Externalities in Payment Card Networks: Theory and Evidence

Sujit Chakravorti

The proliferation of payment cards has dramatically changed the ways we shop and merchants sell goods and services. Today, payment cards are indispensable in most advanced economies. Amromin and Chakravorti (2009) find that greater usage of debit cards has resulted in lower demand for small-denomination bank notes and coins that are used to make change in 13 advanced economies.[1] Recent payment surveys also indicate that consumers are using payment cards instead of checks.

Some merchants have started to accept only card payments for safety and convenience reasons. For example, American Airlines began accepting only payment cards for in-flight purchases on all its domestic routes on June 1, 2009. Also, many quick service restaurants and coffee shops now accept payment cards to capture greater sales and increase transaction speed. Wider acceptance and usage of payment cards suggest that a growing number of consumers and merchants prefer payment cards to cash and checks. In addition, payment cards may allow access to credit that can be used to attract consumers without funds.

Debit, credit, and prepaid cards are three forms of payment cards. Debit cards allow consumers to access funds at their banks (defined broadly as depository institutions) to pay merchants; these are sometimes referred to as "pay now" cards because funds are generally debited from the cardholder's account within a day or two of a purchase.[2] Credit cards allow consumers to access lines of credit at their banks when making payments and can be thought of as "pay later" cards because consumers pay the balance at a future date. Prepaid cards can be referred to as "pay before" cards because they allow users to pay merchants with funds transferred in advance to a prepaid account.[3]

Greater usage of cards has increased the value of payment network operators, such as Visa, Inc., MasterCard Worldwide, Discover Financial Services, and

others. In 2008, Visa had the largest initial public offering (IPO) of equity, valued at close to $18 billion, in U.S. history (Benner, 2008). The sheer magnitude of the IPO suggests that financial market participants value Visa's current and future profitability as a payment network. One potential reason for Visa to change its corporate structure from a card association to a publicly traded company is to reduce antitrust scrutiny by regulators and to lower the threat of lawsuits filed by certain payment system participants (Enrich, 2006). In 2006, MasterCard Worldwide became a publicly traded company. Also, in 2007, Discover Financial Services was spun off by Morgan Stanley.

Some industry observers have suggested that the high profitability of payment card providers has increased scrutiny by public authorities in many jurisdictions.[4] Several U.S. merchants have filed lawsuits against MasterCard and Visa regarding the setting of interchange fees. These fees are paid by the merchant's bank to the cardholder's bank and are set by the network operator.[5] In April 2009, MasterCard reached an interim understanding with the European Commission on interchange fees for cross-border consumer payments in the European Union. Effective July 1, 2009, MasterCard Europe established cross-border interchange fees for consumer card transactions that, on average, do not exceed 30 basis points for credit cards or 20 basis points for debit cards.

To date, there is still little consensus—either among policymakers or economic theorists—on what constitutes an efficient fee structure for card-based payments. In this article, I discuss several types of externalities that are present in payment networks.[6] The first, and perhaps, the most researched, externalities are adoption and usage externalities. In addition to these externalities, underlying fee structures may affect the welfare of individuals or firms participating (or not participating) in the payment network. Finally, I will discuss the limited evidence that exists regarding the effectiveness of some policy interventions.

There are several conclusions that I draw from the academic models, recent interventions in payment card markets, and discussions about potential policy interventions. First, many economic models suggest that the socially optimal interchange fee structure may not be systematically lower than the network profit-maximizing fee. Second, removing merchant pricing restrictions generally improves market price signals. Third, merchant, card issuer, or network competition may result in lower social welfare contrary to generally accepted economic principles. Fourth, if warranted, fees set by the authorities should not only consider costs but also benefits received by consumers and merchants, such as convenience, security, and access to credit that may result in greater sales.

Finally, the motivation for why public authorities intervene differs across jurisdictions. The type of public institution that regulates payment cards also differs. The institution may be an antitrust authority, a central bank, or a court. Often public authorities intervene because the interchange fee is set by a group of competitors and the level of the fee is deemed to be excessive. In other cases, by

mandating fee ceilings, authorities expect a greater number of merchants to adopt payment cards instead of cash.[7] Alternatively, some policymakers argue that lowering card issuers' interchange revenue may reduce incentives to cardholders to use more costly payment cards (for example, credit cards instead of debit cards).

The rest of the article is structured as follows. In the next section, I discuss externalities in payment card markets in the context of theoretical models. I also explore two externalities that have been less researched. In the following section, I investigate market interventions, along with the motivation of the authorities for such interventions and whether they met their objectives. Finally, I offer some concluding remarks.

I. Externalities

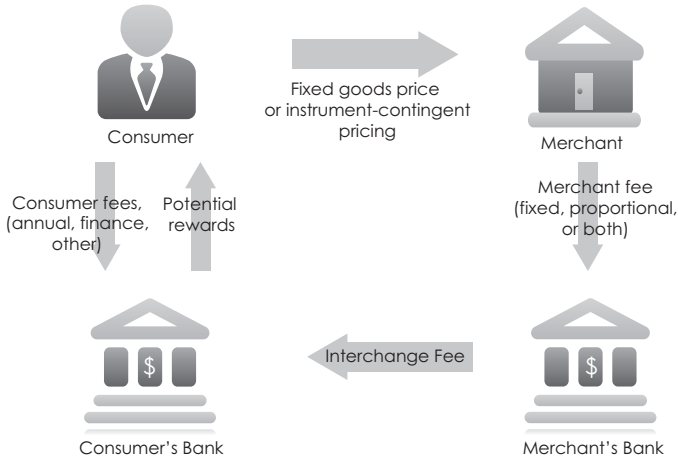
Before discussing the externalities present in payment card networks, let us review the key participants and the monetary transfers among them. Payment networks comprise consumers (more generally, buyers) and their banks (known as issuers), as well as merchants (more generally, sellers) and their banks (known as acquirers), along with the network operator and other participants that facilitate these transactions. Payment card transactions involve a set of interrelated bilateral transactions. First, a consumer establishes a relationship with an issuer and receives a payment card.[8] Second, a consumer makes a purchase from a merchant. Third, if a merchant has established a relationship with an acquirer, the merchant is able to accept payment card transactions. Fourth, the acquirer receives payment from the issuer. A network operator facilitates these bilateral relationships.

In Figure 1, the four key participants and their monetary transfers are diagrammed. When the consumer establishes a relationship with a bank, she agrees to pay an annual fee if one is charged, finance charges if she borrows long term, and other fees. In addition, she may receive per transaction rewards to promote greater usage of the card. When the consumer uses her card to make a purchase, the merchant may impose an additional fee for card acceptance or pass on the cost to all consumers in the form of higher prices. To convert the payment card receipt into a bank deposit, the merchant pays a fee to its bank. In addition to per transaction fees that may be fixed or proportional to the amount of the purchase, the merchant may also pay fixed fees. The merchant's bank pays interchange fees to the cardholder's bank. In this section, I study the effect of a bilateral payment transfer on other bilateral relationships in the network and potential externalities that might arise.

A. Adoption and usage externalities

The two-sided market literature has been used to analyze the structure of fees paid by consumers and merchants. Payment networks are one type of two-sided market.[9] Other types of two-sided market platforms include computer game platforms, newspapers, and online dating sites. These platforms provide goods and services to two or more distinct sets of end-users and must convince all sides to

Figure 1
Payment Card Fees



participate. The price structure or balance is the share that each type of end-user pays of the total price of the payment service.

This literature combines the multiproduct firm literature, which studies how firms set prices on more than one product, with the network economics literature, which studies how consumers benefit from increased participation in networks by other consumers.[10] Rochet and Tirole (2006b) define a two-sided market as a market where end-users are unable to negotiate prices based on costs to participate on a platform and the price structure affects the total volume of transactions.

A key externality examined in the payment card literature is the ability of the network to convince both consumers and merchants to participate in a network. Initially, the literature focused on per transaction fees and ignored fixed costs. In such an environment, there is no distinction between adoption and usage. Baxter (1983) argues that the equilibrium quantity of payment card transactions occurs when the total transactional demand for payment card services, which are determined by consumer and merchant demands jointly, is equal to the total transactional cost for payment card services, including both issuer and acquirer costs, or:[11]

$$f + m = c_I + c_A,$$

where f is the willingness to pay for a consumer, m is the willingness to pay for a merchant when demand for payment services equals the supply of payment services and c_I and c_A are the issuer's marginal cost and the acquirer's marginal cost, respectively. A consumer's willingness to pay is based on her net benefits received. The consumer will participate when her net benefit is greater than or equal to the fee in equilibrium.[12] Similarly, if the merchants' fee, m, is less than or equal to the

net benefits it receives, merchants will accept cards. Pricing each side of the market based on marginal cost—as would be suggested by economic theory for one-sided competitive markets—need not yield the socially optimal allocation. To arrive at the socially optimal equilibrium, a side payment may be required between the issuer and acquirer.

Schmalensee (2002) extends Baxter's (1983) analysis by considering issuers and acquirers that have market power, but still assumes that merchants operate in competitive markets. His results support Baxter's conclusions that the interchange fee balances the demands for payment services by each end-user type and the cost to banks to provide them. Schmalensee finds that the profit-maximizing interchange fee of issuers and acquirers may also be socially optimal.[13]

Given the simultaneous consumption of payment services by consumers and merchants, a side payment may be necessary to get both sides on board if there are asymmetries of demand between consumers and merchants and/or of costs to service consumers and merchants. This result is critically dependent on the inability of merchants to price discriminate between card users and those who do not use cards or among different types of card users. While most economists and antitrust authorities agree that an interchange fee may be necessary, the level of the fee remains a subject of debate.

B. Merchant competition

A common reason given by merchants when asked why they do not reject cards instead of paying high fees to the card networks for accepting them is that they would lose business to their competitors. Some merchants argue that merchants as a whole would be better off by not accepting certain types of payment cards. Some economic models have predicted that merchant competition may increase the ability of networks to set higher interchange fees.

Unlike Baxter (1983) and Schmalensee (2002), Rochet and Tirole (2002) consider strategic interactions of consumers and merchants.[14] They have two main results. First, the interchange fee that maximizes profit for the issuers may be more than or equal to the socially optimal interchange fee, depending on the issuers' margins and the cardholders' surplus. Second, merchants are willing to pay more than the socially optimal fee if they can steal customers from their competitors. However, overall social welfare does not improve when merchants steal customers from their competitors by accepting payment cards.

Wright (2004) extends Rochet and Tirole (2002) by considering a continuum of industries where merchants in different industries receive different benefits from accepting cards. His model is better able to capture the trade-off between consumer benefits and merchant acceptance when the interchange fee is increased because some merchants will not accept cards.[15] Wright concludes that the interchange fee that maximizes overall social welfare may be higher or lower than the interchange fee that maximizes the number of transactions.

These models suggest that merchant competition may actually lead to a greater ability by network operators to extract surplus from them. Furthermore, there is no systematic bias in the social-welfare-maximizing and profit-maximizing interchange fee. In the next section, I explore the ability of merchants to steer consumers to the merchant's preferred payment instrument by using price incentives.

C. Instrument-contingent pricing

The two-sided market literature assumes that end-users are not allowed to negotiate prices of platform services. In many jurisdictions, merchants are not allowed to add a surcharge for payment card transactions because of legal or contractual restrictions.[16] If consumers and merchants were able to negotiate prices based on differences in costs that merchants face and the benefits that both consumers and merchants receive, the interchange fee would be neutral, assuming full pass-through. The interchange fee is said to be neutral if a change in the interchange fee does not change the quantity of consumer purchases and the profit level of merchants and banks. Generally, the merchant charges the same price regardless of the type of payment instrument used to make the purchase. Frankel (1998) refers to merchants' reluctance to set different prices even when they are allowed to do so as price cohesion.

