one of the first to touch on competing networks, albeit briefly. They conclude that when consumers hold only one payment card (single-home), then inter-system competition has no impact on merchant resistance and the networks will continue to choose the highest interchange fee merchants will accept so as to maximize transaction volume and network profit. If consumers choose to hold more than one card (multi-home) from different networks, then merchant resistance can increase but the welfare effects are ambiguous.

This concept is further developed in Rochet and Tirole (2003) as they state that multi-homing on one side of the market intensifies price competition on the other side. Networks will subsequently use low prices to steer end users on the latter side towards an

exclusive relationship. For example, when cardholders multi-home, merchants may choose to accept only the card that is cheapest for them, increasing their merchant resistance. The networks will further entice merchants to accept their card exclusively by decreasing the merchant discount fee and increasing the consumer's card fee. Alternatively, when merchants accept multiple cards, networks will offer low consumer card fees (or high rewards) to induce consumers to choose only their card. Since merchants are multi-homing and their resistance has decreased, they can be charged a larger merchant discount fee. Ultimately, different networks try to induce the side of the market which is single-homing to stay with them exclusively by undercutting their rivals.

Guthrie and Wright (2007, p.39) use their paper to create "...an extension of Rochet and Tirole (2002) [and]...a variation on Rochet and Tirole (2003)." In keeping with other theoretical models, the interests of cardholders are over-represented and merchant resistance is low, creating a situation whereby merchant discount fees and interchange fees tend to be high. Guthrie and Wright do allow for either side of the market to multi-home but approach the issue as having either merchants or consumers play the greater role in determining which payment card will be used prevalently. At one extreme, when consumers play a greater role, they will choose their preferred card while merchants multi-home, causing the network to focus its attention exclusively on attracting cardholders. When merchants play the larger role, the opposite occurs and networks focus on merchant acceptance exclusively.

They also distinguish between a case of homogeneous and heterogeneous merchants with respect to the transactional benefits they derive. In both cases, a single card network will set its interchange fee too high compared to the socially optimal fee

due to over-representation of cardholders within the market. With homogenous sellers (who obtain identical benefits from accepting cards), the interchange fee between competing networks will never be higher than when there is a single card scheme. This is due to the fact that the network has already set the highest interchange fee possible such that merchants will accept their card and so derive no benefit from setting it higher (where they will lose merchant acceptance) or lower (where they would lose consumers). This is not true when the market includes heterogeneous sellers and the interchange fee may rise or fall depending on how heavily buyers' interests are weighted. As long as consumers' interests carry any weight at all, the interchange fee will be higher than the socially optimal one.

Chakravorti and Roson (2006) also consider the effects of network competition but find quite different results. When comparing the case of a monopoly or cartel network to a non-cooperative duopoly, they find that the equilibrium prices for consumers and merchants are the same or lower, indicating that competition may be beneficial. While the previous theoretical models concluded that the effects of competition were ambiguous and could be detrimental to network members, Chakravorti and Roson show that it is unambiguously non-harmful. The difference lies in an assumption of their model which has network profits decreasing with competition as opposed to constant profits for merchants and acquirers and a zero profit condition implied for networks as in other theories.

In all models, the increase in market competition results in downward pressure on prices and is associated with a change in the price structure. The first effect is welfare improving while the second is ambiguous and may be detrimental. In Chakravorti and

Roson, the first effect dominates, making consumers and merchants better off, while in the other models discussed it is the second effect which prevails.

Each decreased price exerts direct and indirect welfare raising effects. Lower consumer fees directly increase consumer welfare but also indirectly increase merchant welfare due to the higher number of consumers using that particular card system. A parallel network effect occurs with lower merchant fees which directly increase merchant welfare and indirectly increase consumer welfare by encouraging more merchants to join the network.

Manenti and Somma (2010) approach the payment card industry by developing a model in which inter-network and intra-network competition can be measured and altered. While most previous papers modeled competition between two networks, this one explicitly includes relative levels of competition between issuers and acquirers in their respective markets (intra-network competition). It also differs from previous papers by having a four-party card association system compete with a three-party proprietary card system.

