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

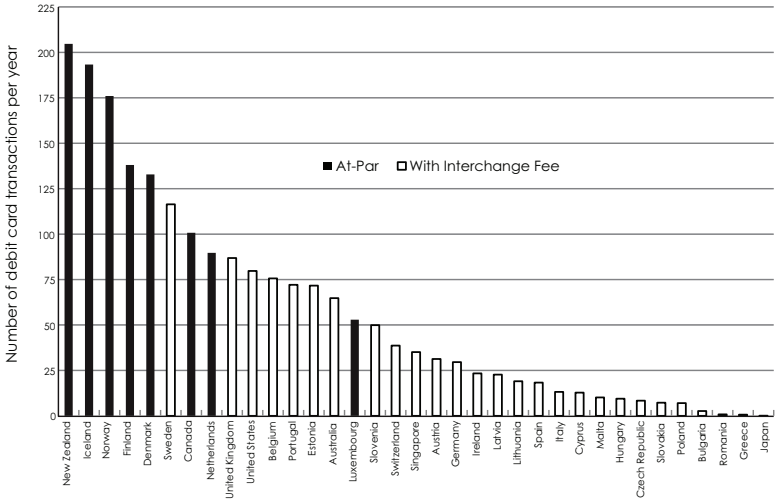
If interchange fees rise, there are several predictable consequences on which there is (or should be) agreement and others on which there is some disagreement. On the agreement side, if interchange fees rise, then in a two-sided market, the cost to the merchant rises and the price that the merchant posts 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.