Even if price differentiation based on the payment instrument used is not common, the possibility to do so may enhance the merchants' bargaining power in negotiating their fees. Merchants can exert downward pressure on fees by having the possibility to set instrument-contingent pricing. Payment networks may prefer non-instrument-contingent pricing because some consumers may not choose payment cards if they had to explicitly pay for using them at the point of sale (POS).

Carlton and Frankel (1995) extend Baxter (1983) by considering when merchants are able to fully pass on payment processing costs via higher consumption goods prices. They find that an interchange fee is not necessary to internalize the externality if merchants set pricing for consumption goods based on the type of payment instrument used. Furthermore, they argue that cash users are harmed when merchants set one price because they subsidize card usage.

Schwartz and Vincent (2006) study the distributional effects among cash and card users with and without no-surcharge restrictions. They find that the absence of pricing based on the payment instrument used increases network profit and harms cash users and merchants.[17] The payment network prefers to limit the merchant's ability to separate card and cash users by forcing merchants to charge a uniform price to all of its customers. When feasible, the payment network prefers rebates (negative per transaction fees) given to card users.[18] Granting such rebates to card users boosts their demand for cards while simultaneously forcing merchants to absorb part of the corresponding rise in the merchant fee, because any resulting increase in the uniform good's price must apply equally to cash users. In this way, the network uses rebates to indirectly extract surplus from cash-paying customers in the form of higher prices.

Gans and King (2003) argue that, as long as there is "payment separation," the interchange fee is neutral regardless of the market power of merchants, issuers, and acquirers. When surcharging is costless, merchants will implement pricing based on the payment instrument used, taking away the potential for cross-subsidization across payment instruments and removing the interchange fee's role in balancing the demands of consumers and merchants. In effect, the cost pass-through is such that lower consumer card fees (due to higher interchange fees) are exactly offset by higher goods prices from merchants. Payment separation can occur if one of the following is satisfied: There are competitive merchants, and they separate into cash-accepting or card-accepting categories, in which each merchant only serves one type of customer and is prevented from charging different prices; or merchants are able to fully separate customers who use cash from those who use cards by charging different prices.

Wright (2003) finds that no-surcharge rules generate higher welfare than when monopolist merchants are allowed to set prices based on the payment instrument used. He argues that merchants are able to extract consumers' surplus ex post from payment card users, while cash users are unaffected. Wright only considers equilibria where merchants will continue to sell the same quantity of goods to cash users at the same price. When merchants are allowed to surcharge, they extract "too much" surplus ex post from customers who use payment cards because merchants set higher prices for card purchases.

Economic theory generally suggests that if merchants were able to recover their payment costs, the impact of the interchange fee would be severely dampened. However, the potential for merchants to charge more than their processing costs exists and consumer welfare could be harmed by such practices. The most interesting puzzle may be why merchants choose not to price differentiate even when they are allowed to do so. Some observers suggest that merchant competition may prevent price differentiation.

D. Network competition

Economic theory suggests that competition generally reduces prices, increases output, and improves welfare. However, with two-sided markets, network competition may yield an inefficient price structure. A key aspect of network competition is the ability of end-users to participate in more than one network. When end-users participate in more than one network, they are said to be "multihoming." If they connect only to one network, they are said to be "singlehoming." As a general finding, competing networks try to attract end-users who tend to singlehome, since attracting them determines which network has the greater volume of business. Accordingly, the price structure is tilted in favor of end-users who singlehome.[19] Even if consumers adopt more than one payment card, Rysman (2007) finds that consumers may have strong preferences to use only one of them.

Some models of network competition assume that the sum of consumer and merchant fees is constant and focus on the price structure.[20] Rochet and Tirole

(2003) find that the price structures for a monopoly network and competing platforms may be the same, and if the sellers' demand is linear, this price structure in the two environments generates the highest welfare under a balanced budget condition. Guthrie and Wright (2007) extend Rochet and Tirole (2003) by assuming that consumers are able to hold one or both payment cards and that merchants are motivated by "business stealing" when deciding to accept payment cards. They find that network competition can result in higher interchange fees than those that would be socially optimal.

Chakravorti and Roson (2006) consider the effects of network competition on total price and on price structure where networks offer differentiated products.[21] Like Rochet and Tirole (2003) and Guthrie and Wright (2007), they find that competition does not necessarily improve or worsen the balance of consumer and merchant fees from the socially optimal one. However, they find that the welfare gain from the drop in the sum of the fees from competition is generally larger than the potential decrease in welfare from less efficient fee structures.

Unlike one-sided markets, competition does not necessarily improve the balance of prices for two-sided markets. Furthermore, if competition for cardholders is more intense because consumers ultimately choose the payment instrument, issuers may provide greater incentives to attract them. If issuers have greater bargaining power to raise interchange fees, they can use this power to partially offset the cost of consumer incentives. I will discuss later the funding of rewards to entice more consumers in the context of the Reserve Bank of Australia's interchange fee regulation.

E. Surplus from revolvers

So far, among the models that I have discussed, the benefits of consumer credit are not considered.[22] Given the high level of antitrust scrutiny targeted toward credit card fees, including interchange fees, this omission in most of the academic literature is rather surprising. In the long run, aggregate consumption over consumers' lives may not differ because of access to credit, but such access may enable consumption smoothing that increases consumers' utility. In addition to extracting surplus from all consumers and merchants, banks may extract surplus from liquidity-constrained consumers.[23] How much surplus can be extracted depends on how much liquidity-constrained consumers discount tomorrow's consumption.

Chakravorti and Emmons (2003) consider the costs and benefits of consumer credit where consumers are subject to income shocks after making their credit card purchases and some are unable to pay their credit card debt.[24] To my knowledge, they are the first to link the insurance aspect of credit cards to their payment component. Observing that over 75 percent of U.S. card issuer revenue is derived from cash-constrained consumers, they consider the viability of the credit card system if it were completely funded by these types of consumers.[25] They find that if consumers sufficiently discount future consumption, liquidity-constrained consumers who do not default would be willing to pay all credit card network costs ex ante, resulting in all consumers being better off than a world with no credit cards.

However, they also find that the inability of merchants to impose instrument-contingent prices results in a lower level of social welfare because costly credit card infrastructure is used for transactions that do not require credit extensions.

Most of the payment card literature ignores consumer finance charges and other types of consumer fees, such as annual, over-the-limit, and cash advance fees. In the United States, the regulation of consumer fees on credit cards has increased and new restrictions have been implemented. Perhaps, with reduced revenue from these sources coupled with greater usage of debit cards, interchange fee revenue may become more critical. Of course, as mentioned previously, these fees continue to face regulatory pressure as well.

F. Merchant fees and consumer credit

Chakravorti and To (2007) consider a scenario with monopolist merchants and a monopolist bank that serves both consumers and merchants where the merchants absorb all credit and payment costs in a two-period dynamic model.[26] Their model yields the following results. First, the merchants' willingness to pay bank fees increases as the number of credit card consumers without income increases. Note that up to a point, merchants are willing to subsidize credit losses in exchange for additional sales. Second, a prisoner's dilemma situation may arise: Each merchant chooses to accept credit cards, but by doing so, each merchant's discounted two-period profit is lower. Unlike the merchants in the previous models, the merchants in this one do not sell the same type of goods and may enjoy significant market power. However, business stealing may occur across merchants that sell different goods across consumption periods.

G. Competition among payment instruments

Most of the payment card literature ignores competition between payment instruments.[27] Furthermore, much of the payment literature focuses on the intensive margin—how fees influence usage—instead of the extensive margin—how fees affect adoption—or does not distinguish the two.[28] Much of the policy debate is about market forces behind consumer choice and merchant acceptance among multiple types of payment instruments.

If consumers carry multiple types of payment instruments, merchants may be able to steer them away from more costly payment instruments. Rochet and Tirole (2007) argue that merchants may choose to decline cards after they have agreed to accept them. They define the "tourist test" as when the merchant accepts cards even when it can "effectively steer" the consumer to use another payment instrument. Rochet has often given the example of an experience that he had in southern Italy, where after having a meal, the restaurant claimed that its payment card terminal was broken and payment had to be made in cash.[29] After visiting a nearby ATM, Rochet paid the bill with cash. In this example, the merchant did not pass the tourist test. The restaurant figured out that being a gentleman, Rochet would not leave the bill unpaid. However, if the consumer is unable to access cash or another form of payment, the merchant would lose the sale.

Merchants may steer consumers through price incentives, if allowed to do so. Bolt and Chakravorti (2008a) study the ability of banks and merchants to influence the consumers' choice of payment instrument when they have access to three payment forms—cash, debit card, and credit card.[30] Unlike most two-sided market models, where benefits are exogenous, they explicitly consider how consumers' utility and merchants' profits increase from additional sales resulting from greater security and access to credit.

Bolt and Chakravorti's (2008a) key results can be summarized as follows. With sufficiently low processing costs relative to theft and default risk, the social planner sets the merchant fee to zero, completely internalizing the card acceptance externality.[31] The bank may also set the merchant fees to zero, but only if merchants are able to sufficiently pass on their payment fees to their consumers or if their payment fees are zero. If the real resource cost of payment cards is sufficiently high, the social planner sets a higher merchant fee than the bank does, resulting in lower card acceptance and higher cash usage. Bolt and Chakravorti (2008a) find that bank profit is higher when merchants are unable to pass on payment costs to consumers because the bank is better able to extract merchant surplus. The relative costs of providing debit and credit cards determine whether the bank will provide both or only one type of payment card.

H. Payment fraud and liability

An aspect of payment networks that has received little attention in the payment network literature is the incentive that each participant has in maintaining the integrity and safety of the system as a whole. An externality arises if one participant on account of negligence and lack of incentives allows a fraudster to gain access to information that may be used to make fraudulent purchases.[32]

For example, consumers often face no liability for fraudulent transactions if proper procedures are followed for payment card transactions. While such a liability waiver encourages greater usage of cards vis-à-vis other payment instruments with less protection, it may also have the unintended consequence of consumers not maintaining appropriate antifraud precautions.[33] Primarily because of this liability shift, the card networks have implemented various fraud prevention strategies, such as real-time verification, the ability to shut down accounts rapidly, and the tracking of spending patterns of cardholders over the last few decades.[34] While U.S. issuers and networks limit consumer liability, consumers may bear losses associated with fraudulent transactions if they do not adopt risk-reducing procedures in other countries. For example, an Italian banker explained to me that most Italian banks shift the liability back to consumers if they do not use the recommended security procedures for Internet card payments. Merchants also enjoy certain protections (though more limited than those for consumers) if they follow set guidelines when accepting payment cards.

Similarly, the lack of merchant and processor data security measures may pose negative externalities. For example, while the cost of not protecting payment

information for an individual entity may be small, its impact on the system as a whole may be significant. Recently, the industry has been exploring various procedures to reduce this risk.

Market participants have expressed the view that better enforcement of current laws regarding payment fraud and greater adoption of existing industry-wide standards would greatly aid in reducing and containing fraud. Some observers have suggested that public authorities should establish standards, provide mechanisms for sharing information on data breaches, and formulate appropriate responses when wide-scale fraud occurs. Understandably, market participants may be reluctant to share or publicize breaches because of the potential loss in future business.