The authors find that an increase in intra-network competition from either the issuer or the acquirer's side of the market causes the optimal interchange fee to increase the price for the less competitive side. In other words, depending on which side of the market is less competitive, merchants or consumers will end up facing higher merchant discount fees and cardholder fees respectively. For example, if the acquirer's side is less competitive, an increase in competition on the issuer's side will lead to an increase in the optimal interchange fee which subsequently causes higher merchant discount fees. If

inter- competition is not too asymmetric then changes in its level will have similar effects on the optimal interchange fee.

Manenti and Somma also find that while interchange fees set by the four-party card association are privately efficient, they are not socially optimal. Furthermore, the socially efficient interchange fee should be set such that the price is minimized on the side of the market with the least intra-network competition.

Though the three-party, proprietary network cannot serve as a model for the debit card industry, the rest of this paper will extend Manenti and Somma's model for the four-party card association. We will use their inter-network and intra-network competition analysis with two four-party, profit maximizing payment card systems then introduce one network as a welfare maximizing system.

4 The Model

4.1 Assumptions

The profit maximizing payment networks can be thought of as joint ventures between their acquiring and issuing members. Following Manenti and Somma, we assume that these parties always remain separate entities and that no affiliated bank can simultaneously be an issuer and acquirer. Each system aims to maximizing the number of transactions handled on its network, thereby maximizing its value, measured by the joint profits of issuers and acquirers. To do so, these members independently set fees for their consumers and merchants, respectively.

The network will also set an interchange fee to structure and coordinate the incentives between the two sides of the market. Normally it flows from acquirers to issuers but we also allow the interchange fee the option to be negative, thus being paid

from issuers to acquirers. We will also normalize to zero any fixed or variable costs that the network may incur directly. This allows interchange fee costs paid from one side of the market to precisely counterbalance the fee revenue received by the other.

Network value is generated by transactions between end users (consumers and merchants). We assume that the population of each of these end users has a mass of one. Furthermore, each buyer-seller pair of can make at most one potential transaction. Both members of the pair must be affiliated with the same network for the transaction to take place. Neither the consumers nor the merchants multi-home so each can only be affiliated with one network. The network also imposes a NSR so the merchant is restricted to one price for each good (and no neutrality result will occur).

The timing of the game is as follows:

(i) The network(s) set(s) the level of the interchange fee(s).

(ii) Issuers and acquirers compete over their consumer card fees and merchant discount fees respectively.

(iii) Consumers decide which network to subscribe to and therefore which payment card to hold while merchants decide which card(s) to accept.

(iv) Merchants set their retail prices and consumers decide whether or not to purchase the good.

Since the interchange fee is set by the network(s) only periodically, it is reasonable to assume that issuing and acquiring members are fully aware of this fee before they set their own prices.

Figure 2 provides a graphical summary of the flow of funds within one profit maximizing network, along with the model's variables. [15]