I. Dynamic efficiency and innovation

Dynamic efficiency and innovation have generally been ignored by economists and policymakers. Some market participants have argued that positive profits are necessary for payment networks to innovate. In other words, regulatory solutions to correct "excessive" interchange fees by using a cost-based approach may stifle future innovation. When general-purpose payment cards were first introduced, issuers and networks faced significant losses and many left the industry to only return later, suggesting that investments in new products and processes may require significant time to recover.

Historically, the card networks have been more innovative than other payment networks, such as those that process checks. In the United States, a law had to be passed relatively recently to facilitate the widespread acceptance of substitute checks instead of the original physical check enabling rapid migration to the truncation of physical checks. In contrast to the networks processing checks, credit card networks were exchanging payment information electronically for more than two decades. In addition, the card networks established real-time authorization systems in the 1970s to combat payment fraud.[35] Interestingly, fees charged by third parties to guarantee checks are pretty close to or sometimes higher than merchant fees for credit cards. When similar protections against payment default are included for checks, the cost of check acceptance with similar protections converges to the cost of payment card acceptance, suggesting that payment instruments may differ with respect to the benefits to merchants. Furthermore, some merchants may be willing to forgo certain benefits because of the type of customers that they serve.

II. Market Interventions

Policymakers in different jurisdictions are encouraging the replacement of cash and checks with electronic substitutes, such as payment cards at the point of sale.[36] In some U.S. municipalities, acceptance of payment cards for cab rides has been mandated. A primary reason cited is the safety of passengers and cab drivers (who are often the targets of muggings). In Mexico, the government gave away terminals to merchants to increase the acceptance of payment cards versus cash (Castellanos et al., 2008). However, forced acceptance of payment cards and

government-subsidized merchant terminals are not common. In this section, I explore several market interventions in various jurisdictions and study the impact of those interventions.[37]

A. Removal of no-surcharge policies

There are several jurisdictions where merchants are able to impose surcharges. Some of the academic research cited previously suggests that if merchants are allowed to surcharge, the level of the interchange fee would be neutral. In this section, I discuss examples where merchants are able to post differentiated prices.

The Australian authorities were concerned about the substitution of credit cards by debit cards; they argued that consumers did not receive the proper price incentives to use debit cards, the less costly payment instrument. The Reserve Bank of Australia (RBA) reported that the average cost of the payment functionality of the credit card was AUS$0.35 higher than a debit card using a consistent AUS$50 transaction size.[38] To encourage better price signals, the RBA removed no-surcharge restrictions in 2002.

While most Australian merchants do not impose surcharges for any type of payment card transaction today, the number of merchants who do are increasing. At the end of 2007, around 23 percent of very large merchants and around 10 percent of small and very small merchants imposed surcharges. Large merchants surcharged around 15 percent of the time. The average surcharge for MasterCard and Visa transactions is around 1 percent, and that for American Express and Diners Club transactions is around 2 percent (Reserve Bank of Australia, 2008a).[39] Using confidential data, the Reserve Bank of Australia (2008a) also found that if one network's card was surcharged more than other networks' cards, consumers dramatically reduced their use of the card with the surcharge. After analyzing consumer surveys, the Reserve Bank of Australia (2008a) noted that nearly 40 percent of credit card convenience users (that is, credit card users who do not need credit to make purchases) did not use a debit card during the time of the survey; this suggests that using credit cards is still preferred by many of those who do not need to borrow.[40]

Some economists have stressed that merchants may surcharge consumers more than their costs. A potential regulatory response is to cap the surcharge. In responding to the 2007/08 review of reforms by the Reserve Bank of Australia, some market participants suggested that merchants might be imposing higher surcharges than their cost to accept payment cards. The RBA has considered setting a limit for the surcharge amount but has not gone ahead with implementing one.

In the United States, merchants are allowed to offer cash discounts but may not be allowed to surcharge credit card transactions. In the 1980s, many U.S. gas stations explicitly posted cash and credit card prices. Barron, Staten, and Umbeck (1992) report that gas station operators imposed these policies when their credit card processing costs were high but later abandoned these policies when acceptance costs

decreased because of new technologies such as electronic terminals at the point of sale. Recently, some gas stations brought back price differentiation based on payment instrument type, citing the rapid rise in gas prices and declining profit margins.

In the Netherlands, Bolt, Jonker, and van Renselaar (2009) study the impact of debit card surcharges. They report that a significant number of merchants are setting different prices, depending on whether cash or a debit card is used. Debit card surcharges are widely assessed when purchases are below 10 euro, suggesting that merchants are unwilling to pay the fixed transaction fee below this threshold. Bolt, Jonker, and van Renselaar find that merchants may surcharge up to four times their fee. In addition, when these surcharges are removed, they argue, consumers start using their debit cards for these small payments, suggesting that merchant price incentives do affect consumer payment choice. Interestingly, in an effort to promote a more efficient payment system, the Dutch central bank has supported a public campaign to encourage retailers to stop surcharging to encourage consumers to use their debit cards for small transactions.

There are instances when card payments were discounted vis-à-vis cash payments. During the conversion to the euro from national currencies, one German department store offered discounts for using cards because of the high initial demand for euro notes and coins to make change for cash purchases (Benoit, 2002). It should be noted, however, that the retailer was in violation of German retailing laws for doing this. In a more permanent move, the Illinois Tollway charges motorists who use cash to pay tolls twice as much as those who use toll tags (called I-PASS), which may be loaded automatically with credit and debit cards when the level of remaining funds falls below a certain level.[41] In addition to reducing cash handling costs, the widespread implementation of toll tags decreased not only congestion at toll booths but also pollution from idling vehicles waiting to pay tolls, since tolls could be collected as cars drove at highway speeds through certain points on the Illinois Tollway. In both of these cases, the benefits of using cards outweighed the costs for society in general. However, benefits from card acceptance vary considerably across merchants.

B. Regulation of interchange fees

There are several jurisdictions where interchange fees were directly regulated or significant pressure was exerted by the public authorities on networks to reduce their interchange fees. In this section, I will discuss the impact of interventions in three jurisdictions—Australia, Mexico, and Spain.

Concluding that surcharges alone would not put sufficient downward pressure on interchange fees, the Australian authorities imposed explicit interchange fee targets for the two large four-party payment networks—MasterCard and Visa—but did not impose any restrictions on three-party networks—American Express and Diners Club.[42] In 2002, the RBA imposed weighted-average credit card interchange fee caps and later imposed per transaction targets for debit cards.

As of April 2008, the weighted-average credit card interchange fees in the MasterCard and Visa networks must not exceed 0.50 percent of the value of transactions. The Visa debit weighted-average interchange fee cap must not exceed 12 cents (Australian) per transaction. The EFTPOS (electronic funds transfer at point of sale) interchange fees for transactions that do not include a cash-out component must be between 4 cents (Australian) and 5 cents (Australian) per transaction.

The Reserve Bank of Australia (2008a) reports that the interchange fee regulation, coupled with the removal of the no-surcharge rule, improved the price signals that consumers face when deciding which payment instruments to use. Specifically, annual fees for credit cards increased and the value of the rewards decreased. The Reserve Bank of Australia (2008a) calculates that for an AUS$100 transaction, the cost to consumers increased from –AUS$1.30 to –AUS$1.10 for consumers who pay off their balances in full every month. A negative per transaction cost results when card benefits such as rewards and interest-free loans are greater than payment card fees.[43]

In its recent five-year review of their payment card policies, the Australian Payments System Board suggested that the explicit regulation of interchange fees be removed subject to certain conditions. In other words, the authorities will remove restrictions if the payment card networks do not raise their fees beyond some threshold. However, the actual threshold is not quantified.

Those who oppose the Australian interchange fee regulation argue that consumers have been harmed by reduced rewards and higher fees and have not shared in the cost savings—in terms of lower prices for goods and services. However, measuring price effects over time of interchange fee regulation is difficult.

Another interesting case where government authorities exerted pressure to decrease interchange fees occurred in Mexico.[44] Similar to the RBA in Australia, the Bank of Mexico—the Mexican central bank—has the authority to regulate retail payment systems throughout the country. Unlike the RBA, the Bank of Mexico used moral suasion to reduce interchange fees. The motivation of the Mexican authorities to reduce interchange fees was to reduce merchant fees that were preventing greater adoption and usage of payment cards in Mexico.

Mexico's Bank Association (ABM) set different interchange fees for debit and credit cards in August 2004; prior to this time, the fees were the same for both types of cards. Interchange fees were set based on a merchant's monthly transaction volume. By August 2005, the debit card interchange fee for the largest merchants fell from 2.00 percent to 0.75 percent while the credit card interchange fee fell from 2.00 percent to 1.80 percent. The category that applied to the smallest merchants was eliminated; as a consequence the interchange fee of this group fell from 3.50% to 1.95% and 3.50% to 2.70% for debit and credit cards, respectively. The ABM also proposed interchange fees based on a formula where the interchange fee balances out the issuing and acquiring banks' profits (net of interchange), and where profits are normalized by revenue (net of interchange). A reference rate is

obtained and specific interchange fee levels are calculated for a number of merchant categories using proxies of the demand elasticity for each category.

In 2008, ABM further reduced debit and credit card interchange fees. The new IF levels implied a reduction in the weighted average of 12.5% and 9% for credit and debit, respectively.[45] As expected, merchant fees also decreased. In order to follow the evolution of merchant fees, Bank of Mexico gathered information from a sample of 1000 firms that accepted card payments. The results are that from 2005 to 2008, the average merchant discount rate has decreased 12.3% and 23.3% for credit and debit, respectively.[46] In addition, the installation of POS terminals was subsidized through a private, nonprofit trust fund called FIMPE that was initially funded by the banks. The banks received a tax credit from the government for their investment. It is important to note that there may be significant fixed and variable costs. As a result, the number of POS terminals installed increased to 446,025 by the end of 2008 compared to 129,971 in 2002. POS transactions increased from 52 million in 2002 to 215 million by the end of 2008 of which 46% were credit card transactions.

Unlike in Australia or Mexico, the antitrust authority, and not the central bank, intervened in payment card markets in Spain. Part of the motivation was based on directives by the European Commission regarding fees that were set by networks that had significant market power. Over the period 1997-2007, the number of debit cards increased by 40.9 percent and the number of credit cards increased by 207.1 percent. During the same period, debit card transactions increased from 156 million to 863 million and credit card transactions increased from 138 million to 1.037 billion. Furthermore, the average number of POS transactions per card per year increased from 7.1 to 27.8 during the same period.

The first intervention occurred in May 1999, when the Spanish government convinced the three Spanish payment card networks to gradually reduce maximum interchange fees from its initial value of 3.5 percent to 2.75 percent by July 2002. These maximum fees varied significantly across merchant categories.

In April 2002, Spain's antitrust authority requested the Spanish networks to provide information on how they determined their interchange fees. From 2003 until 2005, several attempts from the industry to maintain their "special authorization" for the setting of interchange fees were refused. Eventually, the networks were requested to set levels of interchange fees that only reflected operating costs and those due to fraud. In December 2005, the Ministry of Industry, Tourism, and Trade decided that the multilateral interchange fees should not exceed the costs to provide card services.

From January 2006 to December 2008, the highest interchange fee levels were reduced in a stepwise manner. Furthermore, a distinction had to be made between debit card and credit card interchange fees, with the former being a fixed amount per transaction and the latter being a percentage amount per transaction. For merchants with an annual value of less than 100 million euro in POS card payment

receipts, the credit card interchange fee was set to decrease from 1.40 percent per transaction in 2006 to 0.35 percent in 2009; for those same merchants, the debit card interchange fees (regardless of the purchase amount) were reduced from 0.53 euro per transaction in 2006 to 0.35 euro per transaction in 2009. These fees are the maximum allowable, and in some cases the actual fees are lower. Additionally, price differences between debit cards and credit cards, merchant sectors, and intrasystem and intersystem operations should also be progressively reduced.

Carbó Valverde, Chakravorti, and Rodriguez Fernandez (2009) study the effects of interchange fee reductions in Spain from 1997 to 2007. To my knowledge, they are the first to use bank-level data to study the impact of several episodes of interchange fee reductions for debit and credit cards resulting from moral suasion and direct regulation. They find that intense issuer competition coupled with high interchange fees may have made consumers, merchants, and banks worse off. Clearly, merchants benefit from lower fees and consumers benefit when more merchants accept payment cards if the benefit of greater acceptance outweighs any additional cost to payment providers. Surprisingly, they find that revenues increase among the banks in their sample, even though interchange fees decreased. While the effect of these reductions is positive on banks' revenues, their effect on banks' profits could not be determined because of data limitations. Furthermore, there may be a critical interchange fee below which issuer revenue decreases. Unfortunately, their data does not allow them to find this critical interchange fee. Additionally, in the absence of adoption and usage externalities, the level of the interchange fee may not affect social welfare.

C. Honor-all-cards rules

A payment card network may require that merchants that accept one of its payment products accept all of its products. There are different forms of the honor-all-cards rule. The honor-all-cards rule may extend to any payment card that is issued by a member of a network. In other words, if a merchant accepts a network's credit card, it must accept all debit and prepaid cards from that network. Such a rule enables a card network to innovate by producing different products that when introduced will have a large base of merchants that accept them. The introduction of payroll cards, a type of prepaid card, is an example of an innovation that leverages a card network's existing infrastructure.

In the United States, around 5 million merchants sued the two major networks, MasterCard and Visa, over the required acceptance of the network's signature-based debit card when accepting the same network's credit card. The case was settled out of court. In addition to a monetary settlement, MasterCard and Visa agreed to decouple merchants' acceptance of their debit and credit products. While few merchants have declined one type of card and accepted another type, the decoupling of debit and credit card acceptance may have increased bargaining power for merchants in negotiating fees.

As part of the payment system reforms in Australia, MasterCard and Visa were mandated to decouple merchants' acceptance of their debit and credit cards as well. The Payments System Board (Reserve Bank of Australia, 2008b, 16) is unaware of any merchant that continues to accept debit cards but does not accept credit cards from the same network.

A subset of the honor-all-cards rule is the honor-all-issuers rule. In other words, if a merchant accepts a credit card from one issuer, it must also accept credit cards from another issuer within the same network. Such a policy levels the playing field between large and small issuers through a base product, which each issuer can customize. Otherwise, small issuers would not be able to compete with the large issuers. Larger issuers also benefit from the underlying network effects.

Another type of honor-all-cards rule could cover the acceptance of different credit or debit cards from the same issuer. For example, issuers may have a plain vanilla credit card and also have others that earn different types of rewards. While merchants may not care what types of rewards their customers receive from their banks, merchants may pay different fees based on the type of card used by their customers. More recently, policymakers are considering allowing merchants to discriminate within a card classification, such as a credit card, based on differences in interchange fees.

III. Conclusion

In summarizing the payment card literature, I find that no one model is able to capture all the essential elements of the market for payment services. It is a complex market with many participants engaging in a series of interrelated bilateral transactions. Much of the debate over various payment card fees is concerned with the allocation of surpluses from consumers, merchants, and banks, as well as the question of who is able to extract surpluses from whom.

I am able to draw the following conclusions. First, a side payment between the issuer and the acquirer may be required to get both sides on board. However, there is no consensus among policymakers or economists on what constitutes an efficient fee structure for card payments. Second, while consumers generally react to price incentives at the point of sale, merchants may be reluctant to charge higher prices to consumers who benefit from card use. However, surcharging is increasing in jurisdictions where it is allowed. Third, network competition may not improve the price structure but may significantly reduce the total price paid by consumers and merchants. Fourth, both consumers and merchants value credit extended by credit card issuers (along with other benefits such as security), and consumers and merchants are willing to pay for it. Fifth, evidence from recent interventions suggests that market-based fees may not maximize social welfare.

Determining sound public policy regarding the allocation of payment fees is difficult. The central question is whether the specific circumstances of payment

markets are such that intervention by public authorities can be expected to improve economic welfare. Efficiency of payment systems is measured not only by the costs of resources used, but also by the social benefits generated by them. Clearly, further research is warranted to explore the complex market for payment services, and policy recommendations should be based on more in-depth research, especially empirical studies that focus on the effects of government intervention.

Author's Note: I thank Wilko Bolt, Santiago Carbó Valverde, Bill Emmons, Emery Kobor, Francisco Rodriguez Fernandez, Roberto Rosen, and Ted To for shaping my views over the years regarding the economics of retail payment networks. I also thank participants at "The Changing Retail Payments Landscape: What Role for Central Banks," held at the Federal Reserve Bank of Kansas City. I also thank Anna Lunn for excellent research assistance and Han Choi for suggestions to improve the article's readability. The views expressed are my own and do not necessarily reflect those of the Federal Reserve Bank of Chicago or the Federal Reserve System.

Endnotes

[1]Amromin and Chakravorti study 13 countries—Austria, Belgium, Canada, Finland, France, Germany, Italy, Japan, the Netherlands, Sweden, Switzerland, the United Kingdom, and the United States.

[2] There are countries, for example, France, where the cardholder's account is debited much later. These types of cards are referred to as "delayed debit cards." Furthermore, many U.S. debit card issuers extend credit lines as well, primarily as overdraft protection. For more discussion, see Chakravorti (2007).

[3]For a discussion of the economics of prepaid cards, see Chakravorti and Lubasi (2006).

[4]For a summary of antitrust challenges in various jurisdictions, see Bradford and Hayashi (2008).

[5]In Australia, the interchange fee for debit card transactions is paid by the card issuer (banks that issue cards to consumers) to the acquirer (banks that convert payment card receipts into bank deposits for merchants), but this is an exception.

[6]Rochet and Tirole (2006a) provide an overview of some externalities in card systems that I cover in this article.

[7]In addition to cash handling and safekeeping costs, some public authorities may find the inability to trace cash transactions an unattractive feature of cash.

[8]In the case of prepaid cards, the identity of cardholders may not be known to the issuer, but there still exists a relationship.

[9]For a review of the academic literature on two-sided payment networks, see Bolt and Chakravorti (2008b).

[10]For a more general treatment of two-sided markets, see Armstrong (2006), Caillaud and Jullien (2003), Jullien (2001), Rochet and Tirole (2006b), Rysman (2009), and Weyl (2009).

[11]Baxter (1983) considers an environment where consumers are homogeneous, merchants are perfectly competitive, and the market for issuing and acquiring payment cards is competitive.

[12]Net benefits for consumers and merchants are defined by the difference in benefits from using a payment card and using an alternative payment instrument.

[13]Schmalensee defines the socially optimal interchange fee as the one that maximizes the sum of the consumer and merchant surplus. Such a measure is appropriate if card acceptance is not used as a strategic tool to steal customers from another merchant.

[14]Rochet and Tirole consider two identical Hotelling merchants in terms of their net benefits of accepting a payment card for sales and the goods that they sell. Consumers face the same fixed fee but are heterogeneous in terms of the net benefits they derive from using the payment card. They assume that the total number of transactions is fixed and changes in payment fees do not affect the demand for consumption goods.

[15]In Wright's environment, both consumer and merchant fees are per transaction fees. Each consumer buys goods from each industry. Issuers and acquirers operate in markets with imperfect competition. Wright assumes that consumers face the same price regardless of which instrument they use to make the purchase.

[16]No-surcharge restrictions do not allow merchants to impose surcharges for payment card purchases. However, merchants may be allowed to offer discounts for noncard payments. For more discussion about no-surcharge rules and discounts, see Chakravorti and Shah (2003).

[17]Schwartz and Vincent relax the common assumption made in the literature that the demand for the consumption good is fixed. However, they assume that consumers are exogenously divided into cash and card users and cannot switch into the other group.

[18]In this context, a rebate is an incentive for consumers to use their cards—for example, cash back and other frequent-use rewards.

[19]For more discussion, see Evans (2003).

[20]The motivation behind this assumption was based on the earlier cooperative structure of the two large networks. However, the two largest networks changed their structure from associations to for-profit firms.

[21]Chakravorti and Roson only allow consumers to participate in one card network, whereas merchants may choose to participate in more than one network. However, unlike Guthrie and Wright (2007) and Rochet and Tirole (2003), Chakravorti and Roson consider fixed fees for consumers. They compare welfare properties when the two networks operate as competitors and as a cartel, where each network retains demand for its products from end-users but the networks set fees jointly.

[22]I limit my focus here to consumption credit. Payment credit—the credit that is extended by the receiver of payment or by a third party until it is converted into good funds—is ignored. For more discussion, see Chakravorti (2007).

[23]The empirical literature on credit cards has suggested interest rate stickiness along with above-market interest rates, although some have argued that the rate is low compared with alternatives such as pawn shops. For more discussion, see Ausubel (1991) and Brito and Hartley (1995).

[24]All markets for goods and payment services are assumed by Chakravorti and Emmons to be competitive. Chakravorti and Emmons impose a participation constraint on individuals without liquidity constraints such that the individuals will only use cards if they are guaranteed the same level of consumption as when they use cash including the loss of consumption associated with higher prices for consumption goods.

[25]For a breakdown of issuer revenue percentages, see Green (2008).

[26]Chakravorti and To depart from the payment card literature in the following ways. First, similar to Chakravorti and Emmons (2003), rather than taking a reduced-form approach where the costs and benefits of payment cards are exogenously assigned functional forms, they construct a model that endogenously yields costs and benefits to consumers, merchants, and banks from credit card use. Second, their model considers a dynamic setting where there are intertemporal tradeoffs for all participants. Third, they consider consumption and income uncertainty.

[27]Farrell (2006) studies the impact of higher interchange fees on consumers who do not use cards. While the redistributive effects generally do not affect social welfare,

he argues that the impact of pricing of a payment instrument in one network affecting the usage of other payment instruments should be considered by policymakers.

[28]Bedre and Calvano (2009), Bolt and Chakravorti (2008a), and Chakravorti and Roson (2006) are notable exceptions.

[29]I have often had similar experiences at the end of cab rides when I try to pay with my credit card and the driver chooses not to accept it, even though there are multiple signs stating that credit cards are accepted.

[30]In Bolt and Chakravorti's model, consumers only derive utility from consuming goods from the merchant they are matched to. In addition, some consumers prefer to consume before their income arrives. Merchants differ on the types of payment instruments that they accept and type of consumption good they sell. Each merchant chooses which instruments to accept based on its production costs, and each merchant is categorized as cash only, cash and debit card, or full acceptance (cash, debit card, and credit card). Merchant heterogeneity is based on differences in production costs. Bolt and Chakravorti consider the merchants' ability to pass on payment processing costs to consumers in the form of higher uniform and differentiated goods prices.