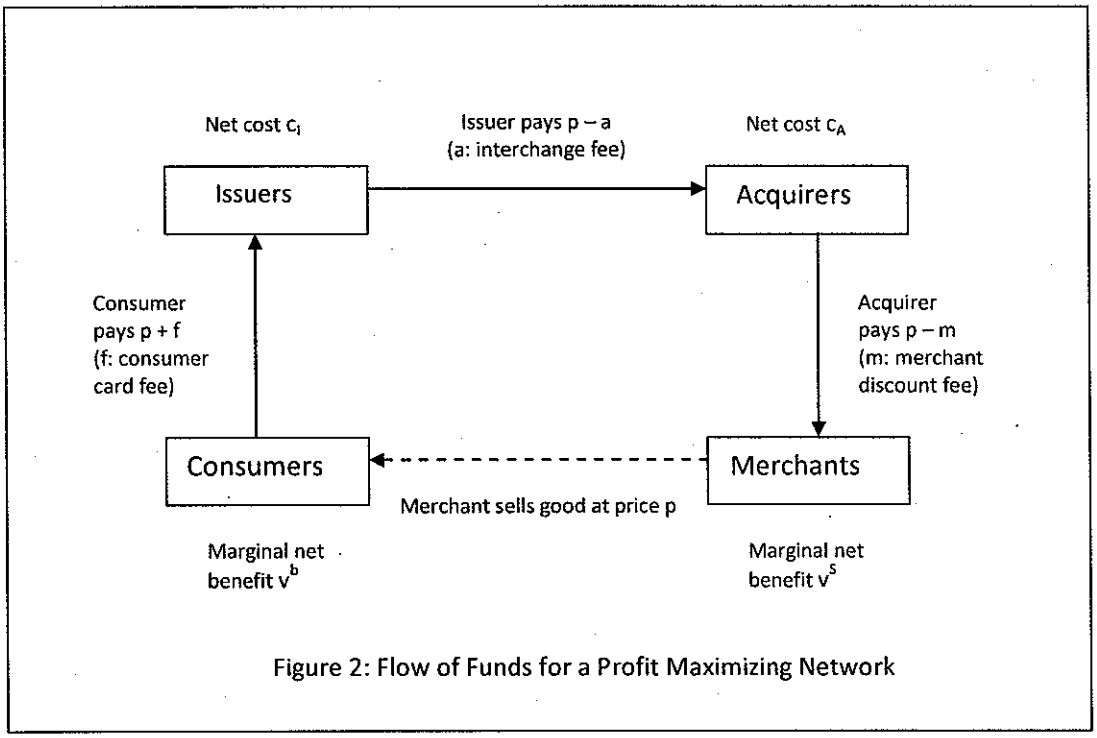

Figure 2: Flow of Funds for a Profit Maximizing Network

4.2 Issuer and Acquirer Behaviour

The network sets an interchange fee, a, while issuers and acquirers compete for end users by setting consumer card fees, f, and merchant discount fees, m, respectively. These parties also face constant marginal costs in the form of c_I for issuers and c_A for acquirers. Since we have assumed the system itself encounters no direct costs, the network's total per-transaction cost is $c = c_I + c_A$.

These network members compete intensely for end users, adapting their prices (f and m) based on the level of competition between issuers and acquirers, on the

[15] Figure adapted from Rochet and Tirole (2002)

interchange fee, a, and on their respective marginal costs, c_i. The simplest approach is to assume that both these members face exogenous price-cost markups.[16]

The merchant discount fee is:

$$m_i = \sigma(c_A + a)$$

where $i = 1,2$ depending on the network in question.

This creates a markup of:

$$\frac{m_i - (c_A + a)}{m_i} = \frac{\sigma - 1}{\sigma}$$

where $\sigma > 1$ is a constant margin representing the degree of intra-network competition for acquirers.

Similarly, the consumer card fee is:

$$f_i = \beta(c_I - a)$$

where $i = 1,2$ which creates a markup of:

$$\frac{f_i - (c_I - a)}{f_i} = \frac{\beta - 1}{\beta}$$

Here, $\beta > 1$ is the constant margin associated with the degree of intra-network for issuers. The larger value of sigma or beta represents the side of the market experiencing less competition (or lower demand elasticity).

4.3 Consumers Behaviour

Both sides of the market are portrayed by a variation of the standard Hotelling model. Merchants and consumers are uniformly distributed along a line segment representing the market while the two networks are located at either end.

[16] This is equivalent to assuming that the equilibrium price f is linear in the issuing bank's marginal cost, as is m for the acquiring bank.

Due to the exogenous price-cost markup in the cardholder fee they face, the gain to consumers from adopting a card depends on their per-transaction benefit. This corresponds to consumers not knowing at the time of card adoption how many transactions they will perform with the network. Thus they face this decision on a per-transaction basis which simplifies our model by removing any uncertainty about each cardholder's number of future purchases.