[31]While default rates and theft will differ across countries, Bolt and Chakravorti provide some estimates. For Italy, Alvarez and Lippi (2009) estimate the probability of being pickpocketed at around 2 percent in 2004. For the United States, Scholtes (2009) reported that credit card default rates hit a record of more than 10 percent in June 2009.

[32]See Amromin and Porter (2009) and Braun et al. (2008).

[33]See Douglass (2009).

[34]See Nocera (1994).

[35]For more discussion about innovations in the payment card market, see Chakravorti and Kobor (2005), Evans and Schmalensee (1999), and Nocera (1994).

[36]In the United States, some payment providers have introduced decoupled debit as a competitor to traditional payment cards. These types of payments use the automated clearinghouse (ACH) network to transfer funds from consumers to merchants for point of sale transactions.

[37]Prager et al. (2009) review the U.S. payment card market and consider potential regulations.

[38]Reserve Bank of Australia (2008a), 17.

[39]Note that in other jurisdictions, card networks may prevent merchants from imposing different surcharges on credit cards from different networks.

[40]Of course, even those credit card users who pay off their balances every month may benefit from short-term loans because of timing asymmetries between their incomes and purchases.

[41]For more discussion, see Amromin, Jankowski, and Porter (2007).

[42]In four-party networks, the issuer and the acquirer need not be the same. In three-party networks, the issuer and acquirer are the same resulting in no explicit interchange fee between issuers and acquirers.

[43]For more discussion about the effect of rewards on card use, see Carbó-Valverde and Liñares-Zegarra (2009) and Ching and Hayashi (2006).

[44]My discussions with Bank of Mexico staff, especially José Luis Negrín, were critical to my understanding of the Mexican payment card market.

[45]The weighted average interchange fee for credit cards decreased from 1.84 percent to 1.61 percent and for debit cards decreased from .78 percent to .71 percent.

[46]From 2005 to 2008, the average merchant fee decreased from 2.85 percent to 2.50 percent and the average debit merchant fee declined from 2.53 percent to 1.94 percent.

References

Alvarez, Fernando, and Francesco, Lippi (2009), "Financial innovation and the transactions demand for cash," *Econometrica*, 77 (2), 363-402.

Amromin, Gene, and Sujit Chakravorti (2009), "Whither loose change?: The diminishing demand for small denomination currency," *Journal of Money, Credit, and Banking*, 41 (2-3), 315-335.

Amromin, Gene, Carrie Jankowski, and Richard D. Porter (2007), "Transforming payment choices by doubling fees on the Illinois Tollway," *Economic Perspectives*, Federal Reserve Bank of Chicago, 31 (2), 22-47.

Amromin, Gene, and Richard D. Porter (2009), "*Economic Perspectives* special issue on payments fraud: An introduction," *Economic Perspectives*, Federal Reserve Bank of Chicago, 33 (1), 2-6.

Armstrong, Mark (2006), "Competition in two-sided markets," *RAND Journal of Economics*, 37 (3), 668-691.

Ausubel, Lawrence M. (1991), "The failure of competition in the credit card market," *American Economic Review*, 81 (1), 50-81.

Barron, John M., Michael E. Staten, and John Umbeck (1992), "Discounts for cash in retail gasoline marketing," *Contemporary Policy Issues*, 10 (4), 89-102.

Baxter, William F. (1983), "Bank interchange of transactional paper: Legal and economic perspectives," *Journal of Law and Economics*, 26 (3), October, 541-588.

Bedre, Ozlem, and Emilio Calvano (2009), "Pricing payment cards," mimeo, Toulouse School of Economics and Harvard University.

Benner, Katie (2008), "Visa's Record IPO rings up 28 percent gain," CNNMoney.com, March 19, available at *http://money.cnn.com/2008/03/19/news/companies/visa_ipo_opens.fortune/index.htm.*

Benoit, B. (2002), "Defiant C&A reignites debate on German shopping laws," *Financial Times*, January 9, 2.

Bolt, Wilko, and Sujit Chakravorti (2008a), "Consumer choice and merchant acceptance of payment media," Federal Reserve Bank of Chicago Working Paper, WP-2008-11.

Bolt, Wilko, and Sujit Chakravorti (2008b), "Economics of payment cards: A status report," *Economic Perspectives*, Federal Reserve Bank of Chicago, 32 (4), Fourth Quarter, 15-27.

Bolt, Wilko, Nicole Jonker, and Corry Van Renselaar (2009), "Incentives at the counter: An empirical analysis of surcharging card payments and payment behavior in the Netherlands," *Journal of Banking and Finance*, forthcoming.

Bradford, Terri, and Fumiko Hayashi (2008), "Developments in interchange fees in the United States and abroad," *Payments System Research Briefing*, Federal Reserve Bank of Kansas City, April.

Braun, Michele, James McAndrews, William Roberds, and Richard Sullivan (2008), "Understanding risk management in emerging retail payments," *Economic Policy Review*, Federal Reserve Bank of New York, 14 (2), 137-159.

Brito, Dagobert L., and Peter R. Hartley (1995), "Consumer rationality and credit cards," *Journal of Political Economy*, 103 (2), 400-433.

Caillaud, Bernard, and Bruno Jullien (2003), "Chicken and egg: Competition among intermediation service providers," *RAND Journal of Economics*, 34 (2), 309-328.

Carbó-Valverde, Santiago, Sujit Chakravorti, and Francisco Rodriquez Fernandez (2009), "The costs and benefits of interchange fee regulation: An empirical investigation," mimeo, Federal Reserve Bank of Chicago and University of Granada.

Carbó-Valverde, Santiago, and José Liñares-Zegarra (2009), "How effective are rewards programs in promoting payment card usage?" *BBVA Foundation Working Paper Series 1*.

Carlton, Dennis W., and Alan S. Frankel (1995), "The antitrust economics of credit card networks," *Antitrust Law Journal*, 63 (2), 643–668.

Castellanos, Sara, Ricardo Medina, Alberto Mendoza, José Luis Negrín, Francisco Solis, and Jean-Charles Rochet (2008), "The role of interchange fees in Mexico's retail payment system: From theory to practice," Washington, D.C.: World Bank.

Chakravorti, Sujit (2007), "Linkages between consumer payments and credit," in *Household Credit Usage: Personal Debt and Mortgages*, Sumit Agarwal and Brent W. Ambrose (eds.), New York: Palgrave MacMillan, 161-174.

Chakravorti, Sujit, and William R. Emmons (2003), "Who pays for credit cards?" *Journal of Consumer Affairs*, 37 (2), 208-230.

Chakravorti, Sujit, and Emery Kobor (2005), "Why invest in payment innovations?" *Journal of Payment Systems Law*, 1 (4), 331-353.

Chakravorti, Sujit, and Victor Lubasi (2006), "Payment instrument choice: The case of prepaid cards," *Economic Perspectives*, Federal Reserve Bank of Chicago, 30 (2), 29-43.

Chakravorti, Sujit, and Roberto Roson (2006), "Platform competition in two-sided markets: The case of payment networks," *Review of Network Economics*, 5 (1), 118-143.

Chakravorti, Sujit, and Alpa Shah (2003), "Underlying incentives in credit card networks," *Antitrust Bulletin*, 48 (1), 53-75.

Chakravorti, Sujit, and Ted To (2007), "A theory of credit cards," *International Journal of Industrial Organization,* 25 (3), 583-595.

Ching, Andrew, and Fumiko Hayashi (2006), "Payment card rewards programs and consumer payment choice," Payments System Research, Federal Reserve Bank of Kansas City, 06-02.

Douglass, Duncan B. (2009), "An examination of the fraud liability shift in consumer card-based payment systems," *Economic Perspectives*, Federal Reserve Bank of Chicago, 33 (1), 43-49.

Enrich, David (2006), "Visa plans restructuring, sees IPO in 12-18 mos.," *Dow Jones Newswires*, October 11.

Evans, David S. (2003), "The antitrust economics of multi-sided markets," *Yale Journal on Regulation*, 20 (2), 325-381.

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

Farrell, Joseph (2006), "Efficiency and competition between payment instruments," *Review of Network Economics*, 5 (1), 26-44.

Frankel, Alan S. (1998), "Monopoly and competition in the supply and exchange of money," *Antitrust Law Journal*, 66 (2), 313-361.

Gans, Joshua S., and Stephen P. King (2003), "The neutrality of interchange fees in payment systems," *Topics in Economic Analysis & Policy*, 3 (1), article 1, available at *www.bepress.com/bejeap/topics/vol3/iss1/art1*.

Green, Jeffrey (2008), "Exclusive bankcard profitability study and annual report 2008," *Card & Payments*, 36-38.

Guthrie, Graeme, and Julian Wright (2007), "Competing payment schemes," *Journal of Industrial Economics*, 55 (1), 37-67.

Jullien, Bruno (2001), "Competing in network industries: Divide and conquer," IDEI (Industrial Economic Institute) and GREMAQ, University of Toulouse, mimeo, July.

Nocera, Joseph (1994), *A Piece of the Action: How the Middle Class Joined the Money Class*, New York, NY: Simon & Schuster.

Prager, Robin A., 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 Finance and Economics Discussion Series*, 2009-23.

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

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

Rochet, Jean-Charles, and Jean Tirole (2007), "Must-take cards and the tourist test," De Nederlandsche Bank Working Paper No. 127.

Rochet, Jean-Charles, and Jean Tirole (2006a), "Externalities and regulation in card payment systems," *Review of Network Economics*, 5 (1), 1-14.

Rochet, Jean-Charles, and Jean Tirole (2006b), "Two-sided markets: A progress report," *RAND Journal of Economics*, 37 (3), 645-667.

Rochet, Jean-Charles, and Jean Tirole (2003), "Platform competition in two-sided markets," *Journal of the European Economic Association*, 1 (4), 990-1029.

Rochet, Jean-Charles, and Jean Tirole (2002), "Cooperation among competitors: Some economics of payment card associations," *RAND Journal of Economics*, 33 (4), 549-570.

Rysman, Marc (2009), "The economics of two-sided markets," *Journal of Economic Perspectives*, 23 (3), 125-143.

Rysman, Marc (2007), "An empirical analysis of payment card usage," *Journal of Industrial Organization*, 55 (1), 1-36.

Schmalensee, Richard (2002), "Payment systems and interchange fees," *Journal of Industrial Economics*, 50 (2), 103-122.

Scholtes, Susanne (2009), "Record credit card losses force banks into action," *Financial Times*, July 1.

Schwartz, Marius, and Daniel R. Vincent (2006), "The no surcharge rule and card user rebates: Vertical control by a payment network," *Review of Network Economics*, 5 (1), 72-102.

Weyl, E. Glen (2009), "A price theory of multi-sided platforms," *American Economic Review*, forthcoming.

Wright, Julian (2004), "The determinants of optimal interchange fees in payment systems," *Journal of Industrial Economics*, 52 (1), 1-26.

Wright, Julian (2003), "Optimal card payment systems," *European Economic Review*, 47 (4), 587-612.

Externalities in Payment Card Networks: Theory and Evidence
Commentary

Dennis W. Carlton

I. Introduction

Antitrust and regulatory issues associated with payment systems continue to occupy legal and regulatory authorities not only in the United States but throughout the world. I comment on some of those issues and expand on some of the themes that Bob raised in his excellent paper (Chakravorti, 2009). Bob's paper provides a clear analysis of the many complicated economic forces at work in payment systems and explains why these sometimes complicated models often cannot give definitive answers to some policy questions. The complexity in modeling payment systems arises in large part because such systems represent two-sided markets. Moreover, the fact that collective action is needed in designing and operating so-called "four-party" payment systems raises the spectre of antitrust harm to the public. I will explain in somewhat simplified terms how the two-sided nature of the industry affects the analysis and why the concept has not always been applied correctly. I will then turn to the thorny issues of surcharge prohibitions and interchange fees.

II. Two-Sidedness

What does two-sidedness mean in a payment system? One simple answer—and I will be more precise in a moment—is that for a payment system to work, merchants require that customers carry the payment card and customers require that merchants accept it. There are two types of relevant externalities that can arise in this situation: the adoption externality and the usage externality.

The adoption externality, sometimes referred to as the "chicken and egg problem," might occur when there are initial setup costs to get one side or the other to participate in the system. But these circumstances arise in many situations throughout the economy other than payment cards. For example, before a consumer will buy a car, he wants to make sure that there are gas stations located conveniently.

When a gas station is built, the gas station provides a benefit to all car manufacturers. Should car manufacturers subsidize gas stations? Should gas stations have the right to negotiate collectively the subsidy level with individual car manufacturers? On the other hand, when a car manufacturer sells a car, this benefits gas station owners. Should a gasoline tax be levied on gas purchases in order to subsidize car sales? Should the car manufacturers be allowed to negotiate collectively with individual gas stations on the size of the tax? The adoption externality logic, which might appear to support such arguments for either taxes or subsidies, is similar to some of the arguments sometimes used to justify interchange fees in payment systems. The fact that one does not often see such schemes, even in markets with "network effects," as pointed out by Liebowitz and Margolis (1994), suggests that the magnitude of this problem is not substantial in most markets. This point may be clearest once the markets have reached some critical size. That is, once markets have developed, there may be no need for ongoing payments from one side of the market to the other and, in the example involving cars, the payment from consumers to the gas stations is sufficient to achieve efficiency.[1] For example, I understand that debit cards in Canada have had no interchange fees since their introduction yet are widely used by consumers and widely accepted by merchants, so one should be skeptical of arguments that interchange fees are now needed there to overcome an adoption externality.

The second type of externality often associated with payment systems is the usage externality. The seminal paper by Baxter (1983) explained this effect. Imagine that credit card customers impose a lower cost on merchants than do cash customers. In such a setting, the merchant would like to charge the customer a lower price if he uses a credit card. But suppose that, for some reason, he cannot—maybe it is too hard (costly) to have two different prices depending on the method of payment or maybe there are some legal restrictions against doing so. In that situation, as Baxter cleverly explains, if there is an interchange fee and competition elsewhere prevails, the money from the interchange fee will be rebated by the credit card company to the credit card customer, thereby lowering the effective price that the credit card customer pays. This allows the merchant to achieve his objective of charging two different (effective) prices—one to the consumer who pays with cash and a lower one to the customer who pays with a credit card. Notice that in Baxter's setup, it is the cash customer who pays the higher effective price than the credit card customer and that the cash price is higher than the price that would otherwise be charged if the merchant could charge only one blended price (which would be determined by the merchant's average costs including the interchange fee).

Why are payment systems a two-sided market? As Rochet and Tirole (2006) point out, a market is "two-sided" when it "matters"—i.e., has real economic effects—how the payments among the parties are structured. To make an analogy to tax incidence, economists know that it does not matter in standard models whether the mechanism to collect a tax works by placing on merchants a $1 tax per unit on some items or by placing the tax on the customers. In either case, the final effective price received by sellers and paid by buyers is identical. In a

two-sided market, this is not true, and it matters which side pays the tax. Imagine, for example, that it is costless for merchants to collect and pay the tax but onerous for consumers to do so (for example, they might forget and incur penalties, they may not have an envelope to send in the payment, etc.). Then whether the tax is placed on merchants or customers will have different economic effects.

In Baxter's case, payment markets are two-sided because he assumes that there can be only one merchant price for cash and credit customers, so the interchange fee matters. In the absence of this assumption, the interchange fee would be redundant and have no real effects given his other assumptions—i.e., the interchange fee would be "neutral."[2] In practice, there are several possible reasons for a lack of neutrality including, importantly, the very rules that Visa and MasterCard have promulgated that prevent or inhibit merchants from charging different prices depending on the method of payment and that restrict the ability of merchants to encourage or "steer" customers to use particular methods of payment.

There are several observations that follow from our discussion of two-sidedness. First, any rules preventing the merchant from charging two different prices to consumers may create a two-sided market where one might not otherwise exist. The consequence of having a two-sided payment system where the interchange fee matters is that there are third-party effects. Specifically, there are third-party effects because as the interchange fee is raised, the merchant price to all customers, cash and credit alike, rises as merchants raise prices to cover their increased costs from the increased interchange fee. Any rebate or reward goes only to credit customers. I have always found it odd that the harmful effect of the interchange fee on cash customers did not receive more attention because cash customers often are poorer than credit customers. (In cases where there are a variety of interchange fees, the consumers whose payment cards have the lowest interchange fees are analogous to cash customers in that they may be harmed as interchange fees associated with other customers rise.)

Second, the rationale to justify rules against surcharging and steering has little, if anything, to do with Baxter's seminal insights. In Baxter's framework, merchants want to charge credit customers lower, not higher, prices so there is no need for credit card companies to prevent merchants from being able to charge two different prices because doing so would benefit, not harm, credit card customers. Hence, in Baxter's setup, merchants want customers to use credit cards so payment systems have no reason to promulgate rules preventing surcharging or prohibiting merchants from steering.

Third, it is possible that competition may not work very well among different card systems in benefiting all consumers, both cash and credit card users.[3] The card systems compete to obtain issuing banks and card customers by increasing interchange fees. This allows issuing banks to obtain more revenue, some of which is used to increase rewards, but also raises overall merchants' costs, resulting in a higher effective price to cash customers. The interchange fees are only partly

returned to credit customers and otherwise retained by card payment networks or issuing banks to fund marketing expenses and generate profits. If competition through interchange fees does not improve overall consumer welfare, then there is the issue as to whether the collective action required to set interchange fees in four-party systems raises antitrust issues in countries where interchange fees are not regulated.

Finally, where merchants are prevented from conveying to consumers the price signals reflecting the merchant's cost for the different payment mechanisms, there is the likelihood that an inefficient payment mechanism will be chosen by consumers. If it is inexpensive for merchants to deal with cash customers or debit card customers, then customers may get the wrong signals about the appropriate payment system to use if surcharging of credit cards is not allowed.

III. The Consequences of Surcharging

What are the consequences if surcharging were allowed? This is a relevant issue because in addition to antitrust and regulatory actions challenging interchange fees, rules prohibiting surcharging have come under attack from antitrust and regulatory authorities around the world and, as a result, have been abolished in some countries.[4] Let me describe some of the consequences.

First, even if surcharging does not occur when allowed, the threat of surcharging can constrain interchange fees. If a payment system knows that an increase in its interchange fee could trigger an increased incidence of surcharging of its payment card, then the payment system may be constrained in its setting of the interchange fee.

Second, there have also been proceedings related to the "honor all cards" rule in which merchants are required to accept all payments cards belonging to the same brand (such as Visa) but having different interchange fees or payment terms (e.g., debit cards, "regular" credit cards, premium credit cards) if the merchant accepts any one card in the brand. With the ability to surcharge, the merchant is protected from being forced to engage in what he deems an uneconomic transaction because he can charge the customer according to the payment card used. Visa and MasterCard have pointed out that such an ability could lead to opportunistic surcharging in which the "best" customers are surcharged. To the extent that such concerns are valid, they could be handled by limiting the amount of the surcharge.[5]

Third, the possibility of surcharging will generally reduce the harm that interchange fees impose on cash customers. The salience of a surcharge also might make consumers more sensitive to the cost of using payment cards and might dissuade their use of the most expensive cards. Usage externalities are completely internalized when the merchant induces the merchant's customers to consider the costs to the merchant of the particular payment system the customer uses.

Fourth, as a practical matter, the ability to surcharge provides some protection to cash customers and therefore should mitigate concerns that interchange fees are harming cash customers. The ability to surcharge does not necessarily eliminate all concerns about interchange fees, because there still is an antitrust issue about whether the collective action to set interchange fees benefits the public even if the extent of any harm from interchange fees is reduced through elimination of the prohibition on surcharging.

Finally, and probably most importantly from the perspective of card networks, the use of surcharging could undo the benefits to the card payment system of interchange fees. As that by itself is such a hotly debated topic, let me turn to it in some detail.

IV. Interchange Fees

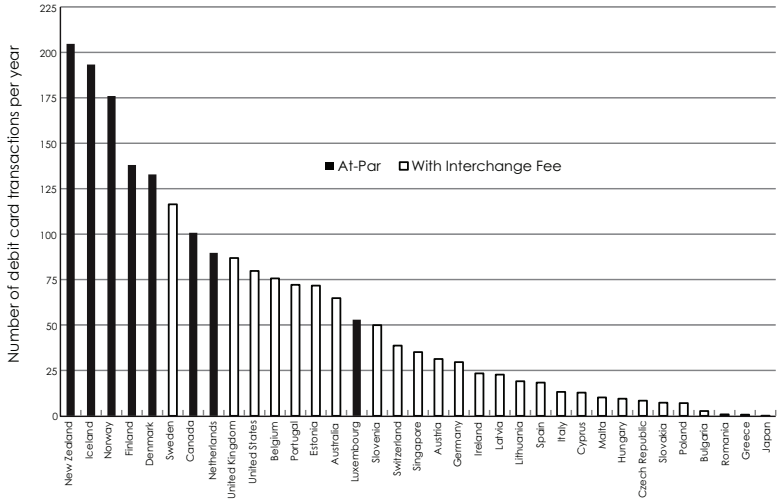
If interchange fees rise, there are several predictable consequences on which there is (or should be) agreement and others on which there is some disagreement. On the agreement side, if interchange fees rise, then in a two-sided market, the cost to the merchant rises and the price that the merchant posts will typically rise. This price increase harms cash customers (and those who use cards with few or no rewards). It may help some card users who may see their rewards rise by more than the interchange fee has increased the merchant price. There likely will be more profit for the issuing bank and more incentive for the issuing bank to spend money on marketing cards to customers.

On the (possible) disagreement side, if interchange fees rise, there will be an incentive for card issuers to compete in order to attract card holders. This competition is, according to some, socially desirable because it creates a benefit to card holders who obtain a sweetened offer from a card issuer. To the extent that this induces more card use, card use could reduce merchant costs. (This is the usage externality discussed earlier in relation to Baxter, 1983.) Furthermore, any constraints on the ability to charge interchange fees could put Visa and MasterCard at a significant disadvantage relative to proprietary systems such as American Express and Discover (who have no interchange fee when they don't rely on outside issuers), thereby harming competition. Let me now evaluate these arguments.