Since each buyer single-homes and can adopt at most one card, they make their choice based on their individual per-transaction utility from using the card on network 1 and 2 respectively:

$$v^b(M_1) - f_1 - kx$$

$$v^b(M_2) - f_2 - k(1-x)$$

where k is the transportation cost facing each consumer, x represents the consumer who is indifferent between both networks, and M_i, $i = 1, 2$ is the expected number of merchants operating on network i. The benefit from paying with a card is $v^b(.)$ which is a positive function, weakly increasing in M_i. This reflects the network externality, indicating that consumers indirectly benefit when more merchants accept their network's card. However, like Manenti and Somma, our model will assume that v^b is independent of M_i for the sake of avoiding cross-market effects and simplifying the algebraic analysis.

The Hotelling model allows us to identify the consumer who is indifferent between joining network 1 or 2 and therefore allows us to find the proportion of consumers who use each system; they are:

$$d_1 = \frac{1}{2} + \frac{f_2 - f_1}{2k} \quad \text{and} \quad d_2 = \frac{1}{2} + \frac{f_1 - f_2}{2k}$$

These become:

$$d_1 = \frac{1}{2} + \frac{\beta(a_1 - a_2)}{2k} \quad \text{and} \quad d_2 = \frac{1}{2} + \frac{\beta(a_2 - a_1)}{2k}$$

The transportation parameter k is an indication of the degree of substitution between the two networks for the issuer.

4.4 Merchant Behaviour

Merchants, on the other hand, do not know which network's card the consumer holds to make a purchase. This leads us to model the gain from card acceptance decision on a per-transaction basis to avoid future uncertainty. The per-transaction utility gained by the merchant from selling the good through networks 1 and 2 are:

$$v^s(D_1) - m_1 - tx$$

$$v_s(D_2) - m_2 - t(1-x)$$

where t is the transportation cost for each merchant, x represents the merchant indifferent between both networks, and D_i, $i = 1,2$ is the expected number of consumers who have adopted cards from network i. The benefit from accepting a card is $v^s(.)$ and, similarly to consumers, we assume away any cross-market effects.

The proportion of merchants who are affiliated with networks 1 and 2 are:

$$e_1 = \frac{1}{2} + \frac{m_2 - m_1}{2t} \quad \text{and} \quad e_2 = \frac{1}{2} + \frac{m_1 - m_2}{2t}$$

These become:

$$e_1 = \frac{1}{2} + \frac{\sigma(a_2 - a_1)}{2t} \quad \text{and} \quad e_2 = \frac{1}{2} + \frac{\sigma(a_1 - a_2)}{2t}$$

Here, t reflects the degree of substitutability between networks 1 and 2 on the acquirers' side.

4.5 Network profits

We have stated that network value and profit depend on the number of transactions which occur between each merchant-consumer pair. We can further describe these transactions as being functions of consumer adoption, d_i, as well as merchant acceptance, e_i, $i = 1, 2$. We can also now define the number of transactions on network i as $h_i(d_i,e_i)$. The simplest form of this function occurs when each consumer who adopts a card from network i makes one transaction with each merchant who accepts the same card. This balanced trading pattern is:

$$h_i(d_i,e_i) = d_i \cdot e_i$$

Given the previous expressions for network prices and transaction amounts, the profits for acquiring and issuing banks are:

$$\pi_{i,A} = (\sigma - 1)(c_A + a)(d_i \cdot e_i)$$

$$\pi_{i,I} = (\beta - 1)(c_I - a)(d_i \cdot e_i)$$

These combine together to form the total profit for the joint venture network:

$$\pi_i = \pi_{1,A} + \pi_{i,I} = G(a)(d_i \cdot e_i)$$

where

$$G(a_i) \equiv [(\sigma - \beta)a_i + (\beta - 1)c_I + (\sigma - 1)c_A]$$

and $i = 1, 2$ depending on the network and $G(a_i)$ represents the total cost margin.

At this point, it is important to provide bounds for the interchange fee to ensure that a network will be unable to price its rival out of the market. To do this we will only allow $a \in [-c_A, c_I]$, otherwise a network's members could receive negative profits.

5 Equilibrium

5.1 A Single Profit Maximizing Network

Proposition 1. A single profit maximizing network can set its interchange fee higher than is socially optimal.

This is one of the more robust results in the literature. Guthrie and Wright (2007) examine this phenomenon with perfectly competitive issuers and acquirers. They note that the socially optimal interchange fee occurs where the joint transactional benefits of consumers and merchants are maximized. This fee also falls between the lowest and highest interchange fees that a privately profitable network would choose.

Guthrie and Wright determine that a single network in the market will set the interchange fee at the highest level such that merchants will still accept the card. In choosing this interchange fee, the network has maximized consumers' transactional surplus and promoted greater card usage. The merchants' card acceptance decision relies on assessing the joint transactional surplus whereas consumers only consider their own surplus when determining their card usage. This allows merchants to internalize the higher consumer surplus and lowers their resistance to a higher interchange fee (and therefore a higher merchant discount fee).

Ultimately, since the profit maximizing interchange fee is at its highest possible level given the participation constraint of merchants, it will be greater than the socially optimal fee. Therefore a single unregulated profit maximizing network will set its interchange fees too high, leading to an overprovision of payment card services.

Rochet and Tirole (2002) also find similar results, even with the addition of market power on the issuer's side. In their model, the level of the privately optimal

interchange fee depends on consumer benefit as well as issuer and acquirer costs, making it possible that the fee could be set higher than is welfare maximizing.

5.2 Competing Profit Maximizing Networks

We are now able to analyze the effect of having two profit maximizing networks competing within the same market. The expanded profit function for network 1 is as follows:

$$\pi_1 = [(\sigma - \beta)a_1 + (\beta - 1)c_I + (\sigma - 1)c_A][\frac{1}{2} + \frac{\beta(a_1 - a_2)}{2k}][\frac{1}{2} + \frac{\sigma(a_2 - a_1)}{2t}]$$

We will begin with the simplest case when competition between acquirers and issuers is symmetric (i.e. $\sigma = \beta$; let us call it β). When we take the first order condition of the profit function above with respect to a_1, we find that $a_1 = \frac{t-k}{2\beta} + a_2$. The analogous result holds when we derive the first order condition of the second network with respect to a_2.

Proposition 2. When $\sigma = \beta$

1. and $t = k$, two competing profit maximizing networks will set their interchange fees at the same level ($a_1 = a_2$).
2. the network's profit function becomes independent of either interchange fee.

It is not surprising that both networks set the same interchange fees considering that they are identical. Taking this a step further and allowing competition to be asymmetric ($\sigma \neq \beta$), we can once again take the first order condition of π_1 with respect to a_1.

$$\frac{\partial \pi}{\partial a1} = (\sigma - \beta)(d_1 \cdot e_1) + [\frac{(t\beta - k\sigma + 2a_2\beta\sigma - 2a_1\beta\sigma)}{4tk}] \cdot G(a_1)$$

We know that $(\sigma - \beta)$ and $G(a_1)$ will be positive when competition among issuers is stronger $(\sigma > \beta)$ and negative when it is weaker $(\sigma < \beta)$ but unfortunately we cannot sign the remainder of the equation.

5.3 Competition between a Profit Maximizing and a Welfare Maximizing Network

We now replace profit maximizing network 1 with a network which maximizes overall market welfare instead. It takes into account the sum of surpluses on both sides of the market as well as profits. The timing of the game also changes slightly as the welfare maximizing network now sets its interchange fee first, followed by the profit maximizing network.

The welfare function for this market is:

$$W = CS_1^b + CS_2^b + CS_1^s + CS_2^s + \pi_1 + \pi_2$$

where CS_i^b and CS_i^s are the surpluses enjoyed by the consumers and merchants respectively on network i. Using individual transactional utility functions we can define consumer surplus for network 1 as:

$$CS_1^b = \int_0^{d_1} (v^b - f_1 - kx)dx = \frac{d_1(2v^b - kd_1 - 2f_1)}{2}$$

Similarly, we can also define CS_2^b, CS_1^s, and CS_2^s:

$$CS_2^b = \frac{d_2(2v^b - kd_2 - 2f_2)}{2} \qquad CS_1^s = \frac{e_1(2v^s - e_1 - 2m_1)}{2} \qquad CS_2^s = \frac{e_2(2v^s - e_2 - 2m_2)}{2}$$