The procompetitive justification for interchange fees is possible theoretically but need not necessarily occur in practice primarily because of the presence of cash customers (or others) whose prices might rise. This means that it is an empirical question whether interchange fees as actually used are helpful or harmful overall to consumers. We do observe that interchange fees exist in payment systems that are much smaller than either Visa or MasterCard, suggesting that such fees can serve some purposes not associated with anticompetitive behavior.[6]

Chart 1 lists the top countries in terms of debit card usage per capita. It turns out that in seven of the eight countries with the highest debit card usage per capita there is no interchange fee, casting empirical doubt on the proposition

Chart 1
Annual Per Capita Debit Card Usage, 2006



Sources: Bank of International Settlements, European Central Bank, Australian Bureau of Statistics and Reserve Bank of Australia, Statistics Iceland, Statistics New Zealand and Reserve Bank of New Zealand, Norges Bank

that interchange fees are necessary to stimulate usage through promotional activity and cross subsidy from the merchant side of the market to the consumer side.[7] Moreover, if you look at the payment system of checks in the United States, it is a system of par clearing (no interchange fee) and, as Frankel (1998) has explained, that par clearing system worked well to reduce the effects of market power in the check payment system.

Finally, as regards the relative harm a restriction on interchange fees imposes on Visa and MasterCard, we now have several empirical experiments where we can see what has happened as a result of regulatory actions that lowered the interchange fee. Australia is the best example. There, the reduction of interchange fees on Visa and MasterCard transactions, together with the elimination of the prohibition on surcharging, forced American Express to lower its merchant fee.[8] After a small initial increase in relative purchase volume by American Express and Diners Club (the proprietary payment systems), the share of purchase volume made on these proprietary systems has now shifted back, so that the relative charge volume of Visa and MasterCard compared to American Express and Diners Club is virtually unchanged from the year prior to the Australian intervention.[9] In no way could one characterize the experience in Australia as confirming the prediction of a "death spiral" that MasterCard and Visa claimed would occur as a result of the lowering of interchange fees.[10]

V. Conclusions

There are two conclusions that everyone involved in these hotly debated issues should be able to agree upon. First, one should be wary of relying on complicated economic models with ambiguous results to justify certain policies. Using such models to justify any particular policy intervention or payment system business practice is fraught with danger because the models often depend in fragile ways on particular assumptions that may be hard to verify. That is why I am skeptical of the theoretical justifications for rules preventing surcharging. But that is why I am also skeptical of arguments that say interchange fees can never be useful to promote competition. Second, in light of the theoretical ambiguity of the consequences of certain practices, one should pay close attention to the empirical evidence, especially that arising from the regulatory interventions into payment systems that are occurring around the world. Only by examining the empirical evidence will we be able to sort out which theoretical models and arguments make reliable predictions. Such empirical evidence should guide our evaluation of the practices of payment systems that are under scrutiny worldwide.

Author's Note: I wish to thank Alan Frankel, Kevin Murphy, Gregory Pelnar, Allan Shampine, and Robert Topel for useful discussions. The views in this paper are mine alone. I have consulted on numerous matters through Compass Lexecon in which I have been adverse to MasterCard and Visa.

Endnotes

[1]This theory is quite similar to Stigler's discussion of the cycles of vertical integration in Stigler (1951). See also Carlton and Frankel (2005).

[2]There is a literature on the neutrality of interchange fees or the lack thereof. See, e.g., Carlton and Frankel (1995) and Gans and King (2003).

[3]See Farrell (2006), Frankel (1998), and Frankel and Shampine (2006).

[4]One sometimes hears the argument that even where surcharging is prohibited, it can still effectively occur as long as it is possible to give a discount for cash. This argument is wrong. A cash discount alone does not allow a merchant to surcharge different payment cards differently depending on their interchange fee. Moreover, if the argument were correct, then presumably neither Visa nor MasterCard would object to dropping the no-surcharge rule in those places where cash discounts are now allowed. I do not understand that to be the position of either Visa or MasterCard.

[5]Another way of viewing payment systems is that they identify buyers with certain desirable buying traits (and influence those buying traits by making payments easier). In this view, Visa, say, approaches each merchant on behalf of a group of specific buyers and asks the merchant for payment for the delivery of these buyers to the store. (In the absence of the merchant agreeing, the buyers may still purchase from the merchant but presumably not to the same degree as if the buyers were using the Visa payment system.) Visa could also engage in some promotional activity to induce buyers to frequent certain stores. In this view, Visa (or its issuers) is getting paid for creating a group of buyers and acting as the bargaining agent for buyers through the interchange fee, some of which it might share with the buyers it represents. Once a bargain is struck between Visa and a merchant, Visa would not want to allow a merchant to undo the bargain by surcharging. The surcharging should then be viewed as a way to breach a contract, but of course, there would be no incentive for the merchant to breach a contract if it was initially in his interest to sign it and he wants it to continue. The interchange fee is then much like a group discount and could raise antitrust issues if Visa represents a large fraction of buyers.

[6]Of course, in the presence of prohibitions on surcharging, issuers favor interchange fees because it increases their revenues. The relevant question is whether there are examples of small payment systems with interchange fees in the absence of prohibitions on surcharging. For purposes of the discussion in the text, I assume that there are such examples.

[7]Countries that reportedly operate debit card systems successfully without interchange fees include Canada, Denmark, Finland, Iceland, Luxembourg, Netherlands, New Zealand, and Norway. In a European Commission investigation, MasterCard claimed that some of the European networks in this list did, in fact, have the economic equivalent of an interchange fee. The Commission reviewed and rejected MasterCard's claim. Commission Decision of 19 December 2007 relating to a proceeding under Article 81 of the EC Treaty and Article 53 of the EEA Agreement (COMP/34.579 MasterCard COMP/36.518 EuroCommerce

and COMP/38.580 Commercial Cards) (Provisional Non-Confidential Version, pp. 555-608).

[8] *Reserve Bank of Australia Bulletin*, Statistical Series C3, Merchant Fees for Credit and Charge Cards, *http://www.rba.gov.au/statistics/bulletin/xls/c03hist.xls.*

[9] *Reserve Bank of Australia Bulletin*, Statistical Series C2, Market Shares of Credit and Charge Card Schemes, *http://www.rba.gov.au/statistics/bulletin/xls/c02hist.xls.*

[10] MasterCard International Incorporated, Submission to Reserve Bank of Australia, June 8, 2001 (as revised, July 20, 2001), pp. 11-12; Visa International Service Association (Prepared by: Network Economics Consulting Group Pty Limited), "Response to the Reserve Bank of Australia's Consultation Document and Report of Professor Michael Katz," March 2002, p. 10. The Australian experience is sometimes used to argue that prices to cash customers did not fall as a result of the reduction in interchange fees, hence the reduction in interchange fees failed to accomplish one of its purposes. I leave a detailed discussion of the Australian experience to another time. I simply point out that most economic models would predict some reduction in cash price in response to the decline in interchange fees and that given the magnitudes involved, identifying a decline in cash prices might be hard to do statistically. But as I explain next, continuing empirical evaluation of interventions such as Australia's are exactly what is needed to resolve some of the concerns associated with payment systems.

References

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

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

Carlton, Dennis W., and Alan S. Frankel. 2005. "Transaction Costs, Externalities, and 'Two-Sided' Payment Markets," 2005 *Columbia Business Law Review* 617.

Chakravorti, Sujit. 2009. "Externalities in Payment Card Networks: Theory and Evidence," prepared for The Changing Retail Payments Landscape: What Role for Central Banks? Conference Sponsored by the Federal Reserve Bank of Kansas City (November 9-10).

Farrell, Joseph. 2006. "Efficiency and Competition Between Payment Instruments," 5 *Review of Network Economics* 26.

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

Frankel, Alan S., and Allan L. Shampine. 2006. "The Economic Effects of Interchange Fees," 73 *Antitrust Law Journal* 627.

Gans, Joshua S., and Stephen P. King. 2003. "The Neutrality of Interchange Fees in Payment Systems," 3 *Topics in Economic Analysis and Policy*.

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

Rochet, Jean-Charles, and Jean Tirole. 2006. "Two-Sided Markets: A Progress Report," 37 *Rand Journal of Economics* 645.

Stigler, George J. 1951. "The Division of Labor Is Limited by the Extent of the Market," 59 *Journal of Political Economy* 185.

General Discussion
Session 3

Mr. Weiner: Thank you, Bob, and thank you, Dennis. I think you have both done a masterful job of summarizing what is a very complex and technical literature. Obviously, a lot of important issues have been raised, a lot of controversial issues, and, all kidding aside, we want to hear all views. Some central banks are actively looking at these markets and have put in place special policies. Others are analyzing them. It is not just central banks, of course. It is also competition authorities and so on. So, this is a very, very important area to be thinking about. And, again, you have done a wonderful job.

First, Bob, is there anything you want to react to?

Mr. Chakravorti: Thanks, Dennis, for those comments. They were great. Let's open it up.

Mr. Levitin: In the United States, at least, the reason we have four-party networks is really a historical matter. We had interstate branch banking restrictions at the time the networks were created, and that's why four-party networks were needed if we were going to have depository institutions involved in them.

Looking ahead, do either of you see an economic case for having four-party networks instead of three-party networks? Do you see benefits to one arrangement or the other? Or, is it like a lot of the questions, just indeterminate in the abstract?

Mr. Chakravorti: If you look at the evolution of the credit card market in the United States, for example, it was Bank of America that started issuing general-purpose credit cards. They realized in order to expand, partly due to branching restrictions, they had to partner with other financial institutions. But, in a global economy, I don't see four-party networks losing their place. That is not to say there shouldn't be three-party networks that co-exist and compete with them. I see a role for four-party networks going forward.

Mr. Carlton: My answer is similar and it has to do with the fact we are not starting from scratch. It would be a different question if you say, "I am going to allow three-party networks and four-party networks. Who is going to win in light of everybody starting at a market share of zero?"

In many countries, there already is very large penetration, say, of Visa and MasterCard and that means—even if you have three-party systems—they have a tough fight on their hands because Visa and MasterCard have an established advantage. So that's why, if you look in, for example, Australia, there has hardly been any significant movement in market shares, even though Visa and MasterCard could claim that the recent regulation of interchange harms them relative to three-party systems, which do not have interchange fees. That disadvantage has not materialized in significant drops in market share. My view is—and the evidence so far is consistent with it—is that Visa and MasterCard will remain important payment systems.

Mr. Bennett: My question is on surcharging. Both Bob and Dennis seem to be pointing toward surcharging being a possible solution to many different issues here.

One of the interesting things we looked at in the OFT was, when we see surcharging, and we are increasingly seeing surcharging in the market, we are very seldom seeing it as being proportional to the actual fees that are charged. Often we are seeing it as significantly higher.

My question is, do the panelists think that is because of a transparency issue and these retail companies are finding another way of extracting greater profits or is there some other type of logical explanation which could account for the fact their surcharging doesn't seem to be proportioned to their costs?

Mr. Chakravorti: There are clear examples of higher surcharges than the retailer's cost to accept payment cards. I agree with Dennis. If you want to rule that out, you could regulate it, but that gets tricky. The U.S. gas station that I mentioned before has been surcharging for years, and they earned revenue from this practice. Are the people who are paying the surcharge harmed? If you compare it with them going to a cash machine and paying the potential surcharge on the cash withdrawal, perhaps not. One has to compare apples to apples. In our Spanish study, we looked at rival ATM density, because those are the ones that could have surcharges on them. So you have to consider all sorts of factors when determining how big or small a payment card surcharge is.

There is also price competition to some extent. The gas station that I mentioned offers among the lowest gas prices in the area. There is a relatively low gas price coupled with debit card surcharges. How do you separate those two effects? It is not clear to me.

If you have really intense competition, you wouldn't expect the surcharge to be higher than the cost to accept payment cards. It depends on the structure of competition to some extent.

Mr. Carlton: To my thinking, I separate things into two parts. The first is the opportunity to charge two different prices, the opportunity to surcharge. Just having that opportunity provides constraints, and my preference is to rely on the market to figure out what it wants to do in terms of whether to levy the surcharge or not.