We begin with the simplest case of symmetric intra-network competition ($\sigma = \beta$; let us call it β) and $t = k$. The first order condition of the derivative of the welfare function with respect to a_1 is:

$$\frac{\partial W}{da_1} = \frac{4k^2\beta + k\beta^2(10a_1 - 16a_2) + 4\beta^2(a_2 - a_1)}{16k^2} + 4\beta^2(\beta - 1)(c_I + c_A)(a_2 - a_1)$$

Therefore when we constrain it such that $a_I = a_2$ we see that:

$$\frac{dW}{da_1} = -\frac{2k\beta^2 a_1}{16k^2} < 0$$

Proposition 3. When a profit maximizing network and a welfare maximizing network compete in a market with a single interchange fee and symmetric competition ($\sigma = \beta$), the welfare maximizing network will set its interchange fee as low as possible. With our constraints, this means that $a = -c_A$.

This result is consistent with previous our intuition. We have already stated that a single profit maximizing network can set interchange fees too high. When we introduce a welfare maximizing network, it sets a lower interchange fee, thereby encouraging the profit maximizing network to do the same and provide a socially efficient level of payment card services.

When we allow competition between issuers and acquirers to be asymmetric ($\sigma \neq \beta$), and keep $t = k$, our first order condition becomes much harder to sign.

$$\frac{\partial W}{\partial a_1} = \frac{4k^2\beta + k\beta^2(2a_1 - 4a_2) + 4\sigma^2(a_2 - a_1) + 8k\sigma^2(a_1 - a_2)}{16k^2} + (\sigma - \beta)(d_1 \cdot e_1)$$

$$+ G(a_1)[\frac{-k\sigma + k\beta + 2\beta\sigma(a_2 - a_1)}{4k^2} + G(a2)[\frac{k\sigma + k\beta + 2\beta\sigma(a_2 - a_1)}{4k^2}$$

Even when we constrain the market to have only one interchange fee we cannot determine the sign of the first order condition.

$$\frac{\partial W}{\partial a_1} = (\sigma - \beta)[\frac{1}{4} + a_1 + 2k\beta] + 4k^2\beta - 2k\beta^2 a_1 - 4k^2\sigma + 2k\beta[(\beta - 1)c_I + (\sigma - 1)c_A]$$

Though we know that the first term is positive if competition among issuers is stronger ($\sigma > \beta$) and negative when it is weaker ($\sigma < \beta$) and all other terms are unambiguously positive or negative the sign of derivative as a whole is uncertain.

6 Conclusion

The aim of this paper has been to extend Manenti and Somma's model to incorporate competition between two profit maximizing networks and to introduce a welfare maximizing network instead of a regulatory body. We began by using conclusions found in Guthrie and Wright (2007) relating to a single profit maximizing network. When a network has a monopoly, it will push the interchange fee as high as it can given that merchants will still continue to accept their cards. This is due to lower merchant resistance and overstated consumer interests in the market. When we introduce another profit maximizing network with symmetric competition between issuers and acquirers, we find that if both transportation costs are the same, both networks will choose identical interchange fees. Finally, in a market with a single interchange fee, a welfare maximizing network will tend towards a lower fee to encourage the profit maximizing network to do the same.

7 References

Arango, Carlos, and Varya Taylor. 2008. "Merchants' Costs of Accepting Means of Payment: Is Cash the Least Costly?" *Bank of Canada Review*, Bank of Canada, vol. 2008 (Winter), pp. 17-25

Baxter, William. 1983. "Bank Interchange of Transactional Paper: Legal Perspectives," *Journal of Law and Economics*, vol. 26, pp. 541-88.

Bergevin, Phillipe. forthcoming. "Competition in the Canadian Debit Card Market: Towards and Efficient and Fair Outcome," C.D. Howe Institute

Canadian Bankers Association. 2009. *Canada's Credit and Debit Card Marketplace*, Submission to the Joint Meeting of the Standing Committee on Finance and the Standing Committee on Industry, Science, and Technology.

Chakravorti, Sujit. 2003. "Theory of Credit Card Networks: A Survey of the Literature," *Review of Network Economics,* vol. 2, pp. 50-68.

_____, and Wilko Bolt. 2008. "Economics of Payment Cards: A Status Report," De Nederlandsche Bank Working Paper No. 193.

_____, and Roberto Roson. 2006. "Platform Competition in Two-Sided Markets: The Case of Payment Networks," *Review of Network Economics*, vol. 5, pp. 118-42.

Carleton, Dennis, and Alan Frankel. 1995. "The Antitrust Economics of Payment Card Networks," *Antitrust Law Journal*, vol. 63, pp.643-68.

DeGennaro, Ramon P. 2006. *Merchant Acquirers and Payment Card Processors: A Look Inside the Black Box*, Federal Reserve Bank of Atlanta.

Department of Finance Canada. 2010. *Code of Conduct for the Credit and Debit Card Industry in Canada.*

Evans, David, and Richard Schmalensee. 2005. *Paying with Plastic: The Digital Revolution in Buying and Borrowing. Cambridge, Mass.: MIT Press, 2nd ed.*

Gans, Joshua, and Stephen King. 2003. "The neutrality of the Interchange Fees in Payment Systems," *Topics in Economic Analysis and Policy*, vol. 3, no.1.

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

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.

Katz, Michael. 2001. "Reform of Credit Card Schemes in Australia II: Commissioned Report," Reserve Bank of Australia Public Document, August.

Manenti, Fabio, and Ernesto Somma. 2010. "Plastic Clashes: Competition Among Closed and Open Systems in the Credit Card Industry," University of Padua Working Paper No. 111

Pancheco, Barbara, and Richard Sullivan. 2006. *Interchange Fees in Credit and Debit Card Markets: What Role for Public Authorities? – A Summary of a Federal Reserve Bank of Kansas City Conference*, Federal Reserve Bank of Kansas City.

Prager, Robin, Mark D. Manuszak, Elizabeth K. Kiser, and Ron Borzekowski. 2009. *Interchange Fees and Payment Card Networks: Economics, Industry Developments, and Policy Issues*, Federal Reserve Board, Washington D.C.

Rochet, Jean-Charles, and Julian Wright. 2009. "Credit Card Interchange Fees," European Central Bank Working Paper No. 1139.

Rochet, Jean-Charles, and Jean Tirole. 2008. "Tying in Two-Sided Markets and the Honor All Cards Rule," *International Journal of Industrial Organization,* vol.26, pp. 1333-47.

_____. 2003. "Platform Competition in Two-Sided Markets," *Journal of the European Economics Association*, vol.1, pp. 990-1029.

_____. 2002. "Cooperation Among Competitors: Some Economics of Payment Card Associations," *RAND Journal of Economics*, vol.22, pp. 549-70.

Reserve Bank of Australia. 2008. *Reform of Australia's Payment System: Conclusions of the 2007/08 Review*, Sydney, September.

_____. 2006. *Reform of the EFTPOS and Visa Debit Systems in Australia: Final Reforms and Regulation Impact Statement,* Sydney, April.

_____. 2005. *Reform of the EFTPOS and Visa Debit Systems in Australia: A Consultation Document*, Sydney, February.

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

Simon, John. "Payment Systems Are Different: Shouldn't Their Regulation Be Too?" *Review of Network Economics*, vol. 4, pp. 364-83

Vickers, John. 2005. *Public Policy and the Invisible Price: Competition Law, Regulation and the Interchange Fee*, Federal Reserve Bank of Kansas City.

Weiner, Stuart, and Julian Wright. 2005. "Interchange Fees in Various Countries: Developments and Detriments," *Review of Network Economics,* vol. 4, pp. 290-323.

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

_____. 2003. "Optimal Card Payment Systems," *European Economic Review,* vol. 47, pp. 587-612.