Whether I then want to go further and either regulate the surcharge or regulate interchange fees raises all these questions about the difficulty of regulating. You can get it right or you can get it wrong. I would rather not have to regulate anything. If you think there's some market failure, for some reason, that is imposing very large costs, then maybe you want to intervene. But my own preference is always to see if the market would solve the problem first.

I don't have a good answer to your question of why there is surcharging in excess of costs other than to say, if you look at a distribution of who is surcharging and who is not surcharging, my suspicion is some people aren't surcharging because it is just not worth their while. And then the people who are surcharging are people who not only don't want to pay the differential costs, but for some other reason have another justification for wanting to charge a higher price, but I have not studied that.

Mr. Kimmet: Dennis, I don't know if you've followed what happened in New Zealand, but it is my understanding there—concurrent with allowing merchants to surcharge on the issuing-bank side—the issuing banks can now negotiate against the four-party systems, say, the Visas and MasterCards of the world, and have that rate set by them be the cap. As I understand, how that has played out in that market is the smaller merchants have started to surcharge, the issuing banks then have started to negotiate with merchants, and the prices are collapsing pretty rapidly. Do you have any insight into that?

Mr. Carlton: Are you talking about this recent New Zealand case that just settled?

Mr. Kimmet: Yes.

Mr. Carlton: I don't think it has gone into effect yet. It is going into effect in January, so I don't know whether there have been any effects yet.

Mr. Kimmit: You are starting to see some conversations happen very quickly.

Mr. Carlton: My understanding of that settlement is that surcharging is going to be allowed absent other contractual arrangements, that a four-party system will not be allowed to pass a rule that prevents surcharging, but individual negotiations with banks are possible such that the interchange fee relevant to that bank's payment cards can be set at or below the cap with the possibility that the bank and merchant could agree to not allow surcharging. I think that is right.

Mr. Kimmit: That is my understanding as well. I would encourage everybody to understand that system.

Mr. Wildfang: Of interest to the group, I am lead counsel for the merchant plaintiffs in the pending litigation.

I have just a couple of observations. One, with respect to real-world evidence: I know the economists here complained that it is hard to get data and evidence. After looking through some 60 million pages of documents in the litigation, I can tell a lot of the assumptions built into a lot of these economic models are inconsistent with the record. One of these days, hopefully, that record will be available, so people can look at it.

Another observation: The rationale for two-sided markets and the need to balance by charging one higher than the other, if that were valid, that would permit issuing banks to also fix the price of interest rates to cardholders or annual fees to cardholders. The economic justification seems identical and yet it seems to be unlikely the Department of Justice would not crack down on an agreement by all the issuing banks to charge the same interest rates or charge annual fees.

Mr. Carlton: I want to make a comment on that. There is an article—I think it's by Liebowitz and Margolis in the 1990s—which talks about the economic literature on network externalities. It makes the point that it is easy to go off the deep end in reading this literature and suggest every single market needs intervention and therefore should get an exemption from some collective-action problem. It makes a very similar point to the one you are making.

Mr. Ruttenberg: In this debate on cards, surcharging, interchange, and all these kinds of things, nobody has explicitly made the point that in the end we are talking about doing payments. If we make a comparison between cash usage versus cards, in the end you would like to replace cash in favor of using cards.

All the debate on interchange—whether or not it is allowed and how high it should be and so on—the point in the end we have all this debate in Europe over the future of the cards business in Europe and yet it is all about the need to have a plastic alternative for cash. If we take this approach, ask yourself the question, why should a merchant pay for my interest-free periods? Why should a merchant pay for the insurance? Why should the merchant pay for the miles I get? In the end, it is just about getting a payment done.

I am in this debate just a little bit. The core function of using a card is doing payments. This approach is just a down-to-earth approach. Is it not also missing in the United States when we discuss cards? It's about doing the payments and not about free miles and all these kinds of things. I would like your reaction for this.

Mr. Chakravorti: For some of those benefits, such as free float, there are merchants in the United States that give you similar terms on their own credit cards they issue, so it is not just bank-issued credit cards that offer some of these benefits.

We can quibble over whether you should get miles, toasters, and other things. In the United States, banks used to give away toasters because they couldn't give interest on checking accounts. We cannot say whether receiving miles for making a purchase at Starbucks with your payment card improves social welfare or not. And there are probably differences in the room about the benefits of rewards, but some essential functions of credit cards are going to be very difficult to mimic with cash.

The second thing I would add is that it's not abundantly clear cards always dominate cash. You could have situations where individuals value their privacy and prefer to use cash. We certainly have heard of southern European examples of tax evasion. Also, like the gentleman on the plane that I mentioned earlier, wanted to use cash because that was his best way to control spending.

Mr. Cook: Bob, you mentioned earlier about the gas station owner in California who surcharged for debit transactions. If I am not mistaken, that is only PIN debit they accept, not signature debit. The two entities that control signature debit don't allow surcharging on those networks. In fact, the network that's owned by one of large schemes only allowed them to surcharge on their network because they were grandfathered in at the time. Is that not correct?

Mr. Chakravorti: There are examples of where the gas station owner actually accepts a credit card that belongs to that chain. So, there are examples where you are right. Although they have to pay a higher fee for credit card transactions, they are not allowed to surcharge for them.

Mr. Cook: But, in the debit card example that you gave, where we're mentioning to the European folks that surcharging does exist in the United States, it is really only in that one example, only in PIN debit, and it's only because of a grandfathered situation.

Mr. Chakravorti: It is true that these are PIN debit transactions because there are network rules that don't allow you to surcharge other types of payment cards.

Mr. Cook: The gentleman from the UK mentioned surcharging may not be a direct correlation to the cost. Would you all like to talk about whether or not you believe interchange is a direct correlation of the financial institutions' cost?

Mr. Chakravorti: The one thing that's clear, I think, is you can't just separate costs and benefits. There is a lot of debate on the cost-based approach to setting fees. Some people argue fees should be purely cost-based. Then the problem arises of trying to figure out what does cost-based mean. So, if a merchant gives out an interest-free loan and then chooses not to give that loan but chooses to accept a four-party credit card, why shouldn't he share in the cost of providing an interest fee loan to move merchandise? There are several arguments that can be made against that—maybe the customer has money in his account. But separating these different types of customers is difficult and trying to figure out the cost is also very difficult in these cases.

Mr. Carlton: In a two-sided market, I think there is agreement an interchange fee by itself isn't necessarily cost-based and it doesn't have to be to create incentives to get the other side on board. That is why regulation of interchange fees is hard if you intend to have a cost-based regulatory system. You have to decide what costs, whose costs.

Mr. Hayes: Dennis, you expressed some skepticism around, would lowering interchange fees curtail efforts to issue new cards into the marketplace? Then, later on in one of your responses, you talk about we're not starting with a clean slate. There are existing competitors here today. As I look at the U.S. market, there are a number of efforts to create new payment mechanisms that do have a much lower interchange rate structure—like Debitman then became Tempo, like Revolution Money, like PayByTouch—all of these were based around being a low interchange cost structure and all of them basically fell by the wayside because they couldn't get cards into the marketplace. They couldn't get the issuing piece figured out.

There does seem to be empirical evidence that a lower revenue proposition of card issuing is insufficient to get consumers to adopt that payment mechanism. I would be curious to get your thoughts—at least in the United States where the pricing mechanism is the way it is today—if there is room for reducing interchange and still growing the card business.

Mr. Carlton: I think you would have to look at when they failed. In other words, if it was at a time (and it must have been) when there still were no-surcharge rules, that's different than if we didn't have no-surcharge rules. So, that's one relevant question.

In the presence of the ability to surcharge, the question is does that so undermine the ability to collect interchange fees, as sometimes Visa and MasterCard allege that it would impair the diffusion of card payment systems. All I'm saying is, based on some of the evidence I gave you, I can see some positive support for the proposition that interchange fees help dissemination of cards. On the other hand, there is a lot of negative support for that proposition. Perhaps one of the simplest implications, as I said in my talk, of the position that interchange fees are critical is that the decision related to the Federal Reserve System to have a par clearing system for checks must have been a mistake. Maybe it was and maybe it wasn't. I won't take a position here. I am just pointing out that it is an *implication* of the position that interchange is critical. That sounds a little strong to me. That's all.

It also sounds a little strong that interchange is critical for promotion when you look at some of the European countries that have zero interchange fees but have very high per capita debit card usage. Now, I'm willing to entertain the possibility that interchange could matter because I believe there are examples of even small systems trying to use interchange. All I am saying is it seems like sometimes people overstate the importance of interchange for card dissemination.

Ms. Masi: My question is to Bob. You mentioned earlier the impact of a two-sided market on innovation. I would like to know more because from the network externality theory and also from reality we know for sure there is some influence of the market power in making an obstacle to innovation. So, it is not so clear but this is traditional theory, a conventional competition theory. My point is just to know your opinion on what the impact is for a two-sided market innovation.

The second one is really a question. Which kind of regulator? Regulation is a word, but regulators are different. As Stu mentioned before, we have a competition authority and we have a standardization authority, which is something we didn't mention before, you know that is implied in this market more than in other sectors. Then there are central bankers. So, tell us more.

Mr. Chakravorti: Let me first say maybe it wasn't clear in the talk. The two-sided market literature doesn't say much on innovation. It's not there. I've had discussants of mine tell me that it is an important issue.

An example of where cost-based pricing was used is the market for electricity in California. As a result of this pricing policy, there wasn't sufficient capacity built up resulting in a crisis in that market. So the notion here is, if you don't have incentives such as the ability to earn profits to innovate, you might not.

One payments example that comes to mind is if you look at checks in the United States, they took a long time to be truncated, partly because of the way they are cleared. An act of Congress was needed to facilitate the mass migration to check truncation. But, in credit cards, you didn't quite see that.

In terms of the other question as to who regulates, that really depends on the country, what's involved and whether the central bank has the authority to regulate retail payment services. Purely for a selfish reason, being an economist, I like the Australian way because comments from all participants are online, whereas in court cases, I have no way of accessing those comments. They are sealed from me for many years.

It seems there may be differences, but in the United States traditionally we've gone through the courts on these challenges, whereas in Australia they went through the central bank. In Spain, they went through the antitrust authority.

The U.S. Department of Justice has had an effect on various types of regulation in terms of the ability to issue non-Visa and MasterCard cards by financial institutions that are members of those networks.

Mr. Hunt: I just have two observations. One is this analogy to par value clearance of checks. I think the analogy to the credit card market is really the honor-all-cards rule, rather than an inference about interchange. We can talk more about that later, if you are interested.

Second, this discussion about surcharging is very interesting. But, in the U.S.

context, it is interesting we do allow a cash discount and yet the implicit assumption here is for some reason that freedom of pricing doesn't seem to work. It is very important to understand the friction, whether it's behavioral or legal or whatever, that prevents the cash discount from behaving in the same way allowing a surcharge would behave.

Mr. Carlton: I think you're right. It is somewhat puzzling there aren't more cash discounts. However, there is a difference between a surcharge and a discount. The surcharge can differentiate among the various credit cards that have different merchant service fees, while a discount for cash cannot.

I also think the salience point people have mentioned earlier would matter. The fact of the matter is, for whatever reason, whoever designed those rules that allow discounting but not surcharging, didn't say, "I allow surcharging too."

So, they must have thought it mattered. My suspicion is it does matter.